Supplemental Appendix

Detection of Collusive Networks in Multistage Auctions

Bruno Baránek[1], Leon Musolff[2], Vítězslav Titl[3]

[1]The Government Analytical Unit, Office of the Government of the Czech Republic; bruno1baranek@gmail.com

[2]University of Pennsylvania; lmusolff@wharton.upenn.edu

[3]Utrecht University School of Economics, Utrecht University; Department of Economics and Empirical Legal Studies, Faculty of Law, Charles University; v.titl@uu.nl

November 2025

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 administered 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 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 remain costly problems. Indeed, only 13.3% of respondents in a 2017 survey agreed that 'the system helps increase competition and achieves value for money' (Open Contracting 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 Equilibrium of ProZorro Auction

In this appendix, we formally derive the equilibrium with two players in a simplified, two-round version of the ProZorro auction. Recall that in a ProZorro auction, players first simultaneously submit initial bids and then get ordered by the competitiveness of their bid.

All initial bids are revealed, and the initial loser immediately gets a chance to publicly lower their bid. Finally, the initial winner can choose to lower his bid in turn. Here, we assume that whenever bids are tied, the latest bid always wins. Specifically, if in the second round the initial loser's updated bid equals the current winning bid, then the initial loser's updated bid prevails. Similarly, when the initial loser's updated bid becomes the current winning bid, the initial winner can regain his position as the final winner by submitting a bid that is 'equal to' or better than this updated bid. In such cases, where a bidder wins by matching the current winning bid, we still refer to this action as "undercutting" the current winning bid. (Intuitively, one can think of this as the bidder offering a bid that is infinitesimally better than the current winning bid.) Finally, we assume the bid update by the winner of the initial round fails to arrive with probability p.

Our proof strategy resembles that of a trembling-hand equilibrium. We first prove equilibrium behavior in the updating round for small but positive bid submission failure probabilities p. Given these equilibrium strategies, we derive player one's ex-ante expected payoff $V_p(\tilde{c}; c_1)$, representing the payoff to a player with actual cost c_1 from imitating the cost type $\tilde{c}$ in the initial bidding phase. This gives us the equilibrium bidding function $b_p(\cdot)$ as the unique solution to the differential equation

$$\frac{\partial V_p(\tilde{c}; c_1)}{\partial \tilde{c}}\Big|_{\tilde{c}=c_1} = 0. \tag{1}$$

While this solution is hard to characterize, the left-hand side of the equation is continuous in p, which implies that for arbitrarily small p, $b_p(\cdot)$ must lie arbitrarily close to $b_0(\cdot)$, which we can (and do) characterize.

Let b_i^r refer to the bid by bidder i in round r (where $r = 1$ is the initial bidding round and $r = 2$ refers to the updating round). Furthermore let w refer to the initial winner and ℓ refer to the initial loser such that $b_\ell^1 > b_w^1$ (ties in the initial round are a zero probability event). We proceed to find a symmetric SPNE in the ProZorro auction by backward induction.

To begin with, we characterize strategies in round $r = 2$, taking initial bids as given. We start with the case of $p = 0$, illustrating why we need $p > 0$ as an equilibrium selection.

Lemma 1. *Suppose ties are broken according to Assumption 1. In any SPNE in which initial bids are placed according to a strictly increasing bidding function $b(\cdot)$, starting at a history of (b_w^1, b_ℓ^1), the initial winner and initial loser choose updating round bids according to*

$$b_w^2(b_\ell^1, b_w^1, b_\ell^2; c_w) = \begin{cases} b_w^1 & \text{if } b_\ell^2 > b_w^1, \\ b_\ell^2 & \text{if } b_\ell^2 \leq b_w^1 \text{ and } b_\ell^2 > c_w, \\ \{b : b > b_\ell^2\} & \text{if } b_\ell^2 \leq b_w^1 \text{ and } b_\ell^2 \leq c_w. \end{cases}$$

and

$$b_\ell^2(b_\ell^1, b_w^1; c_\ell) = \begin{cases} \{b : b > \hat{c}_w\} & \text{if } \hat{c}_w < c_\ell, \\ \hat{c}_w & \text{if } \hat{c}_w \geq c_\ell, \end{cases}$$

where $\hat{c}_w := b^{-1}(b_w^1)$.

Proof. We proceed by backward induction.

1. The last player to move is the initial winner. If he already has the lowest standing bid (i.e., $b_\ell^2 > b_w^1$), he leaves his bid unchanged. Otherwise, he needs to evaluate if catching up to the standing winning bid (i.e., $b_w^2 = b_\ell^2$) is profitable.

   (a) If $b_2^\ell > c_w$, setting $b_w^2 = b_\ell^2$ yields positive profits.

   (b) If $b_2^\ell = c_w$, setting $b_w^2 = b_2^\ell$ yields zero profits, so the player is indifferent. However, updating (or catching up) cannot be part of an equilibrium because, as we will see in 2b below, the initial loser has no best response to updating.[1]

   (c) If $b_2^\ell < c_w$, setting $b_w^2 = b_\ell^2$ yields negative profits. Hence, the initial winner has no bid such that he wins the auction and makes weakly positive profit, and hence is indifferent between all bids that lead to him losing the auction.[2]

   Thus:

[1] The initial loser wants to prevent this undercutting, which he can do by choosing $b_2^\ell = \hat{c}_w - \epsilon$ for arbitrarily small but strictly positive ϵ.

[2] This indeterminacy does not affect either player's payoffs.

$$
b_w^2(b_\ell^1, b_w^1, b_\ell^2; c_w) = \begin{cases} b_w^1 & \text{if } b_\ell^2 > b_w^1, \\ b_\ell^2 & \text{if } b_\ell^2 \leq b_w^1 \text{ and } b_\ell^2 > c_w, \\ \{b : b > b_\ell^2\} & \text{if } b_\ell^2 \leq b_w^1 \text{ and } b_\ell^2 \leq c_w. \end{cases} \tag{2}
$$

2. Before the initial winner moves, it is the initial loser's turn. In equilibrium, as $b(\cdot)$ is strictly increasing and hence invertible, she has a point belief of the initial winner's costs as $\hat{c}_w = b^{-1}(b_w^1)$. Anticipating $b_w^2(b_\ell^1, b_w^1, b_\ell^2; \hat{c})$, the initial loser's behavior depends on whether she believes herself to have lower costs than the initial winner. She believes any bid $b > \hat{c}_w$ will be undercut (or caught up) by the initial winner, i.e. $b_w^2(b_\ell^1, b_w^1, b; \hat{c}_w) = b$ for $b > \hat{c}_w$.

   (a) If $\hat{c}_w < c_\ell$, the bids that will be undercut include all potentially profitable bids $b \geq c_\ell$. Hence, all bids $b > \hat{c}_w$ yield zero payoff while all other bids yield negative profits and we have:

   $$
   b_\ell^2(b_\ell^1, b_w^1; c_\ell) = \{b : b > \hat{c}_w\}. \tag{3}
   $$

   Remark: The indeterminacy in strategies will affect the payoffs of the initial winner – the lower the initial loser bids, the fewer profits the initial winner makes.

