Ownership Networks and Bid Rigging*

Kentaro Asai†

Ben Charoenwong

February 29, 2024

Abstract

Using a dataset of public procurement auctions and registered shareholders of all bidding firms in Singapore, we study the effects of ownership networks on prices and efficiency in product markets. A network-based measure of ownership connections shows that participating bidders with common owners or common owners' owners are more likely to submit identical bids, and identical bids are positively correlated with contract prices. Our structural estimates show that removing ownership network effects significantly improves a contractor's cost efficiency, highlighting how ownership networks can hinder competition. Our findings are robust to falsification tests, bid rounding, placebo tests using other common stakeholder relationships, and sample weighting based on our machine learning prediction of the auction format.

Keywords: Common Ownership, Public Procurement Auctions, Corporate Governance
JEL Classifications: D44, G32, L14, L41, H57

*We thank Sumit Agarwal, Simcha Barkai, Efraim Benmelech, Christopher James, Dirk Jenter, Kei Kawai, Scott Kominers, Alan Kwan, Tong Li, Ron Masulis, Gregor Matvos, Jun Nakabayashi, Hayato Nakanishi, Yoshio Nozawa, Wenlan Qian, David Reeb, Hong Ru, Tianyue Ruan, Amit Seru, Martin Schmalz, Yasutora Watanabe, Bernard Yeung, seminar participants at the Australian National University, the University of Tokyo, and Hitotsubashi University as well as the participants at the 2019 Econometric Society Australasian Meeting, ANU Summer Camp, and the 2019 APIOC for thoughtful comments and suggestions. Ben acknowledges financial support from the National University of Singapore Start Up Grant No. R-315-000-119-133 and the Singapore Ministry of Education AcRF Tier 1 Grant No. R-315-000-122-115, and Kentaro acknowledges financial support from the Australian National University Research School of Finance Actuarial Studies and Statistics. Any remaining errors are ours.

†Kentaro Asai, Email: kentaro.asai@anu.edu.au, Australian National University College of Business and Economics, CBE Building, 26C Kingsley Street, Acton ACT 2601, Australia; Ben Charoenwong, Email: ben.charoenwong@nus.edu.sg, National University of Singapore Business School, 15 Kent Ridge Dr #07-69, Singapore 119245.

1 Introduction

There is a long literature that argues ownership connections between firms facilitate coordination among product market participants. For example, common owners may induce coordination that helps firms reach a collusive arrangement by working as liaisons between them. They may also induce unilateral firm altruism by affecting the objective function of each firm. When a firm's shareholders also own stakes in its rivals, the firm—while upholding its fiduciary duty to shareholders—would aim to maximize the joint surplus of the firm and its rivals.[1] Both mechanisms suggest ownership networks weaken the competitive intensity of an industry. However, econometrically identifying the effect is challenging. Reduced-form approaches linking common ownership to prices are susceptible to interpretation issues depending on the underlying price-setting model, and the approaches that estimate mark-ups based on accounting or production functions are subject to data limitations on capital and are sensitive to the estimates of function parameters.

We overcome this challenge by using coordinated bidding behavior as the measure of anticompetitiveness. We then employ a combination of reduced-form analyses and structural estimation of an auction model to study the effects of ownership connections on prices and cost efficiency in multiple product markets. We use data on public procurement auctions in Singapore to define product markets and information on registered shareholders for all participating firms to construct ownership networks comprising common owners and common owners' owners of firms. Motivated by theoretical results based on first-price auctions, we restrict our sample to the set of auctions in which the lowest bidders win, accounting for approximately half of all procurement auctions.

Public procurement auctions serve as a useful setting to study the effects of ownership networks on product market competition for four reasons. First, a market is well-defined as an auction where bidders are competitors. Second, we observe a large number of auctions tendered by the same procurer around the same time with ample intra-product-type variation

[1] See Bresnahan and Salop (1986) and Reynolds and Snapp (1986), for example.

of ownership concentration. Specifically, different sets of firms participate in different tenders, providing rich variation in ownership concentration across auctions. Third, our sample covers various industries and includes both goods and services, complementing the recent empirical evidence that studies specific industries such as airlines or ready-to-eat cereals (Azar et al., 2018; Backus et al., 2020). Finally, we can also acquire the welfare implications of ownership networks using structural econometric methods in auctions.

The primary measure of anticompetitiveness in our analyses is coordinated bidding behavior. This approach directly addresses concerns raised by Hemphill and Kahan (2020), who argue that mechanisms relating common ownership to coordination between firms have not been empirically shown. Whereas collusion is rarely detected, our sample shows a prevalence of coordinated bidding behavior, allowing us to use it for our empirical analyses.

We first document prevalent identical bidding among firms in our sample—firms participating in the same auction often submit exactly the same bid. In theory, competitive bids have to differ across firms if they have any cost differences. Even if mimicking a rival's bid happens to be an optimal bidding strategy in a competitive setting, we should rarely observe this strategy empirically, as perfectly predicting a rival's bids based only on public information seems technically infeasible in our setting with sealed bids. Instead, firms that submit identical bids are more likely to share information about their bids (Ghosh and Morita, 2017). Therefore, we interpret identical bidding as an indicator of the potential presence of coordination, which is consistent with a view shared among practitioners and academics.[2] In particular, identical bidding facilitates the egalitarian allocation of contracts to coordinating firms, making their agreement sustainable (McAfee and McMillan, 1992).

We find that (1) competing firms with common owners or common owners' owners are more likely to submit identical bids, and (2) the number of identical bids submitted to the

[2]For example, the OECD Competition Committee recommends governments "avoid splitting contracts between suppliers with identical bids and investigate the reasons for the identical bids and, if necessary, consider re-issuing the invitation to tender or award the contract to one supplier only"(OECD, 2009). Moreover, Section 2.3 discusses theoretical foundations behind identical bidding as an effective coordination mechanism in detail.

auction is positively correlated with the normalized contract amount paid by the government. The latter finding is consistent with the anticompetitive aspect of identical bidding and the former suggests the anticompetitive effects of ownership connections. The relations between common owners and common owners' owners with identical bidding persist in auction-level analyses even when accounting for bid rounding and confounding common stakeholder relationships like common corporate secretaries and directors. In pair-level analyses, we find that firms sharing common owners are around 7 to 13 percentage points more likely to submit identical bids in auctions they participate in together, and those sharing common owners' owners are also more likely to submit identical bids in auctions they participate in together.

To quantify how much welfare is lost relative to a scenario where ownership networks do not affect firms' identical bidding behavior, we use a structural estimation to recover the link between the firm's cost and bid. We overcome two empirical challenges: unobserved auction heterogeneity and asymmetric bidders. In particular, our econometric framework allows us to pool our samples with auction heterogeneity that is unobservable for econometricians.[3] The framework is appropriate for our setting because our data do not include project sizes—a source of heterogeneity in public procurement auctions. In addition, it accommodates asymmetric bidders and allows us to consider two different types of bidders—so an identical bid can be distributed differently from a competitive bid. Our simulation results suggest removing ownership network effects on identical bidding reduces spending by the government by 4.2 to 6.0% and the winner's cost by 3.7 to 4.9% of the winning bid.

However, one caveat is that our sample could contain multi-attribute auctions in which the lowest bidders win by chance, despite our focus on first-price auctions. To address this concern, we compute the propensity score of each sample auction being a first-price auction through a machine learning binary classifier.[4] Our main reduced-form results persist when we weigh observations by the propensity score for being a first-price auction. Moreover,

[3] See Li et al. (2000) and Krasnokutskaya (2011) for the framework.

[4] In Section A.2.3 of the appendix, we provide conditions under which our machine learning model predicts auction format with reasonable accuracy. Our machine learning model shows performance metrics consistent with the conditions under which it reasonably predicts auction format.

our baseline simulation results persist when we use only auctions that our machine learning model predicts as first-price auctions. Overall, our results are robust to sample weighting based on our machine learning prediction of the auction format.

For policymakers concerned about competition, our findings suggest regulators should pay attention to ownership connections of various degrees as both common ownership and higher-order connections like common owners' owners affect product market competition. We believe these findings are relevant for other countries because (1) Singapore's procurement practices mirror those in other developed countries like the United States and Germany, and (2) like in Singapore, other wealthy economies' corporate ownership structures feature similar ownership characteristics (La Porta et al., 1999). In particular, the ownership structures in Singapore appear similar to those in Europe but less similar to those in the United States (Kirchmaier and Grant, 2005).

We mainly contribute to the research on the effect of common ownership on product market competition. For example, Azar et al. (2018) provide reduced-form empirical evidence for the anticompetitive effect of common ownership on airline prices. In addition, Gilje et al. (2020) argue that index inclusion or the mergers of financial institutions can reduce managerial incentives to internalize externalities across portfolio firms. However, Lewellen and Lowry (2020) question the validity of the extant empirical identification and argue other factors, such as differential responses of firms to the financial crisis, confound their analyses. Our use of a structural approach is also related to Park and Seo (2019), who adopt a structural estimate to evaluate the welfare impact of common ownership in airlines, and Backus et al. (2020), who adopt a similar approach to study the effects of common ownership in the ready-to-eat cereal industry. Our work is distinct from both in that we use public procurement auctions that contain various products and services instead of a single industry.

We next contribute to the research which has documented the role of inter-firm connections in corporate governance. Engelberg et al. (2012) find that firms that borrow from banks socially connected to them have lower interest rates, but it does not appear related to sweetheart deals, especially because subsequent firm performance improves following a connected deal. Ishii and Xuan (2014) find that mergers with prior target-acquirer social connections lead to more negative abnormal returns to the acquirer upon the merger announcement. Agarwal et al. (2019) show corporate directors of real estate companies visit golf courses together more regularly after land sale announcements and that correlates with lower winning bids and lower revenues to the government.[5]

Lastly, our paper highlights the relative benefit of using identical bidding in the empirical analyses of anticompetitive activities in auctions. The prior literature on the *data-driven* method of collusion detection in auctions shows the presence of bidding behavior inconsistent with the competitive bidding benchmark using the set of bids across auctions. Harrington (2005) reviews methods for detecting cartels and distinguishing collusion from competition.[6] The extant approach is overall convenient for detecting bidding behavior that is inconsistent with a competitive bidding benchmark *in the set of auctions*, but it is not useful for explaining variation in the presence and absence of anticompetitive bidding behavior *across auctions*. Because we can tell the presence and absence of identical bidding at the auction level, we can dissociate auction characteristics that are potentially correlated with a predictor of coordination in our empirical analyses.

[5]Moreover, Matvos and Ostrovsky (2008) and Harford et al. (2011) investigate the effects of ownership networks on shareholder voting and merger decisions, documenting peer effects. Although some research ascribes some of the peer effects to common ownership (He et al., 2019), other research has ascribed them to proxy advisory influence (Iliev and Lowry, 2015; Heath et al., 2021). Even when ascribed to common ownership, He and Huang (2017) argue that the voting behavior fosters product market competition. As an alternative mechanism, Antón et al. (2020) theoretically show that common ownership may increase product prices through reduced managerial incentives and provide empirical evidence supporting the model's predictions. However, Walker (2019) argues existing research linking common ownership to executive pay design is flawed due to measurement and methodological issues.

[6]For example, Porter and Zona (1993) compare the estimates of the parameters that drive a bidder's pricing strategy across various subsets of bids. In the absence of phony bids, there should be no variation in the estimates because seriously submitted bids are likely to be affected by bidder characteristics, such as the distance from the procurement site, in a consistent way. Testing this hypothesis, they identify the presence of bids that are not consistent with competitive ones. Chassang et al. (2019) examine the price elasticity of the empirical probability of winning an auction—it must be bounded above by -1 at any bid in any competitive equilibrium; otherwise, firms would find it profitable to increase their bids.

2 Public Procurement Auctions in Singapore

Since 2015, almost every government agency in Singapore procures goods and services through a one-stop, centralized online portal hosted by the Ministry of Finance called the Government Electronic Business (GeBIZ). Different procurers operate independently on the GeBIZ portal.

Any company (public or private and foreign or domestic) may register as a government supplier subject to due-diligence requirements based on the line of business, a small registration fee, and a minimum net tangible asset value.[7] In each auction, eligible companies submit blind bids. After the solicitation period, bids are finalized, and the GeBIZ platform reveals the winner and the bids submitted by each bidder. The winning bidder provides the good or service to the government at the price it bids. Although almost all government procurements are awarded, a handful is closed with no winners if no qualified bidders participate. This scenario may occur if no bidders are present, if the procurer determines that the number of bidders is insufficient, or if the procurer perceives that none of the bidders in the auction have the capacity to serve the contract fully.

All public procurements are subject to the Government Procurement Act (GPA). In adherence to the Agreement on Government Procurement from the World Trade Organization, the GPA outlines the requirements for public procurements in Singapore that aim to foster fair, competitive, transparent, and non-discriminatory conditions for government purchases of goods, services, and construction works. In addition to the GPA, the Competition Act provides further guidelines and regulations for procurement. Although the GPA does not explicitly place any minimum number of bidders on an auction, Section 34 of the Competition Act prohibits anticompetitive agreements, decisions, and practices. The commission provides a website outlining the basic anticompetitive practices, including price fixing, bid rigging, market sharing, and production control.

These regulations are enforced by the Competition and Consumer Commission of Sin-

[7]See https://www.gebiz.gov.sg/docs/Appln_Guidelines_for_Gov_Supp_Reg.pdf as of June 2022.

gapore (CCCS). The CCCS can commence investigations on its own initiative or if it is provided with reasonable grounds of suspected anticompetitive conduct from other parties through claims by procurers, whistleblowing, or leniency applications. Historically, the CCCS has issued infringing decisions, imposed financial penalties, and jailed convicted individuals (Chan, 2016). The CCCS's official website states price fixing, bid rigging, market sharing, and production control as examples of anticompetitive behavior enforceable under Section 34 of the Competition Act. Further, a CCCS-produced infographic explaining for procurers lists similarities in bids as the top sign to watch out for bid rigging.

2.1 Award Criteria

While price is a key consideration in evaluation, government agencies may check if bids have complied with all the requirements in the tender specifications, as well as evaluate other factors such as quality of the goods and services, timeliness in delivery, reliability, and after-sales support.[8] Roughly half of all procurements are not awarded to the lowest quote. Government agencies do not disclose to the public the evaluation criterion of each tender, but eligible bidders of the tender know the criterion through the tender document in advance.[9]

Since our data do not come with an explicit indicator of auction format, we can at best observe whether the lowest bidder wins a contract. The lowest bidder winning a contract is necessary but not sufficient for a first-price auction. To predict the format of each auction, we introduce an empirical methodology based on machine learning theory (Ghosh et al., 2017) which uses a random forest classifier—a tree-based decision algorithm for which we can account for a noisy measure (whether the lowest bidder wins an auction) of the outcome of interest (auction format). Although the measure used for prediction is noisy, our random

[8]See https://www.gebiz.gov.sg/docs/Supplier_Guide_Summarised.pdf as of June 2022.

[9]In a Feburary 2022 email reply to our inquiry pertaining to the details of awarding criteria, GeBIZ shares two statements. First, government agencies may consider price only or a combination of price and quality criteria in their evaluation of tenders or quotations. Second, the evaluation criteria are indicated in the individual tender or quotation documents, and may differ from tender to tender (or quotation to quotation) due to the specific requirements. However, the tender or quotation documents are only accessible to GeBIZ Trading Partners as they are intended for the sole purpose of allowing suppliers to prepare for their bids. They are not observable by the public and are not part of our dataset.

forest model shows performance metrics suggesting that it predicts auction format. The variable importance plot shows the most important auction characteristic for the prediction of auction format is the number of concurrent auctions. Details on the justification and implementation of the random forest model are discussed in Section A.2.2 and A.2.3 of the appendix. Figure A.2 in Section A.2.2 of the appendix reports the variable importance plot.

2.2 Bid Differences in Auctions

Identical bidding is prevalent in our sample. To show this fact, we plot the distribution of the difference between two randomly sampled bids from the same auction. For each auction in our sample, we randomly sample two bids. To make bids comparable across auctions, we normalize each bid by the standard deviation of all bids in the auction. Then, we compute the difference between the normalized bids.

[Figure 1 Around Here]

The histogram of randomly sampled differences between two bids in each auction shown in Figure 1 documents a probability mass at zero. The density at zero is *discontinuously* larger than the density immediately before and after zero, though the bid difference is smoothly distributed everywhere else. If suppliers independently drew costs from smooth distributions, their cost differences would be smoothly distributed. If cost differences were smoothly distributed, their bid differences would also be smoothly distributed, given that the standard competitive equilibrium bidding strategy suggests their bids are continuous in their costs. In this respect, our finding is inconsistent with a competitive bidding outcome with smoothly distributed costs. The bottom panels of Figure 1 also present the corresponding distribution for the subset of auctions for goods and services, showing similar consistent patterns across both types of auctions. We also find this fact is not driven by scaling, auctions with few bidders, or a particular auction type.[10]

[10] Section A.2.4 of the appendix shows the robustness of this result. See Figure A.3 documenting a similar probability mass at zero.

2.3 Bid Rigging as a Coordination Mechanism

To understand the rationale behind submitting identical bids, we briefly discuss the theoretical basis for identical bidding as a coordination scheme in a first-price auction. Considering a procurer normally awards a contract to one of the suppliers that submit the lowest bids with equal likelihood, identical bidding is a straightforward way of achieving a fair allocation of contract awards for a cartel. Because a cartel's breakdown is often triggered by an internal conflict over how to share the profits (Levenstein and Suslow, 2006), fairly dividing a cartel's spoils through identical bidding is effective for facilitating coordination. Identical bidding — a form of bid rigging — has lower coordination costs compared to bid rotations. Under an identical bidding scheme, only a cartel price must be set. With rotating bids, the allocation of cartel profits must also be agreed upon (Comanor and Schankerman, 1976).

Another possible scheme is *efficient collusion*, where the most productive bidder always wins the project to maximize cartel surplus. This scheme usually requires some side payment mechanism for compensating other firms that agree to lose the auction. Without sophisticated side payments, non-common shareholders of losing firms or any stakeholder whose surplus is correlated with losing firms' profits would have an incentive to cheat on a collusive agreement as their surplus correlates with the revenue of losing firms.[11] Can a cartel facilitate side payments in Singapore?

We believe facilitating side payments is difficult in Singapore due to the risk of prosecution by the government for two reasons.[12] First, the Goods and Services Tax (GST) system in Singapore deters side payments that are not reported to the government. Second, even though side payments might be possible through non-recorded cash payments, mandatory audits, severe penalties, and incentives for whistleblowing deter this practice.

[11] Lambert (2021) argues firm managers' incentives tend to align with those of the bulk of their shareholders in favor of their own-profit maximization. This is because most corporate managers are compensated in part in their company's stock.

[12] The 2018 Transparency International Corruption Perceptions Index ranks Singapore as one of the three least corrupt countries in the world, and the World Economic Forum Global Competitiveness Report ranks Singapore as the most competitive nation in the world, overtaking the United States. Also, see Section A.2.1 of the appendix, where we extensively discuss the difficulty of side payments.

McAfee and McMillan (1992) prove in a static setup that a cartel can do no better than having its members submit identical bids in first-price auctions when side payments are prohibited. Without side payments, only the winning firm can receive the cartel surplus. Hence, cartel members are incentivized to misreport their costs to become the auction's winner, deterring the cartel from identifying the most productive firm. This means identical bidding can be an optimal mechanism for a cartel in the absence of side payments. The caveat is that McAfee and McMillan (1992) focus on all-inclusive cartels. The best mechanism for partial cartels in the absence of side payments remains an open question, though the same logic appears applicable for partial cartels. Incentive compatibility requires cartel bids to be identical to give an equal chance of winning the auction even in the presence of non-cartel bidders.

In addition, the egalitarian allocation of contracts is of particular importance in the presence of a leniency system. Under the leniency system of Singapore, if certain conditions are met, the first coalition member that self-reports its collusive activity to the CCCS will be entitled to immunity from financial penalties or a reduction of up to 100% of the financial penalties. Thus, the leniency system strengthens the incentive of an unsatisfied member to deviate from collusive activity. On the other hand, identical bidding achieves the fair allocation of contracts among members, discouraging an unsatisfied member from betraying the other members.

