Detection of Collusive Networks in E-Procurement *

Bruno Baránek† Leon Musolff‡ Vítězslav Titl§

April 19, 2023

Abstract

We develop a novel collusion detection approach for multistage e-procurement auctions. Our network-based approach uses insights from a theoretical model to identify collusive behavior by analyzing deviations from competitive equilibrium play. In particular, we can convert initial bids to firm costs and detect firm pairs that do not undercut current winning bid against each other even though their costs would allow for it. In detailed procurement auction data from Ukraine, we find model-free evidence of noncompetitive bidding patterns, and subsequently, apply our algorithm to detect suspected collusive links. The soundness of our detection algorithm is validated on a sample of firms penalized by the Anti-monopoly Committee of Ukraine. We also document that the majority of cartel members form a two-member only cartel and often share the same ZIP code.

Keywords: Public procurement, Collusion, Online markets

*We gratefully acknowledge insightful comments from Francesco Decarolis, Adriano De Leverano, Sylvain Chassang, Thomas Fujiwara, Elena Fumagalli, 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 (award number PRN-004) 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. Bruno Baranek conducted this research while at Princeton.

†Electronic address: `bbaranek@princeton.edu`

‡Electronic address: `leonmusolff@microsoft.com`; Microsoft Research New England.

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

1 Introduction

Detecting and studying collusion in auctions is a critical topic for competition policy (Porter and Zona, 1993). There is widespread agreement among scholars and practitioners that collusion has adverse effects on social welfare, leading to higher costs and a reduced supply of public goods. Moreover, these effects have a direct impact on taxpayers. Thus, detecting networks of collusive firms is crucial for improving the efficiency of the public procurement market and public goods provision.

In this paper, we detect and analyze collusion in multistage e-procurement auctions. We develop a theoretical model of the auction mechanism, solve for the equilibrium, and devise a novel collusion detection approach. Our test exploits the fact that we observe each firm's behavior in each auction round to identify pairs and, in turn, rings of colluding firms. Finally, we use unusually detailed data from Ukraine to validate our detection algorithm, verifying the algorithm's accuracy on a sample of firms penalized for collusion by the Antimonopoly Committee of Ukraine.

The multi-round sequential auction mechanism studied in this paper was intentionally designed to inspire oversight by civil society. It proceeds by initially letting interested bidders submit bids simultaneously. Subsequently, all bidders can access an online auction where they compete for the contract by potentially lowering their bids over several rounds.[1] In these rounds, bidding is sequential in an order determined by the previous round's ranking of bids. In particular, the bidder with the highest initial bid starts the bidding in the online auction. This setup can be understood as a mechanism in which 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 can react to her bid. Bidders are, thus, 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 any round.

We show below that if there is a small probability that the current winner fails to update

[1] The Ukrainian auction has three rounds. Neither our collusion detection algorithm nor our other findings depend on the exact number of rounds.

her bid,[2] then it is optimal for the bidder with the second-lowest bid to undercut the current winning bid by a small ϵ (if he can do so while bidding above his costs). Hence, the (repeated) lack of undercutting within firm pairs that submit similar initial bids signals that firms do not behave competitively. Based on this idea and our formal theoretical model, we construct a novel structural algorithm to detect collusion and identify bidding rings in public procurement: network-based ring identification.

Network-based ring identification examines each pair of firms in turn. Using our theoretical model, we compute the probability with which a given firm should underbid a particular competitor. Subsequently, we identify collusive relationships as statistically unlikely deviations from equilibrium play. By identifying all collusive links, we can reconstruct the collusive network and locate bidding rings.

We first present empirical evidence that there is low activity during online auctions. Though (non-cartel) firms might miss out on significant profits by not updating, 52.55% of initial bids are never updated – and in 45.93% of the tenders, *no* participant updates their bid. The lack of undercutting is consistent with a cartel setup where members initially agree on project allocation and only submit initial bids as cover. Indeed, descriptive analysis confirms that close initial bids are particularly frequently followed by a lack of activity.

Subsequently, we run our detection algorithm on all procurement auctions from Ukraine and verify the algorithm's accuracy on a sample of 752 firms penalized for collusion that placed at least one bid in the auctions we consider. These firms participated in 20,340 tenders (4.78% of the total value of the market). The penalization data are unique because they allow us to identify who is colluding precisely – in our terminology, the data shows 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.

The Ukrainian procurement market is a fascinating laboratory for studying collusion for three reasons. First, Ukraine has one of the most modern procurement markets, operating

[2] Our data suggest that this assumption reflects reality.

entirely on an electronic platform that allows unprecedented data access and transparency. Second, the market has suffered from large-scale issues with collusion. Between 11/14/16 and 9/30/19, 1042 firms have been penalized for collusive conduct[3]. Third, the multi-round auction design allows us to propose a novel framework for the identification of collusion.

The same auction mechanism as the one studied in this paper is used on public procurement markets in several other post-Soviet countries such as Georgia, Moldova, Kyrgyzstan, and Ukraine.[4] Similar auctions are also used in real estate auctions, the U.S. government timber rights' sales and art auctions (Engelbrecht-Wiggans, 1988).

Our paper extends and contributes to the literature studying collusion in auctions. We build on seminal work by Porter and Zona (1993) and Porter and Zona (1999) that 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).

