# Rigging the Scores: Corruption through Scoring Rule Manipulation in Public Procurement Auctions

Qianmiao (Michelle) Chen *

February 10, 2025†

Click here for the most recent version ‡

**Abstract**

Public procurement in developing countries remains vulnerable to corruption, even in open auctions. In this paper, I document how procurement officers and firms manipulate scoring rules to predetermine winners, with non-competitive bids meeting minimum requirements. Using rich public procurement and firm data from China, I introduce model-driven statistical tools that detect a 65% corruption rate, validated by a 91% expert recognition rate in an audit survey. Reduced-form regressions suggest China's anti-corruption campaign is inadequate to combat entrenched corruption. Model-based counterfactuals show that anonymous call-for-tender evaluations could improve social welfare by 10%, reducing corruption and encouraging broader participation.

***Keywords*** **: Public Procurement, Open Scoring Auction, Corruption**

***JEL classification***: **H57, D73, D44, L44**

*Research Economist, World Bank DIME. Email: qianmiao@worldbank.org

†I would like to thank Marco Gonzalez-Navarro, Steve Tadelis, Kei Kawai, and Fred Finan for their advice, guidance, and support. I also thank Betty Sadoulet, Cailin Slattery, Daniel Rogger, Edward Miguel, Ethan Ligon, Francesco Trebbi, Jed Silver, Joel Ferguson, Juan Ortner, Karl Schurter, Kirill Borusyak, Thiago Scot, and Guo Xu, as well as numerous seminar participants at NEUDC, DevPEC, IIOC, Conference on Auctions and Firms, and UC Berkeley. Zan Song, Yuning Xue, and Tiffany Liu provided excellent research assistance. All remaining errors are my own.

‡The previous version was circulated under the title "Corruption in Public Procurement Auctions: Evidence from Collusion between Officers and Firms".

# 1 Introduction

Corruption is a pervasive issue that undermines governments worldwide, distorting markets, stalling economic growth, and eroding public trust. Public procurement is especially vulnerable, with bribes in this sector accounting for a significant share of total corruption (OECD, 2016). Corruption in procurement involves illicit practices to manipulate contract awards for goods, services, and works, resulting in inflated costs and placing undue burdens on taxpayers and businesses. This challenge is particularly acute in developing countries, where weak institutional frameworks and inadequate accountability exacerbate the problem.

In response to these challenges, many countries have adopted open tendering systems, following the UNCITRAL Model Law on Public Procurement,[1] with open scoring auctions being one of the most prevalent methods. In these auctions, firms submit confidential bids assessed on quality and price, with contracts awarded based on the highest combined scores.[2] To promote fair competition, a minimum number of bidders, often three, is typically mandated, as in China.[3] However, despite the transparency and competitive intent of open scoring auctions, their effectiveness in curbing corruption is contested. Personal interviews with Chinese procurement officials indicate that corruption persists through the manipulation of scoring rules to favor predetermined suppliers, often accompanied by "zombie" bids that create a facade of competition. Existing auditing and complaint systems are frequently ineffective in detecting these manipulations.[4]

This paper investigates two key questions: **(1)** How prevalent is scoring rule manipulation corruption in open scoring auctions within the Chinese public procurement system? **(2)** What policies can effectively reduce corruption and its associated distortions? To explore these questions, I compiled a novel and comprehensive dataset on public procurement auctions from one province in China, covering all bidding firm profiles and their outcomes. Spanning the period from 2006 to 2021, this dataset includes over 300,000 procurement projects, and represents the first comprehensive analysis of detailed public procurement data within the Chinese context. It is further enriched with administrative firm data and information from corruption investigations.

To answer the first question **(1)**, I begin by documenting two motivating stylized facts in scoring auction public procurement data. First, in over two-thirds of procurement cases, there are only three bidders—the minimum required for an auction to be deemed

[1] United Nations Commission on International Trade Law (UNCITRAL, 2014).

[2] Most countries with modern public procurement systems adopt open scoring auctions for open tendering.

[3] India requires three bidders for limited tenders, while Serbia mandates five.

[4] Detection is challenging because each procurement step adheres to formal regulations, such as selecting open competition and meeting bid thresholds. The low complaint rate is also influenced by firms' reluctance to challenge corruption due to fears of retaliation or the need to maintain positive relations with local governments.

valid—highlighting a lack of robust competition. Second, there are unusually large score gaps between winning and losing bids, particularly in auctions with only three bidders. In these cases, winning bids often stand alone, with no close competitive losing bids to suggest genuine rivalry.

Building on these observations, I propose tools to identify corruption more preciesly through scoring rule manipulation in the standard scoring auction model. Building on the work of Che (1993), Asker and Cantillon (2008), and Hanazano et al. (2020), I construct a measure called "pseudotype," representing the highest scores bidders can achieve while maintaining nonnegative profits. I then demonstrate how scoring rule customization can create bidding patterns inconsistent with competitive behavior. Under the null hypothesis of a competitive procurement auction, the winning score should closely align with the expected pseudotype of the strongest rivals. If the winning score significantly exceeds the expected strongest rivals' pseudotypes, the winner has left a substantial amount of potential profit untouched. At the auction group level, I use a nonparametric maximum likelihood estimation to estimate the proportion of corruption by comparing the mean of winning scores with the mean pseudotype of the strongest rivals. At the individual auction level, I compare the pseudotypes of losing bidders with the expected rivals' pseudotypes given the winning scores. When scoring rule manipulation occurs, a marked discrepancy emerges, resulting in an unusually low average of losers' pseudotypes and rejection of the null hypothesis. Results from the two detection methods closely align, particularly in cases with three bidders, indicating that the null hypothesis of no corruption in scoring rule manipulation is rejected in over 65% of auctions, revealing evidence of scoring rule customization, favoritism, and coordinated non-competitive bids.

To further validate my test, I conducted a procurement expert audit study. Public procurement officials acknowledged the existence of significant opportunities for manipulation in the bidding process(Gong and Zhou, 2015). However, detecting irregularities can be challenging for the public without the necessary expertise to verify the criteria. To address this issue, I engaged five procurement experts from the Province's bid evaluation expert pool, selected for their active involvement in bid evaluation meetings and deep industry and firm knowledge. They evaluated a random sample of 500 procurement projects without knowledge of the bidders or outcomes, identifying signs of criteria customization and assessing the level of competition. The survey results were compared to the predictions made by my model, resulting in high degrees of congruence.[5] When my model predicted a potentially corrupt auction, there was a 91% chance that procurement experts highlighted signs of scoring rule customization, demonstrating the model's effectiveness in detecting suspicious patterns.

[5]Defined as the concordance between corruption test outcomes and the identification of suspicious call-for-tender files by experts, the accuracy rate measures the percentage of cases for which both experts and model predictions correctly classify them as corrupt or not corrupt. This accuracy rate stands at 81%.

To speak to the second question **(2)**, the corruption test results obtained from the previous step were linked to firm and anti-corruption campaign data, enabling a more comprehensive understanding of scoring rule manipulation corruption and its policy implications. First, I explore the disparities between winners in competitive and potentially corrupt scenarios, revealing that winners implicated in suspected cases exhibit characteristics such as lower productivity, stronger ties to state-owned entities, and closer proximity to the local procurement government. Subsequently, I utilize corruption investigation data to analyze the impact of anti-corruption campaign investigations on procurement outcomes. I find that investigations prompt heightened competition and reduce corruption in the short term, but fail to yield a sustained, long-term reduction in corrupt activities. Notably, this finding is predominantly driven by investigations directed at high-level officials.

Lastly, I enhance the scoring auction model by incorporating semi-parametric estimates of firm cost functions to conduct counterfactual policy analyses. I examine two different counterfactual scenarios to study the impact of various policy changes on procurement outcomes. First, I assess how pre-determined winners would bid if anonymously selected experts were tasked with reviewing call-for-tender files prior to auction commencements. If these experts identify any unnecessary rules, scoring rules will be revised accordingly. The elimination of such unnecessary rules leads to an 18% decrease in the winning price and a 3% increase in quality. Collectively, this translates to an 11% increase in social welfare measured by score change. 70% of these welfare gains arise from competitive bidding and the remaining 30% stem from increased entries. Second, considering that corrupt officials often attempt to assign lower weights to transparent factors like price and higher weights to quality in order to manipulate outcomes, I investigate the consequences of increasing the evaluation weights on the price component.

This paper contributes to several strands of literature. First, it contributes to the small but growing body of research on detecting corruption in public procurement by introducing novel methods for detecting corruption in scoring auctions, one of the most widely used procurement methods globally. By linking detailed bidding outcomes with rich firm data, it explores policy implications, offering a comprehensive approach to understanding procurement corruption. Methodologically, the paper employs advanced statistical techniques and integrates an innovative expert audit survey, providing robust validation and deeper insights. Existing studies in this field have measured corruption in procurement auctions through direct indicators such as price signaling (Cai, Henderson, and Zhang, 2013), bid leakage (Andreyanov and Korovkin, 2017), and exaggerated bids (Huang, 2019). Other research has taken an indirect approach by examining procurement thresholds that require public disclosure (Coviello and Mariniello, 2014), the adoption of open auctions (Palguta and Pertold,

2017; Coviello, Guglielmo, and Spagnolo, 2018; Baltrunaite et al., 2021; Szucs, 2024),[6] and audits that reveal bunching behavior aimed at avoiding competition (Gerardino, Litschig, and Pomeranz, 2024). Additionally, several studies focus on favoritism toward politically connected firms in procurement awards (Cao, 2022; Baltrunaite, 2020; Baltrunaite et al., 2021; Colonnelli and Prem, 2021; Brugués, Brugués, and Giambra, 2024). This study distinguishes itself by focusing on corruption mechanisms within scoring rules, moving beyond the conventional binary distinction between open and non-open auctions and uncovering subtler forms of manipulation that do not necessarily involve overt political connections or outright violations of public procurement laws.

Second, this study contributes to the literature on the effects of anti-corruption investigations and audits in reducing corruption by linking corruption detection with anti-corruption campaigns and highlighting the ineffectiveness of these campaigns. Existing research suggests that anti-corruption investigations and audits can decrease corruption (Avis, Ferraz, and Finan, 2018), increase government revenue (Chen and Kung, 2019), reduce expenditure (Olken, 2007; Lichand, Lopes, and Medeiros, 2016; Zamboni and Litschig, 2018), improve firm performance (Colonnelli and Prem, 2021), and enhance electoral accountability (Ferraz and Finan, 2011). However, some studies find that audits are not always effective in reducing corruption and may even have adverse effects due to poor targeting (Duflo et al., 2018), inadequate design (Gerardino, Litschig, and Pomeranz, 2024), or distorted incentives (Khan, Khwaja, and Olken, 2016; Lichand and Fernandes, 2019). This paper demonstrates that, without changes to procurement policies, investigations fail to produce lasting effects. Additionally, I design a targeted audit study to show how improved audit designs can help reduce corruption and enhance procurement outcomes.

Third, this research contributes to the existing literature on scoring auctions by integrating corrupt practices into the auction model, leveraging rich and novel public procurement data, and proposing empirical methods to detect corruption using real-world data. Foundational work by Che (1993) analyzed various types of scoring auctions and identified optimal scoring rules, while subsequent studies like Asker and Cantillon (2008) and Chen-Ritzo et al. (2005) showed that scoring auctions can yield higher payoffs for buyers compared to minimum-quality or price-only auctions. More recent research by Hanazono et al. (2013), Takahashi (2018), Andreyanov (2018), and Hanazano et al. (2020) has further examined equilibrium and mechanism design in scoring auctions, enhancing the understanding of this auction format. Additionally, studies such as Bajari, Houghton, and Tadelis (2014), Ryan (2020), and Kong, Perrigne, and Vuong (2022) have extended the scoring auction model to include contract design in public procurement. Building on this body of work, my re-

[6] A lack of transparency in procurement processes often signals potential corruption, whereas open and transparent processes are presumed to mitigate corrupt practices. Studies have shown that corruption is frequently associated with non-open procurement processes.

search links theoretical insights with empirical evidence, uncovering patterns of corruption within scoring auctions and offering policy implications for reducing corruption in public procurement.

Lastly, this study contributes to the literature on collusion detection in auctions and markets, specifically the detection of non-competitive agreements among two or more bidders.[7] However, corruption originating from the auctioneer or agency side has received comparatively less attention, despite being a significant issue in many developing countries where public procurement laws are poorly enforced. My study takes a novel approach by allowing for both corruption and collusion and investigating their relationship within the context of public procurement in a developing country. As a result, this research not only introduces a data-driven method for detecting corruption but also examines the distortions and policy implications arising from such practices.

The remainder of the paper is organized as follows. In Section 2, I offer background information on the scoring auction procedure and China's public procurement system. Section 3 describes the dataset used for the analysis. Moving forward, Section 4 presents key stylized facts that motivate the investigation. In Section 5, a theoretical model is constructed, laying the foundation for the empirical test conducted in Section 6. To further validate the proposed approach, Section 7 presents the design and results of the expert survey. Section 8 delves into a discussion of the implications of corruption in public procurement by exploring the impact of anti-corruption investigation policies, and conducting counterfactual analyses through the structural model to evaluate different policy scenarios. Finally, Section 9 provides concluding remarks, summarizing the findings and their implications.

# 2 Background

## 2.1 Public Procurement in China

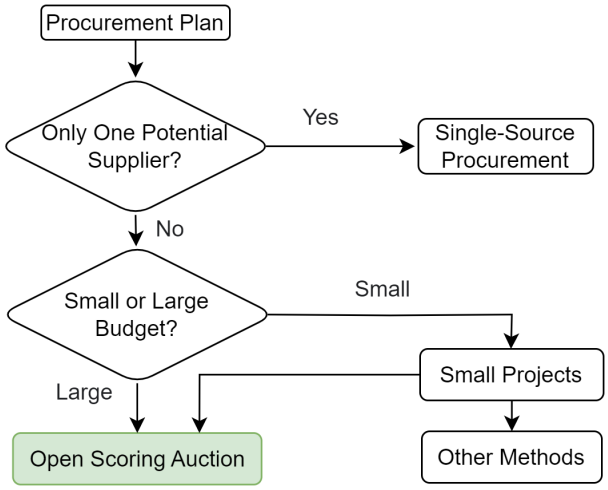
Public procurement projects in China are defined based on the use of public funds and their alignment with legal requirements. According to Article 2 of the *Government Procurement Law of the People's Republic of China*, government procurement involves the acquisition of goods, services, and projects by state agencies, public institutions, and organizations using fiscal funds. These procurements must adhere to centralized procurement catalogs or exceed specified financial thresholds. Only projects meeting these criteria are subject to government procurement regulations, which may include listing on the China

[7] See Porter (2005) for a literature review on collusion in auctions. More recent papers include Conley and Decarolis (2016); Schurter (2017); Chassang et al. (2022); Kawai and Nakabayashi (2022); Kawai et al. (2023); Kawai, Nakabayashi, and Shimamoto (2022).

Government Procurement Catalog.[8]

In contrast, non-government procurement projects refer to those that do not use fiscal funds, fall below the financial thresholds, or are outside the centralized procurement catalogs, thus exempt from government procurement rules. Consequently, most large-scale infrastructure projects, such as high-speed railway construction, are not included in public procurement. These projects are often led by state-owned enterprises like the National Railway and Highway Groups and are funded through a combination of local financing or non-public funds. [9] This distinction underscores the separation between public-funded procurement and alternative financing mechanisms for infrastructure development.

Figure 1: Public Procurement Procedure




**Notes**: The standards for procurement choices are guided by the Regulations for the Implementation of the Government Procurement Law of the People's Republic of China. There are non-open auction methods, such as negotiations, first-price bidding on online platforms, and price solicitations.

Public procurement in China typically progresses through three distinct phases: pre-procurement, procurement, and post-procurement. In the pre-procurement phase, the procuring entity, often a government department or state-owned enterprise, identifies its needs and drafts a detailed procurement plan. This stage involves assessing project feasibility and defining the required goods, services, or construction.

Once the requirements are established, the procuring entity finalizes a procurement plan that outlines key details such as project scope, selected procurement method, budget

[8]Procurement by the Department of Defense is not included in the public procurement system in China, as they have an independent procurement system and regulations.

[9]For instance, *the Notice on Strictly Prohibiting Illegal Financing by Local Governments in the Name of Government Service Procurement (Ministry of Finance [2017] No. 87)* explicitly excludes infrastructure projects in sectors such as railways, highways, airports, telecommunications, utilities (water, electricity, and gas), education, science, healthcare, culture, sports, land development, and agricultural water conservancy from being categorized as government service procurement.

allocation, timeline, and criteria for supplier selection. This plan is then submitted for review and approval by relevant oversight agencies, such as the Audit Bureau and the Bureau of Finance, ensuring compliance with regulatory standards and proper use of fiscal funds.

