# Detection of Collusive Networks in E-procurement *

Bruno Baránek† Leon Musolff‡ Vítězslav Titl§

July 16, 2021

**Abstract**

Collusion likely has adverse effects on social welfare. In this paper, we study collusion in the e-procurement market in Ukraine. We document that the bidding patterns in the data are incompatible with a competitive equilibrium. We develop a novel structural test to detect pairs and, thereby, networks of collusive firms. We validate the soundness of our collusion detection algorithm on a sample of 863 prosecuted collusive firms that participated in 23,515 tenders.

**Keywords:** Public procurement, Collusion, Online markets

*We gratefully acknowledge insightful comments from Adriano De Leverano, Sylvain Chassang, Thomas Fujiwara, Benny Geys, Leonardo Giuffrida, Kate Ho, Adam Kapor, Jakub Kastl, Deni Mazrekaj, Jonas Müller-Gastell, Dina Pomeranz and suggestions from participants of seminars at KU Leuven, Princeton, and Utrecht University. We would like to also thank the Weiss Fund for financial support and Bogdan Salagub for invaluable research assistance. Vitezslav would like to thank the Department of Economics at Princeton University for its hospitality and the Czech Science Foundation – GACR (project number 21-15480M „The Role of Institutional Factors and Information in Public Procurement Markets") for financial support.

†Electronic address: `bbaranek@princeton.edu`; Department of Economics, Princeton University

‡Electronic address: `lmusolff@princeton.edu`; Department of Economics, Princeton University

§Electronic address: `v.titl@uu.nl`; Utrecht University School of Economics, Utrecht University; Department of Economics, Deparment of Economics, Faculty of Law, Charles University

# 1 Introduction

Studying collusion in auctions has been a central topic of economics literature (Porter and Zona, 1993). There is widespread agreement among scholars and practitioners that collusion has adverse effects on social welfare. This paper presents a novel collusion detection approach. We identify networks of collusive firms on bidding data from a large-scale e-procurement platform. The ability to detect networks of collusive firms is likely to be crucial for improving the efficiency of the public procurement market.

We study the procurement market in Ukraine, which is a fascinating laboratory for studying collusion for three reasons. First, Ukraine has one of the most modern procurement markets, operating entirely on an electronic platform that allows for unprecedented levels of data access and transparency. Georgia, Moldova, and Kyrgyzstan[1] have implemented similar multi-round e-procurement systems, which means that our detection mechanism is immediately applicable in these two countries. Second, the market has suffered from large-scale issues with collusion. Since 2016, over 1,200 firms were successfully prosecuted for engaging in collusive conduct. Third, the usage of a multi-round auction design allows us to propose a novel framework for the identification of collusion.

The e-procurement platform uses a multi-round sequential auction mechanism. This mechanism was designed with the intention of introducing high levels of transparency of the behaviour of firms and procuring authorities. The idea was that the transparency and three stages would help with the identification of firms involved in corrupt or collusive behaviour. Therefore, we have access to unusually detailed data. We observe each round of the auction and the behavior of each firm. The mechanism proceeds by initially letting interested bidders submit bids simultaneously. Subsequently, all bidders can access an online auction where they compete for the contract. The online auction allows the bidders to lower their bids in three consecutive rounds sequentially. The bidding order is determined by order of bids in the previous round, and the bidder with the highest initial bid starts bidding in

[1]Georgia, Moldova, Kyrgyzstan, and Ukraine represent about one third of the GDP of the post-soviet countries besides Russia (IMF, 2021).

the online auction. This setup can be understood as a mechanism where bidders initially bid for the order in which they will submit a final bid. Naturally, there is an advantage for the last bidder as nobody is given a chance to react to her bid. We solve this mechanism's equilibrium and show that initial bids in this setup are entirely analogous to bids in first-price procurement auctions. Furthermore, bidders should subsequently always update their initial bids by undercutting each other by a small $\epsilon$[2].

We present empirical evidence that, contrary to theoretical predictions, there is low activity during the online auctions in the data. Even though firms might lose significant profits by not updating, 51% of initial bids are never updated – this is highly sub-optimal behavior (unless firms profit through a cartel). For 45% of all tenders, there are no updates. The lack of undercutting is consistent with a cartel set up where members initially agree on allocating projects and thus lack incentives to underbid each other.

Based on the theoretical discussion, we construct a novel structural algorithm to measure collusion and identify bidding rings in public procurement: the network-based ring identification. This procedure is built on the idea of examining every pair of firms and measuring whether there is a collusive relationship between these firms. By identifying all collusive links, we can locate a bidding ring in a very flexible manner. Specifically, we analyze the probability that a given firm underbids a particular competitor. The theoretical model allows us to compute the predicted probabilities of undercutting a competitor. Comparing observed behavior and model predictions, we find a significant share of firm pairs that are very unlikely to compete against each other, i.e., a collusive link.

We verify our detection algorithm on a sample of 863 prosecuted firms that participated in 23,515 tenders (9.4% of the total value of the market). The prosecution data are unique because they allow us to identify who is colluding precisely – in our terminology, the data shows prosecuted collusive links. One of the significant contributions of this paper is that it presents a method that identifies whole bidding cartels and not only whether specific firms collude or whether there is collusion on the entire market.

[2]This is caused due to an introduction of a low probability of not being able to respond to any submitted bid. Our data suggest that such friction is appropriate.

Our paper extends and contributes to the literature studying collusion in auctions. The pioneering papers by Porter and Zona (1993) and Porter and Zona (1999) identify collusive behavior by contrasting bidding behavior in data to the prediction of a competitive equilibrium. Such an approach was further developed and modified in Bajari and Ye (2003), Conley and Decarolis (2016) and Chassang and Ortner (2019). Our paper significantly contributes to the literature by developing a procedure that can identify pairs of colluding firms and, thereby, whole networks of collusive rings. Currently, the literature has focused on market-wide collusion detection (Porter and Zona, 1993) or testing whether a particular firm behaves in a collusive manner (Kawai and Nakabayashi, 2014). By contrast, we allow a firm to be collusive in some settings but not in others.

Like Asker (2010), Kawai and Nakabayashi (2014), and more recently De Leverano (2019) and Clarka et al. (2020), we also provide an explanation of how current cartels operate based on prosecution data. In a closely related study, Kawai and Nakabayashi (2014) present a data anomaly suggesting the presence of collusion and then verify their empirical approach on a set of prosecuted colluders. We have unusually detailed and large prosecution data – giving us information on who exactly colluded with whom. This prosecution data are used to (successfully) verify the identification power of our algorithm for detecting links of collusive networks.

The remainder of the paper is structured as follows. Section 2 describes the Ukrainian procurement market. In Section 3, we solve for the equilibrium of the sequential auction used on this market. Section 4 shows summary statistics and suggestive evidence that the firms' behavior is suspicious and potentially collusive. In Section 5, we propose a novel method that allows us to identify collusive rings and verify our method on a sample of firms successfully prosecuted for collusion in public procurement. Section 6 concludes.

# 2 Description of the market

Ukraine has battled enormous problems with corruption and collusion in public procurement ever since its independence in 1990. After the Euromaidan revolution in 2014, a group of volunteers started the ProZorro platform in February 2015 to tackle these issues. They successfully promoted this fully online tendering system, and the platform has become compulsory for all public entities in 2016.[3] At its core, ProZorro is *(i)* a unified central database of all public procurement projects conducted in Ukraine, and *(ii)* an API for interacting with this database. The data from this platform will serve as the primary dataset for our analysis.

## 2.1 Exact Tender Procedure

Small purchases[4] can be conducted without an online auction.[5] Larger purchases generally have to be completed as open tenders. Our analysis will focus on competitive tenders both below- and above-threshold. There is a critical difference between below- and above-threshold contracts that will be relevant to our study. Below-threshold agreements do not require an auction in the first place; they can be awarded to the sole bidder should only one bidder participate in the sale. By contrast, above-threshold auctions must be repeated or canceled altogether if there is only one auction participant.

The description of the details of the open tendering procedure for the above-threshold contracts follows. The tender begins with the procuring entity uploading documentation for the tender to ProZorro, at which point the period of proposal submission begins and lasts for at least 15 calendar days. Once the end of the proposal submission period is reached, the tender is automatically canceled if only one proposal has been submitted.[6] If there are

[3] The ownership of the system has been transferred to the state of Ukraine.

[4] A purchase is small if it is *(i)* a good/service bought by an ordinary contracting authority worth less than \$7k, *(ii)* a works purchase purchased by a standard contracting authority worth less than \$53k, *(iii)* a good/service bought by a 'special' contracting authority worth bless than \$35k, and *(iv)* a works purchase by a 'special' CA worth less than \$177k (Supreme Council of Ukraine, 2015).

[5] Though the data has to enter ProZorro as a 'report on concluded agreement.'

[6] As explained above, this does not apply to below threshold auctions.

multiple proposals, the system automatically schedules and runs an online auction, which we discuss below. While the auction is run, the bidders do not yet have access to each others' documents. They are made aware of the number of opposing bidders the moment the online auction starts. Importantly, they remain unaware of their identity and specific proposals until the sale ends.