3 Data & Methodology

3.1 Data and Sample Construction

Our data set comes from merging shareholder registries with GeBIZ procurement auctions. It covers September 21, 2016 to April 2, 2018, starting slightly over a year from the GeBIZ system's initial rollout in 2015. The data contain more than S$16.5 billion (US$12 billion) of expenditures.

We acquire the shareholder registry for all bidders from DC Frontiers Pte Ltd (Handshakes), one of four authorized information resellers licensed by the Accounting and Corporate Regulatory Authority as of April 2018. The data include the names of registered shareholders and officer information for every company registered in Singapore. Registration is required for all shareholders of private companies and for shareholders of public companies with at least 1% ownership or current shareholders that paid in capital at the company inception.[13] Where possible, we also supplement the registered ownership data with ownership of publicly-traded companies in Singapore as of March 31, 2018, based on individual and institutional investor filings (e.g., 13-F in the United States), which include both mandatory and voluntarily disclosed positions from Refinitiv.

We focus on first-price auctions as our analyses are based on a theory for first-price auctions. Since our data do not come with an explicit indicator of auction format, we restrict the sample to approximate the set of first-price auctions by only considering auctions in which the lowest bidders win and with at least two bidders.

3.2 Defining Ownership Networks

To identify ownership networks, we first consider the symmetric non-weighted adjacency matrix A, an $n \times n$ matrix where element $A(i, j)$ is whether entity i is a registered shareholder of entity j, and n is the number of unique entities comprising both individuals and businesses. For simplicity, and because our data spans less than two years, we treat all connections as static and do not consider new or expiring relationships.

We define the degree of connection between two firms i and j as the shortest path from node i to node j. Then, the indicator for i and j to have a k^{th}-degree connection, $S^k(i, j)$, is defined as

$$S^k(i, j) = 1(\inf\{x \geq 1 | A^x(i, j) > 0\} = k),$$

[13] Importantly, companies without updated shareholder registries are unable to file annual returns, which carries a fine. See https://www.acra.gov.sg/legislation/legislative-reform/companies-act-reform/companies-amendment-act-2014/key-amendments-to-the-companies-act/electronic-register-of-members and https://www.acra.gov.sg/announcements/file-your-annual-lodgments-on-time. In addition to shareholders, the data include information on directors and other company officials through time, permitting us to study the impact of different types of connections. We use these non-shareholder connections for robustness tests.

where $A^x(i,j)$ is the element (i,j) of A^x, which captures the number of ways i reaches j with a walk of x steps.[14] We investigate $k \in \{2,4\}$, corresponding to common owners and common owners' owners.[15,16]

3.3 Reduced-Form Methodology

We conduct two empirical analyses documenting the association between ownership networks and identical bidding. First, we estimate the effects of ownership connections among firms participating in the same auction on identical bidding, controlling for auction-specific characteristics and an auction-level measure of bid rounding. Second, we study the relation between the number of ineffective bids (defined as the number of exactly identical bids) and the awarded contract amount.

For each auction l with a set of bidders $B(l)$, arbitrarily indexed from $i = 1, ..., N_l$, we count the number of ineffective bids N_l^I, defined as the number of duplicative bids in the auction. In our setting, each bidder $b \in B(l)$ can only submit one bid, so $N_l^I = N_l - N_l^E$, where N_l^E stands for the number of effective bids, defined as the number of distinct bids in auction l. Under one interpretation, the number of ineffective bids N_l^I reflects the extent

[14]Section A.1 of the appendix shows some examples of the matrix representation of the ownership networks. Figure A.1 in Section A.1 of the appendix visualizes the connection of each degree. For two arbitrary firms i and j, a first-degree connection means i is a shareholder of j, a second-degree connection means i and j have common shareholders, a third-degree connection means i is owned by another firm whose shareholder is also a shareholder of j, and a fourth-degree connection means i and j have common shareholders' shareholders.

[15]Theoretically, our network-based measure also permits chain connections. For example, a two-degree chain captures a relationship between a firm and the subsidiary of its subsidiary. However, we refer to second-degree connections as common ownership and fourth-degree connections as common owner's ownership because the presence of multi-degree chains is exceedingly rare among the participants of an auction. Similarly, we also do not consider first-degree connections due to the lack of data.

[16]The common ownership profit weight might be an alternative meassure of firm altruism, but we consider that it is uninformative about the occurrence of identical bidding. The weight represents the value to the focal firm of a dollar of profit generated for another firm competing with the focal firm, which is the main constituent of the "modified HHI" used in the empirical literature of product market competition. Importantly, conventional computations of profit weights assume proportional control based on the premise of "one share, one vote," and thus they are relevant only when shareholders can exert their influence in accordance with their stakes without being overruled by a manager. In contrast, identical bidding is an optimal coordination mechanism in the presence of agency problem between a manager and a common owner, because it implements fair allocation between connected firms, which is more sustainable for a manager who aims to secure his own firm's surplus rather than a common owner who benefits most from efficient collusion. Thus, when identical bidding is observed, managerial agency problem seems to exist, and hence a common owner is less likely to exert its influence in accordance with its stakes.

of identical bidding in an auction. Under another interpretation, this variable measures the number of firms that do not affect the competitive intensity of bidding.

We count the number of ownership connections of degree k among auction participants by aggregating the indicators $S^k(i,j)$ from every possible pair of bidders in an auction to form an auction-level connectedness. Specifically, we define the number of connections for auction l as

$$NC_l^k = \sum_{(i,j)\in B(l)} S^k(i,j),$$

for $k \in \{2,4\}$ corresponding to common owners and common owners' owners, respectively. We next measure the tendency of bid rounding R_l across all bids at the auction level, which we will use as a control:

$$R_l = \frac{1}{N_l}\sum_{i=1}^{N_l} 1(b_i \equiv 0 \mod 10^{q_l}),$$

where N_l is the number of bidders in auction l, i indexes each bidder with bid b_i, and q_l is the most frequent order of magnitude among all the bids in auction l minus one. This specification of the roundedness assumes the rounding of bids tends to occur at one units below the order of magnitude of the bid. For example, an unrounded bid of S\$14,282 tends to be rounded to S\$14,000 rather than S\$10,000, S\$14,300, or S\$14,280. In accordance with this premise, if auction l has two bids of S\$17,000 and S\$16,800, $q_l = 3$ and $R_l = 0.5$ because S\$17,000 is divisible by 10^3 whereas S\$16,800 is not.[17]

Then, we consider a regression specification of the form:

$$N_l^I = \sum_{k\in\{2,4\}} \beta^k 1\{NC_l^k > 0\} + f(Z_l, R_l) + \epsilon_l, \tag{1}$$

where $1\{NC_l^k > 0\}$ is an indicator at the auction level capturing whether any pair of bidders

[17] Figure A.6 in Section A.5 of the appendix shows the histogram of trailing units. We find that a mass is present at 0, suggesting firms tend to round their bids at least to the nearest 10's unit. In untabulated analyses, we consider additional definitions of rounding, including modifying q_l to the most frequent order of magnitude among all the bids in auction l, one, counting the number of 0's in a bid, counting the number of 5's, and find they do not affect our results.

have a connection with degree k for $k = 2$ and 4, labeled as Common Owners and Common Owners' Owners, respectively. N_l^I is the number of ineffective bids in auction l and $f(Z_l, R_l)$ is specified as the a third-degree polynomial of bid rounding R_l and procurer-by-auction type fixed effects. We cluster standard errors by procurer. β^k is the coefficient of interest and captures the effect of having any connected bidders of degree k on the number of ineffective bids.

The empirical specification in equation (1) uses an indicator because the count measure itself may be problematic. The count measure treats a pair of firms with 40 overlapping owners, each with 1%, as having more connections than one owner owning 40% in each firm. Therefore, the count measure introduces a non-linear bias in the relation between the count measure and ineffective bidding because the single owner with the larger stake probably has a larger capacity to make his portfolio companies coordinate. Consequently, we present the main results using an indicator variable, capturing the relation between the extensive margin of having connected bidders in the auction and identical bidding.

To assess the impact of identical bidding on the award amount, we investigate the link between the number of ineffective bids and the award amount paid by the procurer. Then, we estimate

$$Amount_l = \gamma N_l^I + f(Z_l, R_l) + \epsilon_l, \tag{2}$$

where $Amount_l$ is the realized contract amount in auction l scaled by the median of the bids and captures the planned expenditure of the government and revenue for the bidding firm. The control variables mirror those in equation (1) above. As with equation (1), we cluster standard errors by procurer. The coefficient γ captures the relation between the number of ineffective bids and the awarded contract amount.

3.4 Summary Statistics

Table 1 reports the summary statistics of our sample. We observe 9,087 auctions with bids from 10,000 unique bidder names. The number of bidders per auction is around 5 on average. The average number of ineffective bids in an auction is 0.29, with around 16% of auctions having at least one ineffective bid. The average number of second- and fourth-degree ownership connections per auction is 3.3 percentage points and 20 basis points, respectively. The mean awarded contract amount in our data set is slightly over S$359,000, and the median is S$15,200. A large right skew exists in the data because auctions range from purchasing a single computer to a multi-billion dollar project, such as building new subway systems or public housing units.

In our pair-level data, we observe 99,661 pairs participating in the same auction at least once in our sample. On average, a pair of firms participates in the same auction 1.8 times and submit identical bids for 2.1% of the auctions in which they jointly participate. On average, 9.9 basis points of pairs have common owners, and 0.8 basis points of them have common owner's owners.

[Table 1 Around Here]

4 Reduced-Form Results

4.1 Ownership Connections and Ineffective Bids

Table 2 reports the auction-level regression result of equation (1). Panel A shows our estimates controlling for a third-degree polynomial of the log number of bidders, while Panel B does not control for the number of bidders. We separate these results into a different panel due to the potentially endogenous nature of the number of bidders. Columns (1) and (3) of Panel A show the presence of common owners increases the number of ineffective bids by 0.443, over 1.5 times the unconditional mean of 0.285. Columns (2) and (3) show when bidders share at least one common owners' owner, the number of ineffective bids increases by 0.3, around the size of the unconditional mean. Columns (4) to (6) show results for a propensity-score weighted regression where we weight observations by the propensity score of an auction being price-only computed from our random forest model.[18] Specifically, we use the fraction of trees voting for positive for the propensity score. Weighting by the propensity score allows us to focus on the relationship where the probability of the auction being awarded to the lowest bid is the highest, corresponding with the mechanism proposed in Section 2.3. The results are similar to baseline results for both common owners and common owners' owners. Panel B shows results for corresponding regressions that do not control for the cubic polynomials for the log number of bidders in the auction, showing even larger estimates.

[Table 2 Around Here]

Differences in the easiness of detecting ownership connections may explain the non-trivial effect of common owners' owners on ineffective bidding. On the one hand, all else equal, higher-degree connections correspond with a lower coordination incentive than lower-degree connections. On the other hand, the probability of a regulator detecting ownership connections is likely much lower for higher-degree than for lower-degree connections. For procurers to detect bidders with common owners' owners, they must access the Singapore Enterprise Data Hub (an inter-government agency centralized data portal) to check the shareholder registry individually for all participating firms, and then check the shareholder registry of all those shareholders.

4.2 Pair-Level Analysis

As a robustness check, we also conduct a pair-level analysis of identical bidding behavior for firms that have participated in the same auction at least once since the previous auction-level analysis does not differentiate whether identical bidding occurs among unconnected or connected firms. The pair-level analysis directly tests whether connected firms are more likely to submit identical bids. We regress the probability of submitting identical bids I_p

[18] We impute missing propensity scores with zero, so that the number of observations are unaffected but our results only use the subset of data with valid estimated propensity scores.

for pair p of any (i, j) firms on the dummy variable that indicates the presence of common owners C_p^2 and common owners' owners C_p^4:

$$I_p = \beta^2 C_p^2 + \beta^4 C_p^4 + f_i + f_j + \epsilon_p, \tag{3}$$

where f_i and f_j are firm fixed effects for i and j, respectively. We also report the result where we do not control for firm fixed effects and only include constants.

[Table 3 Around Here]

Table 3 shows that the effects of common ownership and common owner's ownership on identical bidding are positive at the pair-level. In particular, we find that firms sharing common owners are around 7 to 13 percentage points more likely to submit identical bids in auctions they participate in together, and those sharing common owners' owners are also more likely to submit identical bids in auctions they participate in together. This result is consistent with the result of our auction-level analysis reported in Table 2, because the estimated effects of common owners and common owners' owners are all positive, regardless of specifications. In particular, the estimated effects are robust to the control of firm fixed effects. Overall, we find that pairs of firms having common owners or common owners' owners are more likely to submit identical bids than unconnected pairs.

Robustness & Alternative Possibilities We next address three additional concerns and show that our previous results are not due to (1) spurious correlation, (2) the use of indicator variables for measuring ownership connections, or (3) confounding factors. First, our estimates for the effects of ownership connections have relatively small p-values, according to the attained distributions of the corresponding coefficients in a falsification exercise where we randomly assign ownership connections among the participants. Second, our results persist when using the number of connected firms participating in the auction rather than indicator variables for their presence. Third, we obtain quantitatively similar estimates for the effects of ownership connections even when additionally controlling for other common stakeholder relationships apart from shareholder relationships. These results are all reported in Figure A.5, Table A.4, and Table A.5 in Appendix A.4.

But apart from these concerns, our baseline results also appear explainable by two alternative hypotheses centered around the idea that identical bidding may be due to identical cost structures rather than ownership networks. First, identical costs may occur on large infrastructure projects through identical cost estimates or subcontracts. Auctions with larger contract amounts or severe participation constraints that permit only qualified bidders may be those for which respondents may outsource the calculation of cost estimates. If qualified firms acquire the cost estimates from the same engineer or hire the same subcontractor, they may anchor their bids on the same costs, which then generate identical bids. However, this does not appear to be the case since columns (1) and (2) in Table 4, which report the results of robustness tests for equation (1), show that the positive relationship between the number of ineffective bids and ownership networks persists even when considering only smaller auctions whose winning bids are less than S$50,000 (US$36,000) and S$100,000 (US$72,500). Moreover, columns (3) through (5) show that excluding construction auctions, auctions conducted by government agencies related to infrastructure or development, or both do not affect the results qualitatively.

[Table 4 Around Here]

Second, connected firms may have distinct but similar cost levels. If firms simply round their bids, connected firms are likely to submit identical bids even if their costs are distinct. In fact, we find bid rounding is prevalent in general, so some of the estimated effects of ownership connections on the number of ineffective bids could simply reflect cost similarity and rounding. We address this concern by reestimating equation (1) but define the dependent variable as the presence of *non-rounded* ineffective bids. In this set of analyses, the coefficients on ownership networks capture the effects of ownership networks on ineffective bids that are not rounded. Table A.6 in Appendix A.5 shows that our main results persist even when limiting our focus on identical non-rounded bids. Therefore, we believe our baseline estimates of the relationship between ownership connections and ineffective bids are not solely driven by cost similarity and bid rounding.

4.3 Ineffective Bids and Bid Amounts

We document a positive correlation between identical bidding and ownership networks. Because identical bidding is a plausible coordination mechanism for first-price auctions per McAfee and McMillan (1992), we expect the anticompetitive effects of ownership connections to manifest as higher winning bids in first-price auctions. To test this hypothesis, we estimate equation (2) using a baseline as well as a propensity-score weighted regression (Alkurdi and Sizemore, 2019), where the propensity score is calculated from our random forest model predicting auction format.

[Table 5 Around Here]

Odd-numbered columns of Table 5 report the auction-level regression results of equation (2). Column (1) of Panel A shows the effect of the intensive margin of identical bidding; one more identical bid in an auction is associated with an increase in the government expenditure of 4.3% of the median bid. Taking into account the prior result that auctions with common owners are associated with an increase in the number of identical bids by 0.443, the overall estimated effect of common ownership on government expenditure is around 1.9% of the median bid. Even-numbered columns of Table 5 report similar analyses but weigh auctions by the propensity score of an auction being price-only. The estimates do not change much quantitatively. We also report the results for each of the goods and services & construction separately and show similar estimates across both procurement types.

Panel B shows the results for corresponding regressions without controlling for the cubic polynomials for the log number of bidders in the auction. Again, we separate these results into a different panel due to the potentially endogenous nature of the number of bidders. The estimates for the effect of identical bidding are all positive. Moreover, they are statistically significant at the 5% level. Because the number of bidders is negatively associated with the winning bid, excluding the number of bidders from the regression introduces a negative omitted variable bias if the number of ineffective bids is positively associated with the number of bidders. This omitted variable bias can explain the smaller estimates of Panel B relative to the corresponding estimates of Panel A.

5 Ownership Networks and Auction Efficiency

The reduced-form analysis provides suggestive evidence of the role of ownership networks in bid rigging but provides little implication on auction efficiency. Therefore, we use a structural framework to recover the missing link between a firm's cost and bid and identify the distribution of a firm's cost. By simulating the auction outcomes of a counterfactual world with no relationship between ownership connections and bidding behavior, we can assess the effect of ownership networks on the cost of the winning contractor, which is the measure of auction efficiency.

5.1 Counterfactual Simulation

We use the framework by Krasnokutskaya (2011), which permits unobserved heterogeneity across auctions and asymmetric bid distributions. For the appropriate choice of structural framework, we considered two issues. First, our data set consists of auctions with various project sizes. Second, the distribution of competitive bids submitted by non-colluding firms (type-1) and that of identical bids submitted by coalitions (type-2), i.e., a group of firms coordinating with each other, are likely to be asymmetric.[19] Then, we impose the assumptions below:

Assumptions:

1. *A procurer allocates the project to the lowest bidder and randomly allocates the project if there are multiple lowest bidders.*

[19]To show this, we regress effective bids on the indicator for identically submitted bids while controlling for auction fixed effects. Then, we compute the gap between identically submitted and other bids while removing auction-specific effects. Table A.3 in Section A.2.5 of the appendix reports the regression result. We find identically submitted bids are significantly lower than others on average. Our result is consistent with the existing research. For example, Pesendorfer (2000) finds the empirical distribution of cartel bids is first-order stochastically dominated by the empirical distribution of non-cartel bids when studying the asymmetries between cartel and non-cartel bids for school milk contracts in Florida and Texas during the 1980s.

2. *When planning a bid, an auction participant knows the set of rival firms and, in particular, the coalition of firms that submit identical bids.*

3. *An auction participant independently draws the cost, conditional on auction-specific covariates.*

4. *Non-colluding firms and coalition members are ex ante asymmetric. The coalition plans a bid based on the cost randomly selected from members' costs as if a randomly selected member competitively bids against rivals.*

Before presenting the simulation procedure, we briefly discuss each assumption. The former part of Assumption 1 reflects that our sample is the approximate set of first-price auctions. The latter part of Assumption 1 specifies the allocation rule when identically submitted bids are the lowest. We assume a project is randomly allocated to one of the firms that submit the lowest bids, because randomization rationalizes submitting identical bids as the optimal coordination mechanism in the absence of side payments (McAfee and McMillan, 1992) and distributing side payments is difficult under the law enforcement in Singapore.[20] Assumptions 2 and 3 allow us to recover cost distributions under the framework that allows auction heterogeneity and asymmetric bid distributions. Assumption 4 allows us to predict bidding outcomes and winners' costs in our simulation.[21] For the sake of brevity, we provide derivations needed for our structural analysis in Appendix B.1.

In our simulation, we consider model auctions where there are three non-colluding firms and one coalition formed by two coordinating firms that submit identical bids since the median number of effective bids for auctions in which at most two firms submit identical bids is four. For the remaining section, we denote the number of type-i effective bids submitted

[20] There is still a possibility that a procurer selects the winner among the lowest bidders based on some other unobserved attributes under first-price auctions. In this case, we assume that firms find it difficult to predict the attributes the procurer considers and form uniform prior about which firm wins.

[21] One may be concerned with the possibility that non-colluding firms and coalition members are ex ante symmetric, even if their bidding strategies are asymmetric. In particular, the coalition may plan a bid based on the minimum of members' costs as if the most efficient member competitively bids against rivals in accordance with Li and Zhang (2015) and Dalkir et al. (2000). However, in this case, the most efficient member is better off while the least efficient member is worse off, as each member maximizes its expected profits when a coalition bids based on its cost. Because the least efficient member is worst off, a cartel is less likely to be sustained. Considering this matter, we assume that non-colluding firms and coalition members are ex ante asymmetric.

to auction j by $N^E_{j,i}$, analogously to N^I_j, i.e., the number of ineffective bids submittd to the auction.