Specific criteria guide the choice of a procurement method, as shown in Figure 1. When dealing with uniquely qualified suppliers or unforeseen emergencies, single-source procurement, also known as direct contracting, becomes a viable option. This method is also used when it is crucial to maintain project consistency or align with existing services. In other cases, the choice of procurement method depends on the procurement budget. To ensure transparency, projects exceeding a certain threshold are required to use the standard open-scoring auction method.[10] Projects below this threshold have several options, including invited-only auctions, online first-price auctions, and direct price inquiries. However, even for smaller-scale procurement, there is a preference for open-scoring auctions due to the increased scrutiny and audit associated with non-open methods. Various studies (Calvo, Cui, and Serpa, 2019; Decarolis et al., 2020) have found that public procurement using non-open methods often leads to higher costs and greater susceptibility to corruption. In this paper, I specifically focus on the open-scoring auction method highlighted in Figure 1.

Upon approval of the plan, the procuring entity must publicize procurement information on national and local procurement websites. This information includes project details, the chosen procurement method, budget allocation, timeline, and scoring criteria. According to regulations, an open auction requires at least three qualified bidders to be considered valid. If an open auction ends with only two qualified bidders, experts must decide whether to adjust the call-for-tender documents and restart the auction or switch to non-open methods. If only one bidder participates in the open auction, the procuring entity may opt for single-source procurement, provided that experts confirm the call-for-tender documents are fair and compliant with competition regulations.

Because of the complexity of public procurement, corruption within the realm of public procurement has remained a longstanding concern. In the context of open-scoring auctions, a complex network of actors is involved, as shown in Figure A1. This network involves various roles, such as city or county leaders, directors, and their subordinates within government departments responsible for procurement requests (e.g., education, transportation, health), the procurement agencies overseeing bid conferences, and the public procurement experts tasked with bid evaluation.

[10]The thresholds for open scoring auctions in public procurement vary across different provinces in China. In the province under study, a significant change occurred in 2020, setting the benchmark for public bidding on government procurement of goods or services at 4 million yuan across the province. Similarly, the threshold for public bidding in construction projects adheres to both national and provincial regulations. These thresholds have been revised several times since 2007.

## 2.2 Scoring Auction Procedure and Corruption

In addition to the general public procurement procedures, a detailed outline of the open-scoring auction process is presented in Figure 2. Once the procurement request is approved, a procurement agent is chosen. The agent is responsible for drafting the call-for-tender files, which outline the requirements specified by the procuring entity. These files are then published on official websites and newspapers. A noteworthy characteristic of open auctions, as compared to non-open methods, is that any interested companies enter the competition and have the possibility to win the contract. All bidders are given a one-month period to conduct research on the files and submit their proposals, which typically include quality specifications and pricing details.

Figure 2: Open Scoring Auction Procedure




**Notes**: The detailed procedure of an open scoring auction can be found on this website

On the day of the auction, a procurement committee composed of at least five randomly selected experts evaluates the submitted bids according to predetermined scoring rules. The random selection of experts helps minimize the risk of corruption, such as bribery, which could lead to unfairly awarding extra points to certain bidders. Each expert assigns a score to each bid based on the evaluation criteria. These scores are then used to rank the bids and determine the winning bidder. The contract is awarded to the bidder with the highest score.

Typically, the final total score is the weighted sum of the quality score and price score:

$$score = w_q \times Quality\ Score + w_p \times Price\ Score \in [0, 100]$$

where $w_q$ is the weight of the quality score, and $w_p$ is the weight of price score. The quality scores are multi-dimensional, encompassing various aspects such as product parameters, functions, post-purchase services, detailed implementation plans, and demonstration quality. These factors are inherently subjective, as they rely on evaluators' judgment, giving officials significant discretion in determining what they prioritize. This subjectivity creates potential opportunities for manipulation. In contrast, the price score is entirely objective, based solely on the bid price. Firms with higher bid prices receive lower price scores, which are calculated by dividing the minimum bid price among all firms by the individual firm's own bid price.

Given that quality criteria allow for more discretion compared to price, corrupt officials might aim to reduce the weight of the price component to exert greater control over the outcome. To prevent extreme cases where the price weight is entirely disregarded, the Public Procurement Law stipulates that price weights must fall within specific ranges: [30%, 60%] for goods procurement and [10%, 30%] for construction and service procurement. In practice, many open-scoring auctions tend to use the lowest possible price weight, thus providing more leeway for discretion in quality requirements.

One local procurement official interviewed for this research mentioned that many public projects he managed had preselected winners:

> *"'Smart' firms often begin to contact government officials before the bidding process starts. In those cases, only one bidder will meet all the requirements and will get the highest score in the bid evaluation. Other firms participate only to create the appearance of legitimate competition. There was an example when the city leaders and firm owners had a private dinner together. The leader revealed the information about a new construction project. The firm owner expressed the willingness to do the project and then a deal was reached. To pass the check from other authorities, the lower-level officials chose an open auction and disguised it with fake competition. Only the predetermined bidder can meet all the requirements and get the highest score in the bid evaluation. Other zombie bids only create the appearance of legitimate competition."*

Some may be surprised by the seemingly straightforward nature of these corrupt practices. However, this behavior aligns with the audit processes and incentive structures that government officials face. Audits of government procurement primarily focus on budget adherence, execution methods, procurement standards, contract fulfillment, project acceptance, and fund distribution. Given the large sums often involved in procurement projects, those using non-open methods face more frequent audits. Consequently, projects that use open auctions are less scrutinized, making them a more favorable cover for corrupt activities.

# 3 Data

In this section, I describe the datasets I use for the paper to investigate the prevalence of corruption. There are five datasets, public procurement data, corruption investigation data, firm registration data, firm tax data, and expert evaluation data. I leave the expert evaluation data for Section 7.

## 3.1 Public Procurement Data

I collect all the publicly available public procurement data of one province of China from 2006 to January 2021.[11] With increased transparency, all procurement procedures, except those involving national secrets, must be accessible for public tracking.[12] The final public procurement database comprises four primary datasets: procurement plans, procurement announcements, outcomes, and quality checks. The procurement plan dataset includes details such as procurement objects, purchaser names, reserve prices, and procurement methods. The procurement announcement dataset provides information about the bidding process, while the outcomes dataset lists bidder names, bid prices, and scores. The quality check dataset contains a binary variable indicating whether the final delivery passed the quality check.[13] Excluding procurements without bid information, Figure A2 illustrates the distribution of procurement projects with complete information by year.

For procurement projects procured through open scoring auctions within the dataset, as shown in Table 3, 56% are goods procurement, such as various types of vehicles; 35% are service procurement, including professional services and cleaning services; and 9% are construction projects, such as local road maintenance and building electrical systems. The average budget size is 4.46 million RMB, approximately 0.62 million USD.

## 3.2 Firm registration data

The firm registration dataset includes all officially registered firms in China, no matter whether they are still in operation or not. The data are extracted from the Firm Search Platform belonging to Alibaba Group, whose data are from the State Administration for Market Regulation, the Supreme People's Court, and Ali Map. The multi-dimensional data covers firms' names, open dates, industry, type, current status, capital size, and employment

[11] To the best of my knowledge, I am the first to compile such a comprehensive dataset at the provincial level.

[12] However, the procurement information format is often inconsistent and disorganized. This issue is prevalent in many developing countries, where non-transparent data fosters corruption and impedes the government's capacity to leverage big data for detecting collusion or corruption.

[13] The quality check data is less useful in this context, as only 20% of procurement projects can be linked to procurement and bidding information, and fewer than 0.06% of projects failed the quality check.

size. The dataset has been used in Bai et al. (2021) and Chen et al. (2022). I cleaned the names of companies in public procurement data and linked them to firm registration data to obtain the basic registration data of all bidders. The overall successful matching rate is over 99%. I further match the bidders with the government departments initiating procurement and use QGIS to calculate the geographical distances between them. I also obtain the headquarters information of all bidders.

In Figure A3, I plot the distribution of both the number of procurement projects in which firms participate and the number of procurement contracts that firms secure. The data reveal that approximately 80% of firms engage in fewer than five projects, and a mere 5% participate more than 20 times. In terms of victories, 60% of firms never secure a contract, and about 95% win fewer than five contracts. These distributions indicate that the majority of firms have limited participation and success in procurement projects."

### 3.3 Firm Tax data

The firm tax data used in this study were sourced from the Chinese State Administration of Tax (SAT) for the years 2007 to 2016. Serving as China's counterpart to the IRS, the SAT is responsible for tax collection and audit procedures. My dataset is derived from administrative records of enterprise income tax, offering insights into firms' activities over this period. This comprehensive panel data includes critical aspects such as overall production, sales, and input data. Moreover, the inclusion of detailed cost breakdowns provides a nuanced perspective, enabling the assessment of various subcategories within administrative expenses. Using this dataset, I adopt the method employed by Chen et al. (2021) to formulate residualized measures of firm productivity, commonly referred to as Total Factor Productivity (TFP). The mechanics of this TFP measurement process are detailed in Appendix B1. [14]

### 3.4 Corruption Investigations Data

The investigation data from 2011 to August 2016 are obtained from Wang and Dickson (2022), who collected the data from Tencent.Co—the largest Internet company in China. During the anti-corruption campaign, Tencent launched a searchable online database of all corruption investigations across China in 2011. Based on the information provided by Party disciplinary committees, courts, and procuratorates from the central to local levels, Tencent's database includes each official's name, position, locality, rank, and reason for

[14]It's important to note a limitation of this dataset: It doesn't comprehensively include every firm in China for each year. The frequency of the survey varies, with more thorough coverage for larger-scale firms. In contrast, smaller firms undergo random sampling annually. This selective coverage suggests that the TFP-based estimates relate specifically to a subset of procurement projects.

the investigation. As for data after August 2016, since the mainstream aborted the use of the word anti-corruption and Tencent stopped updating the database, I manually collected them from Party disciplinary committees, courts, and the People's Procuratorates from all different government levels.[15] The monthly numbers of corruption investigations in the province I study, starting from 2011 are shown in Figure A4.

# 4 Motivating Stylized Facts

Before discussing how to detect corruption, it's important to note two key stylized facts from the public procurement data. First, approximately 70% of the procurement budget is allocated to contracts with just three bids, the minimum requirement for legal procurement. In contrast, only 7% of projects have fewer than three bidders, and 18% have four bidders. Second, winners in many auctions exhibit significant score gaps compared to losers, raising concerns about the motivation behind the costly participation of the latter, who will never win given such low scores yet still invest time and money to prepare their bids.

## 4.1 Three-Bidder Auctions Dominate

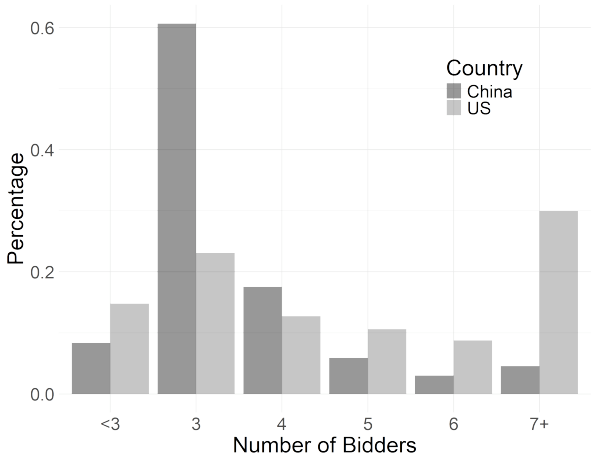
On average, competition within procurement auctions in China remains remarkably low. Figure 3 presents the distribution of bidders in these auctions. The darker bars, which encompass all years, including instances where procurement projects failed due to insufficient bidders, show less than 7% of procurement projects initially fail the requirement of having at least three bidders. Approximately 61% of procurement auctions feature only three bidders, with the percentage decreasing to 18% for those with four bidders. Surprisingly, the proportion of procurement auctions with more than six bidders is less than 10%.

Kang and Miller (2022) discusses several reasons why there is so little competition in U.S. public procurement, including seller homogeneity, and agency information rent. These reasons are possible in my study as well. However, when I examine the distribution of the number of bidders in U.S. data, as illustrated by lighter bars in Figure 3, a striking disparity emerges. Unlike the U.S. data,[16] where the distribution of number of bidders is notably smooth and boasts over 20% of auctions featuring more than seven bidders, the Chinese dataset I've analyzed paints a different picture, with a mere 4% of auctions falling into this category. The predominant presence of three-bidder cases in my dataset contrasts with the general expectation of competition in public procurement.

[15] The People's Procuratorates in China are legal bodies responsible for overseeing the enforcement of laws, safeguarding citizens' legal rights, and prosecuting criminal cases within the Chinese legal system.

[16] Public procurement requested by the Department of Defense (DOD) is excluded from US data to ensure a comparable comparison.

Figure 3: Distribution of number of bidders




**Notes**: The figure illustrates the average percentage of the number of bidders across all years, including procurement projects with fewer than three bidders. The darker bars represent data in China and the lighter bars represent data in the U.S.. Data from both countries exclude the procurement by the Department of Defense.

### 4.2 Large Winning Margins

In gauging the level of competition within procurement auctions, I adopt the concept of the winning margin, akin to the approach employed by Claudio, Frederico, and Dimitri (2015) and Kawai and Nakabayashi (2022). This variable is computed using the equation:

$$Win\ Margin_i = Score_i - \Delta Score_{-i}$$

where $\Delta Score_{-i}$ is the largest score, excluding the bidder $i$ herself. Therefore, winners have positive winning margins, while losers have negative ones. The less the absolute winning margin, the fiercer the competition is.

Table 1: Example of A Corrupt Case

| Company | Quality Score (70) | Price Score(30) | Final Score | Order |
|---|---|---|---|---|
| A | 68.4 | 29.9 | 98.3 | 1 |
| B | 21.2 | 30 | 51.2 | 2 |
| C | 17.2 | 29.5 | 46.7 | 3 |

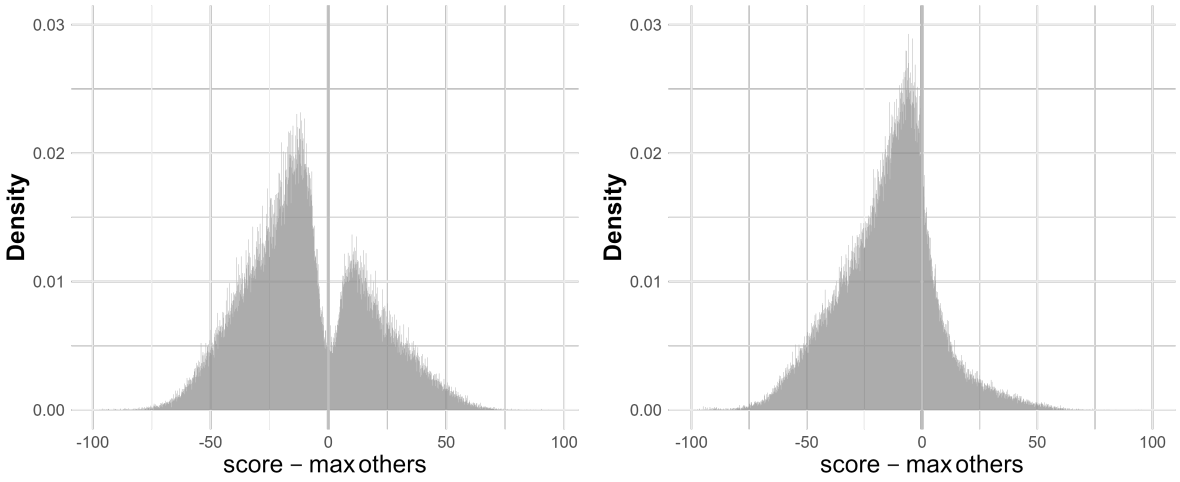
**Notes**: This corruption case is revealed by the judicial judgment. The outcomes of the procurement auction are shown on the public procurement website.

A notable case with significant winning margins is detailed in Table 1. This case involves an organic fertilizer procurement auction, which was part of a modern agricultural sector promotion project with a budget of approximately 5 million RMB in 2013. In this auction,

exactly three bidders met the minimum requirement for participation. Although the three firms had similar price scores, indicating comparable bid amounts, there was a large gap in their quality scores. The quality score for the winning firm, A, was nearly 100%, while the scores for the losing firms, B and C, were very low. This suggests that Firms B and C did not put in the effort to maximize their chances of winning, as they did not invest time and resources into preparing competitive proposals.

This case is a real example of corruption, drawn from retrospective findings on China Judgement Online.[17]The deputy director of a county-level Agriculture and Forestry Department, who managed this auction, was sentenced to seven years in prison for corruption in public procurement. His confession revealed that he had manipulated scoring rules to favor colluding firms. Firm A admitted to promising a 5% kickback to the official upon winning the contract, and Firms B and C acknowledged that Firm A had received insider assistance. These revelations highlight the severity of the corrupt practices, which led to a significant disparity in the scores.