   (b) If $\hat{c}_w \geq c_\ell$, the initial loser can ensure victory by bidding $\hat{c}_w$ as the initial winner will not undercut this bid. As discussed above, the initial winner only has a strict preference not to undercut for any bid strictly smaller than $\hat{c}_w$; for $b_\ell^2 = \hat{c}_w$, he is indifferent between undercutting and not undercutting. However, if the initial winner chose to break his indifference by undercutting, the initial loser would want to bid the largest real number strictly below $\hat{c}_w$, which does not exist. Thus, there cannot be an equilibrium where the initial winner chooses to undercut $\hat{c}_w$. Hence, the initial loser can ensure victory by bidding $\hat{c}_w$. As she expects that any higher bid will be undercut by the initial winner, she will hence bid $\hat{c}_w$ –

unless doing so requires her to place a bid below her cost, in which case she is indifferent between all ineffective bids:

$$b_\ell^2(b_\ell^1, b_w^1; c_\ell) = \begin{cases} \{b : b > \hat{c}_w\} & \text{if } \hat{c}_w < c_\ell, \\ \hat{c}_w & \text{if } \hat{c}_w \geq c_\ell. \end{cases} \tag{4}$$

□

Example. Suppose $b_w^1 = 1$, $b_\ell^1 = 2$ and $c_w = 0.5$. Then the initial loser knows that she can win if she bids $b_\ell^2 \leq 0.5$. Thus,

$$b_\ell^2(b_\ell^1, b_w^1; c_\ell) = \begin{cases} \{b : b > 0.5\} & \text{if } c_\ell > 0.5, \\ 0.5 & \text{if } c_\ell \leq 0.5. \end{cases} \tag{5}$$

The initial winner cannot profitably undercut $b_\ell^2 = 0.5$, so

$$b_w^2(b_\ell^1, b_w^1, b_\ell^2) = \begin{cases} 1 & \text{if } b_\ell^2 > 1, \\ b_\ell^2 & \text{if } 0.5 < b_\ell^2 \leq 1, \\ \{b : b > b_\ell^2\} & \text{if } b_\ell^2 \leq 0.5. \end{cases} \tag{6}$$

Thus, if $c_\ell \leq 0.5$, the initial loser wins at bid 0.5; but if $c_\ell > 0.5$, there is a continuum of equilibria, one for each initial loser's bid $b > 0.5$ with the initial winner winning the auction at $\min\{b, 1\}$.

Remark. While the initial winner can simply follow a "undercut if that yields a bid above my costs" rule, the initial loser needs to use $b^{-1}(\cdot)$ to infer the initial winner's cost. If that cost is sufficiently large, the initial loser can win the auction by bidding the largest bid which cannot be undercut by the initial winner. Otherwise, the initial loser is indifferent between many bids. But this indeterminacy (which affects payoffs as it affects the bid at which the initial winner wins the auction) is unrealistic: if the initial loser expects even an arbitrarily small

$p > 0$ probability of bid submission failure, he wants to submit the largest bid that would ensure victory if the initial winner did not get to respond. As we see in the next Lemma, this resolves the payoff-relevant indeterminacy.

Lemma 2. *Suppose ties are broken according to Assumption 1 and the probability of the initial winner's updated bid failing to arrive is $p > 0$. In any SPNE in which initial bids are placed according to a strictly increasing bidding function $b(\cdot)$, starting at a history of (b_w^1, b_ℓ^1), the initial winner behaves as in Lemma 1 and the initial loser chooses bids in the updating rounds according to*

$$b_\ell^2(b_\ell^1, b_w^1; c_\ell) = \begin{cases} b_w^1 & \text{if } \hat{c}_w < c_\ell \leq b_w^1, \\ \{b : b > b_w^1\} & \text{if } b_w^1 < c_\ell, \\ b_w^1 & \text{if } \hat{c}_w \geq c_\ell \text{ and } b_w^1 \geq \hat{c}_w \text{ and } \mathbb{E}[\pi_\ell | \text{gambling}] \geq \hat{c}_w - c_\ell, \\ \hat{c}_w & \text{if } \hat{c}_w \geq c_\ell \text{ and } b_w^1 \geq \hat{c}_w \text{ and } \mathbb{E}[\pi_\ell | \text{gambling}] < \hat{c}_w - c_\ell, \end{cases}$$

where $\hat{c}_w := b^{-1}(b_w^1)$, and $\mathbb{E}[\pi_\ell|\text{gambling}] \equiv [p + (1-p)\mathbf{1}\{\hat{c}_w = b_w^1\}](b_w^1 - c_\ell)$.

Remark. This amounts to firms undercutting their rival whenever they are not already winning and such a bid is above costs, with two exceptions: if a firm moves first but infers it can make a bid that its rival will not be able to beat, it will make this bid; and if a firm moves first and believes there is a sufficiently high chance that its rival will not be able to respond, it will simply bid the initial winner's bid.

Proof. We proceed by backward induction.

1. This proof of the initial winner's strategy is identical to Lemma 1, Part 1 above.

2. Before the initial winner moves, it is the initial loser's turn. In equilibrium, she has a point belief of the initial winner's costs as $\hat{c}_w = b^{-1}(b_w^1)$. Anticipating $b_w^2(b_\ell^1, b_w^1, b_\ell^2; \hat{c}_w)$, the initial loser's behavior depends on whether she believes herself to have lower costs than the initial winner.

(a) If $\hat{c}_w < c_\ell$, she believes any profitable bid $b \geq c_\ell$ will be undercut by the initial winner, i.e. $b_w^2(b_\ell^1, b_w^1, b; \hat{c}_w) = b$ as $b > \hat{c}_w$; thus, not accounting for bid submission failures, all bids yield at most zero payoff. However, with probability $p > 0$, the initial winner will not have a chance to update their bid. Thus, expecting the winner's failure on her bid, the initial loser's optimal bid is b_w^1 if $b_w^1 \geq c_\ell$ so that winning is at least weakly profitable. On the other hand, if $b_w^1 < c_\ell$, undercutting yields a bid that makes negative profits conditional on victory. Therefore the initial loser is indifferent between all ineffective bid updates, i.e., bids that leave her as the loser even if the initial winner does not update her bid. Hence, conditional on $\hat{c}_w < c_\ell$:

$$b_\ell^2(b_\ell^1, b_w^1; c_\ell) = \begin{cases} b_w^1 & \text{if } b_w^1 \geq c_\ell, \\ \{b : b > b_w^1\} & \text{if } b_w^1 < c_\ell. \end{cases} \tag{7}$$

(b) If $\hat{c}_w \geq c_\ell$, the initial loser can ensure victory by bidding $\hat{c}_w$ as the initial winner will not undercut this bid. As discussed above, the initial winner only has a strict preference not to undercut for any bid strictly smaller than $\hat{c}_w$; for $b_\ell^2 = \hat{c}_w$ he is indifferent between undercutting and not undercutting. However, if the initial winner chose to broke his indifference by undercutting, the initial loser would want to bid the largest real number strictly below $\hat{c}_w$, which does not exist. Thus, there cannot be an equilibrium where the initial winner chooses to undercut $\hat{c}_w$. Hence, the initial loser can ensure victory by bidding $\hat{c}_w$. She will do so unless a high probability of bid submission failure means that her expected payoff from just undercutting the initial winner with the minimal bid b_w^1 exceeds her payoff from securing victory at bid $\hat{c}_w$. Hence, conditional on $\hat{c}_w \geq c_\ell$:

$$b_\ell^2(b_\ell^1, b_w^1; c_\ell) = \begin{cases} b_w^1 & \text{if } b_w^1 \geq \hat{c}_w \text{ and } \mathbb{E}[\pi_\ell|\text{gambling}] \geq \hat{c}_w - c_\ell, \\ \hat{c}_w & \text{if } b_w^1 \geq \hat{c}_w \text{ and } \mathbb{E}[\pi_\ell|\text{gambling}] < \hat{c}_w - c_\ell, \end{cases} \tag{8}$$

where

$$\mathbb{E}[\pi_\ell|\text{gambling}] = [p + (1-p)\mathbf{1}\{\hat{c}_w = b_w^1\}](b_w^1 - c_\ell). \tag{9}$$

Remark: While there is a non-empty set of cost draws c_w for which the initial loser wants to gamble on a bid submission failure for any positive p, this set approaches measure-zero set as $p \to 0$ as long as $\mathbb{P}[\mathbf{1}\{\hat{c}_w = b_w^1\}] = 0$. Hence, gambling on bid submission failures will be a zero-probability event in the limit if $\mathbb{P}[\mathbf{1}\{b(c) = c\}] = 0$, for example, if $b(c) > c$ for all $c < c_{\max}$.