## 2.2 ProZorro Auction

The tendering process's critical element is the online auction, during which bidders compete for the right to complete the contract for the government. First, initial bids are submitted together with technical proposals. Initial bids can thus be understood to be submitted 'simultaneously' because bidders are not aware of each others' bids at this stage. Secondly, bidders enter the online auction. Now they are given a chance to update their bids three times in a unique mechanism[7]:

1. Bidders are ordered in descending order according to their initial bids, and the first updating round begins:
    (a) The bidder with the highest initial bid goes first, observes all initial bids and can update her bid. However, she can only lower her bid.
    (b) After the first bidder moves, the bidder with the second-highest bid observes all bids, i.e., initial bids and the update by the originally highest bidder. This (second highest) bidder then again is given a chance to lower her original bid.
    (c) All the bidders move sequentially in this fashion until the lowest has chosen whether to update her bid.
2. Bidders are ordered based on the size of their updated bids and again sequentially move.

[7] While the mechanism does feel relatively unique, note that Georgia and Moldova use the same multi-round mechanism. Indications are that other countries will follow.

3. Finally, there is a third round of bidding in which bidders are ordered based on their bids from the second round. The bidder with the lowest bid then wins and becomes the vendor of this project.

This mechanism essentially emulates a sequential Bertrand game. As the last-mover is advantaged, bidders are incentivized to submit low initial bids. However, as bids can only be lowered in the updating rounds, there is no incentive to submit arbitrarily low bids in, for example, the initial round. We analyze this auction below and find that it is surprisingly similar to a mix between a first- and a second-price auction, the latter having no impact on equilibrium.

The reader should note that while the initial bid submission period is typically lengthy, the online auction proceeds in about 15 minutes (the exact length depends on the number of participants).

# 3 Model and Equilibrium

We now discuss the intuition behind the equilibrium of the ProZorro Auction, with details and proof relegated to the appendix. To this purpose, consider a simplified version of the auction in which there are only two players and just one updating round. The timing is as follows:

1. Bidders submit their initial bids simultaneously.
2. The initial 'loser' (the agent that submitted the higher bid) is given a chance to update his bid.
3. The initial 'winner' is given a chance to update her bid.

We note that bidders can only update their bids downwards, i.e., initial bids are not just cheap talk.

As the timing clarifies, the equilibrium will hinge on the amount of information revealed in the initial stage of the auction. Therefore, we restrict attention to equilibria in which the initial bid is perfectly revealing (i.e., agents are not randomizing). In such equilibria, agents are fully informed about each other's cost types after the initial stage. This, in turn, generates a potential multiplicity issue: the initial loser may realize that he will lose the overall auction no matter what he bids. We introduce a small probability that any given bid update is not successfully submitted to resolve this multiplicity. This probability captures the natural fact that if you know you will lose if your rival reacts, you may as well bid in such a way as to maximize your surplus if, for some reason, your opponent fails to respond.

Taken together, our assumptions imply that when submitting the last bid in the auction, the initial winner will beat the current standing bid by the minimum amount necessary if doing so is feasible. Before this, the initial loser will predict what sort of bids the initial winner can beat: if there are bids that she cannot beat and give the initial loser a positive surplus, he will make the largest bid satisfying these criteria. If there are none, he will simply update his bid to the current winning bid (hoping for the slight chance that his rival fails to update). More rigorously, if $b(\cdot)$ is the equilibrium bidding function in the initial round, the payoff to type $c_1$ from pretending to be type $\tilde{c}$ in this round is given by

$$
\begin{aligned}
V(\tilde{c}) \quad = \quad & \mathbb{P}\Big(b(\tilde{c}) < b(c_2) \cap c_1 < c_2\Big)\Big(b(\tilde{c}) - c_1\Big) + \\
& \mathbb{P}\Big(b(\tilde{c}) < b(c_2) \cap c_1 > c_2\Big)\Big(b(\tilde{c}) - c_1\Big) + \\
& \mathbb{P}\Big(b(\tilde{c}) > b(c_2) \cap c_1 < c_2\Big)\mathbb{E}[\min\{c_2, b(\tilde{c})\} - c_1 | c_1 < c_2, b(\tilde{c}) > b(c_2)] + \\
& \mathbb{P}\Big(b(\tilde{c}) > b(c_2) \cap c_1 > c_2\Big) \times 0
\end{aligned}
$$

The four lines of this expression correspond to the four cases that could transpire: the agent could pretend to be strong and be strong (first line), he could pretend to be strong and be weak (second line), he could pretend to be weak and be strong (third line) or he could pretend to be weak and be weak (fourth line). The first two lines combine the payoffs from a first-price auction in which each bidder bids according to $b(\cdot)$. The last two lines can be related to the expected payoff from a second price auction so that we can write the overall

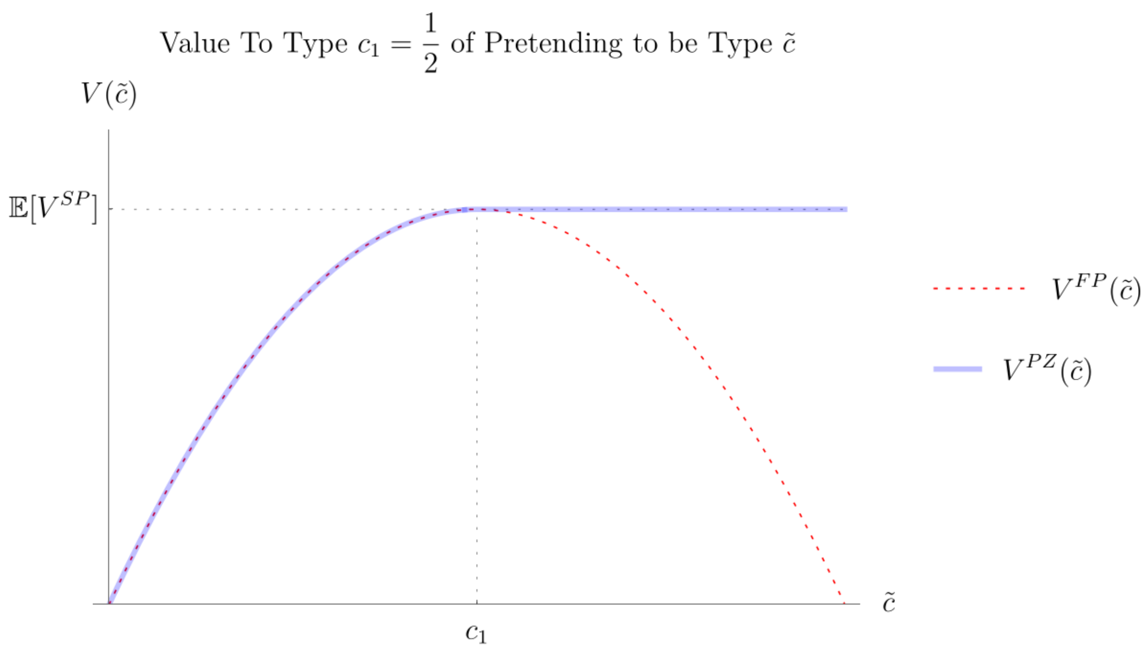



Figure 1: The Equilibrium of the PZA Auction
*Notes*: The expected utility to an agent of a given type from participating in the ProZorro auction reaches its peak at the same time as that of participating in a FP auction, but the second-price component ensures it never drops from this level.

payoff as

$$V(\tilde{c}) = V^{FP}(\tilde{c}) + \mathbb{P}(b(\tilde{c}) > b(c_2))\mathbb{E}[V^{SP}|\underline{c}_{-1} < \tilde{c}]$$

where we use $\underline{c}_{-1} := \min_{j \neq 1} c_j$ as more general notation to emphasize that this way of expressing the payoffs does not depend on the fact that there are exactly two players playing the game; indeed we have the following result:

**Proposition 1.** *In any equilibrium in which initial bids are given by some strictly increasing $b(\cdot)$, the expected payoff from pretending to be type $\tilde{c}$ is given by $V(\tilde{c})$ no matter the number of updating rounds or number of players.*

We illustrate this function in Figure 1 for the case of $c_i \sim U[0, 1]$. Naturally, if $\tilde{c} < c_1$, then the expected value from a second price auction conditional on $c_2 < \tilde{c} < c_1$ is going to be zero, and hence $V^{PZA}(\tilde{c}) = V^{FP}(\tilde{c})$ to the left of $\tilde{c} = c_1$. Furthermore, $V^{FP}(1) = 0$. Thus, $V^{PZA}(1) = \mathbb{E}[V^{SP}]$. But we know that the expected rent that bidders earn in a second-price auction is exactly equal to the expected rent they earn in a first-price auction when pretending to be their true type. Thus, $V^{PZA}(1) = V^{PZA}(c_1)$. It turns out that the effects of decreasing rent from the first-price component of the auction and increasing rent from the second-price component of the auction exactly cancel and hence $V^{PZA}(\cdot)$ is flat to the right of $\tilde{c}$. A more formal version of this heuristic argument in the appendix allows us to conclude:

**Proposition 2.** *The ProZorro auction (with $k \geq 1$ rounds and $n \geq 1$ players) has a unique PBE in which initial bids are given by a strictly increasing $b(\cdot)$. In this equilibrium,*

$$b(c) = \frac{1}{[1 - F(c)]^{n-1}} \int_{c}^{c_{max}} s(n-1)f(s)[1 - F(s)]^{n-2} ds$$

*and bids are decreased by the minimum bid decrement whenever doing so is possible without bidding below one's own cost.*

Thus, we conclude that the initial bids in the ProZorro auction must come from exactly the same bidding function that they would come from in a first-price auction.