Figure 4: Distribution of bid-differences over (bidder, auction) pairs




(a) Three Bidders

(b) Four and More Bidders

**Notes**: The X-axis represents the score gap between a bidder's own score and the highest score among all bidders, excluding the bidder's own score.

To systematically examine how common large winning margins are, I plot the distribution of bid winning margins in Figure 4. Figure 4 visualizes this distribution. Specifically, Figure 4 (a) shows the distribution for procurement auctions with only three bidders. The notable absence of mass around the 0 mark indicates that narrow winning margins are rare in such auctions. This suggests that close victories are uncommon, with most losing bids

[17]This case was propoerly classified as corruption in the corruption test section.

being significantly distant from the winning bid. This pattern reflects a lack of intense competition, even if winning bidders increased their prices. In contrast, Figure 4 (b) presents auctions with four or more bidders, where the previously observed "two-peak" pattern is no longer evident. The addition of more bidders significantly alters the distribution of winning margins, suggesting that the lack of competition is particularly pronounced in auctions with only three bidders.[18]

Despite the contrast between three-bidder and four-or-more-bidder scenarios indicating a reduced prevalence of large winning margins in the latter, it is crucial to note that this discrepancy does not necessarily imply the absence of corruption in the latter case. A case in point, as detailed in Table A1, illustrates another four-bidder procurement auction wherein Company A emerges as the unequivocal victor over Company C. Despite the participation of four bidders, substantial score gaps persist between the winner and the losers, underscoring that the issue of significant winning margins can persist even in scenarios with more bidders.

Aside from the presence of minimal bidders and substantial winning margins, there exists evidence indicating notable disparities in the performance of firms across various auctions. This variance often eludes simple attributions to divergences in procurement prerequisites or evaluation methodologies. Firms undertake diverse roles within this spectrum, including that of predetermined winners, zombie bidders, and genuine competitors. Detailed examinations through a case study and additional data analysis are provided in Appendix B2.

Collectively, these motivating stylized facts offer compelling substantiation for the existence of pervasive corruption within scoring auctions. This malpractice is especially conspicuous in auctions featuring only three bidders.

# 5 Model

This section uses a standard scoring auction model to show how the customized scoring rules result in substantial winning margins inconsistent with the competition null hypothesis.

A government entity seeks to procure a project using a scoring auction. A call-for-tender file is public for all firms, incorporating essential details like the project's budget,[19] location, project timeline, and transaction plan. Of particular importance, it outlines a detailed scoring rule function $S : (p, \boldsymbol{q}) \to \mathbb{R}$, which takes into consideration both the price bid $p$, with $p \in [\underline{p}, \overline{p}]$,[20] and the quality bid $\boldsymbol{q}$, with $\boldsymbol{q} \in \mathbb{R}^L$. More specifically, quality $\boldsymbol{q} = (\boldsymbol{q}^R, \boldsymbol{q}^U)$ consist of both project related quality $\boldsymbol{q}^R$, which is reasonable and necessary for the project, and unrelated quality $\boldsymbol{q}^U$, which is unnecessary and used to exclude other competitors.

[18] Figures for different numbers of bidders are shown in Figure A5.

[19] The budget is publicly disclosed information, known to all firms prior to the submission of bids. It represents the maximum allowable bid, serving as an upper limit on the prices firms are permitted to propose. Consequently, any bids exceeding this budget are deemed invalid.

[20] $\overline{p}$ is the budget of the project, and $\underline{p}$ is the minimum price bound.

In the context of this paper, the final scores are the weighted sum of quality scores and price scores:

$$S(p, \boldsymbol{q}) = \boldsymbol{w_q} \cdot \boldsymbol{q} + w_p \frac{\underline{p}}{p}, \text{ with } \sum_{l=1}^{L} w_q^l + w_p = 100 \text{ and } q^l \in [0, 1]$$

where $\boldsymbol{w_q}$ and $w_p$ represent the weights assigned by the government to the quality and price parts, respectively. The final score $s$ ranges from 0 to 100

The scoring rule reflects the utility function of the government agent.[21] For a competitive auction trying to maximize social welfare, the weights on unrelated quality is $\boldsymbol{w}_{q^U} = \boldsymbol{0}$, while for a corrupt auction trying to customize the scoring rules to favor specific firms, $\boldsymbol{w}_{q^U} > \boldsymbol{0}$.

In a competitive procurement auction, only related quality is required, $\boldsymbol{q} = \boldsymbol{q}^R$. $n$ ex-ante symmetric and risk-neutral firms participate in the auction, denoted by $i = 1, 2, ..., n$.[22] The number of competitors is known. Firm $i$ in the auction $a$ draws its private information, firm type $\boldsymbol{\theta_i} \in \mathbb{R}^M$, independently from a publicly known absolutely continuous distribution function $F(\boldsymbol{\theta})$.[23] Firm $i$ has a cost function $C(\boldsymbol{q_i}, \boldsymbol{\theta_i})$,[24] depending on the quality bid $\boldsymbol{q_i}$ and the firm type $\boldsymbol{\theta_i}$, and a profit function $\pi = P(s, \boldsymbol{q}) - C(\boldsymbol{q}, \boldsymbol{\theta})$.

**Assumption 1.** *Cost function $C(\boldsymbol{q}, \boldsymbol{\theta})$ is continuous, and has a quasilinear format:*

$$C(\boldsymbol{q}, \boldsymbol{\theta}) = \theta^0 + C^0(\boldsymbol{q}, \boldsymbol{\theta}^{-0})$$

*where $\boldsymbol{\theta}^{-0} = (\theta^1, ..., \theta^{M-1})$. For $\forall l$, $C_{q^l}(\boldsymbol{q}, \boldsymbol{\theta}) > 0$, $C_{q^l q^l}(\boldsymbol{q}, \boldsymbol{\theta}) > 0$, and $\forall k$, $C_{\theta^k}(\boldsymbol{q}, \boldsymbol{\theta}) > 0$.*

**Assumption 2.** *The Hessian matrix of $P(s, \boldsymbol{q}) - C(\boldsymbol{q}, \boldsymbol{\theta})$ is negative definite given score $s$, with $\forall l$, $P_{q^l}(s, \boldsymbol{q}) - C_{q^l}(\boldsymbol{q}, \boldsymbol{\theta}) > 0$ at $q^l = 0$, and $P_{q^l}(s, \boldsymbol{q}) - C_{q^l}(\boldsymbol{q}), \boldsymbol{\theta}) < 0$ at $q^l = 1$.*

With the above assumptions, shown in Hanazano et al. (2020), a symmetric pure-monotone equilibrium $s = \sigma(\boldsymbol{\theta})$ exists. The score distribution $G(s)$ is monotone and differentiable. Its density is $g(s)$.

**Theorem 1.** *The competitive equilibrium bidding strategy $(p(\boldsymbol{\theta}), q(\boldsymbol{\theta}))$ is given by*

$$P(\sigma(\boldsymbol{\theta}), \boldsymbol{q}(\boldsymbol{\theta})) = C(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{\theta}) - P_s(\sigma(\boldsymbol{\theta}), \boldsymbol{q}(\boldsymbol{\theta})) \frac{G(\sigma(\boldsymbol{\theta}))}{(n-1) g(\sigma(\boldsymbol{\theta}))}$$

[21] Follow the literature Asker and Cantillon (2008); Takahashi (2018); Huang and Xia (2019); Hanazano et al. (2020).

[22] I also introduce some asymmetry based on firm characteristics $z$ in Appendix C. Then firms with different $z$, $(\theta|z)$ have different bidding strategies. The results are not sensitive to the asymmetry.

[23] If firms are ex-ante asymmetric, the distribution of firm type is $F(\theta|z)$

[24] The assumption here is the cost only depends on the related quality $\boldsymbol{q}^R$ but is independent of the unrelated quality $\boldsymbol{q}^U$.

$$\Delta C_{\boldsymbol{q}}(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{\theta}) = \Delta P_{\boldsymbol{q}}(\sigma(\boldsymbol{\theta}), \boldsymbol{q}(\boldsymbol{\theta}))$$

*Proof.* See Appendix C

□

Following Asker and Cantillon (2008) and Hanazono et al. (2013), I define the pseudotype, the effective cost measured by score, as

$$k(\sigma(\boldsymbol{\theta}), \boldsymbol{\theta}) = \sigma(\boldsymbol{\theta}) - \frac{\pi(\sigma(\boldsymbol{\theta}), \boldsymbol{\theta})}{\pi_s(\sigma(\boldsymbol{\theta}), \boldsymbol{\theta})} \tag{1}$$

$\pi_s(\sigma(\boldsymbol{\theta}), \boldsymbol{\theta})$ represents marginal profit of score and is negative. Pseudotype $k$ can be interpreted as the highest score the firm can achieve with nonnegative profit.

The equilibrium strategy $s = \sigma(\boldsymbol{\theta})$ can be characterized by the order statistics of pseudotype $k$ in the following corollary.

**Corollary 1.1.** *In the competitive equilibrium, the winning scoring bid is the expectation of the strongest rival's pseudotype $k$:*

$$s^{win} = E[k^{rival(1)} | s^{win}] \tag{2}$$

*Proof.* See Appendix C

□

Corollary 1.1 says that under competition, the scores of the winners represent the expected pseudotype of the strongest rivals. For winners, it's unnecessary to submit scores surpassing the expected pseudotype of their strongest rivals. Going beyond the value would involve excessive effort and expenditure and ultimately leave potential extra profits on the table. This is the key to testing the competition null hypothesis.

In a corrupt auction with customization of scoring rules, unrelated quality $\boldsymbol{q}^U$ is required with $\boldsymbol{w}_{q^U} > \mathbf{0}$. [25] The predetermined winners get $\boldsymbol{w}_{q^U}\boldsymbol{q}^U$ advantage over other firms. To make sure the predetermined suppliers win the bids, the corrupt officials usually make the $\boldsymbol{w}_{q^U}$ large enough. Potential non-corrupt competitors see the signal $\boldsymbol{w}_{q^U}$ and understand that the contract is locked by another corrupt firm. They will not enter the auction because they cannot obtain $\boldsymbol{w}_{q^U}\boldsymbol{q}^U$. Even if they lower the price and enhance other quality aspects, they still cannot win the auction. To meet the minimum number of bidders requirement, fake

[25] The corrupt behaviors modeled here are achieved not through a direct increment in the quality score during the bid evaluation process (Burguet and Che, 2004; Huang and Xia, 2019) a stage where the evaluation experts merely add additional points to enable corrupt winners to outperform their competitors. Instead, it is executed through the customization of scoring rules before the initiation of auction competition, impacting both corrupt winners and their competitors.

bidders participate. These fake bidders also cannot obtain $\boldsymbol{w}_{q^U}\boldsymbol{q}^U$. Also, fake bidders do not prepare seriously for price and other quality bids and merely submit the minimum standard bid to fulfill the agreement with the predetermined winners. When there are fake bidders and no strong competitors, the predetermined winners will bid as close as possible to the highest allowable price, which is the project budget, to maximize their profits.

Because of the unrelated quality requirement, pseudotype, the highest possible score of bidders, of both potential competitors and fake bidders in corrupt auctions, $k^{rival}_{corrupt}$, are downgraded compared to that in competitive auctions $k^{rival}$.

**Corollary 1.2.** *With corruptly unrelated quality requirement* $\boldsymbol{w}_{q^U} > \mathbf{0}$*:*

$$E[k^{rival}|s^{win}] > E[k^{rival}_{corrupt}|s^{win}] \tag{3}$$

*and the equation in Corollary 1.1 doesn't hold, instead*

$$s^{win} > E[k^{rival(1)}_{corrupt}|s^{win}] \tag{4}$$

*Proof.* See Appendix C

□

# 6 Empirical Test

The two corollaries provide methods for detecting manipulation and corruption in scoring rules. Equations (2) and (4) offer estimates of group-based corruption proportions, while Equation (3) provides a test for corruption at the auction level.

To proceed with the tests, following Hanazano et al. (2020), Andreyanov (2018), and Huang (2019), I first estimate the pseudotype $k_i$ for each bidder $i$, assuming auctions are competitive. Using equation (1) and the equilibrium conditions,[26] the definition of pseudotype can be rewritten as:

$$k(\sigma(\boldsymbol{\theta}),\boldsymbol{\theta}) = \sigma(\boldsymbol{\theta}) - \frac{\pi(\sigma(\boldsymbol{\theta}),\boldsymbol{\theta})}{\pi_s(\sigma(\boldsymbol{\theta}),\boldsymbol{\theta})} = \sigma(\boldsymbol{\theta}) + \frac{G(\sigma(\boldsymbol{\theta}))}{(n-1)g(\sigma(\boldsymbol{\theta}))} \tag{5}$$

Then the pseudotype can be estimated using:

$$\hat{k}_i = s_i + \frac{\hat{G}(s_i)}{(n-1)\hat{g}(s_i)}$$

The estimated score distribution $\hat{G}(s)$ and density $\hat{g}(s)$ can be obtained non-parametrically

[26] For more details please see Appendix C

following Guerre, Perrigne, and Vuong (2000), Athey and Haile (2002), and Li, Perrigne, and Vuong (2002):

$$\hat{G}_s(s,n,\boldsymbol{x}) = \frac{1}{Th_{g_n}h_{g_x}} \sum_{a=1}^{A} \frac{1}{n_a} \sum_{i=1}^{n_a} \mathbb{1}(s_{ia} \leq s) K_G(\frac{n-n_a}{h_{g_n}}, \frac{\boldsymbol{x}-\boldsymbol{x_a}}{h_{g_x}})$$

$$\hat{g}_s(s,n,\boldsymbol{x}) = \frac{1}{Th_s h_{g_n}h_{g_x}} \sum_{a=1}^{A} \frac{1}{n_a} \sum_{i=1}^{n_a} K_g(\frac{s-s_{ia}}{h_s}, \frac{n-n_a}{h_{g_n}}, \frac{\boldsymbol{x}-\boldsymbol{x_a}}{h_{g_x}})$$

where $n_a$ is the number of bidders; $\boldsymbol{x}$ is a vector of auction characteristics; $\mathbb{1}(.)$ is the indicator function; $A$ is the number of auctions; $K_G$ and $K_g$ are kernels; and $h_{g_n}, h_{g_x}, h_s$ are bandwidths. The estimated distribution of scores is conditional on the number of bidders and other auction characteristics.[27]

### 6.1 Corruption Proportion Estimation

Equation (2) provides the null hypothesis of a competitive auction, and Equation (4) serves as the alternative hypothesis. By taking expectations with respect to winning scores on the right and left sides, the test becomes applicable for a group of auctions:

$$H_0 : E[s^{win}] - E[k^{rival(1)}] = 0 \quad H_1 : E[s^{win}] - E[k^{rival(1)}] > 0$$

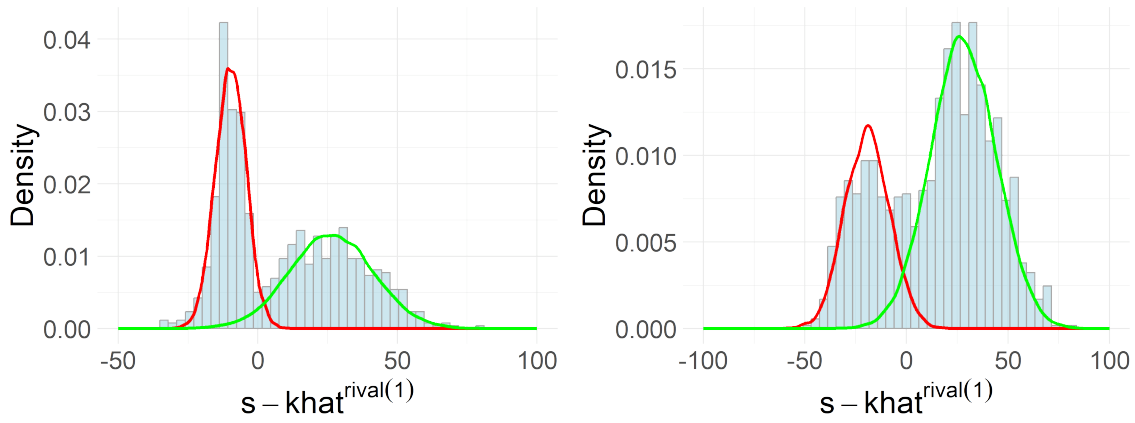
Theoretically, the null hypothesis can be tested directly. However, when corrupt and competitive auctions are mixed in a group, both the scores and the estimated $k^{rival}$ are biased. Generally, $k^{rival(1)}$ is subject to an upward bias in competitive auctions, while it experiences a downward bias in corrupt auctions due to the prevalence of low scores caused by manipulations in the scoring rule and the presence of fake bidders. [28] Also, $E[s^{win}]$ is higher when there are corrupt auctions because the scores of predetermined winners are exaggerated. This results in a separation between competitive and corrupt auctions, with $E[s^{win}] - E[k^{rival(1)}] < 0$ for competitive auctions and $E[s^{win}] - E[k^{rival(1)}] > 0$ for corrupt auctions. For instance, the bar graphs for professional service projects in Figure 5 show two peaks in $s^{win} - k^{rival(1)}$, with one peak clustering on the negative side and the other on the positive side.