The main contribution of our study is a structural test that can identify networks of collusive firms. We explicitly allow firms to be collusive in some settings but not in others and then identify which firm pairs collude, allowing us to reconstruct the collusive network. This detection is made possible by building on (i) a microeconomic equilibrium model and (ii) unusually detailed data on collusive behavior that helps us validate our econometric test. To the best of our knowledge, this paper is the first study that presents a methodology for detecting collusive *networks*. By contrast, Porter and Zona (1993) test whether there is collusion on *the whole market* and Kawai and Nakabayashi (2014) test whether *a particular firm* behaves collusively. Neither isolates which firms are involved in a given cartel.

As firms likely only collude when facing members of their bidding ring, methods that pool data from all auctions a firm participates in will mix information from collusive and non-collusive settings. The results from such biased methods could both over- or underestimate the presence of collusion depending on the setting, an issue our approach avoids. Furthermore,

[3]This figure differs from the 752 firms our analysis is based on as not all firms penalized for collusion appear in our cleaned bidding data.

[4]These four countries represent about one-third of the GDP of the post-Soviet countries besides Russia (IMF, 2021).

our approach can also help prosecute collusive firms: identifying a cartel's exact structure and members can point prosecutors in the correct direction when collecting evidence.

From a theoretical perspective, our setting also allows us to directly solve for optimal equilibrium behavior, which we can compare to the behavior in the data. By contrast, (Kawai and Nakabayashi, 2014) analyze collusive auctions without providing a precise competitive equilibrium.

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. 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, propose a statistical test for collusion, verify our method on a sample of firms penalized for collusion, and characterize networks of colluding firms. Finally, section 6 concludes.

2 Description of the Market

Ukraine has faced corruption and collusion in public procurement 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 became compulsory for all public entities in 2016.[5] 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.

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

2.1 Exact tender procedure

In Ukraine, small ("below-threshold") purchases[6] can be conducted without an online auction.[7] Larger purchases ("above-threshold") generally have to be completed as open tenders. Our analysis will focus on competitive tenders both below- and above-threshold and emphasize a critical difference between the two types of tenders: 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 only one auction participant exists.

We now describe the open tendering procedure for the above-threshold contracts. 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. After this period, the tender is automatically canceled if only one proposal has been submitted.[8] If there are multiple proposals, the system automatically schedules and runs an online auction, which we discuss below. During the online auction, the bidders do not yet have access to each others' documents. Furthermore, while they are informed of the number of opposing bidders when the online auction starts, they remain unaware of their identity and specific proposals until it 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. Second, bidders enter the online auction. Now they are given a chance to update their bids three

[6] 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 'normal' contracting authority worth less than $53k, *(iii)* a good/service bought by a 'special' contracting authority worth less than $35k, and *(iv)* a works purchase by a 'special' CA worth less than $177k (Supreme Council of Ukraine, 2015).

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

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

times in the following mechanism[9]:

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 initial bid observes all bids, i.e., initial bids and the update by the originally highest bidder. This (second highest) bidder is again given a chance to lower her original bid.

   (c) All the bidders move sequentially until the bidder with the lowest initial bid has chosen whether to update her bid.

2. Bidders are ordered based on the size of their updated bids and again move sequentially.

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 at the end of this round wins and becomes the vendor of this project.

This mechanism 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 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).

[9]While the mechanism may seem unusual, Georgia and Moldova use the same multi-round mechanism. Indications are that other countries will follow.

3 Model and Equilibrium

We now discuss the intuition behind the equilibrium of the ProZorro Auction, with details and proof relegated to the appendix. For this purpose, consider a simplified auction with two players and 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) can update his bid.
3. The initial 'winner' can update her bid.

Note that bidders can only update their bids downwards, i.e., initial bids are not 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 information, in turn, generates a potential multiplicity issue: the initial loser may realize that he will lose the overall auction no matter what he bids. To resolve this multiplicity, we introduce a small probability that any given bid update is not successfully submitted. This probability captures the idea that if you know you will lose if your opponent reacts, you may as well bid in such a way as to maximize your surplus if, for some reason, your opponent fails to respond.

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 that still give the initial loser a positive surplus, he will make the highest bid satisfying these criteria. If there are none, he will update his bid to the current winning bid (hoping against hope that she will fail to respond). 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

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

The four lines of this expression correspond to the four possible cases: 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 to yield 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 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 $\underline{c}_{-1} := \min_{j \neq 1} c_j$ to emphasize that this way of expressing the payoffs does not rely on a specific number of players. 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 $V(\tilde{c})$ in Figure 1 for the case of $c_i \sim U[0, 1]$. If $\tilde{c} < c_1$, the expected value from a second price auction conditional on $c_2 < \tilde{c} < c_1$ is zero; hence, $V^{PZA}(\tilde{c}) = V^{FP}(\tilde{c})$ to the left of $\tilde{c} = c_1$. Furthermore, $V^{FP}(1) = 0$. Thus, $V^{PZA}(1) = \mathbb{E}[V^{SP}]$. However, the expected rent that bidders earn in a second-price auction is exactly 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