# 4 Data and summary statistics

We use data from three data sources. First, we use publicly available procurement data from ProZorro. This data contains detailed information about the final price, industry code, delivering firm, procuring authority, and the reserve price for each contract procured by any public entity in Ukraine. Second, we scrape the auction platform employed by ProZorro to complement these covariates with detailed bidding data, including bids and bidders' identities at all stages of each online auction. Third, we obtain data about firms prosecuted for collusive conduct in public procurement from the Ukrainian Anti-Monopoly Agency.

## 4.1 Key features of auction data

### There is low competition in the market

The minimum number of participants for all the above-threshold contracts is two. In the actual data, we do not see much higher levels of competition. The average number of participants is $2.54$ bidders, with the median number of participants of only $2$ bidders. The low number of participants cannot be explained by low competition for small contracts.[8] Indeed, the competition is comparable for big contracts above $25,000,000$ UAH (about $1,000,000$ USD) with an average of $2.59$ bidders and an (unchanged) median of $2$.

### Not updating leaves money on the table

$51\%$ of all submitted initial bids are never updated, leading to no competition during the online auction for $45\%$ of all procurement contracts. For big contracts, the number of auctions without competition during the online phase is even $55\%$. This finding contrasts with our reasoning about the optimal behavior of participants. To further argue that such bidding behavior likely leads to money left on the table by losing bidders, we compute the probability of winning conditional on not being the initial winner and undercutting the initial winner at least once. The conditional probability of winning is $18.8\%$, which shows significant incentives for undercutting the initial winner.

### Realized updates are small

Our equilibrium discussion predicts most updates being very small and equal to the minimum bid decrement. We restrict attention to updates (i.e., situations in which a bidder lowered their bid when compared to their bid in the previous round) and define the relative step size as $\frac{b^{r+1}_{\ell(i)} - b_{w(i)}}{b^{r}_{\ell(i)}}$ where $b^{r}_{\ell(i)}$ is the bid of a firm currently losing an auction $i$ in stage $r$ and $b_{w(i)}$ is the bid of the current winner of auction $i$. Note that a 'positive update' means a bidder lowered their bid, but not by enough to beat the current standing winner; such 'ineffective'

[8]This is often a possible explanation in datasets from other countries that only publish data about larger contracts.

Table 1: Summary of procurement data

| | *Panel A: All contracts* | | | *Panel B: Big contracts* | | |
|---|---|---|---|---|---|---|
| Variable | Mean | Median | Std. Dev. | Mean | Median | Std. Dev. |
| Price | 2.07 | 0.12 | 25.03 | 101.80 | 49.28 | 204.00 |
| Relative price | 0.89 | 0.91 | 0.09 | 0.95 | 0.98 | 0.05 |
| N bidders | 2.54 | 2.00 | 1.04 | 2.59 | 2.00 | 4.15 |
| Is updated | 0.55 | 1.00 | 0.49 | 0.45 | 0.00 | 0.49 |
| Update size | 0.00 | -0.00 | 0.03 | 0.00 | 0.00 | 0.03 |
| Colluder participates | 0.07 | 0.00 | 0.30 | 0.11 | 0.00 | 0.30 |
| $N$ | | 354,642 | | | 4,151 | |

*Notes*: Relative prices are final prices divided by the reserve price. Price is measured in million UAH.

updates make up $17\%$ of all updates. The mass of bid updates is very close to zero with a median update of $-.1\%$ and the median 'effective' update is $-0.4\%$.

**Prosecuted collusion is a widespread phenomenon**

Among vendors of public procurement, the Ukrainian Anti-Monopoly Agency banned 863 firms for collusion since 2015. The majority received the universal penalty for collusion, a three-year ban from participating in procurement tenders. These companies participated in $23,515$ tenders accounting for $7\%$ of all procurement contracts but $9.4\%$ of the total value. Prosecuted colluders also generally update their bids less, $58.4\%$ of their bids never update. Contracts delivered by a prosecuted colluder are on average[9] more expensive: the median contract provided by a colluding firm costs $90.9\%$ of the estimate, whereas contracts delivered by companies not prosecuted cost $89.6\%$.

Overall, these statistics point out the lack of competition in the market for public procurement, which, together with the detailed prosecution data, hint at collusion as a crucial issue.

[9]Note that this is a simple average, i.e., we are not controlling for anything here.

# 5 Empirical Model

## 5.1 Reduced form evidence

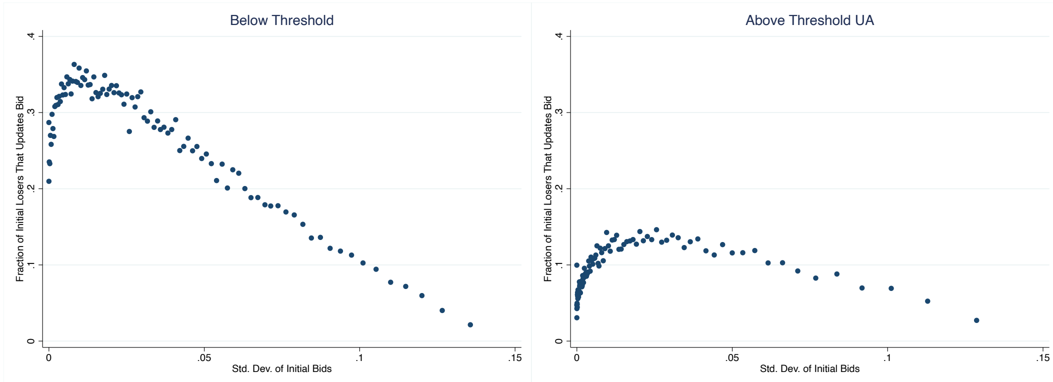
Collusion appears to be a widespread phenomenon throughout the market for public procurement in Ukraine. In the first part of this section, we show suggestive evidence based on patterns in the bidding data that are hard to explain in competitive equilibrium. And in the second one, we present a formal test that will isolate colluding pairs of firms.

Based on our discussion of the equilibrium, initial bids are strictly increasing in the underlying costs. So we can infer that if two bidders submit sufficiently close initial bids, their costs should also be close. Thus, we should see more competition and undercutting in auctions where the initial bids are near. We study whether the data are consistent with this prediction in Figure 2b by plotting the fraction of auctions. The initial loser updates his bid against the difference of initial bids. A competitive model would predict a declining function. This prediction partially holds on the left side of Figure 2b, in which we analyze below-threshold auctions. On the right side of the panel, we observe the opposite. Close initial bids are associated with decreased competition as measured by the likelihood of bid updating. Recall that there are increased incentives for collusion above the threshold due to the automatic canceling of contracts for only one bid.

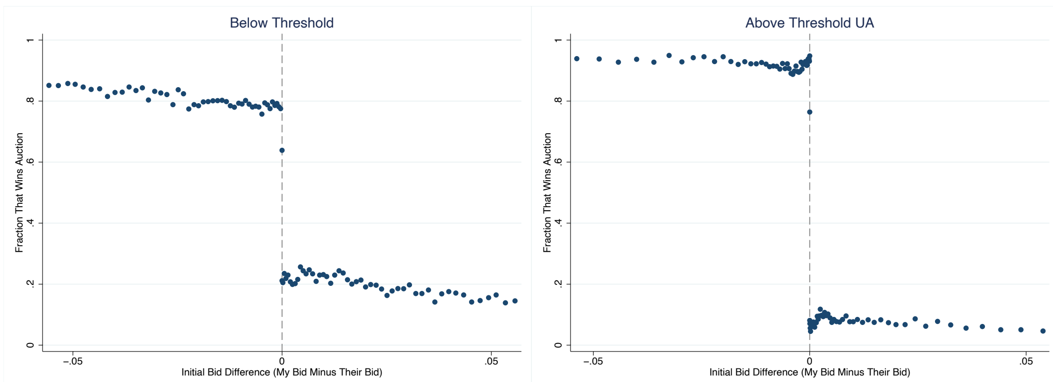
To further demonstrate this pattern, we plot[10] the probability of having the lowest standing bid[11] in the auction against the initial difference of bids in Figure 2a. There is an apparent discontinuity at zero that is explained by the second-mover advantage of the initial winner. Importantly, we document that submitting a bid just slightly above the opponent leads to a lower realized probability of winning than submitting a bid much above the opponent. Such patterns are not consistent with a competitive equilibrium where close cost draws imply increased competition. However, such bidding patterns are consistent with a cartel where

[10]For simplicity, we restrict the sample of auctions for this figure to those with precisely two bidders.