The separation facilitates mixture analysis. I use the nonparametric maximum likelihood estimation (NPMLE) approach to recover the distribution $F$ of $s^{win} - k^{rival(1)}$, following methods outlined in Jiang and Zhang (2009), Koenker and Mizera (2014), Gu, Koenker, and

[27] In the estimation, auction characteristics include procurement categories, project budget, and price weights.

[28] Please see Corollary 1.3 in Appendix C for more discussion.

Figure 5: Estimation of Mixed Analysis and Fitting with Public Procurement Data




(a) Professional Service I (b) Professional Service II

**Notes**: The histograms plot the distribution of the winning score minus the pseudotype of the strongest rival, while the smooth densities represent estimations from the mixed analysis. Figure (a) focuses on professional service projects with 4 bidders, a budget greater than 2 million RMB and less than 3 million RMB, and a price weight of 10 points. Figure (b) features professional service projects with 3 bidders, a budget exceeding 3 million RMB but less than 4 million RMB, and also a price weight of 10 points.

Volgushev (2018), and Gilraine, Gu, and McMillan (2020). These papers establish a general framework for quantifying unobserved heterogeneity without imposing parametric assumptions on it. In the context of corruption within auctions, this framework is applicable as the unrelated quality scores, $\boldsymbol{w}_{q^U}\boldsymbol{q}^U$, remain unobservable. Consequently, $s^{win} - k^{rival(1)}$ can be viewed as a manifestation of unobserved heterogeneity within the auction environments. Although the distribution $F$ of $s^{win} - k^{rival(1)}$ is unobserved, it can be estimated non-parametrically from the data:

$$\hat{F} \equiv argmax_{F\in\mathcal{F}}\{\sum_{a=1}^{\mathcal{A}} log \int \varphi_d((s_a^{win} - k_a^{rival(1)}) - \theta)dF(\theta)\}$$

where $\varphi_d$ represents the standard $d$-dimensional normal density, and $\mathcal{F}$ denotes the set of all probability distributions on $\mathbb{R}$ is the nonparametric maximum likelihood estimator (NPMLE) for $\hat{F}$. Kiefer and Wolfowitz (1956) have demonstrated the consistency of the NPMLE for the mixing distribution $F$. By integrating the density for $\theta > 0$, the proportion of corrupt auctions can be obtained.

The estimated mixing distributions in Figure 5, shown in different colors, illustrate how well the model fits the data. I analyze two groups of professional service projects: one with three bidders and the other with four bidders. The estimation, which models both corrupt

(green) and non-corrupt (red) groups as mixtures, aligns closely with the observed data distribution. To further validate the NPMLE, I conducted Monte Carlo simulations, finding that the estimated proportion of corruption matches closely with the data-generating process (DGP). Additional details are provided in Appendix D1 and D2.

I then apply the method to the entire public procurement dataset. The results are presented in the first part of Table 2. Aggregating the corruption proportions from all subgroups, it is found that 61% of auctions are corrupt. For auctions with three bidders, the estimated proportion is 71%.

### 6.2 Auction level test

Equation (3) provides an action-level corruption test. For each auction $a$:

$$H_0 : E[k_a^{rival}|s_a^{win}] = E[k^{rival}|s_a^{win}] \quad H_1 : E[k_a^{rival}|s_a^{win}] < E[k^{rival}|s_a^{win}]$$

where $E[k_a^{rival}|s_a^{win}]$ is the mean pseudotype of the actual rivals in auction $a$, and $E[k^{rival}|s_a^{win}]$ is expected pseudotype of rivals in competitive auctions conditional on the winning score $s_a^{win}$.

Rejecting the null hypothesis $H_0$ provides evidence in favor of $H_1$, suggesting that rivals' scores are downgraded compared to the winners. This, in turn, leads to the rejection of the hypothesis that the auction is competitive. However, the presence of corrupt auctions threatens the estimation of the distribution of $s$ and introduces bias into the estimation of $k$. Unlike the corruption proportion estimation, the auction-level test requires unbiased pseudotype estimates.

To correct the biased pseudotype, I employ a systematic method similar to the methods such as Iterative Outlier Removal (IOR) (Parrinello et al., 2016) from public health literature, and the Outlier Removal Clustering algorithm (ORC) (Hautamäki et al., 2005), k-means with outlier removal (KMOR) (Gan and Ng, 2017), and mean-shift outlier filtering (Yang, Rahardja, and Fränti, 2021) from the domain of computer science and data mining. The steps are detailed below:

Step 1. Estimate the pseudotypes of all bidders in all auctions using equation (5).
Step 2. Conduct the auction-level test.[29]
Step 3. Filter out the auctions that reject the null hypothesis ($H_0$) and repeat steps 1 and 2 until the proportion of rejected auctions is within the desired level.[30]
Step 4. Use the estimated distribution of $s$ from Step 3 to estimate the pseudotypes of all bidders in all auctions and apply the statistical test to save p-values.

[29]For statistical test methods and significance level choices, please see Appendix D3 for more details.

[30]For main results, I choose $\alpha = 0.1$ to ensure the power of the test, but I also show the results for $\alpha = 0.05$ in Section 6.3.

The structured approach ensures the systematic and iterative application of the pseudotype estimation method, accounting for corrupt auctions. For example, in the "for-sure corrupt case" in Section 4, Table 1, the auction involved three bidders with scores of 98.3, 51.2, and 46.7. Given the winning score of 98.3 and the auction characteristics, [31], $\bar{\hat{k}}^{rival}$ conditional on winning score 98.3 from the iterative method is 90.72. However, in this auction $\bar{\hat{k}}_a^{rival}$ are 53.91. The p-value of the Wilcoxon test is 0.002, leading to the rejection of the null hypothesis. Therefore, the auction is successfully identified as corrupt using my method. The graphical illustration is shown in Figure D7. There are more discussions about the example auctions can be found in Appendix D4.

To demonstrate the efficiency and accuracy of these iterative test steps, I conduct a Monte Carlo Study with simulations in Appendix D3. In the Monte Carlo Study, the approach can achieve an overall 90% accurate prediction rate when mixing 20% to 80% of corrupt auctions. Instead of the test with an iterative way to estimate the corrected pseudotypes, I also propose a method using the biased pseudotypes and the k-means unsupervised learning clustering method (Wu et al., 2008; Hartigan and Wong, 1979; Celebi, Kingravi, and Vela, 2013) to label corrupt and noncorrupt auctions. However, the iterative method overperforms the k-means learning in both simulated data and procurement data in terms of accuracy rate.[32]

Table 2: Corruption Proportion Estimation and Auction-level Test Results

| | All | 3 Bidders | 4 Bidders | 5 Bidders |
|---|---|---|---|---|
| Corruption Proportion Estimation | | | | |
| Corruption Proportion | 60.57% | 70.88% | 46.49% | 26.85% |
| | [60.19, 60.96] | [70.21, 71.55] | [45.95, 47.03] | [24.77, 28.93] |
| Auction Level Test $\alpha = 0.10$ | | | | |
| **Unadjusted P-value** | 64.71% | 69.74% | 51.18% | 47.74% |
| | [64.64, 64.78] | [69.61, 69.87] | [51.03, 51.33] | [47.44, 48.04] |
| Adjusted P-value | 56.83% | 62.36% | 43.39% | 38.32% |
| | [56.78, 56.89] | [62.26, 62.47] | [43.27, 43.51] | [38.17, 38.46] |

**Notes**: 95% confidence intervals from bootstrapping are indicated in brackets. The first two rows of the auction-level test results represent the results using unadjusted p-values, while the following two rows display the adjusted p-values to control the false discovery rate. Unadjusted p-values are more powerful than adjusted p-values but have a higher type I error rate.

Table 2 part 2 summarizes the test results by showing the proportion of procurement auctions that reject the null hypothesis. Overall, approximately 65% of the auctions reject

[31] Auction characteristics include budget, price/variable quality weight, procurement category, and number of bidders

[32] Please see Appendix D5 for more details.

the competitive auction null hypothesis, and even with the conservative approach, this figure is 57%. In three-bidder auctions, the majority reject the null hypothesis. Specifically, in 70% of the 3-bidder auctions, the test results show evidence of a downgrade of the scores of rivals.[33] For four-bidder auctions, although the proportion is smaller than in three-bidder auctions, it is still significant. More conservative results with $\alpha = 0.05$ can be found in Table A2.

The aggregated proportions of corruption derived from auction-level tests closely align with the proportions directly estimated using the nonparametric maximum likelihood estimator (NPMLE).[34] Furthermore, these proportions are similar to those estimated through the k-means clustering method, detailed in Appendix D5.

### 6.3 Discussion

Although rejecting the null hypothesis indicates evidence of the downgrading of rivals' scores, it cannot be concluded that the downgrading is solely due to the manipulation of scoring rules by corruption. One of the most possible ones is collusive bidding by all participants in a cartel group without corrupt officials' involvement. However, several factors make this scenario less likely, including the low co-entry frequency, fragmented markets, and high cost of collusion maintenance under the dominant power of the government.

**Low co-entry frequency.** Cartel members often establish co-entry arrangements to bolster their capacity to manipulate auction outcomes in favor of the cartel (Porter and Zona, 1993; Conley and Decarolis, 2016; Schurter, 2017; Kawai, Nakabayashi, and Ortner, 2021). A pivotal indicator of collusion is the recurrent participation of a group of firms in the same auction. In order to examine how frequently the same pair or trio of firms co-participate in procurement projects, I provide a frequency distribution in Figure 6 that covers all 214,000 procurement instances in my dataset. [35] About 86% of all possible pairs of firms participate in the same procurement equal or less than twice, and only a small minority participate more than twice. For trios of firms, this percentage increases to 96%, indicating that only 4% of three-bidder groups interact more than twice.[36]

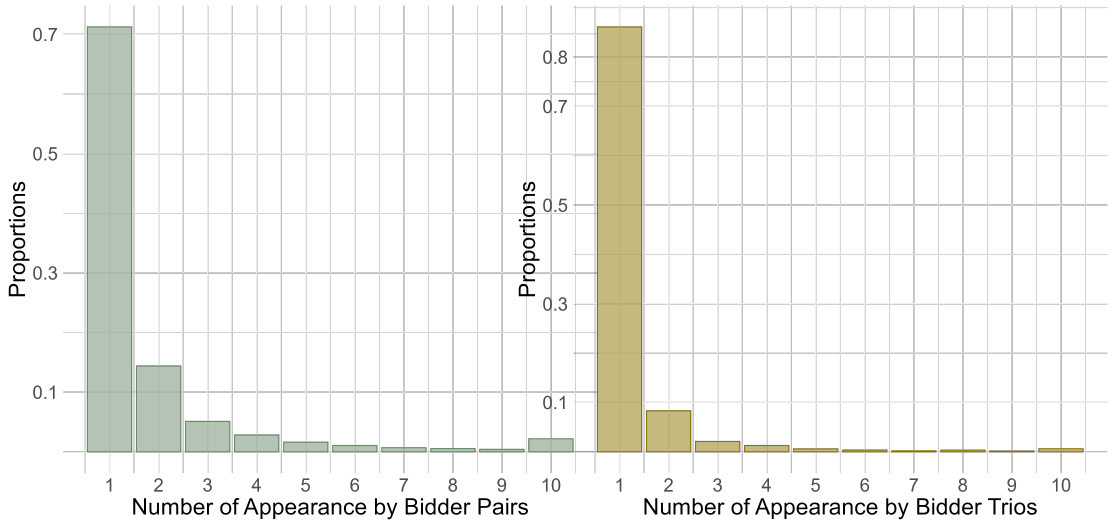
[33] Using adjusted p-values, the proportion diminishes to 63%. Even employing the most conservative approach, utilizing $\alpha = 0.05$ and adjusted p-values, over 50% of auctions consisting of three bidders still reject the null hypothesis.

[34] As shown in the Table A2, for cases with three and four bidders, the results of the proportion estimation and the auction-level test are very similar. However, as the number of bidders increases, the results gradually diverge. This occurs because the proportion estimation only leverages the scores of the winner and the second-highest bidder, without considering other bidders. If the other bidders' scores are clustered closely but at a distance from the winner, the auction test can detect collusion among them.

[35] It is noteworthy that more auctions can offer more comprehensive insights into the firm network. Therefore this analysis uses both open auctions and non-open auctions.

[36] While my dataset doesn't provide comprehensive procurement information for all provinces, it is worth noting that repeated participation is less common in this setting, as firms are less likely to participate in procurement conducted in other provinces.

Figure 6: Frequency of Bidder Pairs and Trios




(a) The Frequency of Bidder Pairs (b) The Frequency of Bidder Trios

**Notes**: The X-axis represents the frequency of occurrence where the same pair or trio of firms participate in the same public procurement auctions. The Y-axis indicates the proportion of occurrences for each frequency.

Furthermore, I applied the participation test using the approach proposed by Conley and Decarolis (2016).[37] The null hypothesis is that the entry of groups of firms is competitive. The results of the test are intriguing, as only a small percentage of groups with two or more members - specifically 7.64% - rejected the null hypothesis. The number was even lower for groups with three or more members, at 2.59%. These figures are significantly lower than the 69% of auctions in Conley and Decarolis (2016) that rejected the null hypothesis. These results indicate that repeated participation, which is a crucial red flag of bid rigging, is not a common occurrence in the data.

**Fragmented Market.** Overall, compared to developed countries, the market in China is notably fragmented and is also characterized by its dynamism, marked by the frequent emergence and exit of firms. To assess the level of concentration [38]. I utilize firm tax data from 2016 to calculate the Herfindahl–Hirschman Index (HHI) by industry and illustrate the distribution of the HHI in Figure A6. With a mean HHI of 335, the market is substantially below the concentrated market threshold of 2500, indicating prevalent market fragmentation.

**High Collusion Maintenance Cost.** Maintaining collusion is costly due to two reasons. First, in economies where government authorities have significant control, they can manipulate auctions to benefit specific bidders, making it difficult for other companies

[37]More details on the methodology and results of the participation test can be found in Appendix D.

[38]Literature (Scherer and Ross, 1990; Motta, 2004) has taken market fragmentation as an important factor undermining the ability of firms to collude

to engage and win contracts. Consequently, firms that manage to forge corrupt connections have a higher chance of securing contracts, eliminating the need for other companies to create collusive groups. Second, even the predetermined winner within a cartel faces the need to bid competitively, incurring significant expenses. This dynamic is made even more complex by the continuous entry of numerous genuine competitors into the market. In such an environment, the presence of these additional competitors and the necessity for competitive bidding undermine the rationale for forming collusive groups, making the maintenance of such groups logically and economically challenging.

The points mentioned above show that it is less likely for simple cartel groups to control procurement auctions and create big differences in scores resulting in the rejection of the test. To further confirm this is true, I conduct an expert audit study in the next section.

# 7 Expert Evaluation Survey

In this section, I utilize data from a dedicated audit study of procurement experts to further demonstrate that the missing mass identified through statistical tests may primarily result from the manipulation of scoring rules.

In interviews with procurement officers, they emphasized that ensuring the predetermined winner secures the auction victory involves crafting customized scoring rules for the favored firm. These rules are deliberately designed with a level of opacity that makes them unfamiliar and unidentifiable to those outside the local industry network. Even government auditing teams are often unable to address this issue because they lack in-depth knowledge of the specific project and industry, making it nearly impossible to find tangible evidence like covert agreements or illicit financial transfers between the corrupt firm and the involved officer. Consequently, annual fiscal auditing reports primarily list procurement cases conducted outside legal procedures, with very few cases addressing the presence of unreasonable scoring rules. Fortunately, government evaluation experts, who are randomly selected to assess bid proposals, possess substantial knowledge about project requirements and the local market structure. They can readily identify any potential superfluous rules in place for the project and determine if these rules have been tailored to favor specific firms.[39]

Drawing from the insights of these procurement experts, I designed the survey as follows. I randomly selected 500 procurement projects from the comprehensive procurement dataset. Table 3 provides a summary of the survey sample alongside all procurement projects. On average, each auction involved 3.68 bidders, with an average project budget of 0.7 million USD. The winning bidder's revenue constituted 97% of the project budget, and the average

[39]Opinions from experts are subjective and not necessarily the ground truth. The survey aims to offer a practical perspective on identifying potential redundant requirements and serves as a lens to evaluate the statistical results from the previous section.