□

Example. Suppose $b_w^1 = 1$, $b_\ell^1 = 2$ and $c_w = 0.5$. Further assume $p < (0.5 - c_\ell)/(1 - c_\ell)$, which rules out gambling on bid submission failure.[3]Then the initial loser knows that she can win if she bids $b_\ell^2 \leq 0.5$. However, if she cannot win, there is still a chance that the initial winner will not get to update her bid. So as long as $b_w^1 = 1 \geq c_\ell$, the initial loser will still update her bid to undercut the initial winner. Thus,

$$b_\ell^2(b_\ell^1, b_w^1; c_\ell) = \begin{cases} 1 & \text{if } 0.5 < c_\ell \leq 1, \\ \{b : b > 1\} & \text{if } c_\ell > 1, \\ 0.5 & \text{if } c_\ell \leq 0.5. \end{cases} \tag{10}$$

The initial winner cannot profitably undercut $b_\ell^2 = 0.5$, so

$$b_w^2(b_\ell^1, b_w^1, b_\ell^2; c_w) = \begin{cases} 1 & \text{if } b_\ell^2 > 1, \\ b_\ell^2 & \text{if } 0.5 < b_\ell^2 \leq 1, \\ \{b : b > b_\ell^2\} & \text{if } b_\ell^2 \leq 0.5. \end{cases} \tag{11}$$

[3]If this is not satisfied, the initial case in $b_\ell^2(\cdot)$ splits into two cases: if the initial loser's costs are sufficiently close to 0.5, then for high enough bid submission failure chances it is worth it for the initial loser to just undercut the initial winner even though the initial winner could undercut the initial loser in turn: if the initial winner does not get a chance to update their bid, the initial loser wins at a high margin.

Thus, there is a p chance that only the initial loser updates her bid and (i) the initial loser wins at price 0.5 if $c_\ell \leq 0.5$, (ii) the initial loser wins at price 1 if $0.5 < c_\ell \leq 1$, (iii) the initial winner wins at price $b_w^1 = 1$ if $c_\ell > 1$. There is a $(1-p)$ chance that the initial winner updates her bid successfully, and (i) the initial loser wins at 0.5 if $c_\ell \leq 0.5$, (ii) the initial winner wins at bid 1.0 if $c_\ell > 0.5$.

We now show that, for sufficiently small bid failure probability p, there exists an equilibrium initial-round bidding function $b(\cdot)$ that satisfies the following two properties.

Condition 1. *$b(c)$ is differentiable and strictly increasing in c.*

Condition 2. *$b(c)$ satisfies $b(c) > c$ for all $c < c_{\max}$.*

To prove this, we adopt a guess-and-verify strategy. Specifically, we begin by conjecturing that the equilibrium initial-round bidding function $b(c)$ satisfies the two conditions stated above — that it is differentiable and strictly increasing, and that it lies strictly above the 45-degree line for all $c < c_{\max}$. We then derive a differential equation that any equilibrium bidding function satisfying the above two conditions must satisfy when the bid failure probability p is small but positive.

This derivation is grounded in the standard approach of computing the expected payoff $V_p(\tilde{c}; c)$ for a bidder with true cost c who deviates and bids as if her cost were $\tilde{c}$. Imposing the first-order condition for optimality, that the bidder's expected payoff is maximized when she bids truthfully, i.e., $\tilde{c} = c$, yields a differential equation characterizing $b(c)$.

Once we obtain this equation, we then verify that, for sufficiently small values of p, the solution $b(c)$ to the derived differential equation indeed satisfies the two required conditions: it remains differentiable and strictly increasing, and it satisfies $b(c) > c$ for all $c < c_{\max}$. This confirms that our conjectured form is consistent with equilibrium behavior under small p, thereby establishing the existence of the desired equilibrium bidding function.

Lemma 3. *Suppose $b(\cdot)$ is an equilibrium initial-round bidding function that satisfies Con-*

ditions 1 and 2 above. Then $b(\cdot)$ must satisfy the following differential equation:

$$b'(c) = \frac{(1-2p)f(c)(b(c)-c)}{p(1-F(b(c)))+(1-p)(1-F(c))-pf(b(c))(b(c)-c)} \tag{12}$$

for all $c \in [c_{\min}, c_{\max})$.

Proof. Let P1 refer to Player 1, and P2 to Player 2, and let c_i be the cost draw of player i. The bid submission failure probability p determines the likelihood with which the first-round winner's bid submission fails. In particular, if Player i won in the first round, with probability p she will not have an opportunity to revise her bid. On the other hand, with probability $(1-p)$, she will successfully revise her bid. In the event of a tie, we assume that the most recently submitted bid always wins. For instance, suppose in the first round the bids are (b_1^1, b_2^1) with $b_1^1 < b_2^1$[4]. Then, if the loser, P2, updates her bid to $b_2^2 = b_1^1$, she will win in the second round if P1 fails to revise her bid (with probability p). However, if Player 1 does have the opportunity to revise, she can outbid Player 2 and retain the win by "updating" her bid to $b_1^2 = b_1^1$.

To establish the differential equation stated in the lemma, we consider a bidding function $b(c)$ that satisfies Conditions 1 and 2 — namely, that it is differentiable and strictly increasing, and that it lies strictly above the 45-degree line on $[c_{\min}, c_{\max})$. We then consider the expected payoff $V_p(\tilde{c}; c_1)$ of P1, whose true cost is c_1, when she deviates and bids as if her cost were $\tilde{c}$. Note that from P1's perspective, $c_2 \sim F$ is a random variable that must be integrated over when computing her expected payoff, whereas c_1 and $\tilde{c}$ are known scalars. Taking the first-order condition for optimality at $\tilde{c} = c_1$ and differentiating $V_p(\tilde{c}; c_1)$ with respect to $\tilde{c}$, we obtain the differential equation that such a function $b(c)$ must satisfy in equilibrium.

We now characterize $V_p(\tilde{c}; c_1)$ by splitting it into two cases: when $\tilde{c} < c_1$, meaning she bids less than prescribed by $b(\cdot)$, and when $c_1 < \tilde{c}$, meaning she bids more than prescribed.

1. Case 1: $\tilde{c} < c_1$.

[4] Under a strictly increasing bidding strategy, which we will assume shortly, ties in the first round occur with probability zero.