[11]Note that this is not precisely the probability of winning the auction as bidders may be disqualified after the auction. To avoid conflating unfair disqualifications with the issue at hand, we plot the implied win probability had there been no disqualifications. The results for the actual win probability are qualitatively similar.




(a) Fraction of Initial Losers That Updates Bid (Binscatter)




(b) Probability Own Bid Should Win Given Initial Bid Difference (Binscatter)

Figure 2: Concerning Patterns in Bidding Data.

*Notes*: Panel (a) is a binscatter of the fraction of initial losers that updates their bid against the standard deviation of initial bids. Panel (b) is a binscatter of the probability of having the lowest standing bid at the end of the auction against the initial bid difference. For easier visibility, we have restricted the sample to an absolute initial bid difference below the 75th percentile.

bidders agree on the winner ex-ante. Members submit bids close to each other and do not compete in the online auction. In the Appendix, we show that a simple collusion model can reproduce such patterns in the data.

## 5.2 Identification of network links

Based on our discussion in the previous sections, we now construct *(i)* a market-wide test for collusion, and *(ii)* develop a methodology that finds collusive networks by identifying collusive links in the network of firm relationships.

To begin with, let us define a function $g$ in the following way: $g(i,l,w)=1$ if bidder $l$ updates his bid against current lowest bidder $w$ in auction $i$ and $g(i,l,w)=0$ otherwise. From the discussion of the equilibrium, we know that bidders should update as long as the standing lowest bid is above their costs.[12] However, as we argue in the equilibrium section, under the assumption of fully revealing initial period bids there is a unique and strictly increasing function $b(\cdot)$ that maps costs to initial bids. Thus, we can use the initial bids to control for costs of all players and write:

$$\begin{aligned} g(i,l,w) &= 1\{b^{-1}(b_{l(i)}) < b_{w(i)}\} \\ &= \phi(b_{l(i)}, b_{w(i)}). \end{aligned}$$

Here, $\phi$ is non-parametric. As argued before, we allow for an idiosyncratic chance of bid submission failure. We also extend the model by allowing a bidder to undercut by mistake. We formalize this by introducing an idiosyncratic shock to undercutting $\epsilon_i$ where $\mathbb{E}[\epsilon_i|\phi(b_l,b_w)]=\alpha$. The bidder then undercuts as long as

$$\alpha + \phi(b_{l(i)}, b_{w(i)}) + \epsilon_i^d \geq 0 \tag{1}$$

with $\epsilon_i^d$ being the demeaned undercutting shock in the auction $i$.

We include a key indicator function tracking a particular link of firms: $\delta_{\ell(i),w(i)}$, which reflects tendency of the current loser $\ell(i)$ to undercut against the current winner $w(i)$. Due to the nature of the data, we observe the bidders in numerous auctions, and we can track their true identities. The panel nature of our data allows us to estimate the size of this coefficient.

In a competitive model, these link indicators are irrelevant as only the costs of bidders matter for their undercutting behavior, and we can control for them.[13] However, this might not be the case in a collusive model. For instance, our data reveal that prosecuted cartels

[12] As in the discussion of the equilibrium, we will assume an arbitrarily small minimum bid decrement.

[13] Also, it is worth stressing the bidders do not know the identities of other participants as they only observe generic names of participants such as *Bidder 1*.

are much less likely to undercut each others' bids in the online auction. In the competitive model, the true value of any pairwise $\delta$ is thus $\delta = 0$. The bidder updates if

$$\alpha + \phi(b^r_{\ell(i)}, b^r_{w(i)}) + \delta_{\ell(i),w(i)} + \epsilon^d_i \geq 0. \tag{2}$$

with $b^r_{w(i)}$ being the winning bid in the $r^{th}$ round of auction $i$ and $b^r_{l(i)}$ the losing bid in the $r^{th}$ round of auction $i$. Equation 2 will be the baseline of our estimation routine. We estimate a linear approximation of Equation 2 via OLS. We linearize the model due to the infeasibility of estimating a nonlinear model with a very high number of fixed effects. The following Proposition summarizes the asymptotic properties of the OLS estimator under the assumption of competitive bidding.

**Proposition 3.** *In data generated by a competitive equilibrium,* $\widehat{\delta^{OLS}_{\ell(i),w(i)}} \sim N(0, \sigma^2)$ *for some* $\sigma^2$.

We provide proof of the Proposition in the Appendix. The intuition is that OLS estimates of $\delta_{\ell(i),w(i)}$ are noisily estimated zero coefficients under the competitive assumption. The linearization of the model does not affect the consistency of the estimation because the true value of the fixed effects $\delta$ is zero. As the fixed effects are not separately identified from the intercept, we normalize their mean to zero.

We estimate the model in (1) and plot the empirical distribution of the estimated links (fixed effects) in Figure 3. When compared to a Normal distribution, the plotted data has an excess mass below the mean. Using the Kolmogorov-Smirnov test, we reject the hypothesis that the distribution is normal. We estimate fixed-effects only for pairs that we repeatedly observe in the data. As the sample size might still be limited for some pairs, we apply Empirical Bayes shrinkage to the fixed effects (Chandra et al., 2016). The additional mass on the left side of the distribution shows a significant number of pairs that are less likely to undercut each other's bid compared to the competitive baseline – an observation that would be expected on the market with lots of collusive firms.

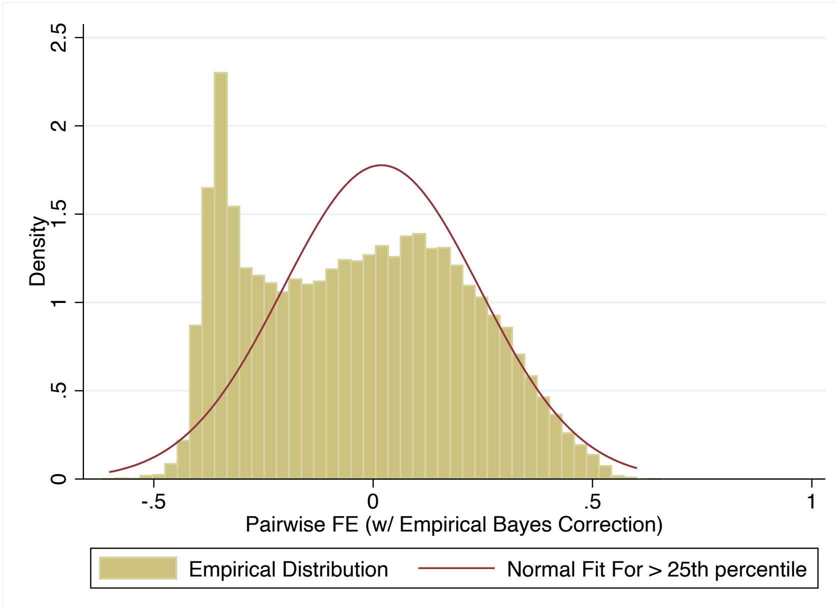



Figure 3: Distribution of Pairwise Fixed Effects

*Notes*: We show the distribution of the pairwise fixed-effects of equation (1). There is excess mass on low fixed-effects, indicating that certain bidder pairs never undercut each other.

### Interpretation of the network links

We based our previous discussions only on the concept of a competitive equilibrium. We now add some simple assumptions on a cartel's behavior that will help us explain the anomalies in data and translate the fixed-effects into probabilities of being in a cartel. We will consider two additional assumptions.

**Assumption 1.** *Cartels collude by not undercutting each other with some probability $p$.*

This assumption formalizes our previous discussion. We assume one of the features (out of potentially many) of a cartel is that the colluding firms do not undercut each other with some probability. This assumption is relatively general as it allows for different cartel strengths, with $p = 1$ being a perfect cartel and $p = 0$ approaching the competitive case. However, it excludes sophisticated behavior on the side of the cartel, where cartels imitate a competitive equilibrium. In this sense, we only identify a subset of all possible cartels, which is in line with classic results suggesting that we can never rule out anti-competitive behavior as sufficiently sophisticated firms could always emulate competitive equilibrium play (Bajari and Ye, 2003).

**Assumption 2.** *The probability of being in a cartel is independent of optimal competitive behavior.*

This assumption will allow us to derive a simple formula for the probability of a pair being in a cartel. In particular, we assume that the auction characteristics do not influence the likelihood of cooperating. We can then write the probability of updating as:

$$\begin{aligned} P(\ell\ updates\ against\ w) &= P(\ell \text{ updates against } w \text{ under competition} \cap = \ell, w \text{ are not a cartel}) \\ &= \phi(b_\ell, b_w) \times P(\ell, w \text{ not in a cartel}) \end{aligned}$$

**Proposition 4.** *Under Assumptions 1-3, we can rewrite the probability of being in a cartel as*

$$P(w, \ell\ in\ a\ cartel\ for\ project\ i\ ) = 1 - \frac{\phi(b_\ell, b_w)}{\phi(b_\ell, b_w) + \delta_{k,l}}.$$

In this case, there is a unique mapping from the fixed effect value to the probability of being a cartel: a lower value of $\delta$ suggests a higher likelihood of being in a cartel. Moreover, this equation holds for each project $i$. Thus, we could estimate the average probability of two bidders being in a cartel by averaging over all observations $i$ on the right-hand side. In Appendix A, we conduct simulations to show how adding only these few simple assumptions on the cartel's behavior helps explain most of the anomalies we observe in the data.