Table 3: Balance Table for the Survey Sample

| | Sample | All | P-Value |
|---|---|---|---|
| Obs | 500 | 186,351 | |
| #Bidders | 3.68 | 3.74 | 0.48 |
| Budget | 4.91 | 4.46 | 0.5 |
| Budget Utilization Ratio | 0.97 | 0.97 | 0.71 |
| Price Weight | 22.76 | 23.36 | 0.29 |
| %Construction | 0.09 | 0.09 | 0.85 |
| %Service | 0.39 | 0.35 | 0.1 |
| Median Year | 2017 | 2017 | |

**Notes**: The unit for the budget is in million RMB. The budget utilization ratio is calculated by dividing the contract value by the budget.

quality score for the winner was 0.858 out of 1. Notably, there were no statistically significant differences between the survey sample and all procurement cases.

Second, I categorized the procurement projects in the sample and distributed the corresponding call-for-tender files to five procurement experts, each based on their area of expertise. In general, two experts focused on procurement for goods, one on construction procurement, and the other two on service procurement. These call-for-tender files included comprehensive project requirements and scoring rules. Without knowledge of the bidders or the outcomes of the procurement sample, the experts were tasked with reading each call-for-tender file and responding to a set of questions. These questions covered various aspects, including the expert's estimation of the general markup,[40] the local market competition, and the design of the RFPs.[41]

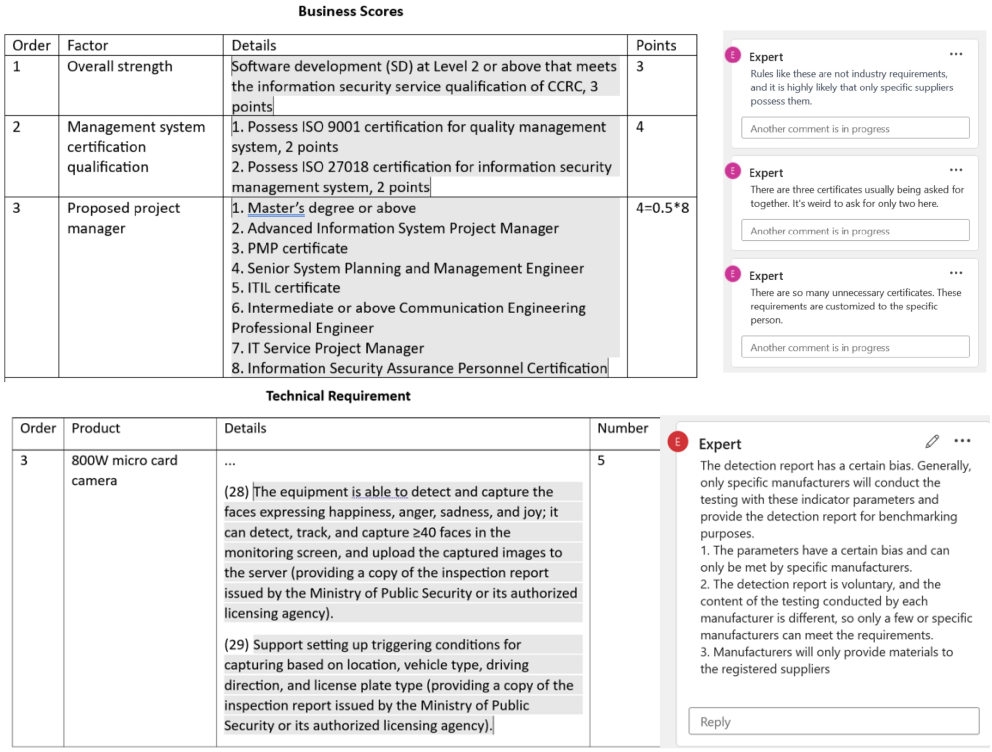
The most crucial part of the survey entailed asking the experts whether they detected any redundant requirements within the scoring rules, particularly in the technical and business sections, that might suggest the customization of scoring rules. If any such instances were identified, the experts were instructed to highlight and provide comments. Prior to distributing the survey, I ensured that none of the experts had participated in the evaluation of the procurement projects they were assigned to review.

Figure 7 provides a snapshot of a call-for-tender file for county-level government procurement of closed-circuit television (CCTV) equipment in 2021. In the business section, the expert identified three significant rule sets. The first set mandated high information security certification for all bidders, a requirement that appeared unnecessary for the project.

[40] Experts are requested to provide approximate estimates of the standard markup for similar projects with which they have had prior experience. The average markup rate provided by the experts is 23%, indicating that the projects are appealing to firms.

[41] An English version of the survey can be found in Appendix B.

Figure 7: An Example of Procurement Survey



**Business Scores**

| Order | Factor | Details | Points |
|---|---|---|---|
| 1 | Overall strength | Software development (SD) at Level 2 or above that meets the information security service qualification of CCRC, 3 points | 3 |
| 2 | Management system certification qualification | 1. Possess ISO 9001 certification for quality management system, 2 points<br>2. Possess ISO 27018 certification for information security management system, 2 points | 4 |
| 3 | Proposed project manager | 1. Master's degree or above<br>2. Advanced Information System Project Manager<br>3. PMP certificate<br>4. Senior System Planning and Management Engineer<br>5. ITIL certificate<br>6. Intermediate or above Communication Engineering Professional Engineer<br>7. IT Service Project Manager<br>8. Information Security Assurance Personnel Certification | 4=0.5*8 |

Expert
Rules like these are not industry requirements, and it is highly likely that only specific suppliers possess them.
Another comment is in progress

Expert
There are three certificates usually being asked for together. It's weird to ask for only two here.
Another comment is in progress

Expert
There are so many unnecessary certificates. These requirements are customized to the specific person.
Another comment is in progress

**Technical Requirement**

| Order | Product | Details | Number |
|---|---|---|---|
| 3 | 800W micro card camera | ...<br>(28) The equipment is able to detect and capture the faces expressing happiness, anger, sadness, and joy; it can detect, track, and capture ≥40 faces in the monitoring screen, and upload the captured images to the server (providing a copy of the inspection report issued by the Ministry of Public Security or its authorized licensing agency).<br>(29) Support setting up triggering conditions for capturing based on location, vehicle type, driving direction, and license plate type (providing a copy of the inspection report issued by the Ministry of Public Security or its authorized licensing agency). | 5 |

Expert
The detection report has a certain bias. Generally, only specific manufacturers will conduct the testing with these indicator parameters and provide the detection report for benchmarking purposes.
1. The parameters have a certain bias and can only be met by specific manufacturers.
2. The detection report is voluntary, and the content of the testing conducted by each manufacturer is different, so only a few or specific manufacturers can meet the requirements.
3. Manufacturers will only provide materials to the registered suppliers
Reply

**Notes**: The graphs displayed are screenshots taken from the expert evaluation survey responses.

The second set drew attention because it is typical to demand compliance with three basic system certifications, whereas this specific file only requested two. The expert suspected that the target firm might possess only two of these certifications. The third set related to the qualifications of team members, encompassing degrees, majors, and examination certificates. Meeting all these stipulations would limit the eligible bidders to just one or two local firms. In the technical section, the expert noted that only one manufacturer could provide the specific test reports requested by the file, and the ultimate winner must be the exclusive local supplier authorized by that manufacturer. Reassuringly, the winning firm and its brand closely aligned with the expert's observations.

Table A3 presents the auction results for the above case. This case rejects the null hypothesis of competitive procurement auctions and is marked as corrupt. The winning bidder achieved a perfect score in the business score section and nearly a full score overall, indicating compliance with all non-essential requirements highlighted by the expert. Consequently, this case falls into the True Positive category.

I collected and digitized responses from the 500 surveys, including the unobserved expert-estimated markup, which is presented in the final row of Table 3. On average, this markup amounted to approximately 23%. This substantial markup figure makes these pro-

curement auctions attractive to competing entities. If experts highlighted any rules in the call-for-tender documents, the corresponding procurement auction is classified as corrupt, as indicated by expert surveys. Conversely, if no such highlights were identified, the auction is considered non-corrupt. The consistency rate is calculated by dividing the number of auctions that statistical tests and surveys give the same classification by the total number of auctions. The results are outlined in Table 4. Remarkably, the consistency rate reaches 82%, with a test power of 84.5% and a test size of 25.6% when assuming the highlights of scoring rules from the expert are the ground truth of corruption.

While the information obtained from the expert survey sheds light on suspiciously corrupt rules, it is important to acknowledge that both binary corruption indicators are only approximations of the 'true' ground truth. This means they are not a perfect, 100% representation of corruption. Notably, among the false positive instances, there are nine cases where experts did not highlight unnecessary rules, but they indicated that the scoring criteria were excessively subjective. This subjectivity creates an environment conducive to score manipulation. Excluding these 9 files results in a test size of 18.4%.

I also provide the same table using adjusted p-values in Appendix Table A4. The overall consistency rate is 79.4%, which is slightly lower than that achieved using unadjusted p-values, but with a lower type I error rate. This comparison further reinforces the decision to use unadjusted p-values as the primary results.

Table 4: Expert Survey and Model Prediction

| Consistency Rate=82% | Total=500 | Model Prediction | | |
|---|---|---|---|---|
| | | Corrupt $H_1$ | Not Corrupt $H_0$ | |
| Expert Survey | Corrupt (P) | 317 | 58 False N | Test Power 84.53% |
| | Not Corrupt (N) | 23+(9) False P | 93 | Test Size 25.6% (18.4%) |

**Notes**: For the expert survey, if any evaluation survey reveals highlighted scoring rules accompanied by expert explanations and concerns, the procurement project is labeled as corrupt. The classification of model prediction uses unadjusted p-values with a significance level of 0.1.

The comparison between test results with the expert audit study shows that in general, the model-based statistical test yields reliable performance. The majority of substantial score disparities identified by the test are attributed to the tailored scoring rules, rather than to other conceivable reasons.

# 8 Policy Implications

In this section, I will first present evidence of the selection of less efficient firms due to corruption in procurement auctions. Then, I will discuss three policies aimed at reducing corruption in scoring auction procurement. The first policy is anti-corruption investigations. I use the recent anti-corruption campaign in China and an event-study design to analyze its effects on corruption. The other two policies are counterfactuals: anonymous evaluation of Requests for Proposals (RFPs) and increasing the weight given to price in procurement decisions. I will use an estimated structural model to evaluate these counterfactual scenarios.

## 8.1 Anti-corruption Investigations

Anti-corruption investigations have been widely recognized as effective tools for reducing corruption (Avis, Ferraz, and Finan, 2018; Chen and Kung, 2019; Colonnelli and Prem, 2021). I examine whether a recent anti-corruption campaign in China has led to changes in the procedures and outcomes of public procurement, potentially reducing corruption. After President Xi came to power in 2014, the Chinese Communist Party launched an anti-corruption campaign aimed at eliminating corruption among officials at all levels. Although the campaign targets general corruption issues beyond just public procurement, recent evidence (Manion, 2016; Wang and Dickson, 2022; Fang, 2023) indicates that it has altered the government's incentive structure and effectively reduced opportunities for bureaucratic corruption.

Unlike much of the prior research, which often relies on indirect measures of corruption, my study uses a direct approach to assess corruption in public procurement. To investigate systematic differences in procurement methods, outcomes, collusion among bidders, and corruption involving both bidders and officials, I analyze variations in exogenous corruption investigations conducted on officials across different government departments. The event study reduced form is presented below:

$$y_{idt} = \alpha_t + \alpha_d + \alpha_0 Procurement_{idt} + \beta_{1T} \sum_{T=-4}^{+5} Time\ to\ Investigation_{idT} + \epsilon_{idt}$$

where the dependent variables $y_{idt}$ for procurement contract $i$ purchased by department $d$ at time $t$ include procurement methods, number of bidders, corruption prediction from the test, and an indicator for whether the winner is a new supplier to the department. $Time\ to\ Investigation_{idT}$ equals 1 if, at time $t$, $T$ quarters have passed before/since a government official in department $d$ was under corruption investigations.[42] I control $\alpha_t$ time year-

[42]The event study specification is similar to the one in Beraja et al. (2023) and consolidates the treatment time period to a quarterly level.

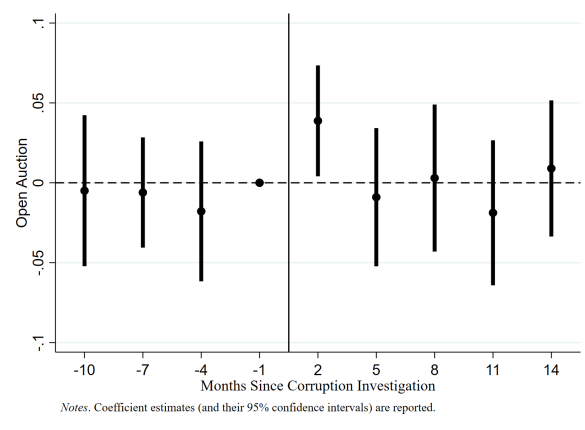
month fixed effects, and $\alpha_d$ buyer department fixed effects. I also control $Procurement_{idt}$, other procurement characteristics, such as procurement of product categories and reserve prices. The analysis uses samples from all departments, regardless of whether corruption investigations occurred or not. Results derived exclusively from departments that underwent investigations are detailed in Appendix A Figure A7.

The analysis focuses on two sets of dependent variables related to the impact of corruption investigations. The first dependent variable examines the selection of auction methods. Specifically, a value of 1 is assigned to $y_{idt}$ if the auction is an open auction, and 0 otherwise. To capture the gradual decline of the shock resulting from the investigation over time, the event study window is limited to one year prior to and one and a half years subsequent to the investigation. The outcomes are illustrated in Figure 8 (a). The results indicate that after the corruption investigations, the departments immediately had 5 percentage points more likely to choose open auctions instead of non-open auction methods. However, the coefficient is only statistically significant in the first following quarter. After the second quarter, this positive effect faded away.

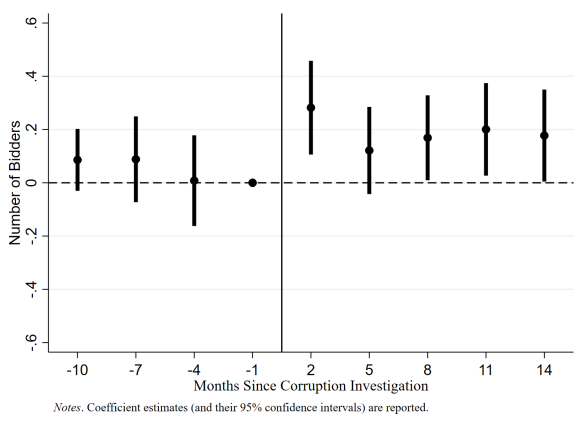
The second set of dependent variables focuses on the outcomes of auctions. Figure 8 (b) examines the number of bidders as the dependent variable. It shows an increase of 0.3 bidders in the first and third quarters following the investigations. Beyond the second quarter, while the coefficients retain their positive values, there's a slight decline. However, they sustain statistical significance at the 0.1 level. The effect indicates a sustained increase in firm entries as a supply-side effect, likely spurred by the corruption investigations. Figure 8 (c) uses the labels of corruption from the statistical test as the dependent variable. The investigation decreases the probability of corruption in a procurement auction of an investigated department, the effect is particularly statistically significant in the post-third quarter, representing a 6 percentage point drop. The delay aligns with the timeline of open scoring auctions, given that the procurement plan is often published several months prior to the actual bid evaluation meeting. Finally, Figure 8 (d) examines the probability of the winning firm being a first-time supplier for the particular local procurement government. The exploration is to see whether there is a new supplier-government network built after the corruption investigation. The results indicate that there is an increase in the probability of new firms winning procurement auctions during the post-investigation period.

The findings from the event study indicate that corruption investigations bolster competition in open auctions by prompting more firms to join the bidding process. Also, these investigations reduce manipulative practices in the design of scoring rules. By referencing the estimated coefficients from the third quarter as depicted in Figure 8 (b) and (c), I can gauge the influence of corruption on the number of entrants using a back-of-the-envelope

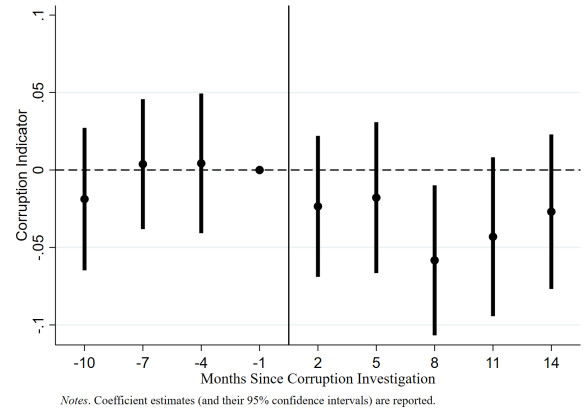
Figure 8: The Effect of Corruption Investigations



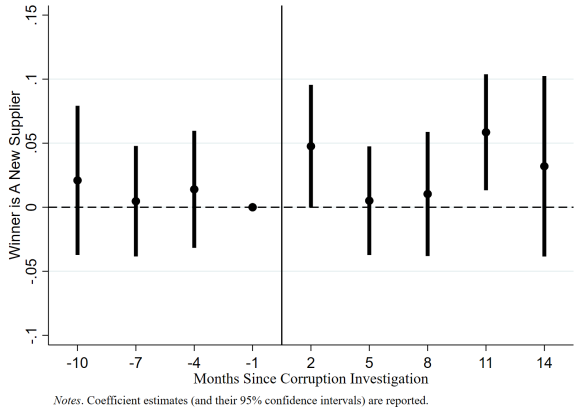

(a) Open Auction




(b) Number of Bidders




(c) Corruption Indicator




(d) New Winning Firm=1

**Notes**: The regression model incorporates time-fixed effects and procurement government department fixed effects. Standard errors are clustered at the department level.

calculation of corruption's marginal effects:

$$\text{Entry Effect of Corruption} = \frac{\Delta Number\ of\ Bidders}{\Delta Corruption\ Probability} = \frac{0.169}{-0.058} = -2.914$$

The calculation exercise can be further used in the counterfactual analysis in Section 8.3 to decompose the effects of reducing unnecessary scoring rules on social welfare.