(a) Case 1A: $c_2 < \tilde{c}$. Hence, P1 is the initial loser, and she (correctly) infers that she has a cost disadvantage. Therefore, she can win (and benefit) only when the initial winner (P2) fails to update (with probability p), and bidding $b(c_2)$ is still beneficial, i.e., $b(c_2) \geq c_1$. The contribution to the expected payoff for P1 from this case is then given by the probability of this case occurring, $F(\tilde{c})$, multiplied by the payoff that she expects if she finds herself in this situation, $V^A(\tilde{c}; c_1)$, i.e.,

$$\begin{aligned} F(\tilde{c}) \times V^A(\tilde{c}; c_1) &= F(\tilde{c}) \times p \times \mathbb{E}[\mathbf{1}(c_1 \leq b(c_2)) \times (b(c_2) - c_1) | c_2 < \tilde{c}] \\ &= F(\tilde{c}) \times p \times \int_{b^{-1}(c_1)}^{\tilde{c}} (b(s) - c_1) \frac{f(s)}{F(\tilde{c})} ds \\ &= p \times \int_{b^{-1}(c_1)}^{\tilde{c}} (b(s) - c_1) f(s) ds \end{aligned} \tag{13}$$

(b) Case 1B: $c_2 > \tilde{c}$. Hence, P1 is the initial winner. P2 thinks she cannot win if P1 can update her bid in the second round. Thus, P2 conditions on an update failure. As a result, P2 will try to beat P1 in the second round when $b(\tilde{c}) \geq c_2$. So P1 loses only when (i) she fails to update her bid, and (ii) $b(\tilde{c}) \geq c_2$. Thus, the contribution to P1's expected payoff from this case is then given by the probability of this case occurring, $1 - F(\tilde{c})$, multiplied by the payoff that she expects if she finds herself in this situation, $V^B(\tilde{c}; c_1)$, i.e.,

$$\begin{aligned} (1 - F(\tilde{c})) \times V^B(\tilde{c}; c_1) &= (1 - F(\tilde{c})) \times (1 - p) \times (b(\tilde{c}) - c_1) \\ &\quad + (1 - F(\tilde{c})) \times p \times \mathbb{P}[c_2 > b(\tilde{c}) | c_2 > \tilde{c}] \times (b(\tilde{c}) - c_1) \\ &= (1 - p) \times (1 - F(\tilde{c})) \times (b(\tilde{c}) - c_1) \\ &\quad + p \times (1 - F(b(\tilde{c}))) \times (b(\tilde{c}) - c_1) \end{aligned} \tag{14}$$

We can now combine the expected payoff contributions from these Cases 1A and 1B to yield an overall payoff that is obtained if $\tilde{c} < c_1$:

$$
\begin{aligned}
V_p(\tilde{c}; c_1) &= F(\tilde{c}) \times V^A(\tilde{c}; c_1) + (1 - F(\tilde{c})) \times V^B(\tilde{c}; c_1) \\
&= p \times \int_{b^{-1}(c_1)}^{\tilde{c}} (b(s) - c_1) f(s) ds + (1 - p) \times (1 - F(\tilde{c})) \times (b(\tilde{c}) - c_1) \\
&\quad + p \times (1 - F(b(\tilde{c}))) \times (b(\tilde{c}) - c_1). \qquad (15)
\end{aligned}
$$

This yields a (left) derivative of

$$
\begin{aligned}
\frac{\partial V_p(\tilde{c}; c_1)}{\partial \tilde{c}} &= p(b(\tilde{c}) - c_1) f(\tilde{c}) \\
&\quad -(1 - p) f(\tilde{c})(b(\tilde{c}) - c_1) + (1 - p)(1 - F(\tilde{c})) b'(\tilde{c}) \\
&\quad -p f(b(\tilde{c}))(b(\tilde{c}) - c_1) b'(\tilde{c}) + p(1 - F(b(\tilde{c}))) b'(\tilde{c}). \qquad (16)
\end{aligned}
$$

2. Case 2: $c_1 < \tilde{c}$. In this case, we can divide the analysis into three subcases.

   (a) Case 2C: $c_2 < c_1 < \tilde{c}$. Then P2 is the initial winner, and P1 is the initial loser. P1's initial bid $b(\tilde{c})$ does not affect her continuation strategy — the only potentially viable second round bid is $b(c_2)$ — and hence perturbing $\tilde{c}$ does not affect her payoffs. In other words, the conditional value $V^C(\tilde{c}; c_1)$ is constant in $\tilde{c}$ over the range $c_1 < \tilde{c}$, and the payoff contribution from this case is:

$$
\begin{aligned}
F(c_1) \times V^C(\tilde{c}; c_1) &= F(c_1) \times p \times \mathbb{E}[\mathbf{1}\{c_1 \leq b(c_2)\}(b(c_2) - c_1) | c_2 < c_1] \\
&\quad +(1 - p) \times \mathbb{E}[\mathbf{1}\{c_1 \leq b(c_2)\}\mathbf{1}\{c_2 \geq b(c_2)\}(b(c_2) - c_1) | c_2 < c_1].
\end{aligned}
$$

   (b) Case 2D: $c_1 < c_2 < \tilde{c}$. Then P2 is the initial winner, but P1 knows she has lower costs than P2. She could bid P2's cost c_2 for certain victory, or she could gamble on a bid submission failure on P2's side and bid $b(c_2)$. Let $G := \{c_2 : [p + (1 - p)\mathbf{1}\{c_2 \geq b(c_2)\}](b(c_2) - c_1) \geq c_2 - c_1\}$, the range of c_2 for which gambling for bid submission failure could benefit P1. Since, (for $c_1 \neq c_{\max}$), $p(b(c_1) - c_1) > c_1 - c_1 = 0$, $c_1 \in int(G)$, i.e., there exists $\eta > 0$, such that

$(c_1 - \eta, c_1 + \eta) \subset G$. Then, for $\tilde{c}$ close to c_1,

$$\begin{aligned}(F(\tilde{c}) - F(c_1)) \times V^D(\tilde{c}; c_1) &= \mathbb{P}[G^c \cap \{c_1 < c_2 < \tilde{c}\}] \times \mathbb{E}[c_2 - c_1 | G^c \cap \{c_1 < c_2 < \tilde{c}\}] \\ &\quad + \mathbb{P}[G \cap \{c_1 < c_2 < \tilde{c}\}] \times p \times \mathbb{E}[b(c_2) - c_1 | G \\ &\quad \cap \{c_1 < c_2 < \tilde{c}\}] \\ &= \int_{G^c \cap \{c_1 < c_2 < \tilde{c}\}} (s - c_1) f(s) ds \\ &\quad + p \int_{G \cap \{c_1 < c_2 < \tilde{c}\}} (b(s) - c_1) f(s) ds \\ &= p \int_{c_1}^{\tilde{c}} (b(s) - c_1) f(s) ds \qquad (17)\end{aligned}$$

(c) Case 2E: $c_1 < \tilde{c} < c_2$. In this case, P1 is the initial winner, but she infers that she has a sufficient cost advantage. Then, analogously to Case 1B,

$$\begin{aligned}(1 - F(\tilde{c})) \times V^E(\tilde{c}; c_1) &= (1 - F(\tilde{c})) \times (1 - p) \times (b(\tilde{c}) - c_1) \\ &\quad + (1 - F(\tilde{c})) \times p \times \mathbb{P}[c_2 > b(\tilde{c}) | c_2 > \tilde{c}] \times (b(\tilde{c}) - c_1) \\ &= (1 - p) \times (1 - F(\tilde{c})) \times (b(\tilde{c}) - c_1) + p \times (1 - F(b(\tilde{c}))) \\ &\quad \times (b(\tilde{c}) - c_1). \qquad (18)\end{aligned}$$

Therefore, if $c_1 < \tilde{c}$ and $\tilde{c}$ sufficiently close to c_1,

$$\begin{aligned}V_p(\tilde{c}; c_1) &= F(c_1) \times V^C(\tilde{c}; c_1) + (F(\tilde{c}) - F(c_1)) \times V^D(\tilde{c}; c_1) + (1 - F(\tilde{c})) \times V^E(\tilde{c}; c_1) \\ &= F(c_1) \times V^C(\tilde{c}; c_1) + p \times \int_{c_1}^{\tilde{c}} (b(s) - c_1) f(s) ds + (1 - p) \times (1 - F(\tilde{c})) \times (b(\tilde{c}) - c_1) \\ &\quad + p \times (1 - F(b(\tilde{c}))) \times (b(\tilde{c}) - c_1), \qquad (19)\end{aligned}$$

where recall $V^C(\tilde{c}; c_1)$ is constant with respect to $\tilde{c}$.