### Illustration of Network Links

To give the reader a better idea of how the network links look like, we plot the links for a sub-sample of firms supplying procurement contracts in the common procurement vocabulary (CPV) category of *"railway and tramway locomotives and rolling stock and associated parts"*.[14] We define suppliers in this category as firms that supply majority of their procurement contracts throughout our full sample in this CPV category. To increase readability,

[14]The CPV code of this particular category is 346 and it has been randomly chosen.

the sample is further reduced to firms with lowest 25 percent lowest pairwise fixed effects. The firms in this sub-sample are more likely to be colluders than the rest of the sample. The network is plot in Figure 4, in which we also distinguish between links of prosecuted firm pairs (dashed violet line) and non-prosecuted pairs (solid orange line). Most links are not (yet) prosecuted even though our detection algorithm suggests that these firms in the network are likely behaving in a collusive way.

**Verification of the Network-Based Identification**

Our algorithm successfully detects prosecuted pairs of firms. In Figure 5, we compare the estimated network links (fixed effects) for firm pairs that the courts identify as colluders and for other firms. The dataset of prosecuted firms is unique because it gives information about identified pairs of indicted firms (not single prosecuted firms). Hence, we can convincingly verify whether our algorithm detects collusive rings. We show that our algorithm's pairs that are predicted to be collusive are indeed much more likely to be prosecuted for collusion by the Ukrainian authorities. We observe that the identified colluding pairs are concentrated in the left part of the distribution with the suspicious additional mass. However, even among firms that were not prosecuted, there is still excess mass on the left, suggesting that the courts did not identify all colluding firms. The figure shows that our algorithm works very well[15] for the subset of identified colluders and could be used for further identification of other likely colluding firms.

# 6 Conclusion

We developed a new algorithm to detect collusion and show that collusion is widespread in the Ukrainian public procurement market. In this algorithm, we exploit the multi-stage nature of the Ukrainian auction mechanism and detect pairs of firms that repeatedly do not behave competitively against each other in procurement auctions. In our setting, this means

[15] The relevant t-statistic from the regression of the dummy of being a collusive pair on the size of the pairwise FE is $-17.06$.

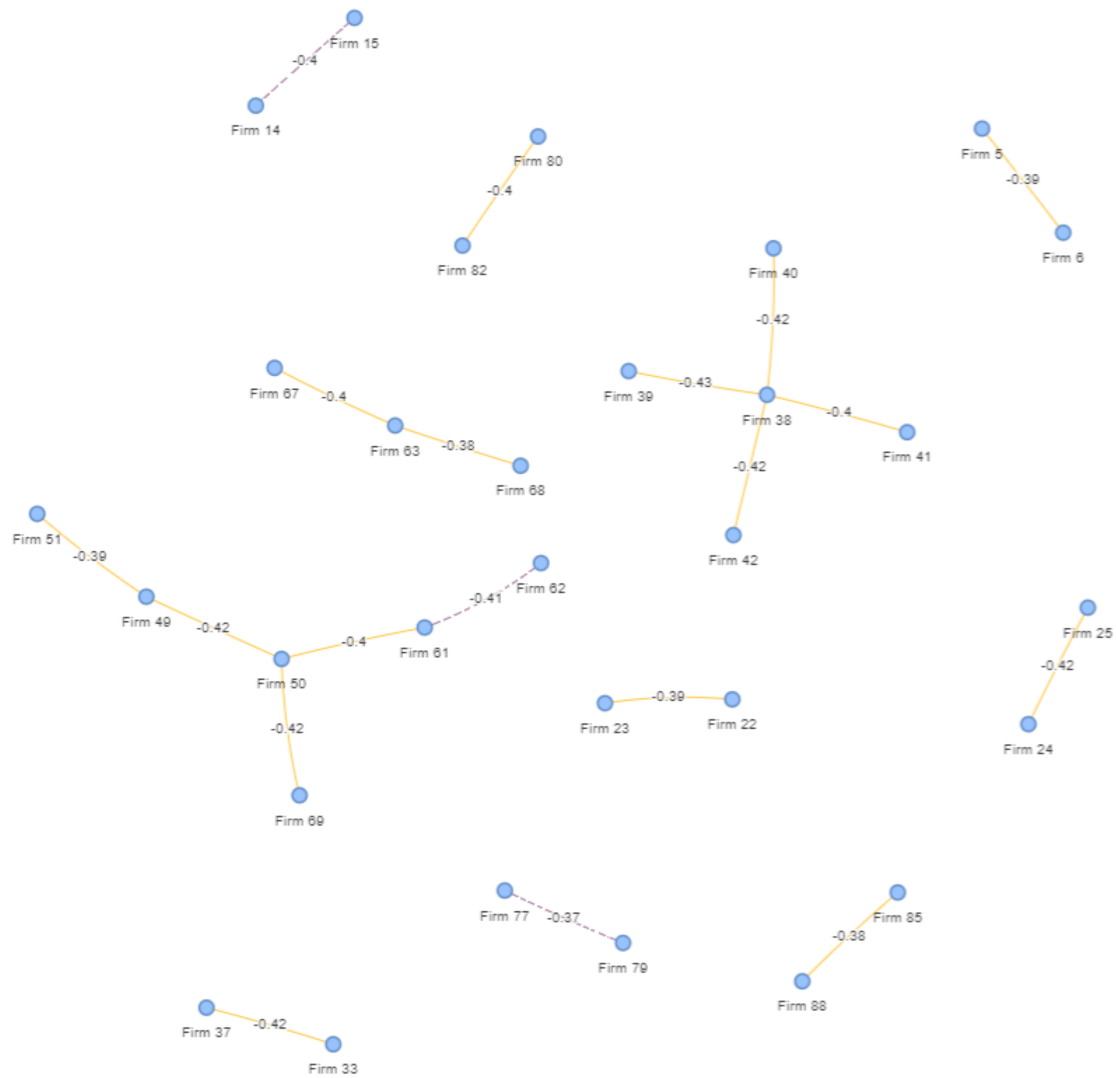



Figure 4: Illustration of cartel network

*Notes*: The figure shows network links among firms that supply procurement contracts in the common procurement vocabulary (CPV) category of *"railway and tramway locomotives and rolling stock and associated parts"*. The sample is reduced to pairs of firms with 25 percent lowest pairwise fixed effects to ensure better readability. The links in dashed violet lines represent already prosecuted link. The solid orange lines represent non-prosecured links. The numbers on the link represent pairwise fixed effects as calculated above.

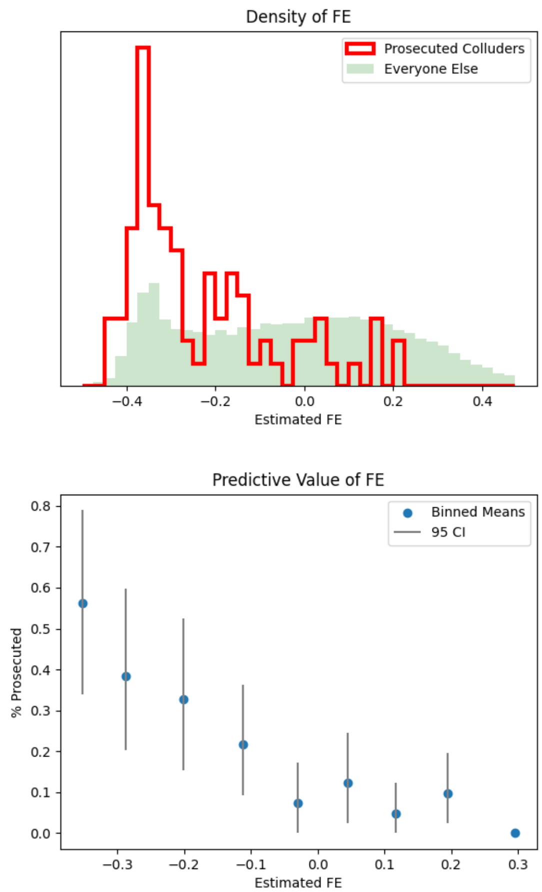



Figure 5: Pairwise FE: Sentenced Colluders vs Non-Sentenced

*Notes*: We break up the distribution of pairwise fixed-effects by whether a pair was mentioned in a successfully prosecuted collusion case or not. Clearly, sentenced colluding pairs are more likely to have lower fixed-effects (the t-stat from a binary linear probability model is 17.06).

that such firms do not update their bids against each other in the subsequent stages of the auctions and only wait till the auction ends. This methodology allows us to estimate the probability of any two firm-pairs colluding and thereby detects whole networks of collusive rings. Furthermore, we successfully verified our detection algorithm's reliability on a dataset of 863 successfully prosecuted firms by the Ukrainian Anti-monopoly Agency.