In our simulation, we draw a cost from the estimated distribution of a type-1 firm's cost for three times and from that of a type-2 firm's cost twice. For the control case, we derive bids from the estimated equilibrium-bid function, where $N^E_{j,1} = 3$, $N^E_{j,2} = 1$, and $N^I_j = 1$.

Our structural analysis uses the reduced-form evidence to link ownership connections and identical bidding behavior. Specifically, for the treatment case, we draw the effect of the presence of k^{th}-degree connections among the participating firms on the number of ineffective bids based on the estimate of equation (1). To account for the standard error in the estimate of this effect, we assume the effect follows the distribution, $\tilde{\beta^k} \sim N\left[\hat{\beta^k}, SE(\beta^k)^2\right]$, where $\hat{\beta^k}$ is the estimate of β^k and $SE(\beta^k)$ is the standard error of the estimate of β^k in equation (1).[22] We do not allow $\tilde{\beta^k}$ to be either negative or above one, because we do not consider either multiple coalitions or new entrants in the counterfactual case. Therefore, we eventually use $\min\left\{\max\left\{\tilde{\beta^k}, 0\right\}, 1\right\}$, denoted by Δ_j. We note this adjustment is very rare, because $\hat{\beta^k}$ is between zero and one and $SE(\beta^k)$ is sufficiently small.[23] Then, Δ_j is the decrement in the number of ineffective bids in the auction due to the removal of the link between k^{th}-degree connections and identical bidding. In our simulation, however, the change in the number of ineffective bids has to be an integer, even if Δ_j is not. To resolve this problem, we also draw a random number y from $U(0,1)$ as a lottery. Based on the lottery outcome, the reduction in the number of ineffective bids becomes one or zero. Specifically, if $y \leq \Delta_j$ ($y > \Delta_j$), the reduction in the number of ineffective bids, δ_j, is one (zero).[24] When the reduced number of ineffective bids is one, a coalition is broken, which increases the number

[22]The standard errors of our estimates for equation (1) are clustered at procurer level. There are over 70 procurers, so the t-distribution with this many degrees of freedom is virtually identical to the normal distribution.

[23]In our structural analysis, we use the first two columns of Table 2 for the estimated effects of the presence of second-degree and fourth-degree connections among the participating firms on the number of ineffective bids.

[24]For example, suppose $\tilde{\beta^k} = 0.6$ and $y = 0.3$. Because $\Delta_j = \tilde{\beta^k}$ and $y \leq \Delta_j$, $\delta_j = 1$. It is trivial to prove the expected reduction in the number of ineffective bids becomes Δ_j.

of originally coordinating but now non-colluding firms by two.[25] On the other hand, when there is no effect on the number of ineffective bids, there would be no change in bidding outcomes between the control and counterfactual cases.

In summary, for the treatment case, we derive bids from the estimated equilibrium-bid function, where $N^E_{j,1} = 3$, $N^E_{j,2} = 1+\delta_j$, and $N^I_j = 1-\delta_j$. We use randomly selected one of two coordinating firms' costs as the input when evaluating a type-2 (coordinated) bid if $N^I_j = 1$.

We then compute the winner's cost for each case. If $N^I_j = 0$, the winning bid cannot be identical to any ineffective bid. Then, the lowest bidder's cost is the winner's cost. If $N^I_j = 1$, the lowest bidder's cost is the winner's cost when the lowest bid is type 1 (competitive), but the winner's cost is randomly drawn from coordinating firms' costs when the lowest bid is type 2 (coordinated).

We simulate 200 times. For each trial, we compute the gap in the winner's bid and cost for the control and treatment cases. To make the gap scaleless, we normalize it by the winning bid for the control case. Then, we take the mean as follows:

$$
\begin{aligned}
dB^* &= \text{mean}\left(\{(B^*_{j,counterfactual} - B^*_{j,control})/B^*_{j,control}\}_{j=1}^{200}\right), \\
dC^* &= \text{mean}\left(\{(C^*_{j,counterfactual} - C^*_{j,control})/B^*_{j,control}\}_{j=1}^{200}\right),
\end{aligned} \tag{4}
$$

where $B^*_{j,s}$ is the winning bid and $C^*_{j,s}$ is the winner's cost for auction j and case s.[26]

[25] We note originally coordinating firms have different technologies than originally non-colluding firms, although they share the same technologies in the previous simulation.

[26] In practice, we compute the following:

$$
\begin{aligned}
db^* &= \text{mean}\left(\{(b^*_{j,counterfactual} - b^*_{j,control})/b^*_{j,control}\}_{j=1}^{200}\right), \\
dc^* &= \text{mean}\left(\{(c^*_{j,counterfactual} - c^*_{j,control})/b^*_{j,control}\}_{j=1}^{200}\right),
\end{aligned}
$$

where $b^*_{j,s}$ is the winning bid and $c^*_{j,s}$ is the winner's cost for auction j and case s, under the assumption that auction-specifc cost shock y_j is equal to one. Although we need to draw an auction-specific cost shock to fully replicate the data-generation process, this process is not required to compute the "normalized" gap in the winning bid and the winner's cost. To see why, notice

$$
\begin{aligned}
dB^* &= \text{mean}\left(\{(y_j b^*_{j,counterfactual} - y_j b^*_{j,control})/y_j b^*_{j,control}\}_{j=1}^{200}\right) = db^*, \\
dC^* &= \text{mean}\left(\{(y_j c^*_{j,counterfactual} - y_j c^*_{j,control})/y_j b^*_{j,control}\}_{j=1}^{200}\right) = dc^*.
\end{aligned}
$$

Finally, to evaluate how the welfare gain is split between the government and firms, we also decompose $-dC^*$ into the following components:

$$\overbrace{\text{Welfare Gain}}^{-dC^*} = \overbrace{\text{Government Savings}}^{-dB^*} + \overbrace{\text{Firm Gains}}^{-dC^*+dB^*}. \tag{5}$$

We report dB^*, dC^*, government savings, and firm gains.

We conduct this procedure twice, once for common owners ($k = 2$) and another time for common owners' owners ($k = 4$). Figure 2a presents the simulation results of removing the effects of common owners on identical bidding. We find the point estimate of dB^* is -0.0601. The 95% confidence interval of estimated dB^* is below -0.03. This result implies that removing the effects of second-degree connections on identical bidding reduces the winning bid and, hence, increases government savings by 6.0% (most likely by at least 3%). We next find the point estimate of dC^* is -0.0491. The 95% confidence interval of estimated dC^* is below -0.02. This result implies removing common ownership effects reduces the winner's cost by 4.9% of the winning bid (most likely by at least 2%). Lastly, we find the point estimate of firm gains is slightly negative (-0.0110), which is not statistically significant. This result implies removing the effect of common owners reduces firms' surplus by 1.1% of the winning bid, if any.

[Figure 2 Around Here]

Figure 2b shows the simulation results of removing the link between common owners' owners and identical bidding. The estimated effects follow a similar pattern as before. We find the point estimate of dB^* is -0.0420 and its 95% confidence interval is below -0.02. This result implies removing the effect of common owners' owners reduces the winning bid and, hence, increases government savings by 4.2% (most likely by at least 2%). Our point estimate of dC^* is -0.0369 and its 95% confidence interval is below -0.01. This result implies removing the effects of common owners' owner connections on identical bidding reduces the winner's cost by 3.7% of the winning bid (most likely by at least 1%). We finally find the point estimate of firm gains is slightly negative (-0.0051), although it is not statistically significant. This result implies removing the effects of common owners' owner connections on identical bidding decreases firms' surplus by 0.5% of the winning bid, if any.

Overall, we confirm that eliminating the link between ownership networks and identical bidding improves contractors' cost efficiency. When multiple firms submit the lowest bid, randomizing contract allocation independently of firm efficiency can raise the winner's cost. Our result highlights this inefficient nature of identical bidding induced by ownership networks. Moreover, breaking a coalition raises competitive pressure by increasing the number of effective bids, which reduces the winning bid. This result is consistent with our reduced-form evidence.

Besides the analysis above, we conduct a robustness check by using only auctions that our machine learning model predicts as first-price auctions among the sample used for our baseline simulation. Section A.3 of the appendix reports the results of this robustness test (Figure A.4). We find the results are similar to our baseline results.

6 Conclusion

We show identical bidding exists in Singaporean public procurement auctions and is positively associated with ownership networks. This result is consistent with firms with shared owners rigging bids to raise a contract's price. Our structural analysis models a firm's strategic bidding behavior and simulates a counterfactual world in which ownership networks are not correlated with identical bidding. We find that removing either common owner or common owners' owner effects on identical bidding can improve the winning contractor's cost efficiency, highlighting how ownership networks hinder competition. More broadly, our findings suggest the need for examining inter-firm networks to predict potential misconduct. Whereas the existing literature tends to focus on firm characteristics alone,[27] this paper

[27] The prior literature focuses on firms' internal structure such as CEOs' connections with top executives and directors (Khanna et al., 2015), board structure (Agrawal and Chadha, 2005), and executive compensation (Efendi et al., 2007) as the driving factors for corporate scandals.

highlights the importance of how firms are connected to each other in networks.

References

Agarwal, S., Qin, Y., Sing, T. F., and Zhang, X. (2019). Informal Interactions among Top Managers on the Golf Courses. *Working Paper*.

Agrawal, A. and Chadha, S. (2005). Corporate governance and accounting scandals. *Journal of Law and Economics*, 48(2):371–406.

Alkurdi, R. and Sizemore, S. (2019). *Matching Models for Causal Inference: A Machine Learning Update*. Humbolt University, Berlin.

Antón, M., Ederer, F., Giné, M., and Schmalz, M. C. (2020). Common ownership, competition, and top management incentives. *Working Paper*.

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

Azar, J., Schmalz, M. C., and Tecu, I. (2018). Anti-Competitive Effects of Common Ownership. *Journal of Finance*, (5):1–79.

Backus, M., Conlon, C., and Sinkinson, M. (2020). Common Ownership and Competition in the Ready-To-Eat Cereal Industry. *NBER Working Paper*, 53(9):1689–1699.

Bajari, P. (2001). Comparing competition and collusion: a numerical approach. *Economic Theory*, 18(1):187–205.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Bresnahan, T. F. and Salop, S. C. (1986). Quantifying the competitive effects of production joint ventures. *International Journal of Industrial Organization*, 4(2):155–175.

Chan, J. (2016). Detecting bid rigging conspiracies.

Chassang, S., Kawai, K., Nakabayashi, J., and Ortner, J. M. (2019). Data driven regulation: Theory and application to missing bids. Working Paper 25654, National Bureau of Economic Research.

Comanor, W. S. and Schankerman, M. A. (1976). Identical bids and cartel behavior. *Bell Journal of Economics*, 7(1):281–286.

Dalkir, S., Logan, J. W., and Masson, R. T. (2000). Mergers in symmetric and asymmetric noncooperative auction markets: the effects on prices and efficiency. *IJIO*, 18(3):383 – 413.

Efendi, J., Srivastava, A., and Swanson, E. P. (2007). Why do corporate managers misstate financial statements? the role of option compensation and other factors. *Journal of Financial Economics*, 85(3):667–708.

Engelberg, J., Gao, P., and Parsons, C. (2012). Friends with money. *Journal of Financial Economics*, 103(1):169–188.

Fibich, G. and Gavish, N. (2011). Numerical simulations of asymmetric first-price auctions. *Games and Economic Behavior*, 73(2):479 – 495.

Ghosh, A., Manwani, N., and Sastry, P. (2017). On the robustness of decision tree learning under label noise. *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 685–97.

Ghosh, A. and Morita, H. (2017). Knowledge transfer and partial equity ownership. *RAND Journal of Economics*, 48(4):1044–1067.

Gilje, E., Gormley, T. A., and Levit, D. (2020). Who's paying attention? Measuring common ownership and its impact on managerial incentives. *Journal of Financial Economics*, 137(1):152–178.

Harford, J., Jenter, D., and Li, K. (2011). Institutional cross-holdings and their effect on acquisition decisions. *Journal of Financial Economics*, 99(1):27–39.

Harrington, J. E. (2005). Detecting Cartels. Economics Working Paper Archive 526, The Johns Hopkins University,Department of Economics.

He, J. and Huang, J. (2017). Product Market Competition in a World of Cross-Ownership: Evidence from Institutional Blockholdings. *Review of Financial Studies*, 30(8):2674–2718.

He, J. J., Huang, J., and Zhao, S. (2019). Internalizing governance externalities: The role of institutional cross-ownership. *Journal of Financial Economics*, 134(2):400–418.

Heath, D., Macciocchi, D., Michaely, R., and Ringgenberg, M. C. (2021). Do Index Funds Monitor? *The Review of Financial Studies*, 00:1–41.

Hemphill, C. S. and Kahan, M. (2020). The Strategies of Anticompetitive Common Ownership. *The Yale Law Journal*, 129:1392–1459.

Hickman, B. R. and Hubbard, T. (2015). Replacing sample trimming with boundary correction in nonparametric estimation of first-price auctions. *Journal of Applied Econometrics*, 30(5):739–762.

Hubbard, T. P. and Paarsch, H. J. (2014). Chapter 2 - on the numerical solution of equilibria in auction models with asymmetries within the private-values paradigm. In Schmedders, K. and Judd, K. L., editors, *Handbook of Computational Economics Vol. 3*, volume 3 of *Handbook of Computational Economics*, pages 37 – 115. Elsevier.

Iliev, P. and Lowry, M. (2015). Are mutual funds active voters? *Review of Financial Studies*, 28(2):446–485.

Ishii, J. and Xuan, Y. (2014). Acquirer-target social ties and merger outcomes. *Journal of Financial Economics*, 112(3):344–363.

Khanna, V., Kim, E. H., and Lu, Y. (2015). Ceo connectedness and corporate fraud. *Journal of Finance*, 70(3):1203–1252.

Kirchmaier, T. and Grant, J. (2005). Corporate Ownership Structure and Performance in Europe. *European Management Review*, 5(2):231–245.

Krasnokutskaya, E. (2011). Identification and estimation of auction models with unobserved heterogeneity. *The Review of Economic Studies*, 78(1):293–327.

La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (1999). Corporate ownership around the world. *Journal of Finance*, 54(2):471–517.

Lambert, T. (2021). Mere common ownership and the antitrust laws. *Boston College Law Review*, forthcoming.

Levenstein, M. C. and Suslow, V. Y. (2006). What determines cartel success? *Journal of Economic Literature*, 44(1):43–95.

Lewellen, K. and Lowry, M. (2020). Does common ownership really increase firm coordination? *Journal of Financial Economics*, forthcoming.

Li, T., Perrigne, I., and Vuong, Q. (2000). Conditionally independent private information in ocs wildcat auctions. *Journal of Econometrics*, 98(1):129 – 161.

Li, T. and Zhang, B. (2015). Affiliation and entry in first-price auctions with heterogeneous bidders: An analysis of merger effects. *American Economic Journal: Microeconomics*, 7(2):188–214.

Matvos, G. and Ostrovsky, M. (2008). Cross-ownership, returns, and voting in mergers. *Journal of Financial Economics*, 89(3):391–403.

McAfee, R. and McMillan, J. (1992). Bidding rings. *American Economic Review*, 82(3):579–99.

OECD (2009). Guidelines for fighting bid rigging in public procurement.

Paarsch, H. and Hong, H. (2006). *An Introduction to the Structural Econometrics of Auction Data*, volume 1. The MIT Press, 1 edition.

Park, A. H. and Seo, K. (2019). Common Ownership and Product Market Competition: Evidence from the U.S. Airline Industry. *Korean Journal of Financial Studies*, 48(5):617–640.

Pesendorfer, M. (2000). A study of collusion in first-price auctions. *Review of Economic Studies*, 67(3):381–411.

Porter, R. and Zona, J. (1993). Detection of bid rigging in procurement auctions. *Journal of Political Economy*, 101(3):518–38.

Reynolds, R. J. and Snapp, B. R. (1986). The competitive effects of partial equity interests and joint ventures. *International Journal of Industrial Organization*, 4(2):141–153.

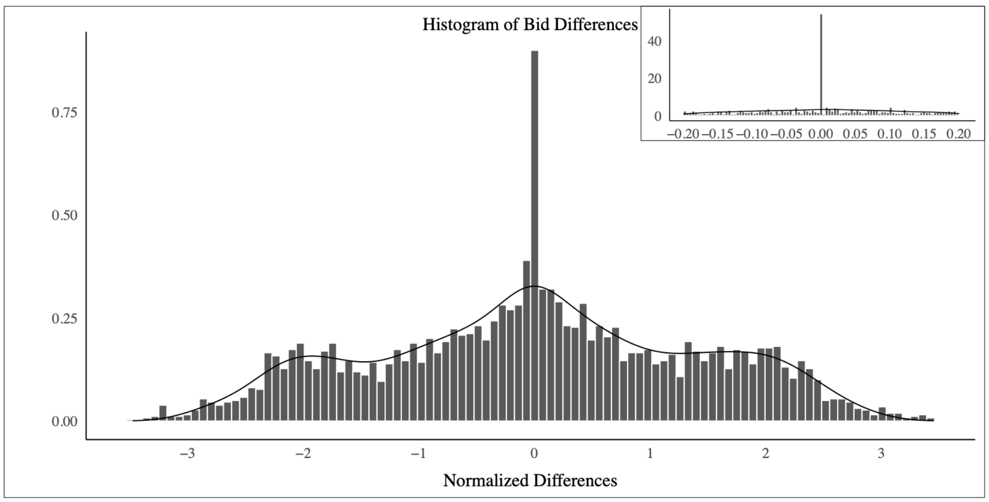
Walker, D. I. (2019). Common ownership and executive incentives: The implausibility of compensation as an anticompetitive mechanism. *Boston University Law Review*, 99(6):2373–2413.

Yang, B.-B., Gao, W., and Li, M. (2019). On the robust splitting criterion of random forest. *2019 IEEE International Conference on Data Mining (ICDM)*, pages 1420–25.

Tables and Figures

Figure 1: Distribution of Bid Differences

The figures below show the distribution of differences from a randomly selected pair of bids from auctions with more than three bids. We normalize each bid by the standard deviation of all bids in the auction. The top panel shows the bid differences across all valid auctions, whereas the second panel splits it into mutually exclusive categories of Goods versus Services. For each histogram, we also show a zoomed-in version restricting the range to be from -0.2 to 0.2.



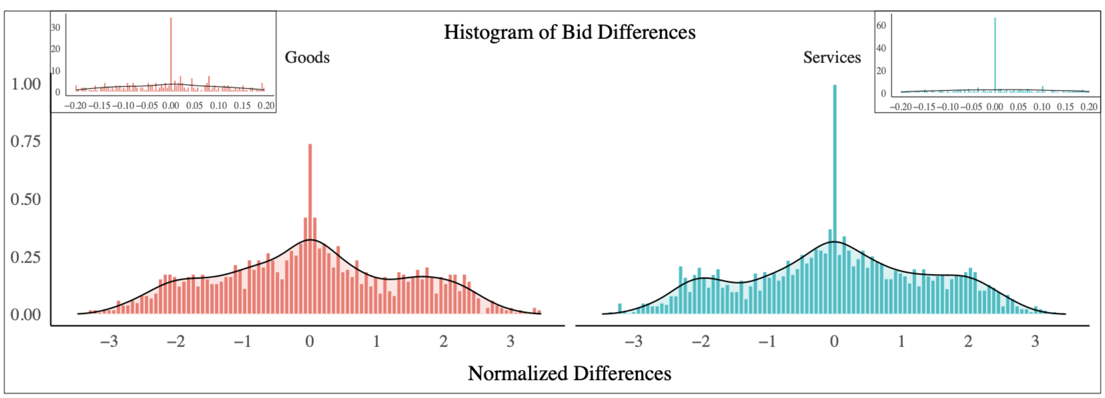


Table 1: Summary Statistics

The table below shows the summary statistics for our main sample. Panel A shows the characteristics and counts of auctions and shows the summary statistics for auction-level observations. Panel B shows summary statistics of all the bids. Panel C considers pairwise ownership connections and shows summary statistics for pair-level observations, where common owners and common owners' owners are scaled to be percentage points (0.01 represents 1 basis point). "N" refers to the number of observations and "SD" stands for standard deviation. Common owners and common owners' owners correspond to the second- and fourth-degree current shareholder connections, respectively, and are all defined as an indicator taking the value of one if there is at least one such owner connection in the auction or in the pair. Num. Common Owners is NC^2 and Num. Common Owners' Owners is NC^4.