Then, if $\tilde{c}$ is sufficiently close to c_1, the right derivative matches with the left derivative:

$$\begin{aligned}\frac{\partial V_p(\tilde{c};c_1)}{\partial \tilde{c}} &= p(b(\tilde{c})-c_1)f(\tilde{c}) \\ &\quad -(1-p)f(\tilde{c})(b(\tilde{c})-c_1)+(1-p)(1-F(\tilde{c}))b'(\tilde{c}) \\ &\quad -pf(b(\tilde{c}))(b(\tilde{c})-c_1)b'(\tilde{c})+p(1-F(b(\tilde{c})))b'(\tilde{c}). \end{aligned} \tag{20}$$

As $b(\cdot)$ is an optimal first-round bidding strategy for both players, $b(c_1)$ has to maximize $V_p(\tilde{c};c_1)$. Note that $V_p(\tilde{c};c_1)$ is continuously differentiable as long as $F(\cdot)$ is continuously differentiable. Then, we have the following first-order condition for all $c_1 \in [c_{\min}, c_{\max}]$:

$$\left.\frac{\partial V_p(\tilde{c};c_1)}{\partial \tilde{c}}\right|_{\tilde{c}=c_1} = 0. \tag{21}$$

Substituting the previously derived expression for the derivative and reorganizing, we obtain the following equation for $c = c_1$:

$$b'(c) = \frac{(1-2p)f(c)(b(c)-c)}{p(1-F(b(c)))+(1-p)(1-F(c))-pf(b(c))(b(c)-c)}. \tag{22}$$

□

Lemma 4. *Let $b(c)$ be a solution to the differential equation* (12) *with the boundary condition $b(c_{\max}) = c_{\max}$. Then there exists $\underline{p} > 0$ such that, for all $p < \underline{p}$, the function $b(c)$ satisfies Conditions 1 and 2; that is, $b(c)$ is differentiable and strictly increasing on $[c_{\min}, c_{\max})$, and it satisfies $b(c) > c$ for all $c < c_{\max}$.*

Proof. We now show that, for sufficiently small values of p, the function $b(\cdot)$, defined as the solution to the differential equation (1) with boundary condition $b(c_{\max}) = c_{\max}$, satisfies Conditions 1 and 2: (i) $b(\cdot)$ is strictly increasing on $[c_{\min}, c_{\max})$, i.e., $b'(c) > 0$ almost everywhere; and (ii) $b(c) > c$ for all $c < c_{\max}$.

We begin with (ii), showing that the bidding function lies strictly above the 45-degree line on the interval $[c_{\min}, c_{\max})$. If $b(c) = c$, then, from the differential equation, $b'(c)$ must be zero. However, $b(c) = c$ and $b'(c) = 0$ can hold only at $c_{\max}$; otherwise, for small $\epsilon > 0$, we would have $b(c+\epsilon) < c+\epsilon$, which must be strictly dominated: it is always better to bid your true cost $c+\epsilon$ than to bid something strictly below it, namely $b(c+\epsilon)$. This confirms that $b(c) > c$ for all $c < c_{\max}$, i.e., (ii).

To prove (i), that $b(\cdot)$ is strictly increasing, now we need to show that the denominator of equation (12) is uniformly positive for sufficiently small p. By dividing both numerator and denominator by $(b(c)-c)$ (which we just found is positive in the preceding paragraph), we have

$$b'(c) = \frac{(1-2p)f(c)}{\frac{p(1-F(b(c)))+(1-p)(1-F(c))}{b(c)-c} - pf(b(c))} \tag{23}$$

Note that, by applying L'Hopital's rule,

$$\begin{aligned}\lim_{c\to c_{\max}} \frac{p(1-F(b(c)))+(1-p)(1-F(c))}{b(c)-c} &= \lim_{c\to c_{\max}} \frac{-pf(b(c))b'(c)-(1-p)f(c)}{b'(c)-1} \\ &= f(c_{\max}) \cdot \frac{-p(b'(c_{\max})-1)-1}{b'(c_{\max})-1} \\ &= f(c_{\max}) \cdot \left(-p - \frac{1}{b'(c_{\max})-1}\right), \end{aligned} \tag{24}$$

Plugging this into (23) evaluated at $c = c_{\max}$ and noting $f(b(c_{\max})) = f(c_{\max})$, we have

$$b'(c_{\max}) = \frac{(1-2p)}{-p - \frac{1}{b'(c_{\max})-1} - p}. \tag{25}$$

Rearranging to isolate $b'(c_{\max})$ yields

$$b'(c_{\max}) = \frac{\sqrt{1-2p} - (1-2p)}{2p} > 0, \tag{26}$$

which is decreasing in p over $[0, \frac{1}{2}]$, with $b'(c_{\max}) = \frac{1}{2}$ at $p = 0$.[5]

Note that, as $b'(c_{\max}) = \sqrt{2}-1$ at $p = \frac{1}{4}$, $b'(c_{\max}) > \sqrt{2}-1$ for all $p < \frac{1}{4}$. Then, by the continuity of $b'(\cdot)$ (which follows from the continuity of $b(\cdot)$ and $f(\cdot)$ given (12)[6]) there

[5] The quadratic equation also gives us another solution with $b'(c_{\max}) < 0$. But this case is not feasible since, if $b'(c_{\max}) < 0$, then $b(c_{\max}-\epsilon) > c_{\max}$ for small $\epsilon > 0$, but strategies bidding more than $c_{\max}$ have to be dominated in equilibrium.

[6] Note that $b(\cdot)$ is continuous as it is differentiable.

exists $\delta > 0$ such that, $b'(c) > \frac{\sqrt{2}-1}{2} > 0$, for all $c \in (c_{\max} - \delta, c_{\max}]$ and $p < \frac{1}{4}$.

Finally, note that, for any $c \in [c_{\min}, c_{\max} - \delta]$, the denominator of (12) satisfies

$$\begin{aligned} p(1-F(b(c))) &+ (1-p)(1-F(c)) - pf(b(c))(b(c)-c) \\ &\geq (1-p)(1-F(c)) - pf(b(c))(b(c)-c) \\ &\geq (1-F(c)) - p\big(1-F(c) + f(b(c))(b(c)-c)\big) \\ &\geq 1 - F(c_{\max}-\delta) - p\big(1 + \max_s f(s) b(c_{\max}-\delta)\big). \end{aligned} \tag{27}$$

The last expression is positive if

$$p < p^{\dagger} = \frac{1 - F(c_{\max} - \delta)}{1 + \max_s f(s) b(c_{\max} - \delta)}. \tag{28}$$

Then, for $p < p^{\dagger}$, the denominator, $p(1-F(b(c)))+(1-p)(1-F(c))-pf(b(c))(b(c)-c)$ is uniformly positive for all $c \in [c_{\min}, c_{\max}-\delta]$, which yields $b'(c) > 0$ for all $c \in [c_{\min}, c_{\max}-\delta]$. Then, for $p < \underline{p} := \min\{p^{\dagger}, \frac{1}{4}\}$, $b'(c) > 0$, for all $c \in [c_{\min}, c_{\max}]$, which proves the presumption of differentiable strictly increasing $b(\cdot)$ for small p less than $\underline{p}$.

□

Lemma 3 and Lemma 4 together establish that, for sufficiently small p, there exists an equilibrium bidding function $b(c)$ which (i) is differentiable and strictly increasing, and (ii) satisfies $b(c) > c$ for all $c < c_{\max}$:

Proposition 1. *There exists a threshold $\underline{p} > 0$ such that, if the probability of bid submission failure by the first-round winner p satisfies $p < \underline{p}$, then there exists an equilibrium in which the first-round bidding function $b(c)$ is differentiable and strictly increasing.*

Finally, we characterize this bidding function $b(\cdot)$. While this function is hard to characterize for $p > 0$, for sufficiently small p it must lie arbitrarily close to the bidding function familiar from a first-price sealed-bid auction as we now prove.