The impacts of "Flies or Tigers" crackdowns exhibit variance across officials at different hierarchical levels. To comprehend which investigations, especially those involving officials at divergent levels, contribute to the observed effects, I categorize the sample into two cohorts: investigations engaging officials who are at least the principal leader of a county-level

department, and those involving officials positioned below this level. [43] The impacts of corruption investigations reveal disparities among officials at varied levels.[44] The comparisons are presented in Figure 9. The positive effects on promoting competition from Figure 8 come from the investigations on high-level officials. Investigations on higher-level leaders reduced the likelihood of corruption in open auctions by 10 percentage points, whereas scrutiny of lower-level leaders did not impact corruption within such settings. Correspondingly, the influence on the preference for open auctions is also predominantly traced back to investigations focused on high-ranking officials. Additional illustrative material is provided in Figure A8.

The comparison between high-level and low-level officials aligns with empirical realities. As officials rise in the hierarchy, their influence and power increase. In contrast, lower-level officials, who hold less authority, often find themselves held accountable and sometimes act as scapegoats for the wrongdoings of their superiors.

Additionally, a more detailed discussion of corrupt firms and variations in city-level corruption is provided in Appendix H.

### 8.2 Counterfactuals: Estimation

Since anti-corruption investigations have failed to generate a sustainable reduction in corruption in public procurement, an alternative approach could be to reform some of the auction procedures. To explore this idea, I go back to the scoring auction model for counterfacutal policy analysis.

I first estimate firm type $\boldsymbol{\theta}$ separately for competitive and corrupt auctions. For competitive auctions, I use the equilibrium conditions provided in Theory 1 to infer firm type $\boldsymbol{\theta}$. With the estimated CDF $G(s)$ and PDF $g(s)$ of the score $s$ from Section 6, the only remaining component that is not yet structured is the cost function. I propose a restricted polynomial form for the cost function with two unknown parameters, $\theta_0$ and $\theta_1$, representing firm type $\boldsymbol{\theta}$. [45]

$$C(q, \theta_0, \theta_1) = \theta_0 + \beta_1(\theta_1 + q)^{\beta_2}$$

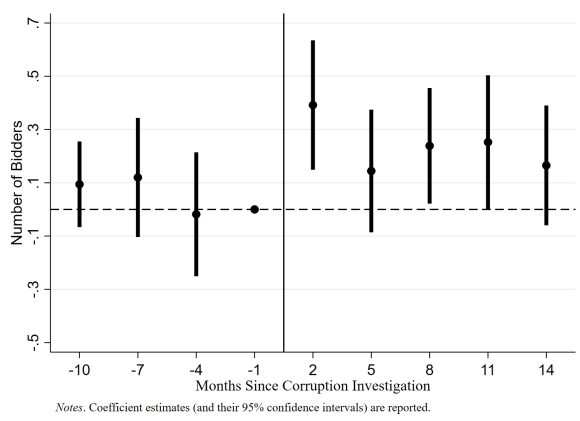
I set $\beta_1 = (4\frac{w_q}{w_p})^2$ to account for the scale differences caused by the weight of price and quality in $P_s$ for different auctions, ensuring that Assumption 2 holds. Additionally, I choose

[43] Officials who are principal leaders of a county-level department, such as the primary director of the education department in M County or the director of the transportation department in B district, G City, are designated as level nine based on China's bureaucratic hierarchy, while officials ranked below are at level ten or lower.

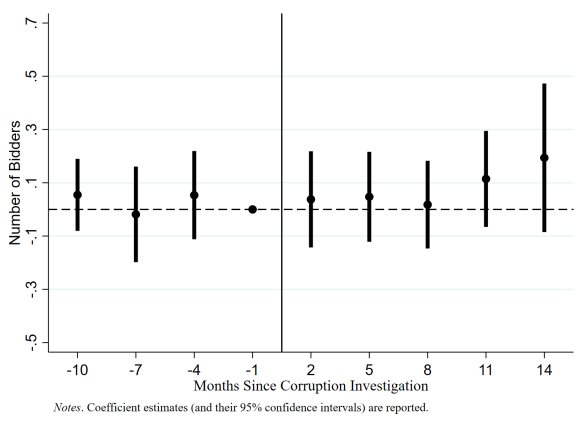
[44] In certain instances, lower-ranking officials are targeted as scapegoats for their higher-ranking counterparts.

[45] I follow existing literature in using a polynomial cost function. Since I only have two equilibrium conditions, I cannot introduce more flexible parameters as coefficients.

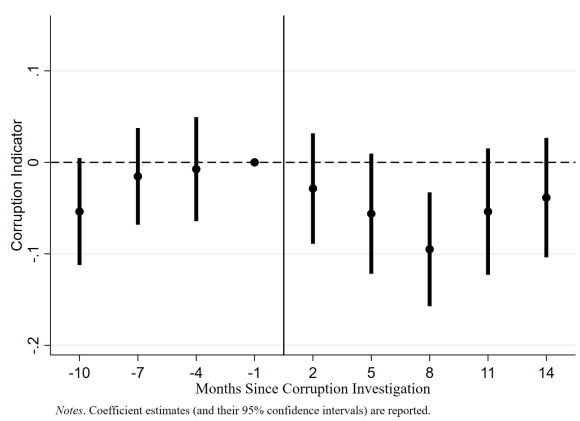
Figure 9: The Effect of Corruption Investigation by Officials with Different Levels



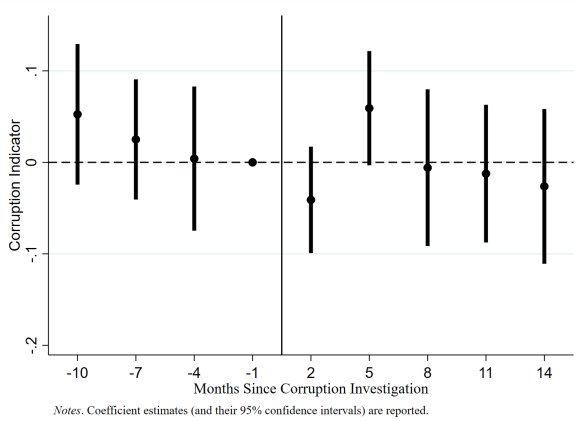

(a) Number of Bidders: High-level Officials




(b) Number of Bidders: Low-level Officials




(c) Corruption Indicator: High-level Officials




(d) Corruption Indicator: Low-level Officials

**Notes**: The regression model incorporates time-fixed effects and procurement government department fixed effects. Standard errors are clustered at the department level.

$\beta_2 = 4$, making the cost function a fourth-order polynomial. [46] This ensures the cost function satisfies Assumption 2, guaranteeing the existence of an optimal choice. With this setup, $\theta_0$ and $\theta_1$ are identified. The 3D estimated joint distribution of $\theta_0$ and $\theta_1$ is shown in Figure A9 (a), while Figure A9 (b) presents the two-dimensional flat density.

However, in the case of corrupt auctions, I am unable to use the first-order competitive equilibrium to estimate the firm type. [47] This is because all bidders achieve systemati-

[46] For the robustness check of the functional form, I set $\beta_1 = (3\frac{w_q}{w_p})^2$ and $\beta_2 = 3$. The counterfactual analysis follows a similar approach to the main version.

[47] I present a comparison of the estimated parameters between corrupt auctions and noncorrupt auctions, assuming they are all estimated from the competitive first-order conditions. I group the $\theta$ values into quintiles for each procurement category. Lower quintile rankings represent higher efficiency. The comparison is shown in Figure A10 and Figure A11.

cally different scores in corrupt auctions. Moreover, I do not have direct observations of the manipulation of the scoring rules $\boldsymbol{w}_{q^U}$. As a result, I estimate the firm types $\boldsymbol{\theta}$ in corrupt auctions by regressing estimated firm type parameters $\theta_0$ and $\theta_1$ on auction and firm characteristics in non-corrupt cases using the following regression specification for winners and losers respectively [48]:

$$\theta_{ia} = \alpha_0 + \alpha_1 Firm\ Characteristics_i + \alpha_2 Auction\ Characteristics_a + \epsilon_{ia}$$

where the dependent variable $\theta_{ia}$ is $\theta_0$ or $\theta_1$ of the bidder $i$ in the auction $a$. *Firm characteristics* include firm capital size, and formal employee size, the exitsing years, and stateowned connection. $Auction\ Characteristics_a$ are procurement auction-level characteristics, including budget size, price weights, category fixed effects, and region fixed effects. The regression outcomes are presented in Table A5 for winners and Table A6 for losers. Then I obtain the predicted $\boldsymbol{\theta}$ for firms in corrupt auctions.[49]

With the estimated firm types $\boldsymbol{\theta}$, I am now equipped to conduct the two counterfactual analyses in the following two subsections.

### 8.3 Counterfactual: Implementing Anonymous Evaluation on RFP

Building on the extensive analysis of procurement data and scoring auctions, this section explores a counterfactual policy intervention aimed at improving the integrity of the Request for Proposal (RFP) process. While measuring actual quality across diverse products and services remains challenging, final scores are commonly used as a proxy for social welfare in the literature on scoring auctions (Hanazano et al., 2020; Takahashi, 2018; Andreyanov, 2018). Despite the availability of binary quality check data, its limited scope and applicability make it less effective for estimation.[50]

The expert audit survey in Section 7 shows an effective way to identify the unnecessary quality requirements in the RFP. Governments could employ a strategy of randomly selecting anonymous experts to assess the RFP prior to the commencement of the auction competition, akin to the expert audit study. Should these anonymous experts identify any superfluous scoring rules, the respective government department would be instructed to revise the RFP. This approach strikes a balance by retaining some discretion in scoring rule design, yet diminishes the chances of tailored scoring rules. Such a policy mirrors the practice

[48] As depicted in Figure A10 and Figure A11, there are significant differences between winners and losers, even in some non-corrupt scenarios.

[49] The predicted $\theta$ values of winners in corrupt cases were then added to the quintile bar plots depicted in Figure A12 and Figure A13, showing that the predicted losing firm types are more efficient the original values suggesting.

[50] The quality check data is less useful in this context, as only 20% of procurement projects can be linked to procurement and bidding information, and fewer than 0.06% of projects failed the quality check.

of randomly selecting bid evaluation experts.

In the majority of procurement auctions, it's unclear which elements are tailored to favor predetermined winners. However, within the expert survey sample, I've requested experts to identify the rules deemed unnecessary. In a counterfactual scenario where no rules are deliberately designed to benefit winners, other requirements would take the place of the current customization. I conduct a counterfactual analysis, eliminating all preferential rules in the quality requirement $\boldsymbol{w}_{q^U} = 0$.

The effects of removing unnecessary scoring rules come from two components. First, as there are no customized rules granted to the predetermined winners, they have to participate in the bidding process competitively. Second, because of the removal of restrictive rules as a signal of corruption (Cai, Henderson, and Zhang, 2013), more firms would enter into the auctions. I set the extra number of bidders as two, as shown in the back-of-envelope calculation in Figure 8 (b). Therefore, I can decompose these two effects. For the two more bidders, they are drawn randomly from the other competitive auctions with the same auction characteristics, such as project type, budget range, and number of bidders. By applying the competitive equilibrium conditions, and the adjusted firm type, I conducted re-simulations of the auctions to estimate the changes in price and quality bids when the predetermined winners engaged in competitive bidding. The results are presented in Table 5.

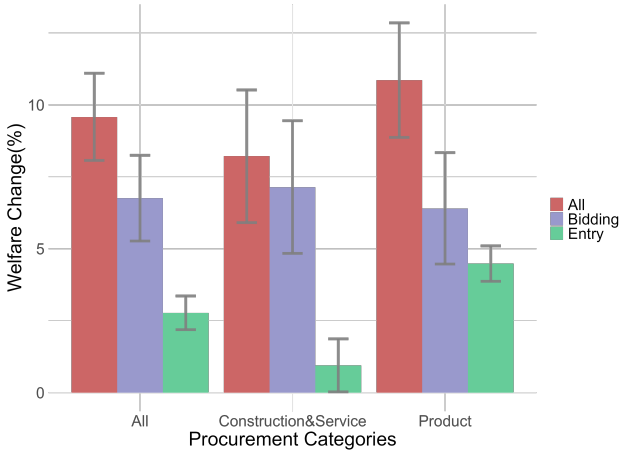
Table 5: Implementing Anonymous Evaluation on RFP

| Counterfactual | ΔPrice | ΔQuality | ΔScore(Welfare) |
|---|---|---|---|
| Competitive Bid | -12.19% | 4.25% | 6.76% |
| | [-13.63%, -10.74%] | [2.27%, 6.23%] | [5.27%, 8.25%] |
| Competitive Bid+Entry | -14.51% | 3.86% | 9.58% |
| | [-16.36%, -12.65%] | [1.84%, 5.88%] | [8.07%, 11.1%] |

**Notes**: In this exercise, the adjusted estimated $\boldsymbol{\theta}$ values are used in the competition model to estimate how predetermined winners bid in a real competitive environment. In the entry exercise, there are two more bidders, the number is estimated from Figure 8. 95% confidence intervals are included in the brackets.

In a counterfactual competitive scenario without entry effects, the predetermined winner tends to submit a price bid that is 12.19% lower than the initial bid, along with a 4.25% increase in quality. This translates to a 6.76% increase in the real final score compared to the original. When accounting for both competitive bidding and entry effects, the price decreases by 14.51% with a 3.86% increase in quality. Increased entries intensify competition over prices, leading to a rise of 9.58% in the final scores. Compared to the scenario without entry effects, 70% of the welfare gains from the introduction of anonymous RFP evaluations stem from removing the unnecessary rules, while the remaining 30% is attributable to more entries, as shown in the first group of bars in Figure 10.

Figure 10: Decomposition of Welfare Changes by Procurement Categories




**Notes**: The first bar in each group represents the score changes when there is an increase in the number of bidders coupled with competitive bidding. The second bar indicates the score changes when the number of bidders remains constant but the bidding becomes competitive and noncorrupt. The last bar in each group shows the difference between the first two bars. All bars are with 95% confidence intervals.

As shown in Table A7 and Figure 10, when segmented by procurement categories, product procurement demonstrates a greater price reduction compared to construction or service procurement. This trend is primarily driven by the higher price weights assigned to product procurement. Additionally, entry effects contribute more to the welfare gains in product procurement than in service or construction procurement. Overall, product procurement benefits more from the combined effects of increased entry and competitive bidding.

Implementing an anonymous evaluation of RFPs offers the dual benefit of retaining a level of discretion while eradicating superfluous scoring rules. In practical application, the selection of experts becomes important. These experts should not only possess the requisite knowledge but also be motivated to provide truthful insights.

### 8.4 Counterfactual: Increasing Price Weights

As highlighted in the background about scoring auctions, public procurement law establishes guidelines on choosing price weights across various procurement categories. This is intended to curb corruption by not emphasizing the quality elements of the auctions too much. Nonetheless, officers frequently choose the lowest available option. Moreover, choosing a price weight above the lowest bound serves as a robust indicator of non-corruption. In this subsection, I explore the counterfactual scenario of elevating price weights. Here, the number of bidders and other variables, are retained as constants, with adjustments being made exclusively to the price weights.