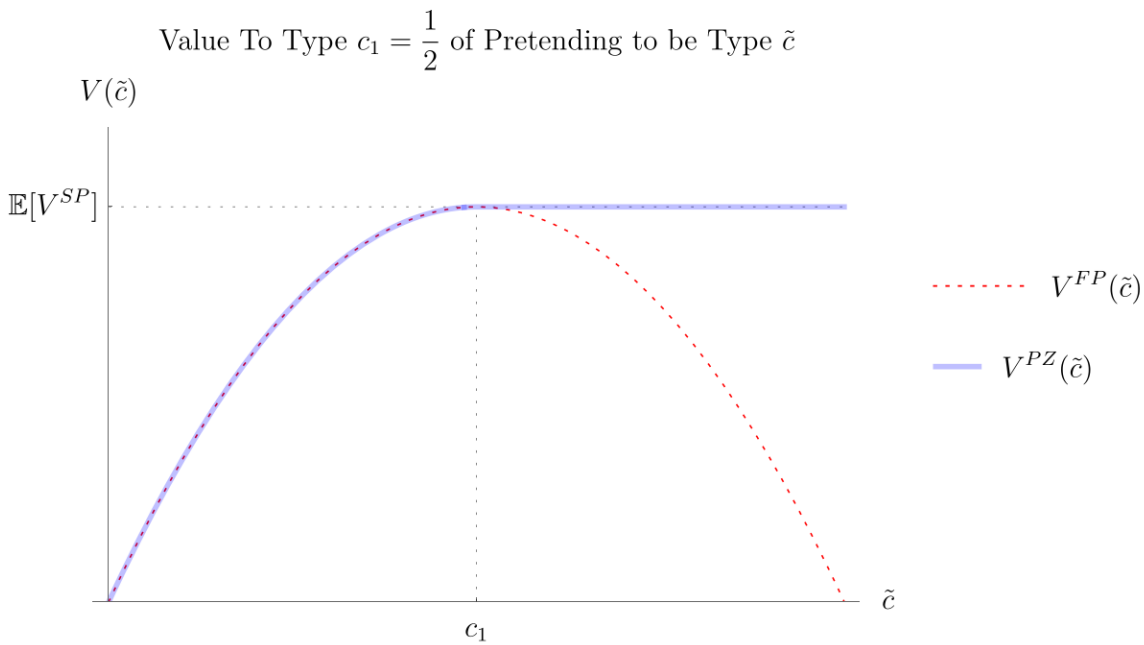


Figure 1: The Equilibrium of the PZA Auction

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 first-price auction, but the second-price component ensures it never drops from this level.

$V^{PZA}(\cdot)$ 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, the initial bids in the ProZorro auction are generated by the same bidding function as in a first-price auction.

4 Data and Summary Statistics

We use data from three data sources, each covering the period of August 2016 to August 2019. First, we use publicly available tender-level 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 bid-level data, including bids and bidders' identities at all stages of each online auction. Third, we obtain data about firms penalized for collusive conduct in public procurement from the Ukrainian Anti-Monopoly Agency.

We provide summary statistics for the tender-level data in Table 1. We draw the readers' attention to a few key stylized facts that become apparent from this table.

There is low competition in the market. The minimum number of participants for all the above-threshold contracts is two. In the data, we do not see much higher levels of competition. The average number of participants is 2.47 bidders (median: 2). While the Ukrainian data contains a large number of small contracts (frequently not reported in

other countries), the low number of participants is not limited to small contracts. Indeed, 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. Of all non-initial bids, only 32% are updates on the bidder's previous bid; indeed, there are zero bid updates in 46% of all online auctions. Low-value auctions do not explain this phenomenon: for big contracts, the fraction of auctions without competition during the online phase is even higher at 50%. This finding contrasts with our reasoning about the optimal behavior of participants. Indeed, the fraction of auctions with undercutting by the initial loser in which the initial loser eventually wins is 25.82% – suggestive evidence that there are strong incentives for undercutting the initial winner.

Realized updates are small. Our equilibrium discussion predicts that most updates equal 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 $\frac{b^{r+1}_{\ell(i)} - b^{r}_{w(i)}}{b^{r}_{\ell(i)}}$ where $b^{r}_{\ell(i)}$ is the bid of a firm currently losing an auction i in stage r and $b^{r}_{w(i)}$ is the bid of the current winner of auction i in stage r. Note that a 'positive update' means a bidder lowered their bid, but not by enough to beat the current standing winner; such 'ineffective' updates make up 15% of all updates. The mass of bid updates is very close to zero with a median update of -0.22%. The median 'effective' update is -0.43%.

Penalization for collusion is a widespread phenomenon. Among vendors of public procurement, the Ukrainian Anti-Monopoly Agency penalized 1,048 firms for collusion between 11/14/16 and 9/30/19. The majority received the universal penalty for collusion, a three-year ban from participating in procurement tenders. These companies participated in 20,340 tenders accounting for 6.21% of all procurement contracts and 4.78% of the total value. Firms penalized for collusion also generally update their bids less, only 26.76% of their non-initial bids are updates. Contracts delivered by a penalized colluder are on average[10] more expensive: the average contract provided by a colluding firm costs 91.29%

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

Table 1: Summary of Procurement Data

	Panel A: All Contracts			*Panel B: Big Contracts*		
Variable	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Bids						
Bid (in million UAH)	2.23	0.13	25.97	101.31	47.98	201.42
Normalized Bid	0.93	0.96	0.09	0.97	0.99	0.06
Was Bid an Update?	0.32	0.00	0.46	0.25	0.00	0.43
Update Size (x100)	-0.01	-0.22	3.38	0.30	-0.08	2.94
Is Update Effective?	0.85	1.00	0.36	0.80	1.00	0.40
Update Size (x100, Effective Only)	-1.04	-0.43	1.74	-0.65	-0.19	1.34
AMCU-Penalized Bidder?	0.03	0.00	0.17	0.06	0.00	0.23
N		3,054,189			356,387	
Tenders						
Mean Normalized Bid	0.93	0.95	0.08	0.97	0.99	0.05
Min. Normalized Bid	0.90	0.92	0.10	0.95	0.98	0.06
Number of Bidders	2.47	2.00	1.01	2.59	2.00	4.63
Any AMCU-Sanctioned Bidders?	0.07	0.00	0.25	0.11	0.00	0.31
Any Updates?	0.54	1.00	0.50	0.49	0.00	0.50
N		308,567			4,018	