Proposition 2. *Let $c_i \sim F(\cdot)$ with $F(\cdot)$ continuously differentiable and $F(c_{max}) = 1$. For arbitrarily small p, the ProZorro auction (with one updating round and two players) has a unique PBE in which initial bids are given by a strictly increasing $b(\cdot)$. This bidding function $b(\cdot)$ is arbitrarily close to*

$$b(c) = \frac{1}{[1 - F(c)]} \int_c^{c_{max}} sdF(s).$$

Proof. For sufficiently small p, the equilibrium initial round bidding function $b_p(\cdot)$ is the unique solution to

$$\frac{\partial V_p(\tilde{c}; c_1)}{\partial \tilde{c}}\bigg|_{\tilde{c}=c_1} = 0, \tag{29}$$

for all c_1. The solution to a parameterized differential equation is continuous in these parameters if the differential equation is continuously differentiable in the parameters (Perko, 1996, p.84).

Hence, the solution for small p must lie close to the $b_0(\cdot)$ which solves the equation at $p = 0$ for all c_1. Note that, at $p = 0$, equation (12) simplifies to

$$b_0'(c) = \frac{f(c)(b_0(c) - c)}{1 - F(c)} \Leftrightarrow (1 - F(c))b_0'(c) - (b_0(c) - c)F'(c) = 0 \tag{30}$$

Together with the boundary condition $b(c_{\max}) = c_{max}$, this differential equation is uniquely solved by

$$b_0(c) = \frac{1}{1 - F(c)} \int_c^{c_{max}} sdF(s), \tag{31}$$

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

□

C 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, leading to the subsequent disqualification of the bids. Because 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 against 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 and 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 not to explicitly model them.

D Detection Validation Robustness

We now conduct a series of robustness checks to validate our detection procedure in Table 2 in the main text. First, Table A.1 investigates what happens if we add fixed effects to the regression that generates the pairwise fixed effects. Third, Table A.2 investigates what happens if we add higher-order inverse bid polynomials to the regression that generates the pairwise fixed effects. Fourth, Table A.3 investigates what happens if we only use the subset of firm pairs that are observed co-bidding frequently. Across all robustness checks and the main analysis, we generate 52 total measures of the ROC AUC, all of which are statistically significantly above 0.5.

We also plot the inverse bid function in Figure A.1.

Table A.1: Detection Validation Robustness: Adding Fixed Effects

	Firm FE: No				**Firm FE: Yes**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Detection Quality							
ROC AUC	0.74	0.73	0.74	0.72	0.63	0.63	0.63	0.63
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N Pairs	237,127	237,127	237,127	237,127	226,021	226,021	226,021	226,021
	Panel B: Control Variables							
Initial Bid	−5.38	−5.42	−5.36	−5.30	−5.38	−5.42	−5.36	−5.30
	(0.55)	(0.55)	(0.55)	(0.53)	(0.55)	(0.55)	(0.55)	(0.53)
Initial Bid Squared	6.87	6.92	6.86	6.69	6.87	6.92	6.86	6.69
	(0.73)	(0.72)	(0.72)	(0.70)	(0.73)	(0.72)	(0.72)	(0.70)
Initial Bid Cubed	−3.33	−3.36	−3.33	−3.22	−3.33	−3.36	−3.33	−3.22
	(0.31)	(0.31)	(0.31)	(0.30)	(0.31)	(0.31)	(0.31)	(0.30)
Log Value	0.68	1.46	1.61	1.53	0.68	1.46	1.61	1.53
	(0.09)	(0.26)	(0.26)	(0.26)	(0.09)	(0.26)	(0.26)	(0.26)
# Projects Bid On	0.43	0.44	0.51	0.55	0.43	0.44	0.51	0.55
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Avg Final Bid	0.21	0.21	0.23	0.22	0.21	0.21	0.23	0.22
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
# Projects Awarded	−0.76	−0.77	−0.75	−0.78	−0.76	−0.77	−0.75	−0.78
	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.09)
Procurer-Tenderer Dist. (km)	−0.04	−0.04	−0.04	−0.05	−0.04	−0.04	−0.04	−0.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	1.03	0.94	0.89	0.90	1.03	0.94	0.89	0.90
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Industry FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Day-of-Week FE	No	No	Yes	Yes	No	No	Yes	Yes
Hour-of-Day FE	No	No	Yes	Yes	No	No	Yes	Yes
Cost Decile FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Procuring Entity FE	No	No	No	Yes	No	No	No	Yes
Firm FE	No	No	No	No	Yes	Yes	Yes	Yes
N Auctions	305,556	305,556	305,556	305,556	305,556	305,556	305,556	305,556
R^2	0.5667	0.5681	0.5687	0.5913	0.5667	0.5681	0.5687	0.5913

Notes: Panel A reports the area under the ROC curve (AUC) when using p-values from pairwise fixed effects as a continuous predictor of collusion penalization. An ROC AUC of 1.0 indicates perfect prediction; 0.5 indicates random guessing. Since AUCs are consistently and significantly above 0.5, the pairwise fixed effects contain information about the likelihood of penalization. Each column reflects a different regression specification used to estimate the pairwise fixed effects, adding controls from left to right. Panel B reports coefficients from these regressions. "Log Value" is the log of the engineer's estimate of the good's value. "# Projects Bid On," "Avg Final Bid," and "# Projects Awarded" measure bidder characteristics in the 180 days before the auction. "Procurer–Tenderer Dist." is the distance between the procurer's and bidder's addresses. Firm fixed effects are collinear with pairwise fixed effects, so specifications (5)-(8) mirror (1)-(4) but remove firm FE from the pairwise FE, changing only the ROC AUC in Panel A, not the coefficients in Panel B.

Table A.2: Detection Validation Robustness: Higher-Order Inverse Bid Polynomials

	Firm FE: No			Firm FE: Yes		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Detection Quality						
ROC AUC	0.72	0.71	0.71	0.63	0.63	0.63
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N Pairs	237,127	237,127	237,127	226,021	226,021	226,021
Panel B: Control Variables						
Initial Bid	−5.26	11.82	−21.76	−5.26	11.82	−21.76
	(0.52)	(1.17)	(1.73)	(0.52)	(1.17)	(1.73)
Initial Bid Squared	6.59	−40.10	114.78	6.59	−40.10	114.78
	(0.68)	(2.76)	(7.06)	(0.68)	(2.76)	(7.06)
Initial Bid Cubed	−3.15	47.36	−246.25	−3.15	47.36	−246.25
	(0.29)	(2.78)	(12.79)	(0.29)	(2.78)	(12.79)
Initial Bid Forth Order		−19.05	230.77		−19.05	230.77
		(1.01)	(10.66)		(1.01)	(10.66)
Initial Bid Fifth Order			−79.19			−79.19
			(3.35)			(3.35)
Constant	1.26	−0.60	1.07	1.26	−0.60	1.07
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Firm FE	No	No	No	Yes	Yes	Yes
N Auctions	305,556	305,556	305,556	305,556	305,556	305,556
R^2	0.5907	0.5917	0.5928	0.5907	0.5917	0.5928

Notes: Panel A reports the area under the ROC curve (AUC) when using p-values from pairwise fixed effects as a continuous predictor of collusion penalization. An ROC AUC of 1.0 indicates perfect prediction; 0.5 indicates random guessing. Since AUCs are consistently and significantly above 0.5, the pairwise fixed effects contain information about the likelihood of penalization. Each column reflects a different regression specification used to estimate the pairwise fixed effects, adding controls from left to right. Panel B reports coefficients from these regressions. Firm fixed effects are collinear with pairwise fixed effects, so specifications (4)-(6) mirror (1)-(3) but remove firm FE from the pairwise FE, changing only the ROC AUC in Panel A, not the coefficients in Panel B.