Panel A: Auctions (N = 9,087)

5,397 service auction; 3,368 goods auction; 322 construction

Auction Type	N	Variables	Mean	SD
Services	2,086	Winning Bid	359,341	8,476,370
Administration & Training	1,920	Num. Bids	5.227	4.071
IT&Telecommunication	1,190	1{Ineffective Bids > 0}	0.157	0.364
Construction	729	Num. Ineffective Bids	0.285	0.897
Facilities Management	670	Common Owners	0.033	0.179
Event Organising, Food & Beverages	630	Common Owners' Owners	0.002	0.042
Dental, Medical & Laboratory	571	Num. Common Owners	0.042	0.271
Transportation	507	Num. Common Owners' Owners	0.002	0.046
Furniture, Office Equipment & Audio-Visual	420	Roundedness	0.203	0.295
Miscellaneous	252			
Workshop Equipment and Services	112			

Panel B: Bids (N = 47,495)		Panel C: Connections (Pair-wise N = 99,661)		
Mean	760,138	Variables	Mean	SD
Standard Deviation	12,195,444	Prob. Submitting Identical Bids	0.021	0.132
10^{th} Percentile	5,382	Num. Joint Participations	1.761	3.052
25^{th} Percentile	10,027	Common Owners (%)	0.099	3.150
Median	21,800	Common Owner's Owners (%)	0.008	0.896
75^{th} Percentile	53,500			
90^{th} Percentile	161,864			

Table 2: Ownership Networks and Identical Bidding

The table below shows the relation between the presence of shareholder connections between participating firms and the number of ineffective bids in an auction. Common owners and common owners' owners correspond to the second- and fourth-degree current shareholder connections, respectively, and are all defined as an indicator taking the value of one if there is at least one such owner connection in the auction. Columns (1) to (3) present results for the baseline specification, whereas columns (4) to (6) present results from an analogous random-forest propensity-score-weighted regression. All regressions include procurer-by-auction type fixed effects as well as cubic controls of the average roundedness of bids in an auction. Panel A includes cubic controls of the log number of bidders and Panel B does not. The coefficients on the controls are suppressed for space. Standard errors are clustered by procurer and shown in parentheses. * denotes p < 0.10, ** denotes p < 0.05, and *** denotes p < 0.01.

Panel A: Controlling for Number of Bidders						
Dependent Variable:	Num. Ineffective Bids					
Specification:	Baseline			P.S. Weighted		
	(1)	(2)	(3)	(4)	(5)	(6)
Common Owners	0.443***		0.443***	0.326**		0.327**
	(0.069)		(0.069)	(0.126)		(0.126)
Common Owners' Owners		0.301***	0.309***		0.324***	0.337***
		(0.085)	(0.082)		(0.088)	(0.087)
Observations	9,087	9,087	9,087	9,087	9,087	9,087
Cubic Num. Bidder Controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.308	0.302	0.308	0.274	0.269	0.275

Panel B: Without Number of Bidder Controls						
Dependent Variable:	Num. Ineffective Bids					
Specification:	Baseline			P.S. Weighted		
	(1)	(2)	(3)	(4)	(5)	(6)
Common Owners	0.749***		0.749***	0.465***		0.466***
	(0.129)		(0.129)	(0.159)		(0.159)
Common Owners' Owners		0.489***	0.491***		0.511***	0.524***
		(0.078)	(0.078)		(0.115)	(0.116)
Observations	9,087	9,087	9,087	9,087	9,087	9,087
Cubic Num. Bidder Controls	No	No	No	No	No	No
R^2	0.183	0.164	0.184	0.186	0.175	0.187

Table 3: Pair-Level Analysis

The table below shows the effects of ownership connections on the probability of submitting identical bids, focusing on pairs of firms that participate in the same auction at least once or five times. Observations are at the pair of participating firms level. Common owners and common owners' owners are indicator variables taking the value of one if there is such connection, respectively. Columns (3) and (4) include firm fixed effects, where within R-squared is reported. Standard errors are in parentheses. * denotes p < 0.10, ** denotes p < 0.05, and *** denotes p < 0.01.

Dependent Variable:	Prob. Submitting Identical Bids			
	(1)	(2)	(3)	(4)
Common Owners	0.134***	0.068***	0.129***	0.069***
	(0.013)	(0.026)	(0.012)	(0.022)
Common Owners' Owners	0.255***	0.172*	0.210***	0.140*
	(0.047)	(0.094)	(0.038)	(0.078)
Firm Fixed Effects	No	No	Yes	Yes
Sample	$N \geq 1$	$N \geq 5$	$N \geq 1$	$N \geq 5$
Observations	99,661	4,924	99,661	4,924
R^2	0.001	0.002	0.001	0.003

Table 4: Robustness Test of Ownership Networks and Identical Bidding

The table below shows the relation between the presence of shareholder connections between participating firms and the number of ineffective bids in an auction. Common owners and common owners' owners correspond to the second- and fourth-degree current shareholder connections, respectively, and are all defined as an indicator taking the value of one if there is at least one such owner connection in the auction. In both panels, columns (1) and (2) exclude auctions where the median bid is more than S$50,000 and S$100,000, respectively, column (3) excludes construction-type auctions, column (4) excludes procurers related to infrastructure and development including the Economic Development Board, Housing and Development Board (which constructs all public housing in Singapore), Building and Construction Authority, Land Transport Authority, Maritime and Port Authority of Singapore, Public Transport Council, and Sentosa Development Corporation, and column (5) excludes both construction-type auctions as well as agencies related to infrastructure and development. All regressions include procurer-by-auction type fixed effects as well as cubic controls of the log number of bidders and the average roundedness of bids in an auction, whose coefficients are suppressed for space. Standard errors are clustered by procurer and shown in parentheses. * denotes $p < 0.10$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Panel A: Intensive Margin (*Dependent Variable: Num. Ineffective Bids*)					
	(1)	(2)	(3)	(4)	(5)
Common Owners	0.482***	0.474***	0.415***	0.483***	0.451***
	(0.104)	(0.083)	(0.069)	(0.079)	(0.087)
Common Owners' Owners	0.271***	0.297***	0.301***	0.294***	0.284***
	(0.086)	(0.084)	(0.086)	(0.084)	(0.088)
Sample	Median Bid	Median Bid	Excl.	Excl. Dev.	Excl. Const. &
	$<$ 50,000	$<$ 100,000	Const.	Agencies	Dev. Agencies
Observations	6,859	8,187	8,358	8,185	7,581
R^2	0.345	0.330	0.327	0.311	0.326
Panel B: Extensive Margin (*Dependent Variable: 1{Ineffective Bids > 0}*)					
	(1)	(2)	(3)	(4)	(5)
Common Owners	0.091***	0.095***	0.083***	0.100***	0.095***
	(0.021)	(0.019)	(0.017)	(0.016)	(0.018)
Common Owners' Owners	0.182***	0.192***	0.195***	0.196***	0.192***
	(0.057)	(0.060)	(0.061)	(0.061)	(0.062)
Sample	Median Bid	Median Bid	Excl.	Excl. Dev.	Excl. Const. &
	$<$ 50,000	$<$ 100,000	Const.	Agencies	Dev. Agencies
Observations	6,859	8,187	8,358	8,185	7,581
R^2	0.324	0.313	0.309	0.304	0.309

Table 5: Number of Ineffective Bids and Contract Price

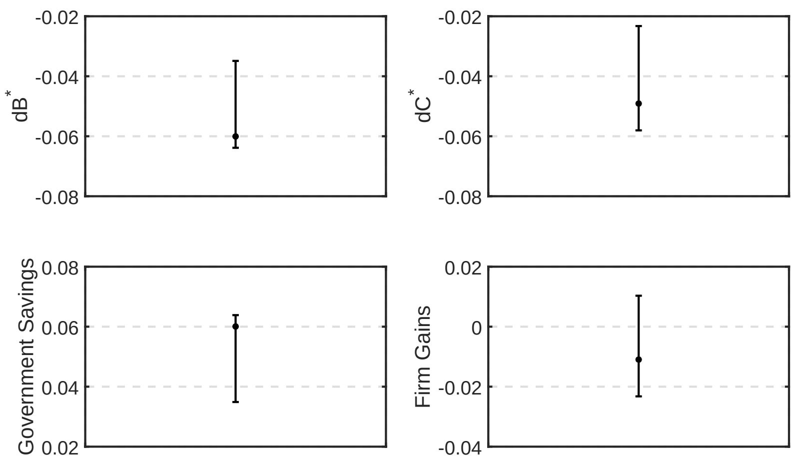
The table below shows the relation between the normalized winning bid and the number of ineffective bids. Columns (1) and (2) present the results for all eligible auctions, columns (3) and (4) only use auctions for goods, and columns (5) and (6) only use auctions for services and construction. Columns (1), (3), and (5) present results for the baseline specification, whereas columns (2), (4), and (6) present results from an analogous random-forest propensity-score-weighted regression. All regressions include procurer-by-auction type fixed effects as well as cubic controls of the average roundedness of bids in an auction. Panel A includes cubic controls of the log number of bidders and Panel B does not. The coefficients on the controls are suppressed for space. Standard errors are clustered by procurer and shown in parentheses. * denotes $p < 0.10$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Panel A: Controlling for the Number of Bidders						
Dependent Variable:	Winning Bid Normalized by Median Bid					
Procurement Type	All		Goods		Services & Construction	
Specification:	Baseline	P.S. Weighted	Baseline	P.S. Weighted	Baseline	P.S. Weighted
	(1)	(2)	(3)	(4)	(5)	(6)
Num. Ineffective Bids	0.043***	0.046***	0.043***	0.047***	0.045***	0.051***
	(0.002)	(0.004)	(0.004)	(0.006)	(0.002)	(0.005)
Observations	9,087	9,087	3,368	3,368	5,719	5,719
Cubic Num. Bidder Controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.291	0.272	0.294	0.306	0.352	0.328
Panel B: Not Controlling for the Number of Bidders						
Dependent Variable:	Winning Bid Normalized by Median Bid					
Procurement Type	All		Goods		Services & Construction	
Specification:	Baseline	P.S. Weighted	Baseline	P.S. Weighted	Baseline	P.S. Weighted
	(1)	(2)	(3)	(4)	(5)	(6)
Num. Ineffective Bids	0.013***	0.017***	0.013**	0.019**	0.014***	0.020***
	(0.002)	(0.005)	(0.006)	(0.008)	(0.002)	(0.006)
Observations	9,087	9,087	3,368	3,368	5,719	5,719
Cubic Num. Bidder Controls	No	No	No	No	No	No
R^2	0.212	0.203	0.238	0.256	0.265	0.249

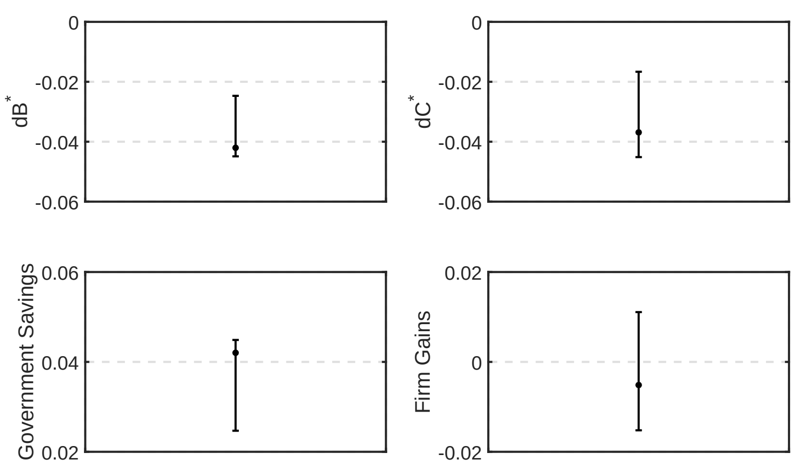
Figure 2: The Simulation Outcome of Excluding Ownership Network Effects

In every panel, black dots stand for point estimates and their associated intervals represent corresponding 95% confidence intervals based on bootstrap. We present the estimated effect of excluding the link between the presence of second-degree or fourth-degree current-shareholder connections and identical bidding on the winning bid in the top left panel, the effect on the winner's cost in the top right panel, the corresponding government savings in the bottom left panel, and the corresponding firm gains in the bottom right panel. Sub-Figure 2a uses the reduced-form evidence in Table 2 Panel A column (1) and Sub-Figure 2b uses the reduced-form evidence in Table 2 Panel A column (2).

(a) Excluding the Effects of Second-Degree Connections



(b) Excluding the Effects of Fourth-Degree Connections



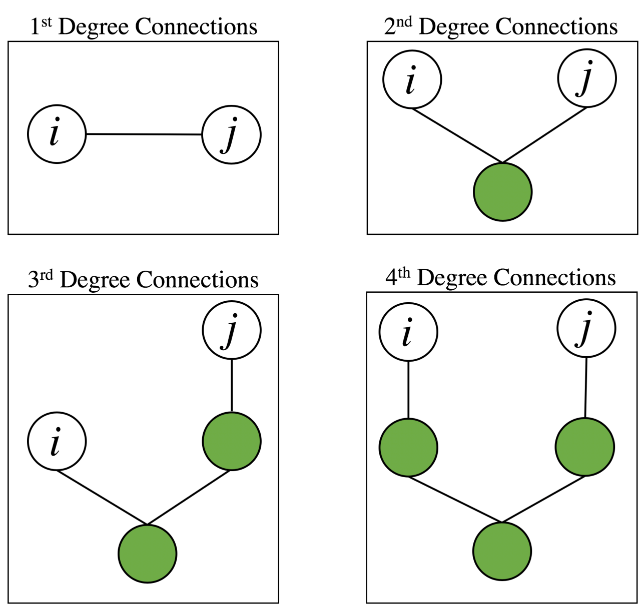
Online Appendix

A Additional Details and Robustness Tests

A.1 Matrix Representation of the Ownership Network

Figure A.1: Connections

We show example connections below for some arbitrary firms i and j. Shaded nodes indicate either an individual or a firm. A first-degree connection requires one firm to own another. A second-degree connection, meaning two firms have common owners, requires either a common individual or firm owner. A third-degree connection requires at least one instance of a firm owning another. The bottom shaded node may be a firm, and the top shaded node must be a firm. A fourth-degree connection, meaning two firms have common owners' owners, also requires at least one instance of a firm owning another. The bottom node may be an individual or company, and the top two shaded nodes must be companies.



In addition to Figure A.1, which shows how our network measure corresponds to labeled ownership relationships, we provide examples of how adjacency matrices, featuring both firms and individuals, represent ownership networks. First, we start from a simple network. Suppose there are two firms 1 and 2 without cross ownership, but there is a common owner that is indexed as investor 3. Let the ownership matrix be A. Entry (i,j) of A represents the presence (= 1) or absence (= 0) of link. There is a link between i and j if i is a shareholder of j or vice versa. As our focus is on whether firms have a connection, we consider an undirected network and, therefore, symmetric entries of A take the same value. Diagonal entries of A are 0 as a firm is not an owner of itself (or vice versa). For this example, entries (1,3) and (3,1) of A are 1. Moreover, entries (2,3) and (3,2) become 1 as well. Then, A is characterized by

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Also,

$$A^2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Entry (1,2) becomes positive for the first time under A^2, and so $\inf\{x \geq 1 | A^x(1,2) > 0\} = 2$. As a result, $S^2(1,2) = 1$, meaning that there is a second-degree connection between firms 1 and 2. In this way, common ownership is represented by a second-degree connection.

Next, we consider a more complex network. Suppose there are six entities 1 through 6. No cross ownership between firms 1 through 3 exists. Firm 4 is a common owner of firms 1 and 2. Also, firm 5 is an owner of firm 3. Finally, investor 6 is a common owner of firms 4 and 5. Then, the ownership matrix A is

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

Entries (1,4), (2,4), and (3,5) are already positive under A, and so $\inf\{x \geq 1 | A^x(i,j) > 0\} = 1$

for $(i,j) \in \{(1,4),(2,4),(3,5)\}$. Consequently, $S^1(i,j) = 1$ for $(i,j) \in \{(1,4),(2,4),(3,5)\}$. Then, there is a first-degree connection between firms 1 and 4, firms 2 and 4, as well as firms 3 and 5. In this way, cross ownership is represented by a first-degree connection. We continue computing higher powers of A:

$$A^2 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 1 & 1 & 1 & 0 & 0 & 2 \end{bmatrix}.$$

Because entries (1,2) and (4,5) become positive for the first time under A^2, $\inf\{x \geq 1 | A^x(i,j) > 0\} = 2$ for $(i,j) \in \{(1,2),(4,5)\}$. As a result, $S^2(i,j) = 1$ for $(i,j) \in \{(1,2),(4,5)\}$. This means that there is a second-degree connection between firms 1 and 2 as well as firms 4 and 5. Again, common ownership is captured by a second-degree connection. Next, we find

$$A^3 = \begin{bmatrix} 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 3 & 3 & 1 & 0 & 0 & 4 \\ 1 & 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 & 3 & 0 \end{bmatrix}.$$

Because entries (1,5), (2,5), and (3,4) become positive for the first time under A^3, $\inf\{x \geq 1 | A^x(i,j) > 0\} = 3$ for $(i,j) \in \{(1,5),(2,5),(3,4)\}$. Therefore, $S^3(i,j) = 1$ for $(i,j) \in \{(1,5),(2,5),(3,4)\}$. This means that the pairs of firms 1 and 5, firms 2 and 5, and firms 3 and 4 have third-degree connections. Last, we have

$$A^4 = \begin{bmatrix} 3 & 3 & 1 & 0 & 0 & 4 \\ 3 & 3 & 1 & 0 & 0 & 4 \\ 1 & 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 0 & 10 & 5 & 0 \\ 0 & 0 & 0 & 5 & 5 & 0 \\ 4 & 4 & 3 & 0 & 0 & 7 \end{bmatrix}.$$

Because entries (1,3) and (2,3) become positive for the first time under A^4, $\inf\{x \geq 1 | A^x(i,j) > 0\} = 4$ for $(i,j) \in \{(1,3),(2,3)\}$. Then, $S^4(i,j) = 1$ for $(i,j) \in \{(1,3),(2,3)\}$. This means that the pairs of firms 1 and 3 as well as firms 2 and 3 have fourth-degree connections. In this way, common owner's ownership is represented by a fourth-degree connection.

In general, entry (i,j) of A^x captures the number of ways entity i reaches entity j with a walk of x steps through paths defined by our links. For example, $A^4(4,5) = 5$. Indeed, there are 5 ways of reaching entity 5 from entity 4 with 4 steps: $4 \to 6 \to 4 \to 6 \to 5$, $4 \to 6 \to 5 \to 6 \to 5$, $4 \to 2 \to 4 \to 6 \to 5$, $4 \to 1 \to 4 \to 6 \to 5$, and $4 \to 6 \to 5 \to 3 \to 5$.

A.2 Singapore Government Procurements

A.2.1 Difficulty of Side Payments in Singapore

We believe facilitating side payments is difficult in Singapore due to the risk of prosecution by the government for two reasons. First, the Goods and Services Tax (GST) system in Singapore deters side payments that are unrecorded by the government. Most sales of goods and services sold within Singapore are subject to this tax. Businesses with taxable turnover in terms of costs or revenues over S$1 million (approximately US$720,000) must register for the GST. These businesses must charge GST on all their sales but may claim input tax credits based on GST incurred on their own purchases and expenses. To help companies check whether they are required to register for GST, the tax authority in Singapore provides an online calculator where firms can input their financial statement information to determine whether they should register.Thus, a medium-sized enterprise making a side payment without proper GST invoicing would not be able to use input tax credits to offset their existing tax burden, and the receiving firm would violate the GST registration. Throughout the sample, this tax is 7% and nonpayment carries large consequences, including a late payment penalty of 5% and subsequently 2% interest per month up to a limit of 50% of the tax outstanding. In addition, the Internal Revenue Authority of Singapore (IRAS) may also consider legal actions, including directly appointing employers, tenants, or lawyers to pay the sum, issuing travel restrictions, and other legal sanctions. For example, from January to September 2019, the tax authority recovered S$175 million in GST from more than 2,000 investigations.