Notes: This table provides summary statistics for bids (top) and tenders (bottom) for all contracts (Panel A) and large contracts (Panel B). We consider any contract with engineer's estimate above 25,000,000 UAH to be a large contract. There are N=3,054,189 bids in our data, with 356,387 of these belonging to large contracts. Note that the bid data includes multiple bids by each bidder for a given tender as the online auction procedes over several stages. There are 308,567 tenders in our data, with 4,018 of these tenders being for large contracts. The normalized bid variable normalizes bids (in UAH) by dividing them by the engineering estimate (also in UAH). The update variables are only defined for non-initial bids.

of the estimate, whereas contracts delivered by companies not penalized cost 90.22%.

Overall, these statistics point out (i) a lack of entrants and (ii) the lack of competitive behaviour conditional on entry in the market for public procurement. These factors, together with the detailed prosecution data, hint at collusion as a crucial issue facing the market.

5 Empirical Model

Collusion appears to be a widespread phenomenon in Ukraine's market for public procurement. 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; in the second one, we present a formal test that will isolate colluding pairs of firms.

5.1 Reduced form evidence

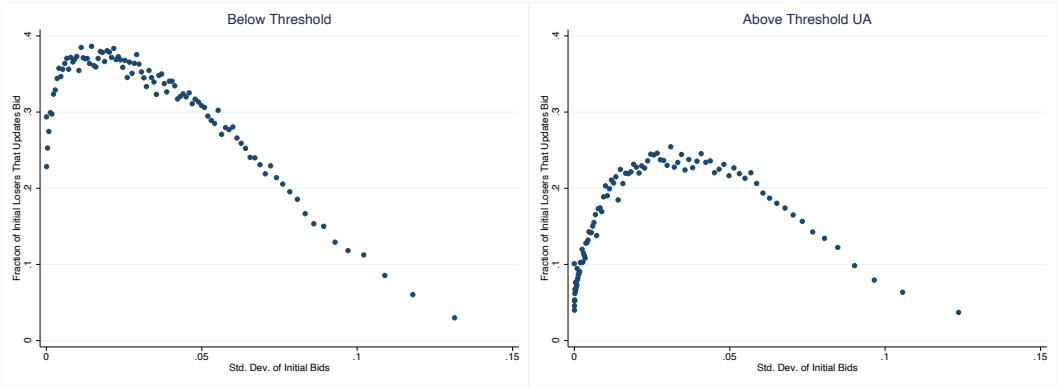
To begin with, recall that our discussion of the equilibrium revealed that initial bids are strictly increasing in the underlying costs. So if two bidders submit sufficiently close initial bids, their costs should also be close. Thus, we should see more undercutting in auctions where the initial bids are near.

We study whether the data are consistent with this prediction in Figure 2a. The figure plots the fraction of auctions in which the initial loser updates his bid (on the vertical axis) against the difference of initial bids (on the horizontal axis). A competitive model would predict a declining function. This prediction partially holds on the left side of Figure 2a, in which we analyze below-threshold auctions. However, we observe the opposite on the right side of the panel: close initial bids are associated with decreased competition as measured by the likelihood of bid updating. One possible explanation for the contrast between the panels could be that there are increased incentives for collusive bidding above the threshold. Recall above-threshold tenders are automatically canceled if only one bid exists. Thus, collusion on below-threshold contracts could take the form of market splitting; by contrast, above-threshold contracts require the submission of bids by at least two cartel members.

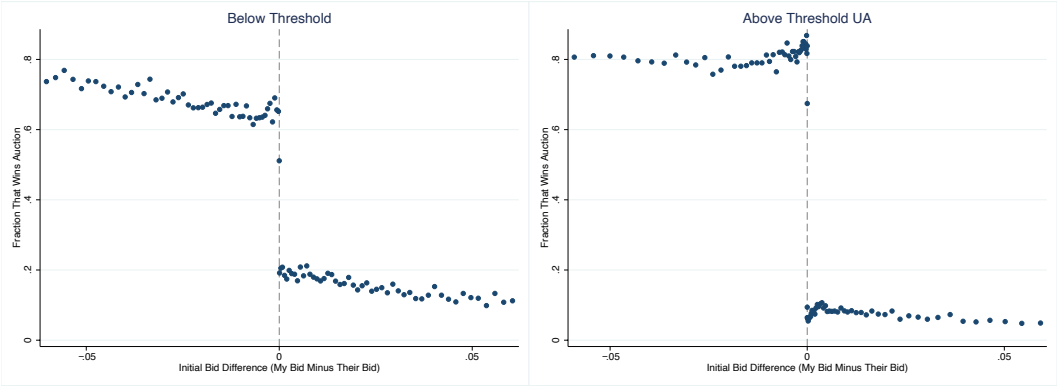
To further demonstrate this pattern, we plot[11] the probability of having the lowest final bid[12] (on the vertical axis) in the auction against the initial difference of bids (on the horizontal axis) in Figure 2b. The second-mover advantage of the initial winner explains the discontinuity at zero. Importantly, we document that for above-threshold contracts, submitting a bid just slightly above the opponent leads to a lower realized probability of winning than submitting a bid much above the opponent (i.e., the graph is locally increasing in a neighbourhood of the discontinuity). Such patterns are inconsistent with a competitive equilibrium where close cost draws imply increased competition. However, they are consistent with a cartel where bidders agree on the winner ex-ante. Cartel members submit bids

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

[12] 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.

close to each other (perhaps to make the auction seem competitive) and do not compete in the online auction. The online appendix shows that a simple collusion model in this vein can reproduce the suspicious patterns of Figure 2.