Our results are relevant outside the Ukrainian context. The Prozorro procurement system was designed with the intention to fight the issues of corruption and collusion, which are common for most countries in the post-soviet region. The idea was that the introduction of high levels of transparency in the behaviour of firms and procuring authorities allow for the identification of firms involved in corrupt or collusive behaviour. Our paper shows that it is indeed possible to credibly detect networks of collusive firms, which makes the system attractive for other countries with similar issues. Moldova and Georgia have already implemented similar multi-round e-procurement systems, which means that our detection mechanism is immediately applicable in these two countries. Georgia, Moldova, Kyrgyzstan, and Ukraine are now seen as leaders in e-governance in the region and the Prozorro system is available for free as an open-source system. This makes Prozorro-like systems even more attractive and easy to adopt in other countries and settings.

# References

Asker, J. (2010). A study of the internal organization of a bidding cartel. *American Economic Review*, 100(3):724–62.

Bajari, P. and Ye, L. (2003). Deciding between competition and collusion. *Review of Economics and statistics*, 85(4):971–989.

Chandra, A., Finkelstein, A., Sacarny, A., and Syverson, C. (2016). Health care exceptionalism? performance and allocation in the us health care sector. *American Economic Review*, 106(8):2110–44.

Chassang, S. and Ortner, J. (2019). Collusion in auctions with constrained bids: Theory and evidence from public procurement. *Journal of Political Economy*, 127(5):2269–2300.

Clarka, R., Coviellob, D., and De Leveranoc, A. (2020). Complementary bidding and the collusive arrangement: Evidence from an antitrust investigation. *SSRN*.

Conley, T. G. and Decarolis, F. (2016). Detecting bidders groups in collusive auctions. *American Economic Journal: Microeconomics*, 8(2):1–38.

De Leverano, A. (2019). Collusion through market sharing agreements: Evidence from quebec's road paving market. *ZEW-Centre for European Economic Research Discussion Paper*, (19-053).

Demokratizatsiya, S. S. (2017). Public procurement reform in Ukraine: the implications of neopatrimonialism for external actors. *demokratizatsiya.pub*.

Greene, W. (2003). Econometrics. *Various Editions*.

Greene, W. H. (1983). Estimation of limited dependent variable models by ordinary least squares and the method of moments. *Journal of Econometrics*, 21(2):195 – 212.

IMF (2021). Report for Selected Countries and Subjects.

Kawai, K. and Nakabayashi, J. (2014). Detecting large-scale collusion in procurement auctions. *Available at SSRN 2467175*.

Partnership, O. C. (2020). Ukraine S&L: MEL Framework. pages 1–20.

Porter, R. H. and Zona, J. D. (1993). Detection of bid rigging in procurement auctions. *Journal of political economy*, 101(3):518–538.

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

Supreme Council of Ukraine (2015). Ukrainian Public Procurement Law. pages 1–41.

Transparency International Ukraine (2017). Co-Creation of ProZorro: An Account of the Process And Actors. Technical report.

Transparency International Ukraine (2019). Behind the Scenes of ProZorro. Does Ukrainian business trust public procurement? Technical report.

World Bank (2020). 2019 Public Expenditure and Financial Accountability (PEFA) Performance Assessment Report.

# A Institutional background

The story of public procurement in Ukraine is long and complicated (for a summary see Transparency International Ukraine, 2017). While a first real effort to develop procurement legislation in 1997 was motivated by the need to harmonize regulations with WTO standards, the resulting law introduced in 2000 was lacking in detail and clarity (Transparency International Ukraine, 2017). The situation deteriorated substantially when the newly established 'Tender Chamber of Ukraine' was put in charge of all public procurement in 2005 and promptly began exercising its power to unduly influence bidder selection (Demokratizatsiya, 2017). An interim period followed in which there were several unsuccessful attempts to fix the system.

In 2013, the suspension of negotiations with the European Union by Ukrainian president Viktor Yanukovych sparked demonstrations. It marked the beginning of a period of political turmoil, the 'Euromaidan.' As protests spread, Yanukovych fled the country, and parliament relieved him of his duty. While an interim government led the country, the head of the Ministry of Economic Development and Trade (MoE) asked volunteers to organize themselves and research possibilities for reforming various governmental institutions. Public procurement was one of them. After meetings with Georgian and EU procurement experts, the volunteers agreed to model their system on the Georgian example.

However, two issues remained. There was a worry that a centrally administrated system would not provide sufficient incentives for ease of use. Furthermore, there was no apparent source of funding for the project: perhaps surprisingly, the official procurement department did not yet support the reform. Ukraine adopted a 'hybrid' system in which access to a central database of procurement contracts is mediated by various marketplaces that are allowed to charge a fee for this access but, in turn, provided initial funding for the development of the system. Transparency International Ukraine agreed to manage the budget during the pilot phase of the project, collected the contributions from the marketplaces, and selected a company for the necessary software development.

With initial funding secured, a pilot of what would eventually become the ProZorro e-procurement system was live in February 2015. However, at this stage, the project was still entirely a volunteer-led reform initiative: things only changed when a volunteer representative became the head of the Department of Public Procurement Regulation in March 2015. Thus, the status of the project was elevated, parliament passed new legislation in November 2015, and new funding from multiple international organizations allowed various refinements of the pilot necessary for full deployment. Finally, in April (August) 2016, the use of ProZorro became compulsory for many (all) public entities.

At its core, ProZorro is *(i)* a unified central database of all public procurement projects conducted in Ukraine and *(ii)* an API for interacting with this database. Appropriate legislation ensures that procurers post all public tenders to this database, and (crucially!) read-only access (e.g., for monitoring or research) is always free. Procuring entities and tenderers interact with the database via one of several profit-oriented marketplaces that allow the (free) posting of and (fee-incurring) participation in tenders via their unique interfaces. However, the 'auctions' themselves are run by the central database so that marketplaces cannot unduly influence their result.

The marketplaces (or the whole system) are often referred to as 'eBay for public procurement' in the media. Such simplification, however, falsely suggests that the main innovation of the system is the easy access to new tenderers through the use of information technology. While this plays a part in the success of ProZorro, the platform's primary purpose is better described by its name: 'transparency.' By design, all the information that exists about a tender is readily available publicly. All interested parties can, therefore, easily monitor procurement contracts.

The fact that transparency was the primary purpose of the development of ProZorro becomes even more salient when we examine several initiatives built to complement and support the platform. Firstly, the 'analytics module' allows quick access to summary statistics; the module is sufficiently interactive to allow for productive exploration of the data at a journalistic level. Furthermore, the MoE and ProZorro have introduced several procurement qualifications. While the university courses mainly aim at teaching potential future civil servants how to *run* successful tenders, there are also online courses with a more explicit focus on monitoring, for example, the aptly named 'Monitoring of Public Procurement; Or: How To Look for Betrayal.'

The introduction of ProZorro has been widely lauded as a highly positive step for public procurement in Ukraine. Indeed, ProZorro has received several awards (such as being rated #1 by the World Procurement awards 2016 in the Public Sector nomination). The World Bank in 2020 assigned the Ukraine letter-grades of A in nearly all scored dimensions of public procurement. The sole exception was the 'procurement methods' score since only 78.1% of the total cost of all public procurement covered by the relevant law in 2018 was tendered in competitive procedures (World Bank, 2020).

In the main text, we argue that while the formal institutions in Ukraine have greatly improved, a closer examination of the bidding suggests that collusion and shill-bidding have become costly problems. Indeed, only 13.3% of respondents in a 2017 survey agreed that 'the system helps increase competition and achieves value for money' (Partnership, 2020). When asked a similar question in 2019, this number improved, and 46.3% of respondents said that the level of corruption in public procurement had slightly or significantly decreased after the launch of ProZorro (though 12.2% said it had increased) (Transparency International Ukraine, 2019). However, 24.2% still stated that they had personally encountered situations in which they were 'forced to pay a bribe or resort to nepotism after ProZorro was launched, and 34.2% say that corruption is the most severe problem facing the platform. Our analysis supports the public perception of widespread collusion.

# B Simulated model of collusion

Here we present a model of collusion that likely reflects the behavior of actual cartels in our data. We will conduct simulations for auctions with precisely two players. There are two possibilities: either the auction is competitive, or there is a collusive pair participating.

In the competitive situation, the bidders act in line with their equilibrium strategies.

Collusive pairs, on the other hand, are modeled in the following way. The pair designates a winner ex-ante; the designated loser just exists to submit a phantom bid, thereby causing the auction to go ahead (and perhaps ensuring regulators do not investigate the relevant market). Thus, the designated loser simply submits a bid $(1 + a)b_{c,w}$, where $a$ is a small constant and $b_{c,w}$ is the bid of the designated winner. The selected winner submits a bid above the equilibrium competitive bid. Thus, collusive bidders do not undercut each other as they have no incentive to do so.

We present results of a simulation where bidders have uniform costs, i.e., $c_i \sim U[0, 1]$. We furthermore parametrize the model for our simulation. We simulate 400 bidders, each of which interacts with all other bidders in exactly two auctions. Of all bidding pairs, 3% are in a cartel.