Table 6 reports the results of increasing price weights by 10 percent points. The first

Table 6: Increasing Price Weights by 10pp

| Counterfactual | ΔPrice | ΔQuality | ΔScore(Welfare) |
|---|---|---|---|
| Noncorrupt | -2.88% | -6.51% | -3.88% |
| | [-4.79%, -0.96%] | [-6.79%, -6.24%] | [-4.24%, -3.51%] |
| Corrupt | -8.17% | -1.53% | 0.92% |
| | [-9.68%, -6.67%] | [-3.52%, 0.46%] | [-0.56%, 2.4%] |
| Overall | -6.59% | -3.24% | -0.67% |
| | [-7.82%, -5.37%] | [-4.58%, -1.9%] | [-1.69%, 0.35%] |

**Notes**: In this exercise, I examine the change in bid behaviors by increasing the price weights while holding other conditions constant. For predetermined winners in corrupt cases, I utilize the adjusted $\theta$ and conditions from the corruption model, assuming all previous bidders were non-fake bidders. The results are presented with 95% confidence intervals in brackets.

column presents the average change in the price bid of the winner. The second column shows the average change in the quality bid of the winner. In the final column, I report the mean change in the final score, representing social welfare as per Asker and Cantillon (2008) and Hanazano et al. (2020). These values are computed using the original scoring rules.

It is not surprising that when procurement officers place more emphasis on price, bidders attempt to lower their bid prices. With a 10-point increase in the weight assigned to price, corrupt winners reduce their bids by approximately 8% to gain an advantage in the price score. Competitive winners also lower their bids, but the reduction is much smaller than in corrupt cases. Regarding quality bids, predetermined winners in corrupt auctions do not make significant adjustments to their actual quality offers, likely due to their advantage under manipulated scoring rules. Their quality tends to be considerably lower than that of genuine bidders, allowing them to reduce prices without compromising much on quality. Meanwhile, in competitive cases, with a lower weight on quality, bidders invest less in quality, leading to a 6% decrease in their quality scores.

From a social welfare perspective, increasing the weight on price does not necessarily result in statistically or economically significant improvements. The outcome largely depends on the level of corruption and the extent of the increase in price weight. In non-corrupt auctions, any welfare gains from lower prices are often offset by a corresponding drop in quality, potentially resulting in a net decrease in overall welfare. The scenario with a 20-point increase in the price weight is shown in Table A8.

# 9 Conclusion

In this paper, I developed statistical tools based on the standard scoring auction model to identify and quantify corruption in scoring auctions, where corrupt procurement officers

manipulate scoring rules to favor predetermined winning firms. The results from the model-based tests show that approximately 65% of the procurement auctions reject the null of no corruption. To further validate my model's predictions, I conducted an expert survey to evaluate the call-for-tender files and identify instances of scoring rule manipulation. By comparing the survey results with my model's predictions, I found that the model achieves a high accuracy rate of 81%.

To study the distortions introduced by corruption, I merged firm administrative data with public procurement data. The results indicate that firms with lower efficiency, smaller sizes, and stronger connections to state-owned entities are more likely to be corrupt winners. Which policies can help mitigate corruption in public procurement? I explored the exogenous shocks resulting from corruption investigations during China's most extensive Anti-corruption campaign. I found that these corruption investigations only had a temporary impact on enhancing competition and reducing corruption, lacking a lasting impact. Moreover, most of these effects can be attributed to the involvement of officials from higher bureaucratic levels in the investigations.

In subsequent analyses, I parametrically estimated firm types from the scoring auction model and conducted a counterfactual analysis. In a counterfactual scenario with anonymous evaluations of RFPs, social welfare increases by 11%. Of this increase, 70% arises from compelling firms to bid competitively, while 30% stems from increased participation. However, increasing price weights doesn't lead to a significant improvement in social welfare.

Looking forward, there are several paths for future research. Firstly, researchers can focus on specific industries with direct and measurable quality metrics, such as the life-cycle cost of products or service satisfaction, to evaluate the real-world impact on social welfare. Additionally, studies could assess the effectiveness of related public procurement policies, such as set-aside programs and green energy innovation promotion initiatives. Furthermore, more rigorous and experimental studies are needed to examine bidder responses to the policy experiments discussed in this paper and to uncover potential adaptive strategies that bidders might use to evade detection.

# References


**Andreyanov, Pasha.** 2018. "Mechanism choice in scoring auctions." *Working Paper*.

**Andreyanov, Pasha, Alec Davidson, and Vasily Korovkin.** 2017. "Detecting auctioneer corruption: Evidence from Russian procurement auctions." *Working Paper*.

**Asker, John, and Estelle Cantillon.** 2008. "Properties of scoring auctions." *The RAND Journal of Economics*, 39(1): 69–85.

**Athey, Susan, and Philip A Haile.** 2002. "Identification of standard auction models." *Econometrica*, 70(6): 2107–2140.

**Avis, Eric, Claudio Ferraz, and Frederico Finan.** 2018. "Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians." *Journal of Political Economy*, 126(5): 1912–1964.

**Bai, Chong-En, Ruixue Jia, Hongbin Li, and Xin Wang.** 2021. "Entrepreneurial Reluctance: Talent and Firm Creation in China." National Bureau of Economic Research.

**Bajari, Patrick, Stephanie Houghton, and Steven Tadelis.** 2014. "Bidding for incomplete contracts: An empirical analysis of adaptation costs." *American Economic Review*, 104(4): 1288–1319.

**Baltrunaite, Audinga.** 2020. "Political contributions and public procurement: evidence from Lithuania." *Journal of the European Economic Association*, 18(2): 541–582.

**Baltrunaite, Audinga, Cristina Giorgiantonio, Sauro Mocetti, and Tommaso Orlando.** 2021. "Discretion and supplier selection in public procurement." *The Journal of Law, Economics, and Organization*, 37(1): 134–166.

**Brugués, Felipe, Javier Brugués, and Samuele Giambra.** 2024. "Political connections and misallocation of procurement contracts: evidence from Ecuador." *Journal of Development Economics*, 170: 103296.

**Burguet, Roberto, and Yeon-Koo Che.** 2004. "Competitive procurement with corruption." *RAND Journal of Economics*, 50–68.

**Cai, Hongbin, J Vernon Henderson, and Qinghua Zhang.** 2013. "China's land market auctions: evidence of corruption?" *The Rand journal of economics*, 44(3): 488–521.

**Calvo, Eduard, Ruomeng Cui, and Juan Camilo Serpa.** 2019. "Oversight and efficiency in public projects: A regression discontinuity analysis." *Management Science*, 65(12): 5651–5675.

**Cao, Yiming.** 2022. "Institutional Weakness and Societal Vulnerability: Evidence from the Sichuan Earthquake."

**Celebi, M Emre, Hassan A Kingravi, and Patricio A Vela.** 2013. "A comparative study of efficient initialization methods for the k-means clustering algorithm." *Expert systems with applications*, 40(1): 200–210.

**Chassang, Sylvain, Kei Kawai, Jun Nakabayashi, and Juan Ortner.** 2022. "Robust screens for noncompetitive bidding in procurement auctions." *Econometrica*, 90(1): 315–346.

**Chen, Qianmiao, Qingyang Huang, Chang Liu, and Peng Wang.** 2022. "Career Incentives of Local Leaders and Crisis Response: A Case Study of COVID-19 Lockdowns in China." *European Journal of Political Economy*, 75: 102180.

**Chen-Ritzo, Ching-Hua, Terry P Harrison, Anthony M Kwasnica, and Douglas J Thomas.** 2005. "Better, faster, cheaper: An experimental analysis of a multi-attribute reverse auction mechanism with restricted information feedback." *Management science*, 51(12): 1753–1762.

**Chen, Ting, and James Kai-sing Kung.** 2019. "Busting the "Princelings": The campaign against corruption in China's primary land market." *The Quarterly Journal of Economics*, 134(1): 185–226.

**Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu.** 2021. "Notching R&D investment with corporate income tax cuts in China." *American Economic Review*, 111(7): 2065–2100.

**Che, Yeon-Koo.** 1993. "Design competition through multidimensional auctions." *The RAND Journal of Economics*, 668–680.

**Claudio, Ferraz, Finan Frederico, and Szerman Dimitri.** 2015. "Procuring Firm Growth: The Effects of Government Purchases on Firm Dynamics."

**Colonnelli, Emanuele, and Mounu Prem.** 2021. "Corruption and Firms." *The Review of Economic Studies*, 89(2): 695–732.

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

**Coviello, Decio, and Mario Mariniello.** 2014. "Publicity requirements in public procurement: Evidence from a regression discontinuity design." *Journal of Public Economics*, 109: 76–100.

**Coviello, Decio, Andrea Guglielmo, and Giancarlo Spagnolo.** 2018. "The effect of discretion on procurement performance." *Management Science*, 64(2): 715–738.

**Decarolis, Francesco, Raymond Fisman, Paolo Pinotti, and Silvia Vannutelli.** 2020. "Rules, discretion, and corruption in procurement: Evidence from Italian government contracting." National Bureau of Economic Research.

**De Winter, Joost CF.** 2019. "Using the Student's t-test with extremely small sample sizes." *Practical Assessment, Research, and Evaluation*, 18(1): 10.

**Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan.** 2018. "The value of regulatory discretion: Estimates from environmental inspections in India." *Econometrica*, 86(6): 2123–2160.

**Fang, Hanming.** 2023. "Measurements, Determinants and Causes of Corruption: Lessons from China's Anti-Corruption Campaign."

**Fay, Michael P, and Michael A Proschan.** 2010. "Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules." *Statistics surveys*, 4: 1.

**Ferraz, Claudio, and Frederico Finan.** 2011. "Electoral accountability and corruption: Evidence from the audits of local governments." *American Economic Review*, 101(4): 1274–1311.

**Gan, Guojun, and Michael Kwok-Po Ng.** 2017. "K-means clustering with outlier removal." *Pattern Recognition Letters*, 90: 8–14.

**Gerardino, Maria Paula, Stephan Litschig, and Dina Pomeranz.** 2024. "Distortion by audit: Evidence from public procurement." *American Economic Journal: Applied Economics*, 16(4): 71–108.

**Gilraine, Michael, Jiaying Gu, and Robert McMillan.** 2020. "A new method for estimating teacher value-added." National Bureau of Economic Research.

**Gong, Ting, and Na Zhou.** 2015. "Corruption and marketization: Formal and informal rules in Chinese public procurement." *Regulation & Governance*, 9(1): 63–76.

**Gordon, A.D.** 1999. *Classification, 2nd Edition. Chapman & Hall/CRC Monographs on Statistics & Applied Probability*, CRC Press.

**Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong.** 2000. "Optimal nonparametric estimation of first-price auctions." *Econometrica*, 68(3): 525–574.

**Gu, Jiaying, Roger Koenker, and Stanislav Volgushev.** 2018. "Testing for homogeneity in mixture models." *Econometric Theory*, 34(4): 850–895.

**Hanazano, Makoto, Yohsuke Hiroze, Jun Nakabayashi, and Masanori Tsuruoka.** 2020. "Theory, Identification, and Estimation for Scoring Auctions." *Working Paper*.

**Hanazono, Makoto, Jun Nakabayashi, Masanori Tsuruoka, et al.** 2013. "Procurement auctions with general price-quality evaluation." *Working Paper*.

**Hartigan, John A, and Manchek A Wong.** 1979. "Algorithm AS 136: A k-means clustering algorithm." *Journal of the royal statistical society. series c (applied statistics)*, 28(1): 100–108.

**Hautamäki, Ville, Svetlana Cherednichenko, Ismo Kärkkäinen, Tomi Kinnunen, and Pasi Fränti.** 2005. "Improving k-means by outlier removal." 978–987, Springer.

**Hodges Jr, Joseph L, and Erich L Lehmann.** 1956. "The efficiency of some nonparametric competitors of the t-test." *The Annals of Mathematical Statistics*, 324–335.

**Hsieh, Chang-Tai, and Peter J. Klenow.** 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics*, 124(4): 1403–1448.

**Huang, Yangguang.** 2019. "An empirical study of scoring auctions and quality manipulation corruption." *European Economic Review*, 120: 103322.

**Huang, Yangguang, and Jijun Xia.** 2019. "Procurement auctions under quality manipulation corruption." *European Economic Review*, 111: 380–399.

**Jiang, Wenhua, and Cun-Hui Zhang.** 2009. "General maximum likelihood empirical Bayes estimation of normal means."

**Kang, Karam, and Robert A Miller.** 2022. "Winning by Default: Why is There So Little Competition in Government Procurement?" *The Review of Economic Studies*, 89(3): 1495–1556.

**Kawai, Kei, and Jun Nakabayashi.** 2022. "Detecting large-scale collusion in procurement auctions." *Journal of Political Economy*, 130(5): 1364–1411.

**Kawai, Kei, Jun Nakabayashi, and Daichi Shimamoto.** 2022. "A Study of Bid-rigging in Procurement Auctions: Evidence from Indonesia, Georgia, Mongolia, Malta, and State of California." National Bureau of Economic Research Working Paper 30271.

**Kawai, Kei, Jun Nakabayashi, and Juan M Ortner.** 2021. "The value of privacy in cartels: an analysis of the inner workings of a bidding ring." National Bureau of Economic Research.

**Kawai, Kei, Jun Nakabayashi, Juan Ortner, and Sylvain Chassang.** 2023. "Using bid rotation and incumbency to detect collusion: A regression discontinuity approach." *The Review of Economic Studies*, 90(1): 376–403.

**Khan, Adnan Q, Asim I Khwaja, and Benjamin A Olken.** 2016. "Tax farming redux: Experimental evidence on performance pay for tax collectors." *The Quarterly Journal of Economics*, 131(1): 219–271.

**Kiefer, Jack, and Jacob Wolfowitz.** 1956. "Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters." *The Annals of Mathematical Statistics*, 887–906.

**Koenker, Roger, and Ivan Mizera.** 2014. "Convex optimization, shape constraints, compound decisions, and empirical Bayes rules." *Journal of the American Statistical Association*, 109(506): 674–685.

**Kong, Yunmi, Isabelle Perrigne, and Quang Vuong.** 2022. "Multidimensional auctions of contracts: An empirical analysis." *American Economic Review*, 112(5): 1703–1736.

**Lichand, Guilherme, and Gustavo Fernandes.** 2019. "The dark side of the contract: do government audits reduce corruption in the presence of displacement by vendors?" *University of Zurich, Switzerland, mimeo.*

**Lichand, Guilherme, Marcos FM Lopes, and Marcelo C Medeiros.** 2016. "Is corruption good for your health." *Job Market Paper.*

**Li, Tong, Isabelle Perrigne, and Quang Vuong.** 2002. "Structural estimation of the affiliated private value auction model." *RAND Journal of Economics*, 171–193.

**Manion, Melanie.** 2016. "Taking China's anticorruption campaign seriously." *Economic and Political Studies*, 4(1): 3–18.

**Motta, Massimo.** 2004. *Competition policy: theory and practice.* Cambridge university press.

**OECD.** 2016. "Preventing Corruption in Public Procurement." https://www.oecd.org/gov/ethics/Corruption-Public-Procurement-Brochure.pdf.

**Olken, Benjamin A.** 2007. "Monitoring corruption: evidence from a field experiment in Indonesia." *Journal of Political Economy*, 115(2): 200–249.

**Palguta, Ján, and Filip Pertold.** 2017. "Manipulation of procurement contracts: Evidence from the introduction of discretionary thresholds." *American Economic Journal: Economic Policy*, 9(2): 293–315.

**Parrinello, Christina M, Morgan E Grams, Yingying Sang, David Couper, Lisa M Wruck, Danni Li, John H Eckfeldt, Elizabeth Selvin, and Josef Coresh.** 2016. "Iterative outlier removal: a method for identifying outliers in laboratory recalibration studies." *Clinical chemistry*, 62(7): 966–972.

**Porter, Robert H.** 2005. "Detecting collusion." *Review of Industrial Organization*, 26: 147–167.

**Porter, Robert H, and J Douglas Zona.** 1993. "Detection of bid rigging in procurement auctions." *Journal of Political economy*, 101(3): 518–538.

**Ryan, Nicholas.** 2020. "Contract enforcement and productive efficiency: Evidence from the bidding and renegotiation of power contracts in India." *Econometrica*, 88(2): 383–424.

**Scherer, Frederic M, and David Ross.** 1990. "Industrial market structure and economic performance." *Working Paper*.

**Schurter, Karl.** 2017. "Identification and inference in first-price auctions with collusion." *Working Paper*.

**Szucs, Ferenc.** 2024. "Discretion and favoritism in public procurement." *Journal of the European Economic Association*, 22(1): 117–160.

**Takahashi, Hidenori.** 2018. "Strategic design under uncertain evaluations: structural analysis of design-build auctions." *The RAND Journal of Economics*, 49(3): 594–618.

**UNCITRAL.** 2014. "UNCITRAL model law on public procurement."

**Wang, Yuhua, and Bruce J Dickson.** 2022. "How corruption investigations undermine regime support: evidence from China." *Political Science Research and Methods*, 10(1): 33–48.