Second, even though side payments might be possible through non-recorded cash payments, mandatory audits, severe penalties, or incentives for whistleblowing deter this practice. Indeed, side payments for larger firms are less plausible because audits are required for firms meeting two of the three following conditions: more than S$10 million in revenues, more than S$10 million in assets, or more than 50 employees. If not caught by the auditors, the detection of unaccounted side payments by the IRAS is subject to large penalties for tax evasion, carrying a penalty of up to 400% of the missing tax, fines of up to S$50,000, and imprisonment up to several years for the guilty person. To deter tax evasion, the IRAS makes public all detected violations, their case, and the penalty. In 2018, there were 15 detected and prosecuted cases of tax evasion, ranging from sole proprietors to medium corporations. In addition, the whistleblower program awards 15% of all taxes recovered, capped at S$100,000 (US$73,500) to informants who provide a lead to recovery. Thus, we believe the combination of transparent business practices, high penalties of deviations, and incentives to report misbehavior deter illicit side payments, and the assumption of no side payments is reasonable.

A.2.2 Government Awarding Criteria

The Singapore government's guidelines for the awarding process of public procurement auction stipulate that price is one important criterion in the awarding rule, but other dimensions such as quality, ability to service a contract, and whether a procurer concurrently has multiple open auctions also play a role.[28] The lowest bidders win around half of the public procurement auctions in Singapore. We estimate several statistical models to predict whether the lowest bidder wins an auction purely based on publicly available data and find predictability between 60% to 80% across different models.

In our empirical exercise, we consider 9 auction features from point-in-time publicly available data. We include the third-order polynomial of the log number of bidders, the auction type, the procurement method (open quotation or tenders), the log number of concurrent auctions, the log number of concurrent auctions in the same auction type, and the log number of concurrent auctions in the same auction type that the same procurer has. Each of the latter three is measured over the past 30 days based on an auction's open date. To maximize our sample, we do not drop auctions at the beginning of our sample and assume there were no overlapping auctions before that.

Table A.1: Performance Metrics of Models Predicting the Lowest Bid Winning

	Performance Metrics				
	Correct	Correct Positive	Correct Negative	False Positive	False Negative
Model		(Sensitivity)	(Specificity)	(Type I)	(Type II)
Logit	0.619	0.297	0.844	0.156	0.703
Probit	0.620	0.299	0.844	0.156	0.701
LASSO	0.620	0.298	0.845	0.155	0.702
Random Forest	0.765	0.578	0.896	0.104	0.422

We consider four models, using logistic and probit regressions as a benchmark, a LASSO specification with 10 fold cross-validation for variable selections, and finally a random forest

[28] Section 2 discusses in detail the framework provided by the Singapore government for awarding public procurement contracts.

with 2 variables to try at each tree split, with 500 trees in the forest. Other hyperparameters are the default from the *ranger* package in R, kept for simplicity. Nodes are split based on Gini impurity, and the model uses bagging to prevent overfitting (Breiman, 2001).

Table A.1 reports the average performance of these models that are constructed to predict an auction's winner. The random forest model performs the best when using the whole sample with a correct probability of around 80%, while the other models have a correct probability of around 62%. The improved performance suggests that nonlinear interactions of the observed variables are statistically important. The random forest model's correct positive probability is around 60% and that of correct negative is around 90%.

Thus, to understand contract rewarding criteria, the random forest provides a better approximation to the actual decision process. We show the importance chart in Figure A.2, based on the mean decrease in Gini impurity. The log number of concurrent auctions and the log numbrer of bidders are the most important variables.

Crucially, Table A.1 also shows that the percentage of correct negative is always higher than that of correct positive. If our empirical setting is such that the lowest bidder wins if and only if (1) the auction is a first-price auction (FPA) or (2) the auction is a multi-attribute auction but the lowest bidder wins by chance, then even a perfectly precise model in terms of predicting auction format would predict auctions in which the lowest bidders do not win with higher predictability, and those in which the lowest bidders win with lower predictability, because it cannot predict the case of (2) above. Then, the probability of correct positive is lower, that of false positive is small, that of correct negative is higher, and that of false negative is modest. This is close to what we find. Therefore, our results above are consistent with each statistical model predicting auction format.

Finally, to show the effects of different procurement types, Table A.2 shows the estimated average marginal effects in linear models. The number of bidders is negatively correlated with the indicator for the lowest bidder winning the auction. We also find that the lowest bidders tend to win construction, facilities, transportation, furniture, and information &

technology auctions. These auction types are more commodity products compared to other types like health and services, which may have more heterogeneity in quality. We note that despite construction having a lot of heterogeneity in quality, in Singapore, all construction companies registering for GeBIZ require certification, so the heterogeneity is limited.

Figure A.2: Random Forest Variable Importance Plot

The figures below visualize the variables in the random forest based on a decreasing order of importance. We show the variable importance plot based on Gini impurity. The variables denoted as either categorical or numerical. “Auction (Type)” considers auctions with the same auction type and “Auction (Type, Procurer)” considers auctions with the same auction type and procurer. “w/i” denotes “within”.

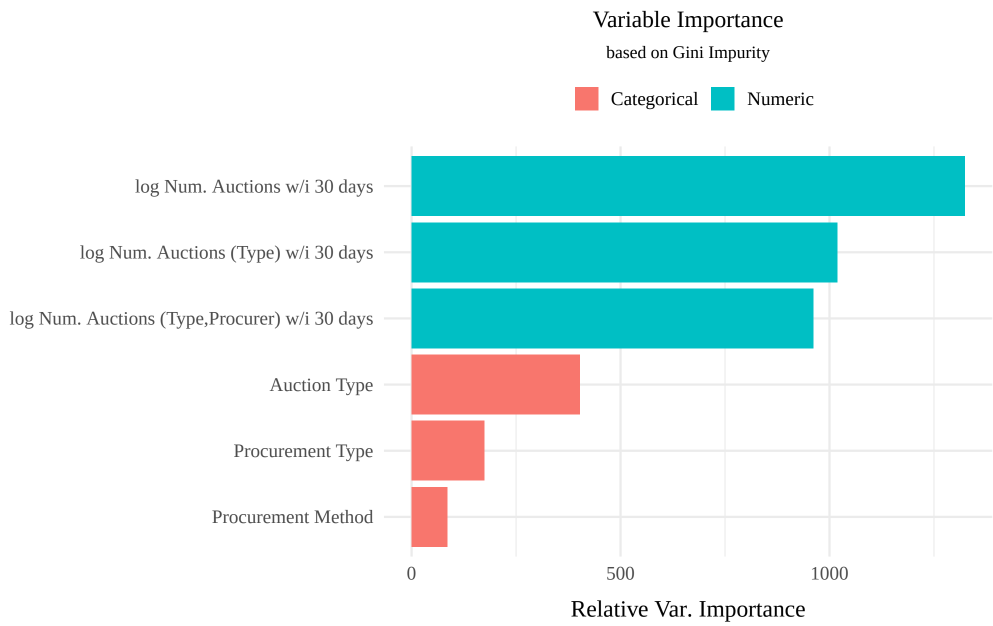


Table A.2: Linear Models for the Lowest Bid Winning

The table below presents the estimates from a linear fixed-effects model, probit, and a logistic model to show what variables correlate most with whether the lowest bid in an auction won the auction. For comparability, the average marginal effects are shown for the probit and logistic models. We cluster standard errors by procurer. We suppress the estimates for the constants and denote control variables in the "Controls" row for space. The sample are all auctions in our data. Standard errors are in parentheses after the point estimate.

Dependent variable:	Lowest Bidder Wins			
Model	Linear		Probit	Logistic
	(1)	(2)	(3)	(4)
log(1+Number of Auctions Overlapping)	-0.042 (0.014)***	-0.010 (0.022)	-0.040 (0.005)***	-0.040 (0.005)***
log(1+Number of Auctions in Same Auction Type and Procurer Overlapping)	0.006 (0.009)	0.012 (0.006)**	0.005 (0.004)	0.005 (0.004)
log(1+Number of Auctions in Same Auction Type Overlapping)	0.055 (0.016)***	0.025 (0.025)	0.054 (0.008)***	0.053 (0.008)***
Auction Type = Construction	0.267 (0.042)***	0.199 (0.049)***	0.262 (0.021)***	0.261 (0.021)
Auction Type = Health	0.224 (0.023)***	0.131 (0.042)***	0.221 (0.020)***	0.220 (0.020)***
Auction Type = Event Organizing	0.091 (0.042)**	0.054 (0.045)	0.086 (0.017)***	0.084 (0.016)***
Auction Type = Facilities	0.311 (0.025)***	0.232 (0.039)***	0.306 (0.020)***	0.305 (0.020)***
Auction Type = Furniture	0.209 (0.024)***	0.145 (0.044)***	0.204 (0.023)***	0.204 (0.023)***
Auction Type = Information & Technology	0.255 (0.019)***	0.208 (0.034)***	0.249 (0.015)***	0.248 (0.015)***
Auction Type = Miscellaneous	0.147 (0.057)**	0.028 (0.071)	0.142 (0.027)***	0.140 (0.027)***
Auction Type = Services	0.097 (0.022)***	0.091 (0.023)***	0.093 (0.009)***	0.093 (0.009)***
Auction Type = Transportation	0.201 (0.023)***	0.129 (0.044)***	0.196 (0.021)***	0.196 (0.021)***
Auction Type = Workshop	0.419 (0.075)***	0.286 (0.093)***	0.408 (0.041)***	0.405 (0.040)***
Control: Procurement Method (Open Quotation or Tender)	✓	✓	✓	✓
Control: Procurement Type (Goods or Services)	✓	✓	✓	✓
Fixed Effects		Procurer		
Observations	22,098	22,098	22,098	22,098
R^2	0.049	0.084	0.037	0.037

A.2.3 Identifying Award Criteria via Machine Learning

A random forest algorithm is different from a basic decision tree algorithm in that it is a set of randomized tree algorithms. As in a usual tree algorithm, each tree starts from the root node and moves down to the branches of the internal nodes by a splitting process, that is, the process of partitioning the data set into subsets. The split is determined by a particular criterion that aims to reduce a measure of inaccuracy or "impurity" in the grouping (*split criterion*). The prediction is made at the terminal nodes, or leaves, that are polarized enough that they can be labeled as either one of the outcome classes in line with a particular rule (*labeling rule*). One unique feature of the algorithm is that each tree selects a random set of features used for splitting (*random split selection*). Moreover, it is an ensemble prediction *(bagging)*. Specifically, each tree learns from the training data set sampled from the original data set with replacement. The final prediction is obtained by the majority vote of the trees.

In particular, because we do not observe the auction format—whether a contract is awarded based on price only or multiple attributes—but only observe whether the lowest bidder wins an auction, we frame our random forest analysis as one under label noise. We consider the observed indicator for whether the lowest bidder wins or not ($\tilde{y}_i = \pm 1$) as the noisy label for the unobserved indicator of auction format ($y_i = \pm 1$). If auction i is a first-price auction ($y_i = +1$), the lowest bidder wins ($\tilde{y}_i = +1$) for sure, so there is no noise. However, if it is a multi-attribute auction ($y_i = -1$), the lowest bidder still wins ($\tilde{y}_i = +1$) with some probability η, where $0 < \eta < 1$. In other words, conditional on $y_i = -1$, the true outcome y_i is not equal to the observed indicator $\tilde{y}_i$ with probability η. If this mismatch is independent of auction features x_i, conditional on $y_i = -1$, label noise in our setting becomes *class-conditional.*

The recent literature on machine learning investigates whether the random forest classifier learnt from noise-free sample matches that learnt from the sample with label noise.[29] The literature suggests that classifiers trained on noise-free and noisy data coincide when the noise

[29]See Ghosh et al. (2017) and Yang et al. (2019), for example.

is class-conditional. In our setting, this is equivalent to considering whether our random forest classifier actually predicts auction format y_i, instead of the lowest bidder winning $\tilde{y}_i$. Because random splitting and bagging are immune to label noise (Breiman, 2001), the literature focuses on the noise-tolerance of split criterion and labeling rule used in each tree of the forest.

Therefore, we theoretically characterize whether and when a random forest classifier (one that is trained to predict whether the lowest bidder wins an auction based on ex ante observed auction characteristics) would align with the correct classifier of whether an auction is price-only. We show that the node splitting process that minimizes Gini impurity under noise-free samples is same as that which minimizes it in the presence of label noise in our setting. Moreover, we show that majority voting at a leaf node is also robust to label noise in our setting if certain conditions are met. To state these claims formally, we list a couple of assumptions.

Assumptions:

1. *There is a unique split rule minimizing Gini impurity at each node.*
2. $\eta < 0.5$.
3. *Let X denote the feature space. Let F denote a set of split rules. Let f_0 denote the split rule that perfectly predicts auction format y_i*, where $y_i = f_0(x_i), \forall x_i \in X$. $f_0 \in F$.

Assumption 1 is a regulatory condition used in the proof (Ghosh et al., 2017). Assumption 2 assures that the probability of the lowest bidder winning a multi-attribute auction is smaller than one-half. As discussed later, we evaluate the applicability of this assumption based on the performance metrics of the estimated model. Under Assumption 3, the auction format is predictable by features at hand. Then, we derive the following propositions.

Propositions:

1. *Noise-tolerance of split criterion: If Assumption 1 is met, in the large sample limit,* $\arg\min_{f \in F} C(f) = \arg\min_{f \in F} C^{\eta}(f)$*, where $C(f)$ is Gini impurity for a split rule $f \in F$ on noise-free data and $C^{\eta}(f)$ is corresponding Gini impurity for f under label noise.*
2. *Noise-tolerance of labeling rule: If Assumptions 1, 2, and 3 are met, in the large sample limit, the class predicted by the leaf node under label noise is the same as the one under noise-free case if the majority vote is used for labeling.*

3. *Equivalence: If Assumptions 1, 2, and 3 are met, in the large sample limit, the prediction made by the tree learnt under label noise is the same as the one predicted by the tree with depth 1 characterized by split* f_0.

4. *Observed performance in the large sample limit:*

   (a) *Correct positive (sensitivity):* $P[f_0(x_i) = +1|\tilde{y}_i = +1] < 1$.

   (b) *Correct negative (specificity):* $P[f_0(x_i) = -1|\tilde{y}_i = -1] = 1$.

5. *Predictability of auction format in the large sample limit:* $P[y_i = f_0(x_i)] = 1$.

Proof for Proposition. We follow the procedure and most of the notations used by Ghosh et al. (2017). Let the noise-free observations at a node v be $\{(x_i, y_i), i = 1, ..., n\}$. Under label noise, the observations at this node would become $\{(x_i, \tilde{y}_i), i = 1, ..., n\}$. Suppose in the noise-free case a split rule f sends n_l of these n observations to the left child v_l and $n_r = n - n_l$ to the right child v_r. Note that a split rule is a function of only the feature vector. Since the split rule depends only on the feature vector and not the labels, the points that go to v_l and v_r would be the same for the noisy observations also. Thus, n_l and $a = n_l/n$ would be the same in both cases.

Let n^+ and $n^- = n - n^+$ be the number of observations of the two classes at node v in the noise-free case. Similarly, let n_l^+ and $n_l^- = n_l - n_l^+$ be the number of observations of the two classes at v_l and define n_r^+, n_r^-, similarly. Let the corresponding quantities in the noisy case be $\tilde{n}^+, \tilde{n}^-, \tilde{n}_l^+, \tilde{n}_l^-$ etc. Define random variables, W_i, $i = 1, ..., n$ by $W_i = 1$ if $\tilde{y}_i \neq y_i$ and $W_i = 0$, otherwise. Thus, W_i are indicators of whether or not the label on the i^{th} example is corrupted. In our setting, $W_i = 0$, conditional on $y_i = +1$, and W_i are i.i.d. Bernoulli random variables with expectation η, conditional on $y_i = -1$.

Let $p = n^+/n, q = n^-/n = (1 - p)$ be the fractions of the two classes at v under noise-free case. Let p_l, q_l and p_r, q_r be these fractions for v_l, v_r. Let the corresponding quantities for the noisy observations case be $\tilde{p}, \tilde{q}, \tilde{p}_l, \tilde{q}_l$ etc. Let p^η, q^η be the values of $\tilde{p}, \tilde{q}$ in the large sample limit and similarly define $p_l^\eta, q_l^\eta, p_r^\eta, q_r^\eta$.

The value of $\tilde{n}^+$ is the number of i such that $\tilde{y}_i = +1$. Similarly, the value of $\tilde{n}_l^+$ would be the number of i such that x_i is in v_l and $\tilde{y}_i = +1$. Hence we have

$$\tilde{p} = \frac{\tilde{n}^+}{n} = \frac{1}{n}\left(\sum_{i:\tilde{y}_i=+1} 1\right) = \frac{1}{n}\left(\sum_{i:y_i=+1} 1 + \sum_{i:y_i=-1} W_i\right),$$

$$\tilde{p}_l = \frac{\tilde{n}_l^+}{n_l} = \frac{1}{n_l}\left(\sum_{i:x_i\in v_l,\tilde{y}_i=+1} 1\right) = \frac{1}{n_l}\left(\sum_{i:x_i\in v_l,y_i=+1} 1 + \sum_{i:x_i\in v_l,y_i=-1} W_i\right).$$

All the above expressions involve sums of independent random variables. Hence the values of the above quantities in the large sample limit can be calculated, by laws of large numbers, by essentially replacing each W_i by its expected value. Thus, from the above, we get

$$p^\eta = p + q\eta = p(1 - \eta) + \eta; p_l^\eta = p_l + q_l\eta = p_l(1 - \eta) + \eta.$$

First, we prove Proposition 1. For a node v, the Gini impurity is $G_{Gini}(v) = 2pq$ under noise-free case. Under label noise, Gini impurity becomes

$$G^{\eta}_{Gini}(v) = 2p^{\eta}q^{\eta} = 2(p(1-\eta)+\eta)(1-\eta)q = G_{Gini}(v)(1-\eta)^2 + 2\eta(1-\eta)q.$$

Similar expressions hold for $G^{\eta}_{Gini}(v_l)$ and $G^{\eta}_{Gini}(v_r)$. The large sample value of impurity gain of f under label noise can be written as

$$\begin{aligned} gain^{\eta}_{Gini}(f) &= G^{\eta}_{Gini}(v) - \left[aG^{\eta}_{Gini}(v_l) + (1-a)G^{\eta}_{Gini}(v_r)\right] \\ &= gain_{Gini}(f)(1-\eta)^2 + 2\eta(1-\eta)\left[q - aq_l - (1-a)q_r\right] \\ &= gain_{Gini}(f)(1-\eta)^2. \end{aligned}$$

Thus, if $gain_{Gini}(f^1) > gain_{Gini}(f^2)$, then $gain^{\eta}_{Gini}(f^1) > gain^{\eta}_{Gini}(f^2)$. This means that a maximizer of impurity gain (minimizer of Gini impurity) under the noise-free case will be also a maximizer of gain under label noise in the large sample limit. Therefore, Proposition 1 holds under Assumption 1.

Second, we prove Proposition 2. Note $gain_{Gini}(f) \leq G_{Gini}(v)$ as $G_{Gini}(v_l) \geq 0$ and $G_{Gini}(v_r) \geq 0$. Because $gain_{Gini}(f_0) = G_{Gini}(v)$, f_0 maximizes impurity gain at any node under the noise-free case. Under Assumptions 1 and 3, Proposition 1 suggests f_0 also maximizes gain at any node under label noise in the large sample limit. Set v be the parent node for the leaf node. When f_0 is chosen, $(p_l, p_r) = (1, 0)$ or $(0, 1)$. Without loss of generality, set $(p_l, p_r) = (1, 0)$. In the large sample limit, $(p^{\eta}_l, p^{\eta}_r) = (1, \eta)$. Assumption 2 suggests $\eta < 0.5$. Therefore, the majority vote will assign a positive label to the left leaf node and a negative label to the right leaf node as it will do so under the noise-free case.