Table A.3: Detection Validation Robustness: Frequent Co-Bidding

	Firm FE: No				**Firm FE: Yes**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample: Pairs w/ $\geq$ X auctions	≥ 3	≥ 5	≥ 10	≥ 20	≥ 3	≥ 5	≥ 10	≥ 20
Panel A: Detection Quality								
ROC AUC	0.73	0.74	0.71	0.71	0.60	0.61	0.62	0.65
	(0.03)	(0.03)	(0.05)	(0.12)	(0.02)	(0.03)	(0.04)	(0.07)
N Pairs	36,805	17,574	6,589	2,418	36,320	17,404	6,562	2,413
Panel B: Control Variables								
Initial Bid	−6.00	−5.73	−6.19	−5.87	−6.00	−5.73	−6.19	−5.87
	(0.60)	(0.63)	(0.84)	(1.02)	(0.60)	(0.63)	(0.84)	(1.02)
Initial Bid Squared	8.23	7.86	8.48	7.93	8.23	7.86	8.48	7.93
	(0.80)	(0.84)	(1.12)	(1.35)	(0.80)	(0.84)	(1.12)	(1.35)
Initial Bid Cubed	−4.15	−4.01	−4.27	−3.99	−4.15	−4.01	−4.27	−3.99
	(0.34)	(0.37)	(0.48)	(0.59)	(0.34)	(0.37)	(0.48)	(0.59)
Log Value	1.45	1.48	1.79	1.96	1.45	1.48	1.79	1.96
	(0.27)	(0.30)	(0.35)	(0.44)	(0.27)	(0.30)	(0.35)	(0.44)
Constant	1.17	1.10	1.17	1.10	1.17	1.10	1.17	1.10
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
Firm FE	No	No	No	No	Yes	Yes	Yes	Yes
N Auctions	208,558	171,634	127,858	91,922	208,558	171,634	127,858	91,922
R^2	0.4355	0.4167	0.4075	0.4126	0.4355	0.4167	0.4075	0.4126

Notes: Panel A reports the area under the ROC curve (AUC) when using p-values from pairwise fixed effects as a continuous predictor of collusion penalization. An ROC AUC of 1.0 indicates perfect prediction; 0.5 indicates random guessing. Since AUCs are consistently and significantly above 0.5, the pairwise fixed effects contain information about the likelihood of penalization. Each column reflects a different regression specification used to estimate the pairwise fixed effects, adding controls from left to right. Panel B reports coefficients from these regressions. "Log Value" is the log of the engineer's estimate of the good's value. Firm fixed effects are collinear with pairwise fixed effects, so specifications (1)-(4) mirror (5)-(8) but remove firm FE from the pairwise FE, changing only the ROC AUC in Panel A, not the coefficients in Panel B.

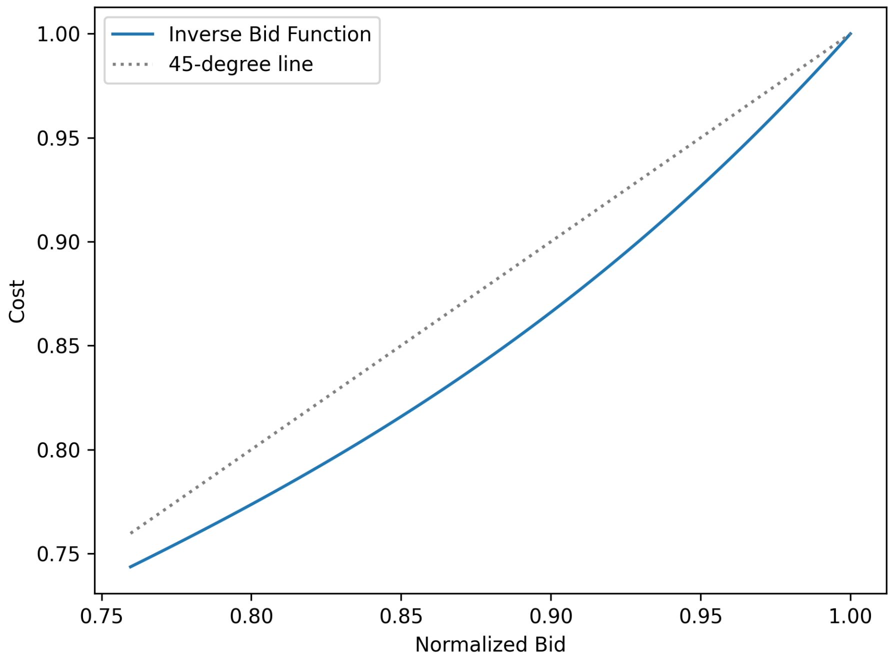


Figure A.1: Inverse Bid Function

Notes: This figure plots the inverse bid function $\phi(\cdot)$ implied by our estimates in the main analysis, i.e., Column 1 (and equivalently 7) of Table 2 and the restriction that $\phi(1) = 1$ (we need this restriction to distinguish a constant in $\phi(\cdot)$ from a market-wide tendency to undercut). We exhibit the function for the range of bids that are observed in the data, i.e., we restrict the horizontal axis to lie between the 5th and 95th percentiles of the distribution of normalized bids, which are 0.76 and 1, respectively. The function is strictly increasing and lies below the 45-degree line, as required by theory.

E Cost Analysis Robustness

E.1 Balance Test

In Section 5.5, we examine the correlation between normalized prices and a dummy variable indicating whether a firm pair identified as collusive by our algorithm participated in a given auction. To rule out that our findings are driven by observable omitted covariates, we verify that the characteristics of auctions with and without collusive participants are similar on observables once we condition on enough fixed effects (see Table A.4). While there are significant differences in many variable in the non-conditional comparison, in the conditional comparison, the only remaining significant difference is in the logarithm of the engineer's

		Control		Treatment – Control			
				Unconditional		Conditional	
	Obs.	Mean	SD	Diff.	P-value	Diff.	P-value
Log Value	308,566	11.7962	2.1829	0.0107	< 0.0001	0.0077	0.0014
Defense	308,566	0.0175	0.1311	0.0006	< 0.0001	0.0003	0.4806
No Bids	308,566	2.4696	0.8671	0.0043	< 0.0001	.	.
Construction	308,566	0.1023	0.3031	0.0015	0.1229	.	.
Capital	308,566	0.0402	0.1964	0.0009	< 0.0001	.	.
Below Threshold	308,566	0.5537	0.4971	0.0024	< 0.0001	.	.

Table A.4: Balance Test - Auctions with and without Collusive Participants

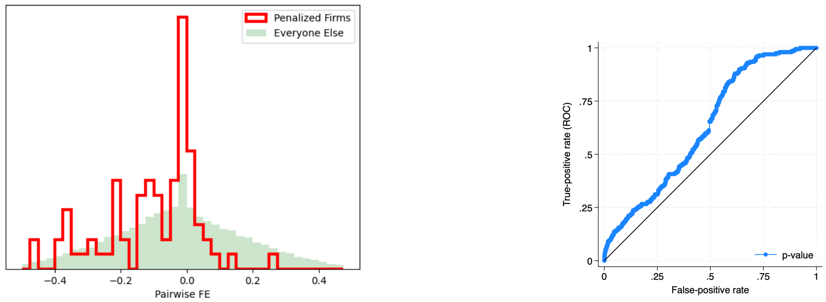
Notes: We compare the characteristics of procurement auctions with and without collusive participants, both before and after conditioning on the same set of fixed effects as in the most saturated cost regression analysis (i.e., industry, year, procurer, and firm fixed effects as well as the control for the number of bids and the below the threshold procedure). While the differences between the two groups are statistically significant before conditioning on fixed effects, they are not statistically significant after conditioning on fixed effects.

estimate of the value of the contract, which we use to normalize the dependent variable. This provides reassurance that the findings in the cost analysis are not driven by these observable differences.