5.2 Testing for collusion

Based on our discussion in the previous sections, we now *(i)* develop a market-wide test for collusion, *(ii)* verify our detection on a sample of firm pairs fined by the Anti-Monopoly Committee of Ukraine, *(iii)* present a statistical test of collusive behaviour for individual

pairs of firms, and *(iv)* discuss characteristics of the detected cartels.

To begin with, define an indicator u_i that equals one if the initial loser $l(i)$ updates his bid against initial winner $w(i)$ in any round of auction i and equals zero otherwise. From the discussion of the equilibrium, we know that bidders should update as long as the standing lowest bid is above their costs.[13] 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[14] the initial bids by the initial loser $b_{l(i)}$ and initial winner $b_{w(i)}$ to control for the costs of all players in auction i. In particular, the model implies

$$u_i = 1\{b^{-1}(b_{l(i)}) < b_{w(i)}\}.$$

While our empirical specification builds on this model prediction, it extends it in several ways to account for the fact that $b^{-1}(\cdot)$ is unobserved, bidding mistakes, bid submission failures, and potential collusion:

$$u_i = 1\{b_{w(i)} - \phi(b_{\ell(i)}) + \delta_{\ell(i),w(i)} + \epsilon_i \geq 0\}. \tag{1}$$

We now discuss the differences between the model prediction and (1) in turn.

Firstly, the equilibrium bidding function $b(\cdot)$ is not observed by the econometrician; hence we proxy $b^{-1}(b_{l(i)})$ with a non-parametric function of the initial loser's bid $\phi(b_{l(i)})$. In practice, $\phi(\cdot)$ is a third-degree polynomial.

Secondly, we are interested in the possibility of collusion and hence allow the likelihood of undercutting to depend on a pairwise[15] fixed effect $\delta_{\ell(i),w(i)}$ that reflects the tendency of the initial loser $\ell(i)$ to undercut against the initial winner $w(i)$. In a competitive model, the fixed effects are irrelevant as only the costs of bidders matter for their undercutting behavior.

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

[14] 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*.

[15] Note that our estimation imposes $\delta_{a,b} = \delta_{b,a}$ for maximum power.

Hence, under competition, the true value of any pairwise fixed effect is $\delta_{a,b} = 0$. However, this might not be the case in a collusive model. For instance, our analysis above revealed that penalized cartels are much less likely to undercut each others' bids in the online auction.

Finally, we introduce an idiosyncratic exogenous auction-specific shock ϵ_i to undercutting. This shock accounts for the previously referenced idiosyncratic chance of bid submission failure, but also further extends the model by allowing a bidder to undercut by mistake.

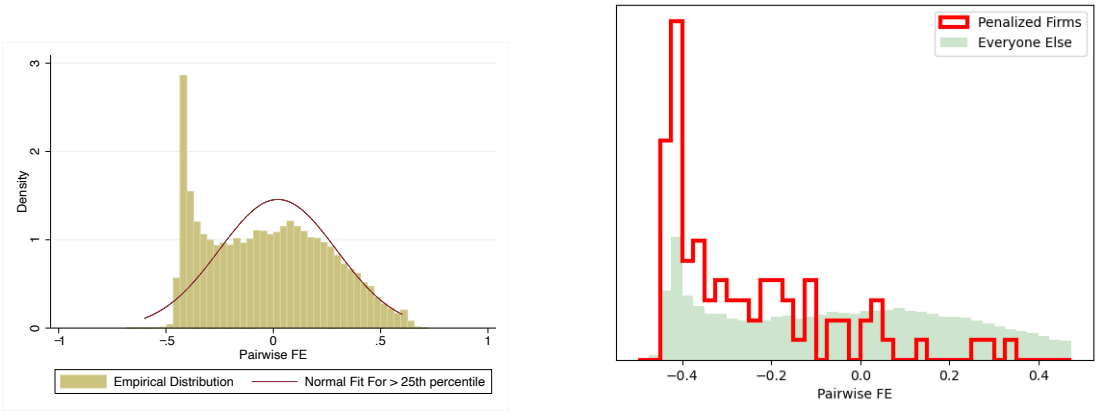
To complete the model, it remains to specify a distribution for ϵ_i. To keep estimation feasible and side-step a potential incidental parameter problem, we specify ϵ_i in such a way as to yield a linear probability model[16]. With this error distribution, (1) can be consistently estimated via OLS, which allows us to handle high-dimensional fixed effects without encountering computational constraints. Furthermore, we note that if our data had been generated by competitive equilibrium play, this would imply $\delta \equiv 0$ and hence we would have $\hat{\delta} \sim N(0, \sigma^2)$ for some σ^2 (Kwon, 2023).