Last, we prove Proposition 3. Under Assumptions 1 and 3, Propositions 1 suggests f_0 maximizes impurity gain at any node both in the presence and absence of label noise in the large sample limit. Therefore, in the large sample limit, at the root node, split f_0 is selected both in the presence and absence of label noise. In the absence of label noise, each child node of the root node contains only positive or negative examples. Because the above analysis suggests $gain_{Gini}(f_0)$ is maximum impurity gain at each node under the noise-free case, $gain_{Gini}(f) \leq gain_{Gini}(f_0) = 0$ for any f at each child node. Therefore, the tree will not grow further and stops at depth 1 under the noise-free case. Without loss of generality, consider the right child node contains only negative examples. In the presence of label noise, further split of the right child node may improve Gini impurity as the node contains negative examples mistakenly labelled as positive (the left child node only contains positive examples labelled as positive, and there is no impurity gain by further split). However, the above analysis suggests, at the right child node, $gain^{\eta}_{Gini}(f) = gain_{Gini}(f)(1-\eta)^2 \leq gain_{Gini}(f_0)(1-\eta)^2 = 0$ for any f, meaning that further split does not improve Gini impurity in the large sample limit. Therefore, in the large sample limit, the tree will not grow further and stops at depth 1 under label noise as well. If Assumptions 1 through 3 are met, Proposition 2 suggests, in the large sample limit, the label of the leaf node predicted by the majority vote under label noise is the same as the one under the noise-free case. Thus, in the large sample limit, the prediction made by the tree with depth 1 characterized by split f_0 is the same as the one by the tree learnt under label noise.

Propositions 1 through 3 suggest that if Assumptions 1 through 3 are met, each tree in a random forest classifier based on Gini impurity is the same as the one that perfectly predicts the auction format, at least in the large sample limit. Although some of these assumptions are not directly testable, we can still check whether the result of our random forest analysis appears consistent with Propositions 4. Table A.1 shows that the correct positive for a random forest model lower than the correct negative rate (which is close to 1). This result is consistent with Proposition 4.[30] The result of our random forest analysis hence appears consistent with the model implication under these assumptions. Given Proposition 5, it is plausible that auctions predicted by our random forest classifier as the one in which the lowest bidder wins are first-price auctions.

A.2.4 Results Using Alternative Scaling

We here show the bid difference histogram using different subsets of sample auctions and scalings. First, the top panel in Figure A.3 shows similar histograms when looking across auctions with different numbers of bids. Second, the bottom panels in Figure A.3 show similar patterns without any scaling or with scaling bid differences by the difference between the maximum and minimum bids (as opposed to the standard deviation).

A.2.5 Coordinated Versus Competitive Bids

We also consider the relation between identical bidding and bid levels. We regress all bids on the indicator for identically submitted bids while controlling for auction fixed effects. Then, we compute the gap between identically submitted and other bids while removing auction-specific effects. Table A.3 reports the regression results. We find identically submitted bids are around S$17,400—or 14%— lower than others. For both specifications, our estimates

[30] We can also examine whether the estimate of η based on observed sensitivity (73%) and fraction of auctions predicted as the one in which the lowest bidder wins (40%) is consistent with Assumption 2. Notice $P[f_0(x_i) = +1|\tilde{y}_i = +1] = P[f_0(x_i) = +1]/(P[f_0(x_i) = +1] + \eta P[f_0(x_i) = -1])$. Therefore, $\eta = P[f_0(x_i) = +1](1/P[f_0(x_i) = +1|\tilde{y}_i = +1] - 1)/P[f_0(x_i) = -1] = 0.40(1/0.73 - 1)/0.6 = 0.247$. The estimate of η is 24.7% and smaller than 50%. This result is consistent with Assumption 2.

are statistically significant at the 5% level.

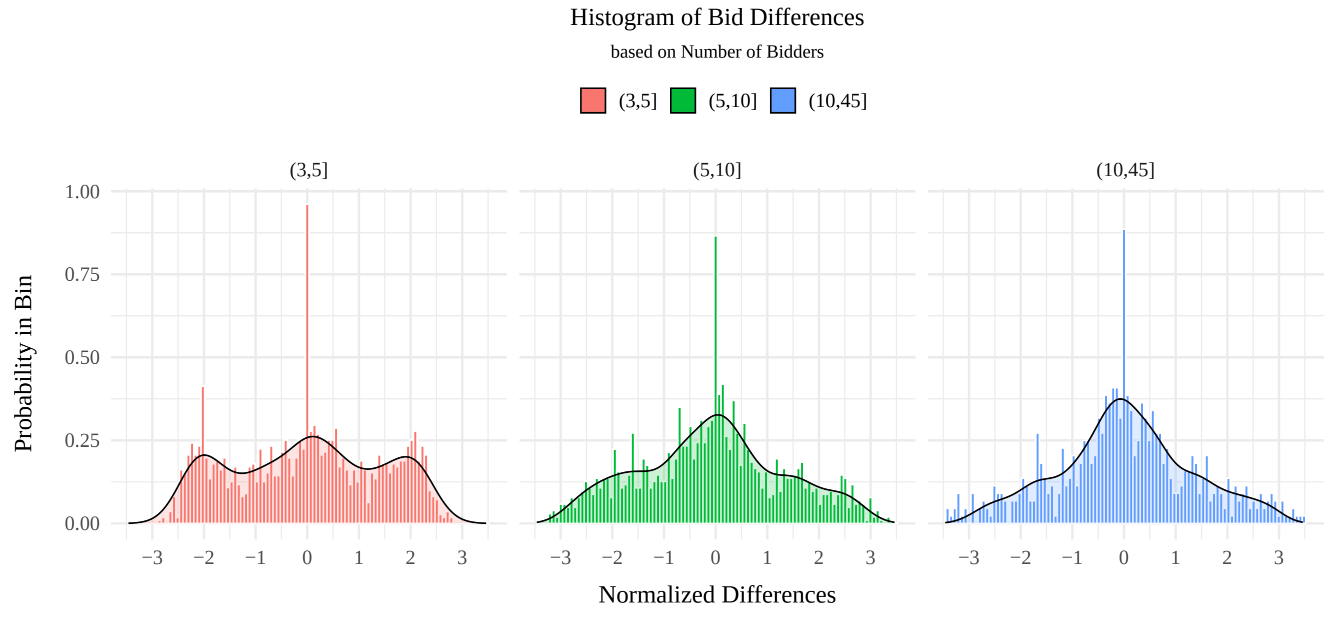
Table A.3: Coordinated Versus Competitive Bids

The table below shows the relation between identical bidding and bid levels. Observations are at the bid level. The dependent variable in column (1) is the bid amount in dollars and for column (2) is the log bid amount. Identical is an indicator variable taking the value of one if a bid is duplicated in an auction. All dependent variables are winsorized at the 1% level. All regressions include auction fixed effects. Standard errors are clustered by auction and shown in parentheses. * denotes $p < 0.10$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$. The coefficient and standard error in column (1) is rounded to the nearest integer to save space.

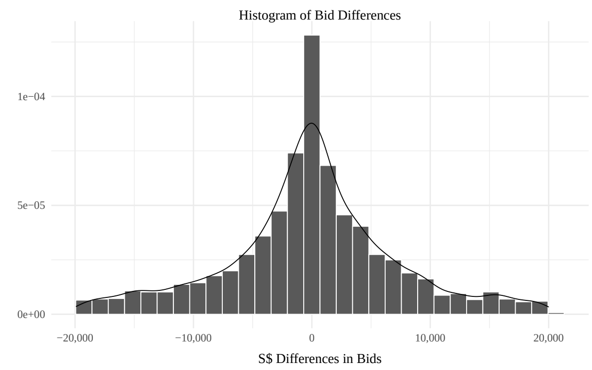
Dependent Variable:	Bid Amount	Log Bid Amount
	(1)	(2)
Identical	-17,415**	-0.136***
	(8,465)	(0.016)
Observations	47,495	47,495
R^2	0.950	0.920

Figure A.3: Distribution of Bid Differences with Different Number of Bidders and Normalizations

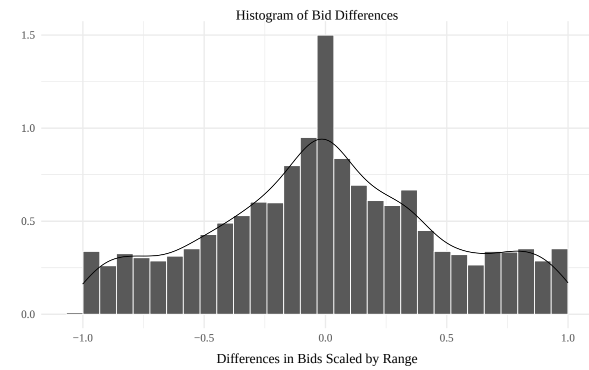
The top figure below shows the histogram based on auctions with different numbers of bidders. The two bottom figures below show the histogram based on auctions having more than three bidders with different scalings. We compute the difference between two sampled bids and make it either unscaled or scaled by the range of bids, where the range of bids is the maximum of all bids in the auction minus the minimum of all bids in the auction.



(a) Unscaled Raw Dollar Differences



(b) Scaling by Bid Range



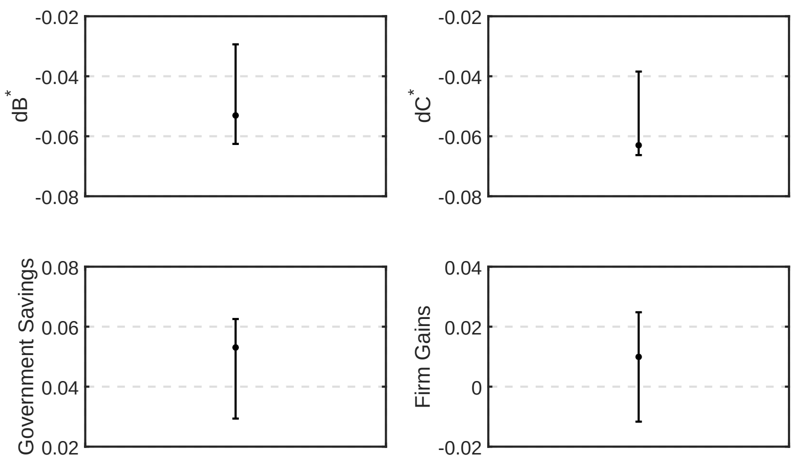
A.3 Robustness for Structural Analysis

For robustness check, we use the subsamples that are predicted by our random forest model as first-price auctions and allocate contracts to lowest bidders for our structural analysis. Except for this change, we follow the same procedure used in our benchmark analysis. Figure A.4 reports the results based on the alternative subsamples. We attain similar results as benchmark ones.

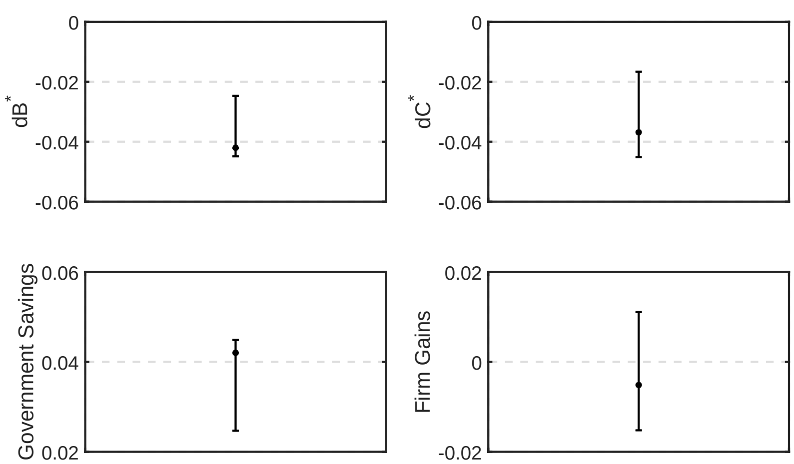
Figure A.4: The Simulation Outcome of Excluding Ownership Network Effects (Robustness)

In every panel, black dots stand for point estimates and their associated intervals represent corresponding 95% confidence intervals based on bootstrap. We present the estimated effect of excluding the link between the presence of second-degree or fourth-degree current-shareholder connections and identical bidding on the winning bid in the top left panel, the effect on the winner's cost in the top right panel, the corresponding government savings in the bottom left panel, and the corresponding firm gains in the bottom right panel. Sub-Figure A.4a uses the reduced-form evidence in Table 2 Panel A column (1) and Sub-Figure A.4b uses the reduced-form evidence in Table 2 Panel A column (2).

(a) Excluding the Effects of Second-Degree Connections



(b) Excluding the Effects of Fourth-Degree Connections



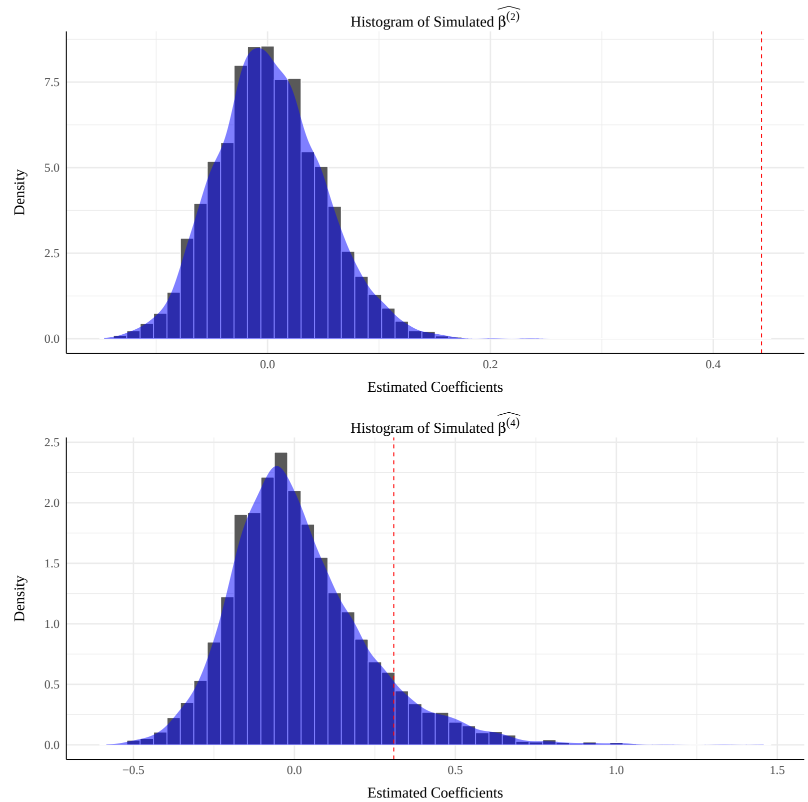
A.4 Falsification Test and Confounding Variables

We (1) conduct a falsification test by randomizing ownership connections within procurement type (goods, services, or construction), (2) show empirical results using the number of ownership connections rather than the indicators for the presence of at least one pair of connected firms, and finally, (3) consider additional variables addressing whether common owners or other relations appear to generate our observed results.

Figure A.5 shows that the estimated effect reported in the paper, denoted by the dotted vertical line, is statistically significant at all standard levels compared to the histogram of estimates from the falsification exercise for common owners, and it has a boot-strapped two-sided p-value of 0.12 for common owners' owners. Taken together, this suggests that our main estimated effects do not appear simply due to spurious correlation.

Figure A.5: Falsification

The figure below shows the histogram based on 5,000 simulated effects for Common Owners and Common Owner's Owners corresponding to $\widehat{\beta^{(2)}}$ and $\widehat{\beta^{(4)}}$, respectively. Each simulation draws connections for every auction from the same procurement type with replacement.



Next, Table A.4 reports the empirical results using the number of ownership connections rather than the indicators for the presence of at least one pair of connected firms and shows the results are similar to those in Table 2. We continue to find statistically significant relationships between the number of pairs of firms in an auction with common owners and the number of ineffective bids, and also between the number of pairs of firms in an auction with common owners' owners and the number of ineffective bids. This suggests that our main estimated effects are robust to the alternative measure of ownership connections.

Table A.4: Number of Ownership Connections and Identical Bidding

The table below shows the relation between the number of shareholder connections between participating firms and the number of ineffective bids in an auction. Num. Common Owners and Num. Common Owners' Owners correspond to the second- and fourth-degree current shareholder connections, respectively, and are all defined as the number of such pairwise owner connections in the auction. Column (1) uses all eligible auctions, columns (2) and (3) exclude auctions where the median bid is more than S$50,000 and S$100,000, respectively, column (4) excludes construction-type auctions, column (5) excludes procurers related to infrastructure and development including the Economic Development Board, Housing and Development Board (which constructs all public housing in Singapore), Building and Construction Authority, Land Transport Authority, Maritime and Port Authority of Singapore, Public Transport Council, and Sentosa Development Corporation, and column (6) excludes both construction-type auctions as well as agencies related to infrastructure and development. All regressions include procurer-by-auction type fixed effects as well as cubic controls of the log number of bidders and the average roundedness of bids in an auction, whose coefficients are suppressed for space. Standard errors are clustered by procurer and shown in parentheses. * denotes $p < 0.10$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Dependent Variable:	Num. Ineffective Bids					
	(1)	(2)	(3)	(4)	(5)	(6)
Num. Common Owners	0.407***	0.463***	0.431***	0.383***	0.417***	0.393***
	(0.046)	(0.066)	(0.046)	(0.042)	(0.042)	(0.044)
Num. Common Owners' Owners	0.325***	0.273***	0.307***	0.327***	0.307***	0.306***
	(0.081)	(0.062)	(0.073)	(0.080)	(0.078)	(0.077)
Sample	Full	Median Bid $< 50{,}000$	Median Bid $< 100{,}000$	Excl. Const.	Excl. Dev. Agencies	Excl. Const. & Dev. Agencies
Observations	9,087	6,859	8,187	8,358	8,185	7,581
R^2	0.315	0.353	0.337	0.333	0.317	0.331

Last, Table A.5 reports the estimated relationship between the number of ineffective bids and the presence of common owners as well as other common stakeholder relationships. Other common stakeholder relationships correlate negatively with the number of ineffective bids, while other common stakeholders' stakeholders do not show any statistically significant relationship with identical bidding. However, the main estimated effects of common owners and common owners' owners appear similar to those in Table 2. This suggests that our main estimated effects are not driven by confounding common stakeholder relationships.

Table A.5: Non-Shareholder Connections and Identical Bidding

The table below shows the relation between the number of stakeholder connections of different types between participating firms and the number of ineffective bids in an auction. Common owners and common owners' owners correspond to the second- and fourth--degree current shareholder connections, respectively, and are all defined as an indicator taking the value of one if there is at least one such owner connection in the auction. Other common stakeholders include common corporate directors, auditors, and secretaries and are also defined as an indicator taking the value of one if there is at least one such relationship. All regressions include procurer-by-auction type fixed effects as well as cubic controls of the log number of bidders and the average roundedness of bids in an auction, whose coefficients are suppressed for space. Standard errors are clustered by procurer and are shown in parentheses. * denotes $p < 0.10$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

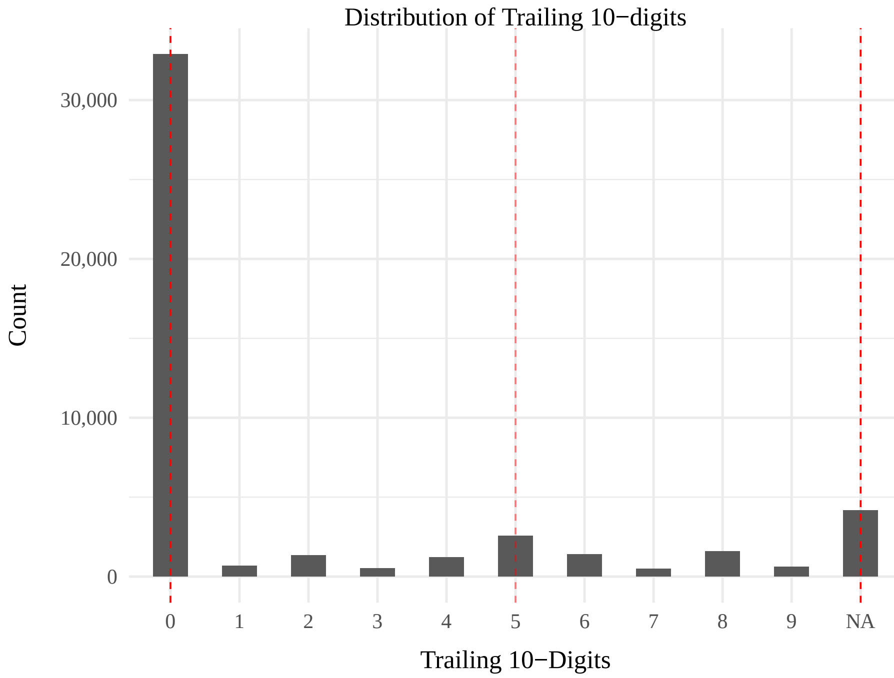
Dependent Variable:	Num. Ineffective Bids		
	(1)	(2)	(3)
Common Owners	0.444***		0.444***
	(0.069)		(0.069)
Other Common Stakeholders	-0.208***		-0.211***
	(0.053)		(0.053)
Common Owners' Owners		0.301***	0.317***
		(0.085)	(0.081)
Other Common Stakeholders' Stakeholders		-0.030	-0.020
		(0.175)	(0.127)
Observations	9,087	9,087	9,087
R^2	0.309	0.302	0.309

A.5 Rounding

We present summary statistics on the rounding of bids to different digits in Figure A.6. We find clustering at 0's and 5's. This summary provides the frequency of rounding. We find that while 69% of bids above S$10 appear rounded to the nearest S$10, almost 60% of the bids above S$100 are not rounded to the nearest S$100, suggesting that a large number of bids actually appear to have fairly detailed bids.