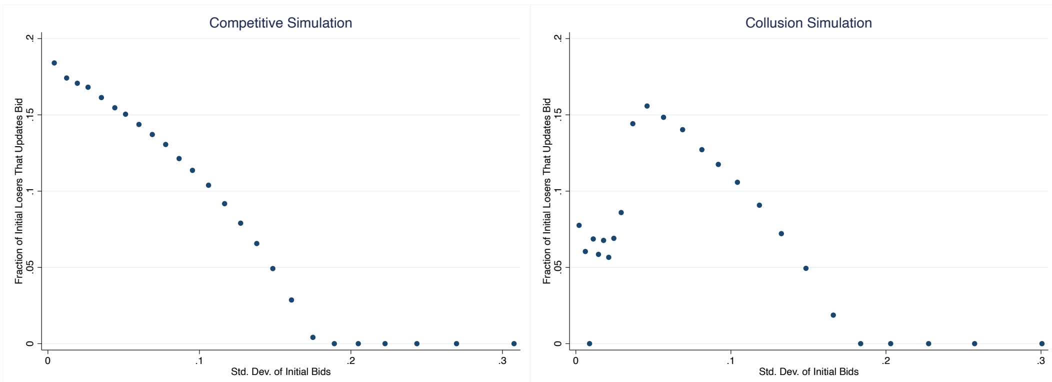
Using data from this model we reproduce Figure 2a and 2b. We see that patterns in these figures can be explained with this simple model of collusion. The above-threshold auctions have a higher incentive for collusion and are comparable to the simulation of collusive behavior. By contrast, below-threshold auctions are comparable to the competitive ones.

It is hard to argue that this model of collusion is a unique model as collusive agreements can have many different forms. However, it is striking that we can explain all anomalies in our data with such a simple model.

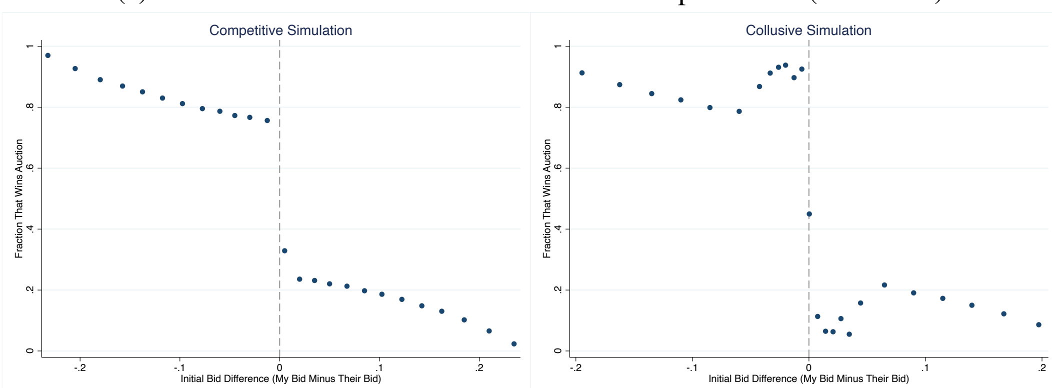
# C Motivating examples and evidence

A brief review of tenders on the online platform reveals suggestive evidence that collusion is an issue on the market. In Figure 7, we show a screenshot from the online platform showing bidding on a tender in which two bidders submitted virtually identical bids - 31,864,899.19 UAH[16] and 31,864,900.00 UAH. The losing bidder did not update her bid in the following rounds even though the difference in bids was about 0.81 UAH, i.e., 3 cents. This behavior is suspicious. We use the fact that many firms on the market behave in such a noncompetitive

[16] Roughly 1,138,000 USD.




(a) Simulation: Fraction of Initial Losers That Updates Bid (Binscatter)




(b) Simulation: Probability Own Bid Should Win Given Initial Bid Difference (Binscatter)

Figure 6: Results of Bidding Simulation.

*Notes*: Panel (a) is a binscatter of the fraction of initial losers that updates their bid against the standard deviation of initial bids. For easier visibility, we have restricted the sample to an absolute initial bid difference below the $75^{th}$ percentile. Panel (b) is a binscatter of the probability of having the lowest standing bid at the end of the auction against the initial bid difference. For easier visibility, we have restricted the sample to an absolute initial bid difference below the 75th percentile.

way only when competing with a particular firm to identify colluding pairs of firms.

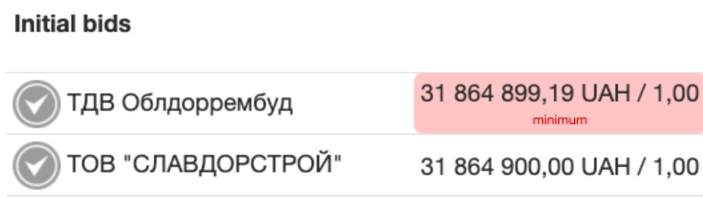



Figure 7: Suspiciously close bids and no undercutting

*Notes*: These are two initial bids in an auction where no subsequent bids happened.

# D Equilibrium of ProZorro Auction

**Proposition 1.** *In any equilibrium in which initial bids are given by some strictly increasing $b(\cdot)$, the expected payoff from pretending to be type $\tilde{c}$ is given by $V(\tilde{c})$ no matter the number of updating rounds or number of players.*

*Proof.* We consider the PZA auction with $k+1$ rounds (i.e., $k$ updating rounds) and $n$ players; we will index rounds by $r$ and players by $i$. We will refer to the bid by player $i$ in round $r$ as $b_i^r$ and use $\underline{b}_i^r$ to notate the standing lowest bid *before* $i$ moves in round $r$. Note that bidding in updating rounds is not (necessarily) in order of player indices as updating priority is based on the ranking of the bids from the previous round; hence, we also introduce $\sigma(r,t)$ as notation for the index of the player that moves in position $t = 1, \ldots, n$ in round $r$. Thus, for example, $b^1_{\sigma(2,3)}$ refers to the first round bid by the player who moves third in the second round.

We assume that initial bids are fully revealing, and hence can let $\hat{c}_i := b^{-1}(b_i^1)$ be the shared (point-)belief of $j \neq i$ about the cost type of player $i$. As we are considering only deviations by P1 (wlog), we have $\hat{c}_i = c_i$ for all $i \neq 1$. This also implies that all bidders but P1 move in order of their costs in the first updating round; the position of P1 is determined by the cost-type he chooses to imitate.

We will regularly need to refer to the optimal bid of a player $i$ who anticipates that no firm moving after her is capable of beating a standing bid of $x$ but at least one is capable of beating all higher bids. Such a player would like to bid $x$, but may be constrained by her own cost. If $x$ is below her cost, the player – anticipating that she will be beaten – would be indifferent between all other bids were it not for the possibility of bid submission failure. As it is, however, there is a small but positive probability $p > 0$ that any given subsequent bid submission attempt will fail. This gives her a chance to nevertheless win the auction: for instance, if there is just one player to move after her that could beat $y > c_i$, she could submit $y$ and hope that this player will fail to submit his bid. Even if she believes that all players to move after her can beat her own cost $c_i$ (as all players believe in equilibrium), there is still a chance that they all fail (repeatedly) at submitting their bids, in which case she can win by undercutting the standing winning bid by $\Delta$. More generally, we will refer to the optimal undercut as $\Delta^*(i, r)$ without characterizing it further and introduce the following notation for the optimal bid of a player $i$ in round $r$ who anticipates that he would not be beaten if she bid $x$:

$$g_i^r(x) = \begin{cases} \max\{b : b \leq x, b \leq \underline{b}_i^r - \Delta, b \geq c_i\} & \text{if this set is nonempty,} \\ \underline{b}_i^r - \Delta^*(i,r) & \text{o/w and if } \underline{b}_i^r - \Delta^*(i,r) \geq c_i \\ b_i^{r-1} & \text{o/w.} \end{cases}$$

It of course remains to characterize the value of $x$ after each history, which we will now do by proceeding with backward induction. Firstly, noting that $\sigma(k+1, n)$ is the last player to move in the last round, we claim that in any SPE,

$$b_{\sigma(k+1,n)}^{k+1} = \begin{cases} b_{\sigma(k+1,n)}^{k} & \text{if } b_{\sigma(k+1,n)}^{k} = \underline{b}_{\sigma(k+1,n)}^{k+1} \\ g_{\sigma(k+1,n)}^{k+1}\left(\underline{b}_{\sigma(k+1,n)}^{k+1}\right) & \text{o/w} \end{cases}$$

Thus, the last agent to move will simply undercut by as much as necessary in order to win the contract (assuming this yields positive profit). Anticipating this, all other agents in the last round would like to scoop, i.e., ensure that their bid cannot be undercut by anyone moving after them. Hence, they will anticipate that they can win if and only if they bid no more than the bid decrement $\Delta$ above the cost of whoever they believe to be the lowest cost agent moving after them. They thus bid

$$\forall t < n : b^{k+1}_{\sigma(k+1,t)} = g^{k+1}_{\sigma(k+1,t)}\Big(\min\{\hat{c}_{\sigma(k+1,s)}|s > t\} + \Delta\Big).$$

It should be noted that if $k = 1$, this implies that all players but P1 and $\sigma(k+1, n)$ will simply undercut the current standing bid by $\Delta$ (if possible without going under their cost). This is because the order in which players are moving is exactly the order of player strength given their beliefs: hence they anticipate never being able to win the auction if no bid submission failure occurs. The same is true for P1 as long as he is pretending to be either a stronger type than he actually is or his true type. If he is pretending to be a weaker type, then and only then can he successfully 'scoop'.