We estimate the linear probability corresponding to (1) and plot the empirical distribution of the estimated links (fixed effects) in Figure 3a. We estimate fixed-effects only for pairs that we repeatedly observe in the data. 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 ($p < 0.001$). Furthermore, we also reject the (weaker) hypothesis that the distribution is symmetric ($p < 0.001$). 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.

Verification of the network-based identification

Our algorithm successfully detects penalized pairs of firms. In Figure 3b, we compare the estimated pairwise fixed effects for pairs of firms that the Anti-Monopoly Committee penalized for collusion and for other firms. The dataset of penalized firms is unique as it

[16]This implies $\epsilon_i \sim U[-1/2, 1/2]$. The mean is required to be zero; the size of the support plays the same role as the variance of the error term in a probit model, i.e., it is not identified and has to be normalized.



(a) Distribution of Pairwise Fixed Effects

(b) Histogram by Penalization Status

Figure 3: Pairwise Fixed Effects Reveal Collusion.

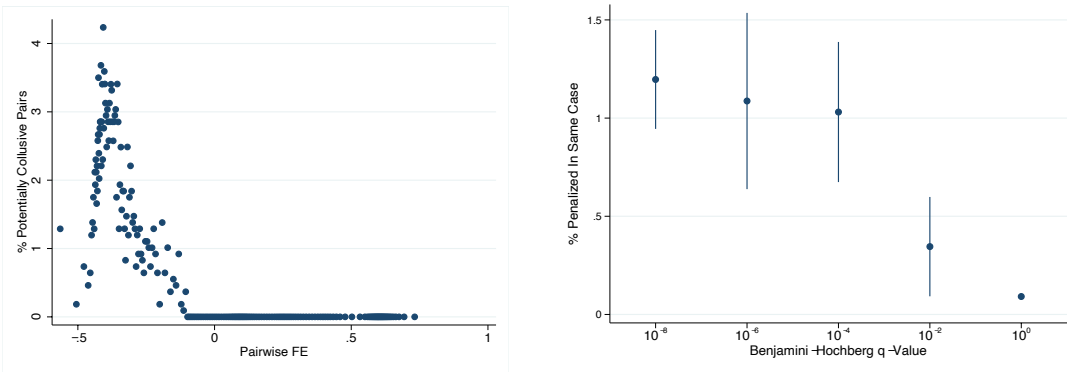
Notes: In 3a, we show the distribution of the pairwise fixed-effects of (1). As a visual guide to emphasize departure from normality, we also exhiibt a normal fit for the data above the 25th percentile. There is excess mass on low fixed-effects, indicating that certain bidder pairs never undercut each other. In 3b, we break up the distribution of pairwise fixed-effects by whether a pair was penalized in a collusion case by the Ukrainian Anti-Monopoly Committee or not and plot the distributions of pairwise FE for penalized and non-penalized firms separately.

gives information about identified pairs of penalized firms. Hence, we can examine whether our algorithm detects collusive rings. 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 penalized, 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[17] for the subset of identified colluders and could be used for further identification of other likely colluding firms.

Detecting collusive pairs

Pairwise fixed effects, our measure of competitive behaviour within pairs, are coefficients from a regression with standard errors and other statistics, which means that we can apply standard statistical tests on them. First, note that under the null of competitive behaviour,

[17]The relevant t-statistic from the regression of the dummy of being a collusive pair on the size of the pairwise FE is -14.63.



(a) Rejection of Null of Competition

(b) Fraction of Pairs Penalized

Figure 4: Statistical Test For Collusion

Notes: In 4a, we see that for the intermediate negative fixed-effects, we can frequently reject the null of competition. Note this is not true for the most extreme negative fixed-effects: these are mostly driven by sampling error. In 4b, we show that pairs with lower q-values (i.e., those our test detects as colluding) are also more likely to have been penalized in the same case for collusive conduct (the t-stat from a binary linear probability model regressing an indicator for penalization on the q-values is -9.68).

pairwise fixed effects we would have $\delta_{ab} = 0$ for all firms a and b. Second, suppose that cartels collude by not undercutting each other with some positive probability. Under this assumption, $\delta_{ab} < 0$ if *(a)* a and b are members of the same cartel and *(b)* at least one of a and b has counter-parties with which it does not collude. This observation means by testing the hypothesis that $\delta_{ab} < 0$, one also examines whether the firms in the pair behave uncompetitively against each other (provided that *(b)* is satisfied).

Accordingly, we form a one-sided test of the null hypothesis that the fixed effect does not lie significantly below zero for each pairwise fixed effect δ_{ab}. Formally, let n be the number of observations and k the number of regressors (including the fixed effects) in the linear regression correspond to 1. Define

$$t_{ab} := \frac{\hat{\delta}_{ab}}{se(\hat{\delta}_{ab})} \sim t_{n-k}.$$

Then, if F is the CDF of a t_{n-k} distribution, we can find the p-value associated with t_{ab}:

$$p_{ab} = 1 - F(-t_{ab}).$$

This gives us a (large) set of p-values. To avoid the classic multiple hypothesis testing problem, we will choose not to control the size of each individual test, but rather the false discovery rate, i.e., the expected fraction of discovered collusive links that turn out to be false positives. To do so, we follow the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995) to transform the p-values into an associated set of q-values, and finally reject the null hypothesis only for links for which $q_{ab} < 0.05$. These links are the ones declared collusive by our procedure. To further verify the soundness of our detection algorithm, we show that pairs with low q-values –i.e., those our test detects as colluding – are more likely to have been penalized in the same case for collusive conduct by the Anti-monopoly Committee of Ukraine. We present these findings in Figure 4b. The relevant t-statistics from a binary linear probability model regressing an indicator for penalization on the q-values is -9.68.