E.2 Including Bidder Fixed Effects

To further check robustness of our findings, we rerun the main analysis while including bidder fixed effects in Equation 5. In Figure A.2, we present the main findings. First, we show the distributions of pairwise fixed effects separately for the pairs penalized by the Antimonopoly Committee of Ukraine for collusion and those that were not penalized (Figure A.2a). And second, we show the relationship between q-values (i.e., the measure of being detected as a colluder by our algorithm) and the penalization status (Figure A.2b). In this estimation, we filter out any firm-specific average inactivity. Both figures are substantively similar to the findings in the main analysis.

Finally, we use the new sample of auctions detected as collusive in the estimation including bidder fixed effects and rerun the cost analysis as in Section 5.5. The findings are reported in Tables A.5 and A.6. In our preferred specification (Column (5)), the point



(a) Histogram by Penalization Status (b) Receiver Operating Characteristic Curve

Figure A.2: Pairwise Fixed Effects & Statistical Test For Collusion with Bidder FEs
Notes: In A.2a, we show the distribution of the pairwise fixed effects of (5) broken up by whether a pair was penalized in a collusion case by the Antimonopoly Committee of Ukraine or not and plot the distributions of pairwise FE for penalized and non-penalized firms separately. In A.2b, we show the receiver operating characteristic curve measuring the quality of the p-values of the pairwise fixed effects as a signal of being penalized for collusion.

estimates for the association between cartel participation and costs are qualitatively similar in magnitude and remain strongly significant, as in the main analysis without firm fixed effects.

F Network structure

To illustrate the network structure, we plot the links for a subsample of firms supplying procurement contracts in the Common Procurement Vocabulary (CPV) category of "laboratory instruments for checking physical characteristics." We define suppliers in this category as firms that, over the full sample period, obtain the majority of their procurement contracts in this CPV category. To focus on the structure of links between potentially collusive firms, the sample is restricted to firm pairs with q-values below 0.05. The resulting network is shown in Figure A.3.

	(1) Normalized Price	(2) Normalized Price	(3) Normalized Price	(4) Normalized Price	(5) Normalized Price
Cartel Participation	0.00298	0.0143	0.0139	0.0145	0.0185
	(0.0018)	(0.0020)	(0.0019)	(0.0012)	(0.0014)
No of Bids	-0.0219	-0.0229	-0.0228	-0.0212	-0.0181
	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Below-the-Threshold	-0.0501	-0.0385	-0.0371	-0.0271	-0.00720
	(0.0012)	(0.0012)	(0.0011)	(0.0011)	(0.0011)
Industry FE	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
Procurer FE	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes
N	291368	291368	291368	291368	291368

Table A.5: Cartel Participation and Normalized Price (Adding Firm Fixed Effects In Detection Regressions)

Notes: The dependent variable is the ratio of the final price of a contract over the estimated cost. The main independent variable $Cartel\ Participation$ is a dummy variable equal to one if a pair of firms is identified as (statistically significantly) collusive by the version of our algorithm that uses firm fixed effects as a control; otherwise, zero. $No\ Bids$ is the number of firms that participated in the auction at the auction. $Below\ the\ Threshold$ is a dummy variable equal to one if the auction's estimated cost is below the national Ukrainian threshold; otherwise, zero. The estimates are obtained as means from the 10-fold cross-validation procedure. The number of observations is lower than in the full sample, as we fix the observations across all five specifications to those used in the most saturated specification in Column (5). Standard errors clustered at the procuring entity level are in parentheses. The analysis is identical to the analysis in Table 3 except that firm fixed effects are included as a control in the detection regression.

	(1) Normalized Price	(2) Normalized Price	(3) Normalized Price	(4) Normalized Price	(5) Normalized Price
Cartel Participation	-0.0111	0.00367	0.00323	0.00484	0.0106
	(0.0022)	(0.0023)	(0.0022)	(0.0016)	(0.0018)
Cartel Alone	0.0198	0.0148	0.0150	0.0136	0.0116
	(0.0019)	(0.0018)	(0.0018)	(0.0018)	(0.0019)
No of Bids	-0.0213	-0.0224	-0.0223	-0.0208	-0.0178
	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Below-the-Threshold	-0.0500	-0.0384	-0.0370	-0.0270	-0.00710
	(0.0012)	(0.0012)	(0.0011)	(0.0011)	(0.0011)
Industry FE	No	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
Procurer FE	No	No	No	Yes	Yes
Firm FE	No	No	No	No	Yes
N	291368	291368	291368	291368	291368

Table A.6: Cartel Participation with and without Competitors (Adding Firm Fixed Effects In Detection Regressions)

Notes: The dependent variable is the ratio of the final price of a contract over the estimated cost. The main independent variables are $Cartel\ Participation$, which is a dummy variable equal to one if a pair of firms identified as (statistically significantly) collusive by the version of our algorithm that uses firm fixed effects as a control, otherwise 0, and $Cartel\,Alone$ is a term indicating whether a cartel is alone in the procurement auction. $No\ Bids$ is the number of firms that participated in the auction at and $Below\ the\ Threshold$ is a dummy variable equal to one if the auction's estimated cost is below the national Ukrainian threshold, otherwise zero. The estimates are obtained as means from the 10-fold cross-validation procedure. The number of observations is lower than in the full sample, as we fix the observations across all five specifications to those used in the most saturated specification in Column (5). Standard errors clustered at the procuring entity level are in parentheses. The analysis is identical to the analysis in Table 4 except that firm fixed effects are included as a control in the detection regression.

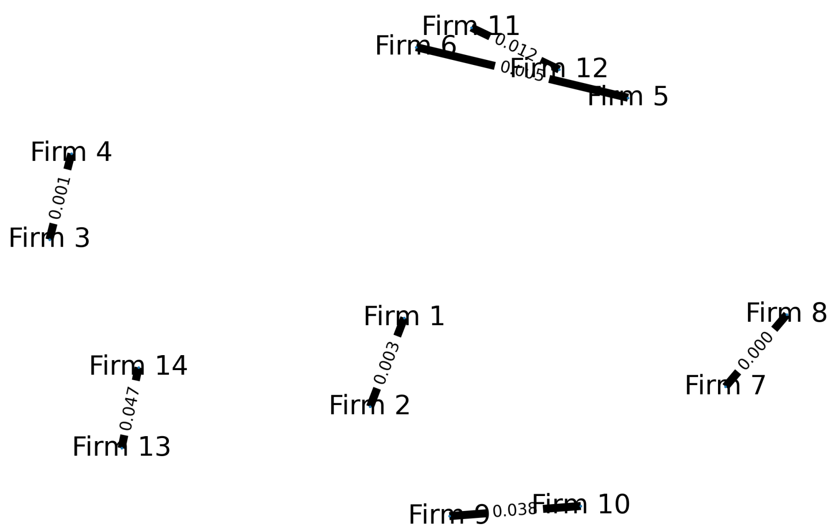


Figure A.3: Illustration of cartel network

Notes: The figure depicts network links among firms supplying procurement contracts in the Common Procurement Vocabulary (CPV) category of "laboratory instruments for checking physical characteristics." The sample is restricted to pairs of firms with q-values below 0.05. The numbers on the links represent the corresponding q-values.

References

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

Open Contracting Partnership (2020). Ukraine S&L: MEL Framework. Technical report, Open Contracting Partnership.

Perko, L. (1996). *Differential Equations and Dynamical Systems*. Texts in Applied Mathematics. Springer New York.

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

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

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