Figure A.6: Round Numbers in Bids

The figure below shows the histogram of trailing numbers in bids, shown for all bids. A 10-digits unit value of "NA" means the bid did not end in an integer. In addition, the table shows the number of bids greater than a particular digit and the probability that a bid appears rounded to that digit. For example, for digit = 10,000, the table shows the number of bids above S$10,000 and counts the probability that a bid greater than S$10,000 ends with "0,000."



X (digits)	Pr (Bid Rounded to Nearest X | Bid > X)	Num. Bid > X
10	0.696	47,222
100	0.417	47,098
1,000	0.161	46,347
10,000	0.043	35,791
100,000	0.039	6,265
1,000,000	0.020	2,044

We also examine the possibility that the observed link between ownership connections and identical bidding can be explained by bid rounding. If connected firms have distinct but similar costs and round bids to the same extent, identical bidding can occur mechanically. To investigate this possibility, Table A.6 presents the estimates for equation (1) while redefining ineffective bids as those explicitly not being rounded. If bid rounding drove our main results, selecting outcome variables based on ineffective bids without rounding would generate null results. However, we continue to observe results qualitatively similar to those in Table 2.

Table A.6: Presence of Ineffective Bids, Rounding, and Ownership Networks

The table below shows the relation between the presence of shareholder connections between participating firms and the presence of ineffective bids in an auction. Common owners and common owners' owners correspond to the second- and fourth-degree current shareholder connections, respectively, and are all defined as an indicator taking the value of one if there is at least one such owner connection in the auction. In column (1), the dependent variable is the presence of ineffective bids that are not rounded to the nearest S$100 and the sample comprises auctions whose median bids are greater than S$100. In column (2), the dependent variable is the presence of ineffective bids that are not rounded to the nearest S$1,000 and the sample comprises auctions whose median bids are greater than S$1,000. In column (3), the dependent variable is the presence of ineffective bids that are not rounded to the nearest S$10,000 and the sample comprises auctions whose median bids are greater than S$10,000. All regressions include procurer-by-auction type fixed effects as well as cubic controls of the log number of bidders and the average roundedness of bids in an auction, whose coefficients are suppressed for space. Standard errors are clustered by procurer and shown in parentheses.

Dependent Variable:	$1\{\text{Ineffective Bids} > 0\}$		
Ineffective Type =	Not Rounded to S$100	Not Rounded to S$1,000	Not Rounded to S$10,000
	(1)	(2)	(3)
Common Owners	0.076***	0.076***	0.064***
	(0.015)	(0.015)	(0.024)
Common Owners' Owners	0.199***	0.199***	0.396***
	(0.059)	(0.060)	(0.092)
Sample	Med. Bid > 100	Med. Bid > 1,000	Med. Bid > 10,000
Observations	9,003	8,898	6,961
R^2	0.305	0.306	0.289

B Details of Structural Framework

B.1 Structural Analysis

First, we non-parametrically identify model primitives using subsample auctions in which two firms submit identical bids and the other two submit competitive bids.[31] Details of selecting subsample auctions are described in Appendix B.2.

We use Krasnokutskaya (2011), where the cost to the type-i bid is characterized as the product of a common component Y that is known to all firms and an individual component c^i that is privately observable, i.e., $Y \times c^i$. Under this condition, the equilibrium-bid function takes a special functional form. In particular, for type i, the equilibrium-bid function is

$$B^i \quad = \quad Y \times \sigma_i(c^i), \tag{6}$$

where $\sigma_i(.)$ denotes the equilibrium-bid function for type i where Y is 1. Let b^i be the corresponding equilibrium bid where Y is 1, that is, $b^i = \sigma_i(c^i)$. Meanwhile, Paarsch and Hong (2006) suggest:

$$b^i \quad = \quad c^i + \frac{(1 - G_0^i(b^i))(1 - G_0^{-i}(b^i))}{N_{-i}^E g_0^{-i}(b^i)(1 - G_0^i(b^i)) + (N_i^E - 1) g_0^i(b^i)(1 - G_0^{-i}(b^i))}, \tag{7}$$

where N_i^E (N_{-i}^E) is the number of effective bids for type i (the rival type of i), G_0^i (G_0^{-i}) is the CDF for the equilibrium-bid distribution of type i (the rival type of i) where Y is 1, and g_0^i (g_0^{-i}) is the corresponding PDF for G_0^i (G_0^{-i}).

Unfortunately, the presence of unobserved heterogeneity Y means G_0^i is unobservable and only the joint distribution of B^1 and B^2 is observable. So, we start with removing Y from B^1 and B^2 to identify g_0^1 and g_0^2. Denoting the joint-characteristic function of $\ln(B^1)$ and

[31] Fixing the number of bidders for each type is needed for structural estimation. Observed bid distributions vary by the size of a coalition and the number of competitive bids.

$\ln(B^2)$ by $C(.,.)$, $C(.,.)$ is

$$C(\tau_1, \tau_2) = E\left\{\exp\left[\boldsymbol{i}\tau_1 \ln(B^1) + \boldsymbol{i}\tau_2 \ln(B^2)\right]\right\}, \tag{8}$$

where $\boldsymbol{i}$ denotes the imaginary number $\sqrt{-1}$. Then, the deconvolution method allows us to derive the characteristic function of $\ln(Y)$, $\ln(b^1)$, and $\ln(b^2)$ by

$$\begin{aligned} C_{\ln(Y)}(\tau) &= \exp\left[\int_0^\tau C_1(0, u_2)/C(0, u_2) du_2 - \boldsymbol{i}\tau E\{\ln(b^1)\}\right], \\ C_{\ln(b^1)}(\tau) &= C(\tau, 0)/C_{\ln(Y)}(\tau), \\ C_{\ln(b^2)}(\tau) &= C(0, \tau)/C_{\ln(Y)}(\tau), \end{aligned} \tag{9}$$

where $C_1(\cdot,\cdot)$ is the partial derivative of $C(\cdot,\cdot)$ with respect to the first argument. Without loss of generality, we normalize $E\{\ln(b^1)\}$ to zero. Then, the PDF of $\ln(Y)$, $\ln(b^1)$, and $\ln(b^2)$ is recovered by

$$\begin{aligned} f_{\ln(Y)}(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-\boldsymbol{i}\tau x) C_{\ln(Y)}(\tau) d\tau, \\ f_{\ln(b^i)}(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-\boldsymbol{i}\tau x) C_{\ln(b^i)}(\tau) d\tau. \end{aligned} \tag{10}$$

Through a change of variable formula, we obtain:

$$\begin{aligned} f_Y(y) &= \frac{1}{y} f_{\ln(Y)}(\ln(y)), \\ g_0^i(b^i) &= \frac{1}{b^i} f_{\ln(b^i)}(\ln(b^i)). \end{aligned} \tag{11}$$

Hence, f_Y, g_0^1, and g_0^2 are identified from the joint distribution of B^1 and B^2. F_Y, G_0^1, and G_0^2 are constructed from f_Y, g_0^1, and g_0^2. Given g_0^1, g_0^2, G_0^1, G_0^2, N_1^E, and N_2^E, equation (7) allows us to identify the PDFs and CDFs of the distributions for c^1 and c^2: f_0^1, f_0^2, F_0^1, and F_0^2. We follow the method of non-parametric estimation developed by Hickman and Hubbard (2015) that employs boundary-correction techniques. Estimated distributions of g_0^i, f_0^i, G_0^i, and F_0^i are denoted by $\hat{g}_0^i$, $\hat{f}_0^i$, $\hat{G}_0^i$, and $\hat{F}_0^i$. Appendix B.3 provides additional details on the estimation procedure, and Figures B.2 and B.3 present the results of estimations.

Second, following Hubbard and Paarsch (2014), we compute the inverse equilibrium-bid function as the solution of the system of ordinary differential equations for auction j:

$$\begin{bmatrix} \frac{d\sigma_1^{-1}(b)}{db} \\ \frac{d\sigma_2^{-1}(b)}{db} \end{bmatrix} = \begin{bmatrix} \frac{1-\hat{F}_0^1(\sigma_1^{-1}(b))}{\left(N_{j,1}^E+N_{j,2}^E-1\right)\hat{f}_0^1(\sigma_1^{-1}(b))} \frac{N_{j,2}^E(\sigma_2^{-1}(b)-\sigma_1^{-1}(b))+b-\sigma_2^{-1}(b)}{(b-\sigma_2^{-1}(b))(b-\sigma_1^{-1}(b))} \\ \frac{1-\hat{F}_0^2(\sigma_2^{-1}(b))}{\left(N_{j,2}^E+N_{j,1}^E-1\right)\hat{f}_0^2(\sigma_2^{-1}(b))} \frac{N_{j,1}^E(\sigma_1^{-1}(b)-\sigma_2^{-1}(b))+b-\sigma_1^{-1}(b)}{(b-\sigma_1^{-1}(b))(b-\sigma_2^{-1}(b))} \end{bmatrix}, \tag{12}$$

with the initial and boundary value conditions: $\sigma_1^{-1}(\bar{c}) = \sigma_2^{-1}(\bar{c}) = \bar{c}$ and $\sigma_1^{-1}(\underline{b}) = \sigma_2^{-1}(\underline{b}) = \underline{c}$, where $\underline{c}$ ($\bar{c}$) denotes the estimate for the common lower (upper) bound of the pseudo-cost distribution,[32] $\underline{b}$ is the common lower bound for the equilibrium bid, and $N_{j,i}^E$ denotes the number of effective bids for type i. We solve the system using the numerical methods developed by Bajari (2001), first using the shooting algorithm (the "first" method of Bajari (2001)) to acquire the initial proposal for the subsequent routine and then using the projection algorithm based on polynomial approximation (the "third" method of Bajari (2001)) with the initial proposal we acquired in the previous step. We solve the system for every pair of $N_{j,1}^E$ and $N_{j,2}^E$, where $3 \leq N_{j,1}^E \leq 5$ and $0 \leq N_{j,2}^E \leq 2$. Appendix B.4 provides more detail on the estimation procedure.[33]

B.2 Subsample Selection

To utilize Krasnokutskaya (2011), we use subsamples in our sample auctions for our structural estimation. Among our sample auctions in which at most two firms submit identical bids, we focus on auctions in which submitted values are moderate by excluding auctions where the winning bid is below S\$10,000 or above S\$30,000. In addition, we restrict our focus on "Open Quotation" tenders to homogenize participation constraints across auctions in our subsamples. Then, we restrict to $N_1^E = 2$ and $N_2^E = 1$, leaving us 65 subsample auctions.

[32] As in Bajari (2001) and Hubbard and Paarsch (2014), we assume that the support of the pseudo-cost distribution is identical across types.

[33] As an example, Figure B.4 shows the inverse equilibrium-bid functions when $N_{j,1}^E = 4$ and $N_{j,2}^E = 1$.

The median winning bid for our subsample auctions is S$16,000, which almost matches the median winning bid for our sample auctions. After this procedure, we exclude outliers from the sample as in Asker (2010). Figure B.1 shows outliers in the sample. After excluding outliers, our subsamples comprise 62 auctions.

B.3 Estimation Procedure

Our estimation procedure consists of the following steps:

1. $B^i_{m_i,l}$ denotes the m_i-th bid in auction l for type i that we observe from our bid data. The log transformation of bid data is performed to obtain $LB^i_{m_i,l} = \ln(B^i_{m_i,l})$, where $m_1 \in \{1, ..., N^E_1\}$ and $m_2 \in \{1, ..., N^E_2\}$, for each auction l we use to recover the pseudo cost distributions. We rank $B^i_{m_i,l}$ such that $B^i_{x,l} < B^i_{x',l}$, where $x < x'$.

2. The joint-characteristic function of an arbitrary pair $(LB^1_{m_1,l}, LB^2_{m_2,l})$ is estimated by

$$\hat{C}(\tau_1, \tau_2) = \frac{1}{N^E_1 N^E_2} \sum_{1 \le m_1 \le N^E_1, 1 \le m_2 \le N^E_2} \frac{1}{L} \sum_{l=1}^{L} \exp(\boldsymbol{i}\tau_1 LB^1_{m_1,l} + \boldsymbol{i}\tau_2 LB^2_{m_2,l}),$$

where L is the number of auctions we use for estimation. Then, we acquire the estimates of characteristic functions as

$$\begin{aligned} \hat{C}_{\ln(Y)}(\tau) &= \exp\left[\int_0^\tau \hat{C}_1(0, u_2)/\hat{C}(0, u_2) du_2\right], \\ \hat{C}_{\ln(b^1)}(\tau) &= \hat{C}(\tau, 0)/\hat{C}_{\ln(Y)}(\tau), \\ \hat{C}_{\ln(b^2)}(\tau) &= \hat{C}(0, \tau)/\hat{C}_{\ln(Y)}(\tau). \end{aligned}$$

3. The inversion formula is used to estimate densities $\hat{f}_{\ln(Y)}$, $\hat{f}_{\ln(b^i)}$, $i = 1, 2$, as

$$\begin{aligned} \hat{f}_{\ln(Y)}(u_y) &= \frac{1}{2\pi} \int_{-T_{\ln(Y)}}^{T_{\ln(Y)}} d_{T_{\ln(Y)}}(\tau) \exp(-\boldsymbol{i}\tau u_y) \hat{C}_{\ln(Y)}(\tau) d\tau, \\ \hat{f}_{\ln(b^1)}(u_1) &= \frac{1}{2\pi} \int_{-T_{\ln(b^1)}}^{T_{\ln(b^1)}} d_{T_{\ln(b^1)}}(\tau) \exp(-\boldsymbol{i}\tau u_1) \hat{C}_{\ln(b^1)}(\tau) d\tau, \\ \hat{f}_{\ln(b^2)}(u_2) &= \frac{1}{2\pi} \int_{-T_{\ln(b^2)}}^{T_{\ln(b^2)}} d_{T_{\ln(b^2)}}(\tau) \exp(-\boldsymbol{i}\tau u_2) \hat{C}_{\ln(b^2)}(\tau) d\tau, \end{aligned}$$

where $u_y \in [\underline{LY}, \bar{LY}]$, $u_i \in [\underline{Lb}, \bar{Lb}]$, $i = 1, 2$, and $T_{\ln(Y)}$, $T_{\ln(b^1)}$, $T_{\ln(b^2)}$ are smoothing parameters. Following Krasnokutskaya (2011), we introduce a damping factor $d_T(\tau)$ defined as $d_T(\tau) = \max\{1 - |\tau|/T, 0\}$ in the inversion formula. Through a change in the variable formula, we obtain

$$\begin{aligned} \hat{f}_Y(y) &= \frac{1}{y}\hat{f}_{\ln(Y)}(\ln(y)), \\ \hat{g}_0^i(b^i) &= \frac{1}{b^i}\hat{f}_{\ln(b^i)}(\ln(b^i)). \end{aligned}$$

Then, $\hat{G}_0^i$ is constructed from $\hat{g}_0^i$. We estimate the inverse equilibrium-bid function as

$$\hat{\xi}_i(b^i) = b^i - \frac{(1 - \hat{G}_0^i(b^i))(1 - \hat{G}_0^{-i}(b^i))}{N_{-i}^E \hat{g}_0^{-i}(b^i)(1 - \hat{G}_0^i(b^i)) + (N_i^E - 1)\hat{g}_0^i(b^i)(1 - \hat{G}_0^{-i}(b^i))}.$$

To implement the above estimation, we need to determine the smoothing parameters and the common support of bid distributions. As in Krasnokutskaya (2011) and Asker (2010), we choose the smoothing parameters and the common support of bid distributions based on the moment-matching method. Because the mean of $\ln(b^1)$ is zero by normalization, the estimates for the means and variances of $\ln(Y)$, $\ln(b^1)$, and $\ln(b^2)$ are

$$
\begin{aligned}
\hat{\mu}(\ln(Y)) &= \frac{1}{N_1^E L} \sum_{m_1=1}^{N_1^E} \sum_{l=1}^{L} LB^1_{m_1,l}, \\
\hat{v}(\ln(Y)) &= \frac{1}{2(N_1^E L - 1)} \sum_{m_1=1}^{N_1^E} \sum_{l=1}^{L} \left[LB^1_{m_1,l} - \frac{\sum_{m_1=1}^{N_1^E} \sum_{l=1}^{L} LB^1_{m_1,l}}{N_1^E L} \right]^2 + \\
& \frac{1}{2(N_2^E L - 1)} \sum_{m_2=1}^{N_2^E} \sum_{l=1}^{L} \left[LB^2_{m_2,l} - \frac{\sum_{m_2=1}^{N_2^E} \sum_{l=1}^{L} LB^2_{m_2,l}}{N_2^E L} \right]^2 - \\
& \frac{1}{2(N_1^E N_2^E L - 1)} \times \\
& \sum_{1 \le m_1 \le N_1^E, 1 \le m_2 \le N_2^E} \sum_{l=1}^{L} \left[\begin{array}{c} \left(LB^1_{m_1,l} - LB^2_{m_2,l} \right) - \\ \frac{\sum_{1 \le m_1 \le N_1^E, 1 \le m_2 \le N_2^E} \sum_{l=1}^{L} \left(LB^1_{m_1,l} - LB^2_{m_2,l} \right)}{N_1^E N_2^E L} \end{array} \right]^2, \\
\hat{v}(\ln(b^1)) &= \frac{1}{N_1^E L - 1} \sum_{m_1=1}^{N_1^E} \sum_{l=1}^{L} \left[LB^1_{m_1,l} - \frac{1}{N_1^E L} \sum_{m_1=1}^{N_1^E} \sum_{l=1}^{L} LB^1_{m_1,l} \right]^2 - \hat{v}(\ln(Y)), \\
\hat{\mu}(\ln(b^2)) &= \frac{1}{N_2^E L} \sum_{m_2=1}^{N_2^E} \sum_{l=1}^{L} LB^2_{m_2,l} - \hat{\mu}(\ln(Y)), \\
\hat{v}(\ln(b^2)) &= \frac{1}{N_2^E L - 1} \sum_{m_2=1}^{N_2^E} \sum_{l=1}^{L} \left[LB^2_{m_2,l} - \frac{1}{N_2^E L} \sum_{m_2=1}^{N_2^E} \sum_{l=1}^{L} LB^2_{m_2,l} \right]^2 - \hat{v}(\ln(Y)).
\end{aligned}
$$

Our choices for the smoothing parameters and the common support of bid distributions are set to replicate these moments.