There are four main caveats to our procedure. Firstly, we rely on a colluder's behaviour differing across the firms she competes with and those she colludes with; our procedure has no power to detect any collusion between players that each collude with all firms they encounter across all auctions. Secondly, our procedure aims to be conservative by controlling for the rate of false positives: thus, it may miss collusive links for some pairs of players that have not faced each other often (as their absence of undercutting could be purely random). Thirdly, our procedure is unable to detect sophisticated behavior on the side of the cartel including, for instance, a cartel that imitates competitive play under inflated costs. 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). Finally, as highlighted by the same paper, our procedure jointly tests the null of our model being correctly specified and firms acting competitively; hence, our tests of competitive behaviour could reject for no other reason than a misspecified model.

Characteristics of network links

In this section, we present a descriptive analysis of the network characteristics exhibited by colluding firms. Our investigation begins by examining the geographical distribution of these firms based on their registered office locations. It is reasonable to assume that cartel-involved firms are more likely to be situated within the same geographical region, as this may facilitate the development of trust and coordination of illicit activities. To approximate firm locations, we utilize ZIP codes as a proxy measure.

Our findings reveal that approximately 5.9% of non-collusive firm pairs are situated within the same ZIP code area, whereas this proportion increases to 11% for collusive firm pairs. This pattern persists when evaluating larger geographical areas, such as those sharing the same first 2, 3, or 4 digits of ZIP codes. For instance, in the largest geographical areas defined by the first two digits of ZIP codes, we find that 18.5% of non-collusive firm pairs and 31.2% of collusive firm pairs are located within the same region. The observed differences in means are statistically significant across all examined ZIP code variations, with corresponding t-statistics of -9.5, -10.3, -13.7, and -14.3 for the full code and each truncated version, respectively.

Furthermore, we observe these disparities across various industries, with the most pronounced difference occurring in the "Food, beverages, tobacco and related products" sector. In this industry, 7% of non-collusive firm pairs are located within the same ZIP code area, compared to 17% of collusive firm pairs—a statistically significant difference, as evidenced by a t-statistic of -5.7. Overall, our findings suggest a notable geographical clustering of collusive firms, which may reflect the inherent need for proximity to establish trust and coordinate illegal activities within cartels.

Subsequently, we investigate the dimensions and structures of collusive networks. Our analysis identifies 2,371 unique firms participating in collusion, forming 1,919 distinct firm pairs. A majority of these firms (1,730 or 73%) engage in collusion with a single partner, while the remaining 641 firms with multiple connections contribute to the formation of 319 cartels. It is important to note that some firms with only one connection are also involved in

these 319 cartels.

The most extensive cartel comprises 46 firms connected through a chain of potentially varying lengths. Within this network, 22 cartel members possess a single connection, the average number of links per firm is 2.3, and the most connected firm has 15 links to other firms. Across the 319 cartels, the average number of firms per cartel is 3.4. In these cartels, 62% of members maintain a single link to another firm, and the average maximum number of links constitutes only 6.8% of the cartel's size. This proportion is significantly lower than the 32.6% (15/46) observed for the largest cartel, suggesting that in most cartels, firms typically collude with a few other firms rather than relying on a central cartel member for cooperation — evidence against a hub-and-spoke cartel structure.

Additionally, we analyze differences across industries, with a notable example being the "Medical equipment, pharmaceuticals, and personal care products" sector. In this industry, characterized by an unusual prevalence of large cartels, 288 firms participate in collusion. The five largest cartels within the sector have 18, 17, 13, 11, and 11 members, respectively.

6 Conclusion

We developed a new algorithm to detect collusion in multistage auctions. Our findings show that collusion is widespread in the Ukrainian public procurement market. In this algorithm, we exploit the multistage 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 that such firms do not update their bids against each other in the subsequent stages of the auctions, instead simply waiting until the auction ends. This methodology allows us to estimate the probability of any two firms colluding, thereby detecting whole networks of collusive rings. Furthermore, we verified our detection algorithm's reliability on a list of 752 firms penalized for collusion by the Ukrainian Anti-monopoly Agency that have bid on contracts in our data.

Our results are relevant outside the Ukrainian context. The ProZorro procurement system

was designed 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 behavior of firms and procuring authorities allow for the identification of firms involved in corrupt or collusive behavior. Our paper shows that it is possible to credibly detect networks of collusive firms, which makes the system attractive to other countries with similar issues. Georgia, Kyrgyzstan, and Moldova have already implemented similar multi-round e-procurement systems, which means that our detection mechanism is immediately applicable in these three 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

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

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, 57(1):289–300.

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.

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

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

Engelbrecht-Wiggans, R. (1988). On a possible benefit to bid takers from using multi-stage auctions. *Management Science*, 34(9):1109–1120.

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*.

Kwon, S. (2023). Optimal shrinkage estimation of fixed effects in linear panel data models. Working Paper.

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 went 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.