If $k > 1$, the argument in the preceding paragraph still applies as long as the order of players hasn't changed between updating rounds. However, it may change due to the behavior of P1. Nevertheless, the strategies stated above are still optimal.

Moving backward, given the situation in the last updating round, all agents anticipate that the agent with the lowest cost will win. Thus, all agents (including P1) are in the same situation in round $k$ as in $k + 1$, and hence they will play essentially the same strategies: all players but P1 undercut in the hope of a bid submission failure, and P1 scoops if he is actually the lowest type but was initially pretending not to be. Why does P1 scoop 'early' rather than 'late'? By scooping early, he guards against the fact that his own late scooping bid may not go through. Thus, strategies in earlier updating rounds are mostly unchanged from later updating rounds:

$$\forall\, 1 < r < k + 1 : \forall t : b^r_{\sigma(r,t)} = g^r_{\sigma(r,t)}\Big(\min\{\hat{c}_{\sigma(r,s)}|s = 1, \ldots, n\} + \Delta\Big).$$

Finally, note that if $\tilde{c} \leq c_1$, then $\Delta^*(i, r) \equiv \Delta$ as all agents (including P1) anticipate that all agents moving after them can beat their own costs. If $\tilde{c} > c_1$, this is not true anymore:

in particular, P1 may anticipate that some firms that will get to update their bid after him cannot beat his costs. However, either $c_1$ is the lowest cost draw or not. If it is, then P1 will never be forced to contemplate the case in which he relies on bid submission failure to win, and as $p \to 0$, his payoff from pretending to be $\tilde{c}$ will converge towards that he would get if there was no bid submission failure chance. If it is not the lowest cost, then with probability approaching one, P1 will not win the auction. Hence, his payoff will be zero, no matter what complicated undercutting strategies $\Delta^*$ he employs in the meantime.

Thus, as we take the limits $p \to 0$, $\Delta \to 0$, the strategies derived in this proof imply the following payoff from pretending to be type $\tilde{c}$ in the initial round (when your true type is $c_1$):

$$\begin{aligned} V(\tilde{c}) \;\; = \;\; & \mathbb{P}\Big(b(\tilde{c}) < \min_{j \neq 1} b(c_j)\Big)\Big(b(\tilde{c}) - c_1\Big) + \\ & \mathbb{P}\Big(b(\tilde{c}) > \min_{j \neq 1} b(c_j) \cap c_1 < \min_{j \neq 1} c_j\Big) \times \\ & \mathbb{E}[\min\{c_j : c_j < \tilde{c}, j \neq 1\} - c_1 | c_1 < \min_{j \neq 1} c_j, b(\tilde{c}) > \min_{j \neq 1} b(c_j)] \end{aligned}$$

□

**Proposition 2.** *The ProZorro auction (with $k \geq 1$ rounds and $n \geq 1$ players) has a unique PBE in which initial bids are given by a strictly increasing $b(\cdot)$. In this equilibrium,*

$$b(c) = \frac{1}{[1 - F(c)]^{n-1}} \int_c^{c_{max}} s(n-1)f(s)[1 - F(s)]^{n-2} ds$$

*and bids are decreased by the minimum bid decrement whenever doing so is possible without bidding below one's own cost.*

*Proof.* We have

$$\begin{aligned} V(\tilde{c}) \;\; = \;\; & \mathbb{P}\Big(b(\tilde{c}) < \min_{j \neq 1} b(c_j)\Big)\Big(b(\tilde{c}) - c_1\Big) + \\ & \mathbb{P}\Big(b(\tilde{c}) > \min_{j \neq 1} b(c_j) \cap c_1 < \min_{j \neq 1} c_j\Big) \times \\ & \mathbb{E}[\min\{c_j : c_j < \tilde{c}, j \neq 1\} - c_1 | c_1 < \min_{j \neq 1} c_j, b(\tilde{c}) > \min_{j \neq 1} b(c_j)] \end{aligned}$$

Say $c_i \sim F(\cdot)$ with $\max \operatorname{supp} c_i = c_{max}$. We will use

$$G(\tilde{c}) = 1 - [1 - F(\tilde{c})]^{n-1}$$

as a short-hand to refer to the distribution of the minimum of the $n-1$ other costs. Then

$$\begin{aligned} V(\tilde{c}) &= [1 - G(\tilde{c})]\big(b(\tilde{c}) - c_1\big) + [1 - G(c_1)] \times \\ &\quad \max\Big\{\frac{G(\tilde{c}) - G(c_1)}{1 - G(c_1)}, 0\Big\}\Big(\frac{1}{G(\tilde{c}) - G(c_1)} \int_{c_1}^{\tilde{c}} cdG(c) - c_1\Big), \end{aligned}$$

where we used the fact that $\min\{c_j : c_j < \tilde{c}\} = \min_{j\neq 1} c_j$ given that $\tilde{c} > \min_{j\neq 1} c_j$.

Although on first glance it may seem[17] like $V(\tilde{c})$ is not differentiable at $\tilde{c} = c_1$, this is in fact wrong because the potentially non-differentiable part of $V(\tilde{v})$ is multiplied by the expected rent from a second price auction conditional on your strongest opponent having a cost draw below $\tilde{c}$, which tends to zero as $\tilde{c} \to c_1$. After recognizing this, it is easy to see that

$$V'(c_1) = (1 - G(c_1))b'(c_1) - (b(c) - c)g(c),$$

where $g(c) = G'(c)$. Together with the boundary condition $b(c_{max}) = 0$, this differential equation is uniquely solved by

$$b(c) = \frac{1}{1 - G(c)} \int_c^{c_{max}} sdG(s),$$

which is just the classic first-price auction equilibrium bidding strategy. $\square$

[17] As $\max\Big\{\frac{G(\tilde{c})-G(c_1)}{1-G(c_1)}, 0\Big\}$ is not differentiable at this point.

# E Data manipulation

### Some bids are implausibly low

We note that a large share of bids is 'too good to be true'. Such bids are likely to be provided without showing that the company is reliably able to deliver the demanded project which leads to the subsequent disqualification of the bids. As other firms can see such low bids at the start of the auction and anticipate that the suspiciously low bidder will be disqualified. In such cases, the optimal behavior would change and the bidders would only compete against other bidders and not the low bidder. To alleviate this problem we conduct our analysis only on the sample without very low bids, which we define as any auction where the lowest bid is below a conservative threshold of $80\%$ of the highest bid of other participants. This leads to omitting around 35% of all auctions. Our results are robust to both using the specified sub-sample or the whole sample of all auctions. There are also other reasons why a firm might get disqualified but as these are not easily predicted both from the data available to companies before the auction starts and also from the ex-post data available to researchers we choose to not explicitly model them.

# F Estimation

**Proposition 3.** *In data generated by a competitive equilibrium,* $\widehat{\delta^{OLS}_{\ell(i),w(i)}} \sim N(0, \sigma^2)$ *for some* $\sigma^2$.

*Proof.* $\widehat{\delta_{\ell(i),w(i)}}$ is the OLS estimate of $\delta_{\ell(i),w(i)}$. Recall that the bidder undercuts if and only if

$$\alpha + \phi(b^1_{\ell(i)}, b^1_{w(i)}) + \delta_{\ell(i),w(i)} + \epsilon^d_i \geq 0, \tag{3}$$

where the true value of the $\delta_{\ell(i),w(i)} = 0$ in a competitive model and $\epsilon^d$ is a random shock with $\mathbb{E}(\epsilon_i | X, \delta) = 0$. We estimate a linear analog of this equation:

$$\alpha^{OLS} + \phi(b^1_{\ell(i)}, b^1_{w(i)}) + \delta^{OLS}_{\ell(i),w(i)} + \epsilon^{OLS}_i \tag{4}$$

The parameters $\delta^{OLS}$ and $\alpha^{OLS}$ are not separately identified as the fixed effects cannot be identified separately from $\alpha^{OLS}$ by the standard argument. For the estimation we impose the additional constraint that $\sum_{\ell(i),w(i)} \widehat{\delta_{\ell(i),w(i)}^{OLS}} = 0$. Now, it follows directly by Greene (1983) that the OLS estimator of a coefficient from a probit model will converge in probability to the true value multiplied by a scalar $a$. So $\widehat{\delta_{\ell(i),w(i)}^{OLS}} \rightarrow a \cdot (\delta_{\ell(i),w(i)} - \overline{\delta_{\ell(i),w(i)}})$. But recall that the true values $\delta_{\ell(i),w(i)} = 0$ and $\overline{\delta_{\ell(i),w(i)}} = 0$ implying that $\widehat{\delta_{\ell(i),w(i)}^{OLS}} \rightarrow 0$ proving the consistency of this estimator. Normality then follows by the central limit theorem (Greene, 2003). $\square$