We also require the estimates of the inverse equilibrium-bid functions to be increasing in bids for both types. Furthermore, we consider the estimated inverse equilibrium-bid functions to be inadmissible if estimated pseudo costs are negative. Given these considerations, our choices for $\left\{\underline{Lb}, \bar{Lb}, T_{\ln(b^1)}, T_{\ln(b^2)}\right\}$, $\left\{\hat{\underline{Lb}}, \hat{\bar{Lb}}, \hat{T}_{\ln(b^1)}, \hat{T}_{\ln(b^2)}\right\}$, satisfy

$$\begin{aligned}\left\{\hat{\underline{Lb}}, \hat{\bar{Lb}}, \hat{T}_{\ln(b^1)}, \hat{T}_{\ln(b^2)}\right\} \quad &\in \quad \arg \min_{\left\{\underline{Lb}, \bar{Lb}, T_{\ln(b^1)}, T_{\ln(b^2)}\right\}} M_1^T M_1 \\ &\text{s.t.} \quad \sum_{i=1}^{2}\sum_{n=0}^{K}\left[\mathbf{1}\left\{\hat{g}_0^i(t_n)<0\right\}+\mathbf{1}\left\{\hat{\xi}_i(t_n)<0\right\}+\mathbf{1}\left\{d\hat{\xi}_i(t_n)<0\right\}\right] \leq 0,\end{aligned}$$

where we define the grid $\exp(\underline{Lb}) = t_0 < t_1 < \ldots < t_{K-1} < t_K = \exp(\bar{Lb})$ ($K = 100$ for our reported results) and the moment gap M_1 as

$$M_1 \quad = \quad \begin{bmatrix} \int_{\underline{Lb}}^{\bar{Lb}} u_1 \hat{f}_{\ln(b^1)}(u_1)du_1 - 0 \\ \int_{\underline{Lb}}^{\bar{Lb}} \left[u_1 - \int_{\underline{Lb}}^{\bar{Lb}} u_1 \hat{f}_{\ln(b^1)}(u_1)du_1\right]^2 \hat{f}_{\ln(b^1)}(u_1)du_1 - \hat{\nu}(\ln(b^1)) \\ \int_{\underline{Lb}}^{\bar{Lb}} u_2 \hat{f}_{\ln(b^2)}(u_2)du_2 - \hat{\mu}(\ln(b^2)) \\ \int_{\underline{Lb}}^{\bar{Lb}} \left[u_2 - \int_{\underline{Lb}}^{\bar{Lb}} u_2 \hat{f}_{\ln(b^2)}(u_2)du_2\right]^2 \hat{f}_{\ln(b^2)}(u_2)du_2 - \hat{\nu}(\ln(b^2)) \end{bmatrix}.$$

The objective function captures the gap between predicted and observed moments. The first term in the constraint represents the penalty against negative values of estimated densities for b^1 and b^2. The second term is the penalty against negative values of estimated pseudo costs, whereas the third term represents the penalty against decreasing inverse equilibrium-bid functions. Given the estimate $\left[\hat{\underline{Lb}}, \hat{\bar{Lb}}\right]$, the consistent estimator for the support of $[\underline{LY}, \bar{LY}]$ becomes $\left[\min_{1\leq l\leq L}\{LB_{1,l}^1\} - \hat{\underline{Lb}}, \max_{1\leq l\leq L}\{LB_{2,l}^1\} - \hat{\bar{Lb}}\right]$. Let this interval be $\left[\hat{\underline{LY}}, \hat{\bar{LY}}\right]$.

Then, our choice for $T_{\ln(Y)}$, $\hat{T}_{\ln(Y)}$, satisfies

$$\begin{aligned}\hat{T}_{\ln(Y)} \quad &\in \quad \arg\min_{T_{\ln(Y)}} M_2^T M_2 \ , \\ &\text{s.t.} \quad \sum_{n=0}^{K}\mathbf{1}\left\{\hat{f}_Y(t_n)<0\right\}+\mathbf{1}\left\{T_{\ln(Y)} \geq 50\right\} \leq 0,\end{aligned}$$

where we define the grid $\exp(\hat{\underline{LY}}) = t_0 < t_1 < \ldots < t_{K-1} < t_K = \exp(\hat{\bar{LY}})$ ($K = 100$ for our reported results) and the moment gap M_2 as

$$M_2 = \begin{bmatrix} \int_{\underline{\hat{LY}}}^{\hat{\bar{LY}}} u_y \hat{f}_{\ln(Y)}(u_y)du_y - \hat{\mu}(\ln(Y)) \\ \int_{\underline{\hat{LY}}}^{\hat{\bar{LY}}} \left[u_y - \int_{\underline{\hat{LY}}}^{\hat{\bar{LY}}} u_y \hat{f}_{\ln(Y)}(u_y)du_y\right]^2 \hat{f}_{\ln(Y)}(u_y)du_y - \hat{v}(\ln(Y)) \end{bmatrix}.$$

As in the previous optimization problem, the objective function captures the gap between predicted and observed moments. The first term in the constraint represents the penalty against negative value of the estimated density for $\ln(Y)$. The second term in the constraint represents the penalty against large smoothing parameters that may result with the estimated PDF possessing a wavy tail.

For the first optimization problem, we use $[-0.4, 0.4]$ for the initial value of $\left[\underline{Lb}, \bar{Lb}\right]$ and arbitrary one of $\{3, 5, 7, 9, 11\}$ for the initial value of $T_{\ln(b^i)}$. Regarding the second optimization problem, we use arbitrary one of $\{5, 10, 15, 20\}$ for the initial value of $T_{\ln(Y)}$. For each optimization, we first try every possible set of initial values and obtain local optima. We then choose the one that minimizes the objective function without violating the constraint.

From this analysis, we find $\left[\underline{\hat{LY}}, \hat{\bar{LY}}\right] = [9.71, 10.61]$, $\left[\underline{\hat{Lb}}, \hat{\bar{Lb}}\right] = [-0.45, 0.40]$, $\hat{T}_{\ln(Y)} = 18.83$, $\hat{T}_{\ln(b^1)} = 11.00$, and $\hat{T}_{\ln(b^2)} = 11.00$. Figure B.2 shows the distributions of auction-specific and individual components of bid distributions based on this deconvolution process.

4. Using the estimated inverse equilibrium-bid function $\hat{\xi}_i(b^i)$, we estimate pseudo costs. Then, $\hat{f}^i_{0,raw}$ is the boundary-corrected KDE for the pseudo costs. $\hat{F}^i_{0,raw}$ is constructed from $\hat{f}^i_{0,raw}$. The support of $\hat{F}^i_{0,raw}$ is $\left[\hat{\xi}_i\left(\exp\left(\underline{\hat{Lb}}\right)\right), \hat{\xi}_i\left(\exp\left(\hat{\bar{Lb}}\right)\right)\right]$. However, considering the pseudo cost distributions have the common support, we replace the support of each distribution by

$$[\underline{c}, \bar{c}] = \left[\max_{1\le i\le 2}\left\{\hat{\xi}_i\left(\exp\left(\underline{\hat{Lb}}\right)\right)\right\}, \min_{1\le i\le 2}\left\{\hat{\xi}_i\left(\exp\left(\hat{\bar{Lb}}\right)\right)\right\}\right].$$

Then, we adjust densities by $\hat{f}^i_0(c^i) = \hat{f}^i_{0,raw}(c^i) / \left(\hat{F}^i_{0,raw}(\bar{c}) - \hat{F}^i_{0,raw}(\underline{c})\right)$. $\hat{F}^i_0$ is adjusted in accordance with $\hat{f}^i_0$. Figure B.3 shows the estimated pseudo cost distributions of type-1 and type-2 bids on $[\underline{c}, \bar{c}]$ before this adjustment.

B.4 Numerical Method for Asymmetric Auctions

We combine the "first" and "third" methods of Bajari (2001) to solve equation (12). First, we use the shooting algorithm to obtain the initial proposals for the subsequent estimations. Let $\{s_1(b; \underline{b}), s_2(b; \underline{b})\}$ be the solution of the system where $\sigma_1^{-1}(\underline{b}) = \sigma_2^{-1}(\underline{b}) = \underline{c}$. Then, the shooting algorithm consists of the following steps:

1. Fix $b_{low} = \underline{c}$ and $b_{high} = \bar{c}$.

2. Set $b_{guess} = \frac{1}{2}(b_{low} + b_{high})$.

3. Determine whether the system $\{s_1(b; b_{guess}), s_2(b; b_{guess})\}$ diverges, that is, whether it is in S^2, where $S = \{s : \ s \text{ is } C^1,\ s : [\underline{c}, \bar{c}] \to [\underline{c}, \bar{c}] \text{ and } s(b) < b \text{ for all } b < \bar{c}\}$.

4. If $\{s_1(b; b_{guess}), s_2(b; b_{guess})\}$ is in S^2, set $b_{high} = b_{guess}$.

5. If $\{s_1(b; b_{guess}), s_2(b; b_{guess})\}$ is not in S^2, set $b_{low} = b_{guess}$.

6. If $b_{high} - b_{low} < \epsilon$, stop. Otherwise, go to step 2.

7. After the stop, set $b_{min} = b_{high}$ and $b_0 = \frac{1}{2}(b_{low} + b_{high})$.

Although Bajari (2001) proves $\{s_1(b; b_{min}), s_2(b; b_{min})\}$ converges to the solution of the system as $\epsilon \to 0$, this shooting mechanism is inherently unstable. The instability cannot be eliminated by changing the numerical methodology of the solver (Fibich and Gavish, 2011) and becomes severe when the number of effective bids in the auction is large. Indeed, $\{s_1(b; b_{min}), s_2(b; b_{min})\}$ deviates from $\bar{c}$ as $b \to \bar{c}$ when the number of effective bids is large, as presented in the left panel of Figure B.4. We fix this problem by using the projection algorithm based on polynomial approximation. Specifically, we approximate the inverse equilibrium-bid function by the polynomial of degree 4. The approximated inverse equilibrium-bid function is

$$\hat{\sigma}_i^{-1}(b;\alpha,\underline{b}) = \sum_{k=1}^{4}\alpha_{i,k}(b-\underline{b})^k+\underline{c},$$

where $\alpha=\{\alpha_{i,k}\}_{1\le i\le 2, 1\le k\le 4}$. Then, the projection algorithm consists of the following steps.

1. Acquire the coefficients on $b-b_0$ for a polynomial of degree 4 that is a best fit for $s_i(b;b_{min})-\underline{c}$ for each i. In this approximation, we use the closed interval on which $\frac{ds_i(b;b_{min})}{db}\ge 0$ and $s_i(b;b_{min})\le b$ for all b on the interval. The interval starts at b_{start}, where $b_{start}=b_0$ or $s_i(b_{start}-\epsilon;b_{min})>b_{start}-\epsilon$ or $\frac{ds_i(b_{start}-\epsilon;b_{min})}{db}<0$ for any $0<\epsilon<\bar{\epsilon}$ ($\bar{\epsilon}$ is a certain threshold), and ends at b_{end}, where $b_{end}=\bar{c}$ or $s_i(b_{end}+\epsilon;b_{min})>b_{end}+\epsilon$ or $\frac{ds_i(b_{end}+\epsilon;b_{min})}{db}<0$ for any $0<\epsilon<\bar{\epsilon}$ ($\bar{\epsilon}$ is a certain threshold). If multiple such intervals exist, we use the one that starts from the smallest b_{start}. From this approximation, we obtain the initial proposal $\{\alpha_0,b_0\}$.

2. We define the grid $\underline{b}=t_0<t_1<...<t_{K-1}<t_K=\bar{c}$ ($K=50$ for our reported results). Using $\{\alpha_0,b_0\}$ as the initial proposal, we solve the estimates of $\{\alpha,\underline{b}\}$, $\{\hat{\alpha},\hat{\underline{b}}\}$, satisfying

$$\{\hat{\alpha},\hat{\underline{b}}\} \in \arg\min_{\{\alpha,\underline{b}\}}\sum_{n=0}^{K}D_n^TD_n+P\sum_{i=1}^{2}(\hat{\sigma}_i^{-1}(\bar{c};\alpha,\underline{b})-\bar{c})^2,$$

where we define the gap between the right hand and left hand sides for the empirical counterpart of equation (12) as

$$D_n = \begin{bmatrix}\frac{d\hat{\sigma}_1^{-1}(t_n;\alpha,\underline{b})}{db}\\ \frac{d\hat{\sigma}_2^{-1}(t_n;\alpha,\underline{b})}{db}\end{bmatrix}-\begin{bmatrix}\frac{1-\hat{F}_0^1(\hat{\sigma}_1^{-1}(t_n;\alpha,\underline{b}))}{\left(N_1^E+N_2^E-1\right)\hat{f}_0^1(\hat{\sigma}_1^{-1}(t_n;\alpha,\underline{b}))}\frac{N_2^E(\hat{\sigma}_2^{-1}(t_n;\alpha,\underline{b})-\hat{\sigma}_1^{-1}(t_n;\alpha,\underline{b}))+t_n-\hat{\sigma}_2^{-1}(t_n;\alpha,\underline{b})}{(t_n-\hat{\sigma}_2^{-1}(t_n;\alpha,\underline{b}))(b-\hat{\sigma}_1^{-1}(t_n;\alpha,\underline{b}))}\\ \frac{1-\hat{F}_0^2(\hat{\sigma}_2^{-1}(t_n;\alpha,\underline{b}))}{\left(N_2^E+N_1^E-1\right)\hat{f}_0^2(\hat{\sigma}_2^{-1}(t_n;\alpha,\underline{b}))}\frac{N_1^E(\hat{\sigma}_1^{-1}(t_n;\alpha,\underline{b})-\hat{\sigma}_2^{-1}(t_n;\alpha,\underline{b}))+t_n-\hat{\sigma}_1^{-1}(t_n;\alpha,\underline{b})}{(t_n-\hat{\sigma}_1^{-1}(t_n;\alpha,\underline{b}))(t_n-\hat{\sigma}_2^{-1}(t_n;\alpha,\underline{b}))}\end{bmatrix}.$$

We choose a sufficiently large P such that the estimated inverse equilibrium-bid function for each type converges enough to $\bar{c}$ as $b\to\bar{c}$ ($P=5,000$ for our reported results). $\{\hat{\sigma}_1^{-1}(b;\hat{\alpha},\hat{\underline{b}}),\hat{\sigma}_2^{-1}(b;\hat{\alpha},\hat{\underline{b}})\}$ is the final estimate for the inverse equilibrium-bid function. We estimate the inverse equilibrium-bid functions for every pair of N_1^E and N_2^E, where $3\le N_1^E\le 5$

and $0 \leq N_2^E \leq 2$.

The right panel of Figure B.4 shows that the final estimates of the inverse equilibrium-bid functions converge enough to $\bar{c}$, even when the number of effective bids in the auction is relatively large. This finding suggests the instability problem of the shooting algorithm is fixed by the projection algorithm based on polynomial approximation.

B.5 Confidence Intervals

We bootstrap confidence intervals by randomly drawing auctions from the original subsample with replacement with the size equal to the number of auctions in the original subsample. Then, we estimate $\hat{f}_Y, \hat{g}_0^i, \hat{f}_0^i, \hat{\sigma}_i^{-1}(b; \hat{\alpha}, \hat{\underline{b}}), \forall i$, using each resample. We repeat this process 500 times.

To reduce the computational burden, we use the same smoothing parameters and bid bounds $(\hat{\underline{Lb}}, \hat{\bar{Lb}}, \hat{T}_{\ln(b^1)}, \hat{T}_{\ln(b^2)}, \hat{\underline{LY}}, \hat{\bar{LY}}, \hat{T}_{\ln(Y)})$ for the estimates of $\hat{f}_Y, \hat{g}_0^i, \hat{f}_0^i, \forall i$ and also skip to the second step using the initial value used for the original subsample $\{\alpha_0, b_0\}$ for the estimates of $\hat{\sigma}_i^{-1}(b; \hat{\alpha}, \hat{\underline{b}}), \forall i$. If there remain resamples from which the minimized value of the objective function minus the penalty term on the deviation from the terminal condition exceeds 100, we use the initial value $\{\alpha_0, b_{j+1}\}$ for these resamples, where $b_{j+1} = b_j + 0.05$. We repeat this process three times.

Given $\hat{f}_Y, \hat{g}_0^i, \hat{f}_0^i, \hat{\sigma}_i^{-1}(b; \hat{\alpha}, \hat{\underline{b}}), \forall i$, we compute test statistics for each resample. Obtaining the distributions of test statistics, we are able to construct their confidence intervals.

Figure B.1: Outliers in the Subsample

The left (right) plot is the scatter plot of log type-1 bids that are the lowest (highest) among the two bids and log type-2 bids. The outliers are represented by red circles.

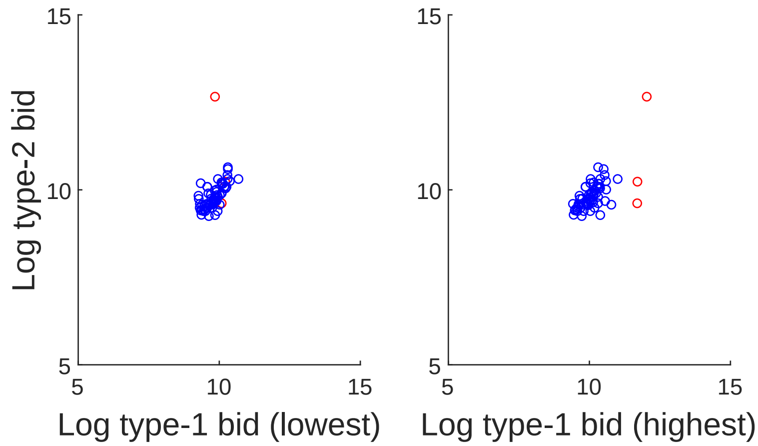


Figure B.2: Auction and Individual Components of Bid Distributions

The top panel plots $\hat{f}_Y$. The bottom left panel plots $\hat{g}_0^1$, whereas the bottom right panel plots $\hat{g}_0^2$. The solid lines depict the point estimates of the PDFs. The dotted lines show 5% and 95% point-wise quantiles of the estimated distributions.

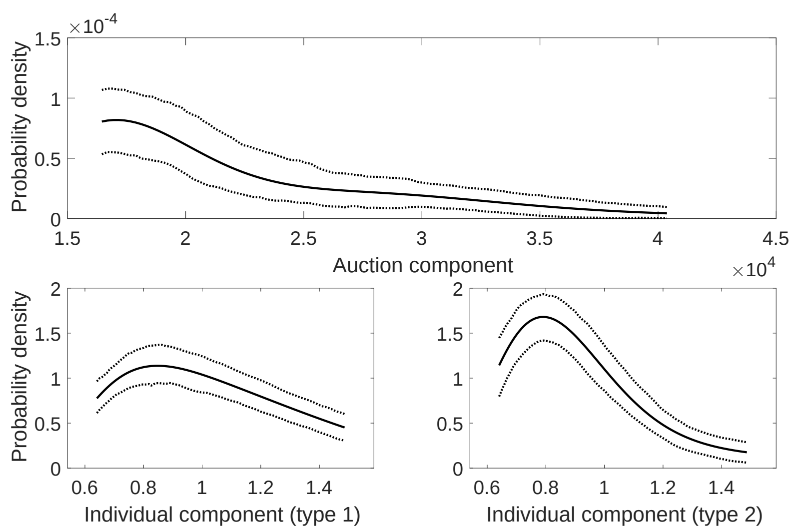


Figure B.3: Estimated Cost Distributions

The left panel plots $\hat{f}^1_{0,raw}$, whereas the right panel plots $\hat{f}^2_{0,raw}$. The solid lines depict the point estimates of the PDFs. The dotted lines show 5% and 95% point-wise quantiles of the estimated distributions.

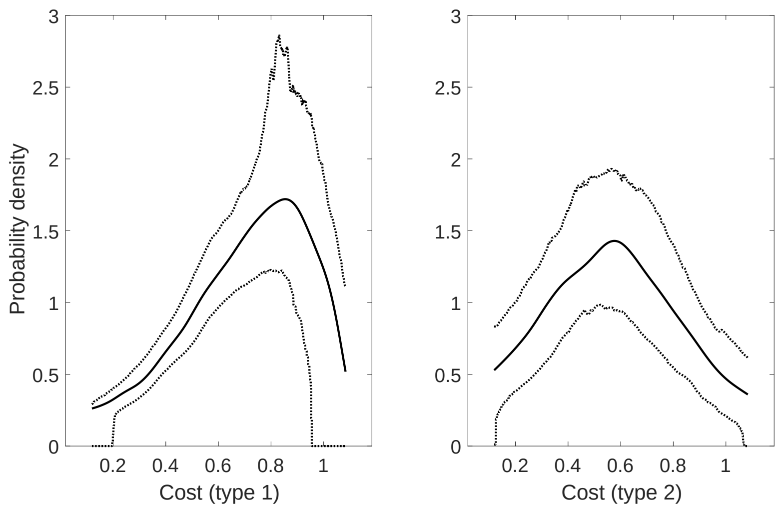


Figure B.4: Inverse Equilibrium-Bid Functions

The left panel plots the initial proposals for the inverse equilibrium-bid functions, whereas the right panel plots the final estimates for them ($N_1^E = 4$ and $N_2^E = 1$).