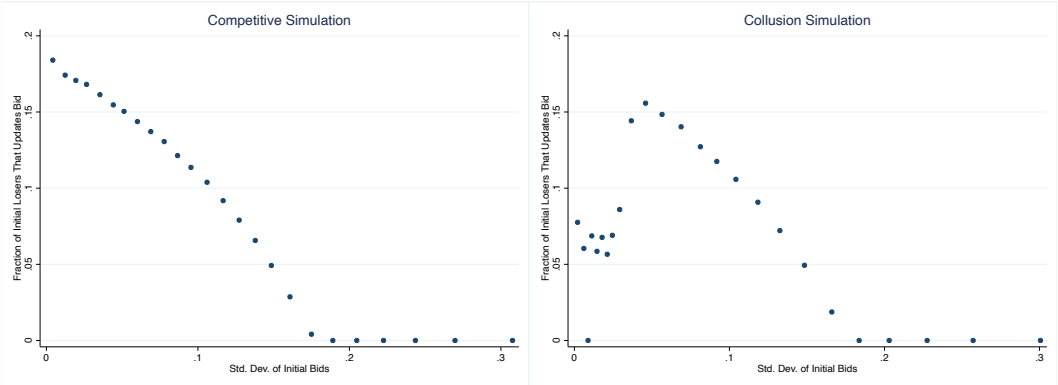
Using data from this model we reproduce Figure 2b and 2a. We see that patterns in these figures can be explained with this simple model of collusion. The above-threshold auctions have a higher incentive for collusion and are comparable to the simulation of collusive behavior. By contrast, below-threshold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

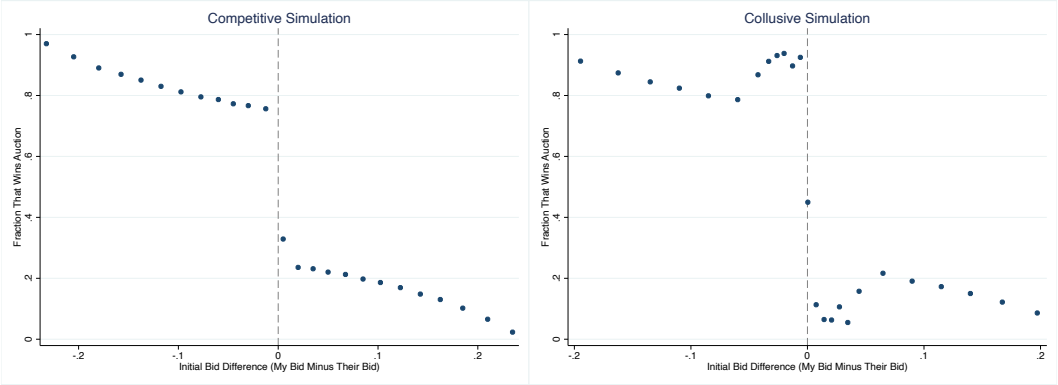
C Motivating Examples and Evidence

A brief review of tenders on the online platform reveals suggestive evidence that collusion is an issue on the market. In Figure 6, 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[18] 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

[18]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 5: 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 75th 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.

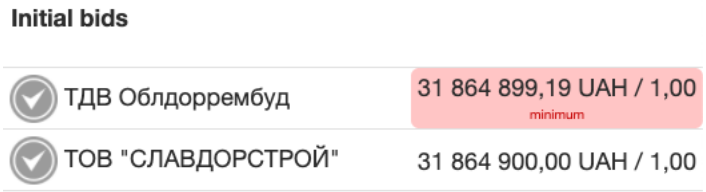


Figure 6: Suspiciously close bids and no undercutting

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 Δ. 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 Δ 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 Δ (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 Δ^* 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):

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

□

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

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

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

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

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[19] 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

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

which is just the classic first-price auction equilibrium bidding strategy. □

E Data Manipulation

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

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

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 of auctions without very low bids, which we define as containing 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 or from the ex-post data available to researchers we choose to not explicitly model them.

F Functional Form

In the main analysis, we run a linear probability regression model corresponding to Equation 1 to obtain pairwise fixed effects. There is potential sector, time, and procuring entity heterogeneity in the procurement tenders in our dataset. To test for robustness of our findings, we estimate the same regressions with sector and time fixed effects as well as procuring entity fixed effects. In Table 2, we present the findings using the baseline specification (*Baseline*), the specification with sector[20] and time fixed effects (*Sector and time FEs*), the specification with procuring entity fixed effects (*Entity FEs*), and the specification including all three sets of fixed effects (*Sector, time, and entity FEs*). In all specifications, the coefficient is negative and strongly significant. The negative sign shows that with smaller pairwise fixed effects, the chances that a company was penalized increases. This confirms that the pairwise fixed effects are good predictors of actually penalization by the Anti-Monopoly Committee of Ukraine.

[20]Sectors are implemented as four digits CPV codes.

Table 2: Summary of procurement data

	Coefficient	t-stats	p-vals
Baseline	-.183	-14.635	0
Sector and time FEs	-.195	-15.087	0
Entity FEs	-.180	-13.610	0
Sector, time, and entity FEs	-.192	-13.988	0

Notes: This table shows the coefficients, t-statistics, and p-values from the regressions of the dummy of being a collusive pair according to the Anti-Monopoly Committee of Ukraine on the size of the pairwise fixed effects. Each row in the table represents the relevant statistics from the regressions using different specifications to obtain the pairwise fixed effects. The firs one is from the specification used in the main analysis (the linear regression corresponding to Equation 1). In the following three rows, we use the specifications with sector and time fixed effects, the specification with procuring entity fixed effects, and the specification including all three sets of fixed effects, respectively.