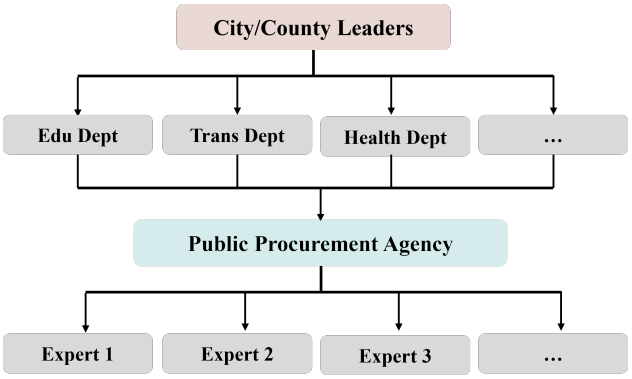
**Wu, Xindong, Vipin Kumar, J Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J McLachlan, Angus Ng, Bing Liu, Philip S Yu, et al.** 2008. "Top 10 algorithms in data mining." *Knowledge and information systems*, 14: 1–37.

**Yang, Jiawei, Susanto Rahardja, and Pasi Fränti.** 2021. "Mean-shift outlier detection and filtering." *Pattern Recognition*, 115: 107874.

**Zamboni, Yves, and Stephan Litschig.** 2018. "Audit risk and rent extraction: Evidence from a randomized evaluation in Brazil." *Journal of Development Economics*, 134: 133–149.

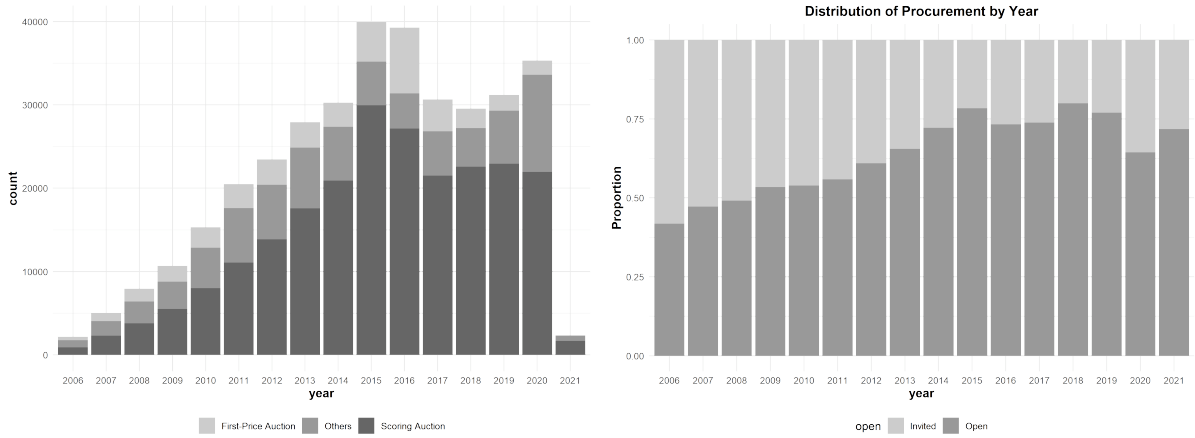
# Appendix A: Supplementary Tables and Figures

Figure A1: Different Roles Involved Public Procurement




**Notes**: Government departments primarily initiate public procurement for public projects. In autocratic regimes, lower-level government agencies are accountable to higher-level authorities rather than to the voters.

Figure A2: The Distribution of Public Procurement by Year

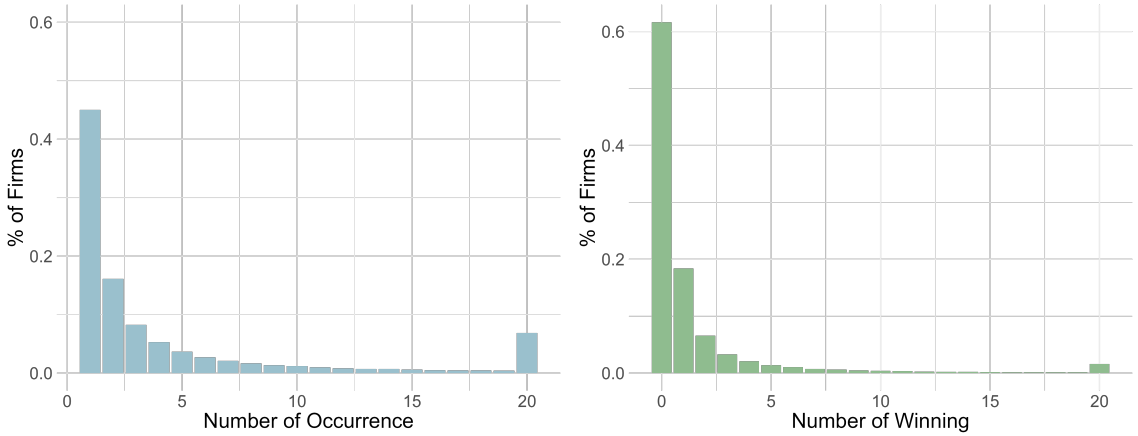



(a) All Procurement

(b) Open V.S. Invited Auctions

**Notes**: (a) presents the number of procurement projects in the dataset spanning from 2006 to 2021, categorized by various procurement auction methods such as first-price auction, scoring auction, and other methods. (b) illustrates the proportion of procurement projects utilizing an open auction format.

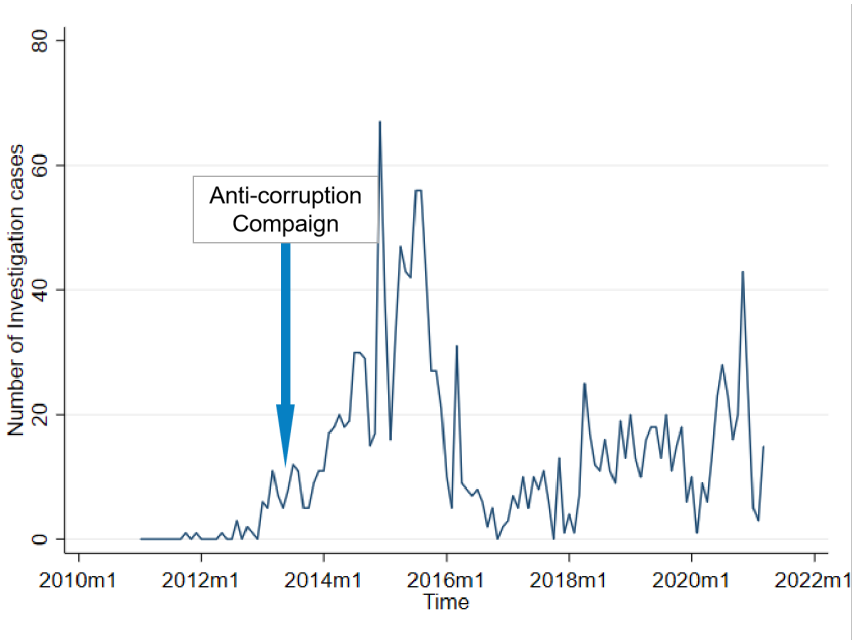
Figure A3: How frequently do firms participate in public procurement




(a) Distribution of Participation Times (b) Distribution of Winning Times

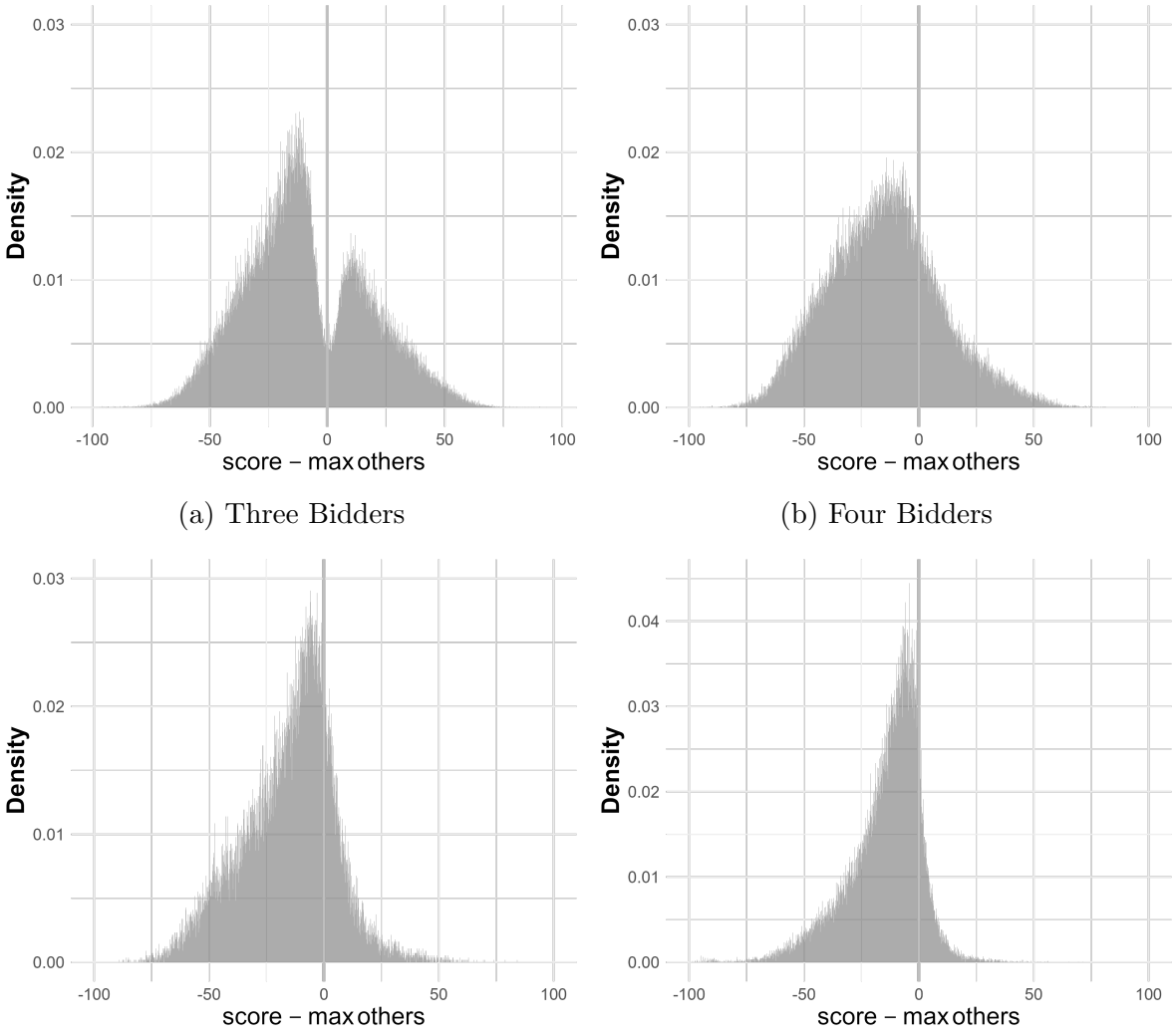
**Notes**: (1) displays the distribution of the number of firms participating in procurement project competitions. (2) illustrates the distribution of the number of procurement contracts firms secure. These graphs encompass all procurement projects, extending beyond open-scoring auctions to incorporate those from non-open procurement competitions.

Figure A4: Number of Corruption Investigation Cases




**Notes**: The y-axis of the figure represents the total number of anti-corruption investigations per month in the Province from which the public procurement data was collected. The investigation data prior to August 2016 are sourced from Wang and Dickson (2022), while the data from August 2016 onwards have been collected manually by me, encompassing all levels of Commissions for Discipline Inspection.

Figure A5: Distribution of bid-differences over (bidder, auction) pairs




(a) Three Bidders

(b) Four Bidders

(c) Five Bidders

(d) Six and More Bidders

**Notes**: The X-axis represents the score gap between a bidder's own score and the highest score among all bidders, excluding the bidder's own score.

Table A1: A Four-Bidder Case

| Company | BusiTech Score (65) | Price Score(35) | Final Score | Order |
|---|---|---|---|---|
| A | 63.6 | 34.65 | 98.25 | 1 |
| D | 26.2 | 35 | 61.2 | 3 |
| E | 28.2 | 34.53 | 62.73 | 2 |
| C | 22.6 | 34.56 | 57.16 | 4 |

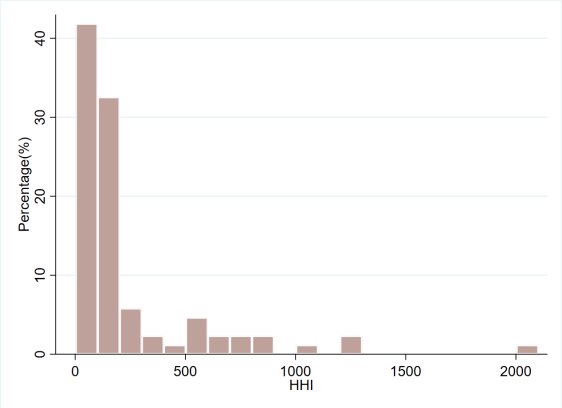
**Notes**: The outcomes of this procurement auction are shown on the public procurement website.

Table A2: Corruption Proportion Estimation and Auction-level Test Results

| | All | 3 Bidders | 4 Bidders | 5 Bidders |
|---|---|---|---|---|
| Corruption Proportion Estimation | | | | |
| Corruption Proportion | 60.57% | 70.88% | 46.49% | 26.85% |
| | [60.19, 60.96] | [70.21, 71.55] | [45.95, 47.03] | [24.77, 28.93] |
| Auction Level Test $\alpha = 0.05$ | | | | |
| Unadjusted P-value $\alpha = 0.05$ | 54.43% | 59.95% | 40.7% | 34.45% |
| | [54.37, 54.49] | [59.84, 60.05] | [40.58, 40.82] | [34.29, 34.61] |
| Adjusted P-value $\alpha = 0.05$ | 45.23% | 50.06% | 33.71% | 27.18% |
| | [45.19, 45.27] | [49.99, 50.13] | [33.61, 33.82] | [27.03, 27.33] |

**Notes**: 95% confidence intervals from bootstrapping are indicated in brackets. The first two rows of the auction-level test results represent the results using unadjusted p-values, while the following two rows display the adjusted p-values to control the false discovery rate. Unadjusted p-values are more powerful than adjusted p-values but have a higher type I error rate.

Figure A6: Distribution of HHI for Industries




**Notes**: HHIs are derived from the 2016 Firm Tax data. Typically, an HHI exceeding 2500 signifies a highly concentrated market.

Table A3: Bidding Results

| Company | Tech Score (50) | Busi Score (20) | Price Score(30) | Final Score | Order |
|---|---|---|---|---|---|
| M | 49.86 | 20 | 27.13 | 96.99 | 1 |
| N | 21.57 | 15.5 | 30 | 67.07 | 2 |
| O | 11.43 | 9 | 27.06 | 47.49 | 3 |

**Notes**: The outcomes of this procurement auction are shown on the public procurement website.

Table A4: Expert Survey and Model Prediction

| Accuracy Rate=79.4% | Total=500 | Model Prediction | | |
|---|---|---|---|---|
| | | Corrupt (PP) | Not Corrupt (PN) | |
| Expert Survey | Corrupt (P) | 300 | 75 False N | Test Power 80% |
| | Not Corrupt (N) | 28 False P | 97 | Test Size 22.4% |

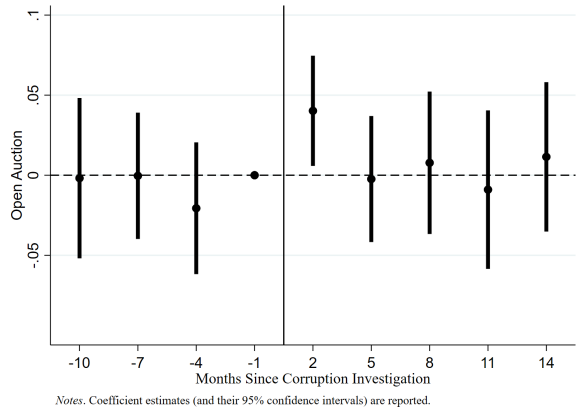
**Notes**: For the expert survey, if any evaluation survey reveals highlighted scoring rules accompanied by expert explanations and concerns, the procurement project is labeled as corrupt. The classification of model prediction uses adjusted p-values with a significance level of 0.1.

Table A5: Firm Characteristics and $\theta$ of Bidders

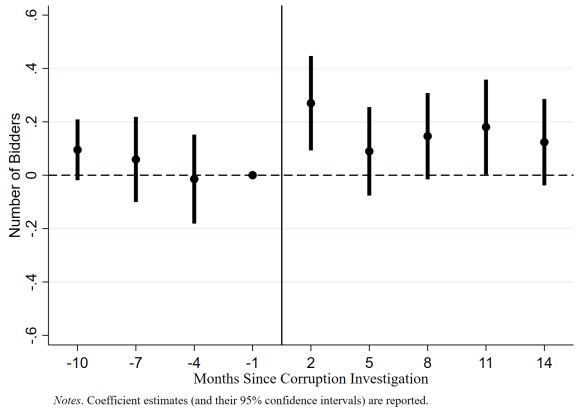
| | Winners: Noncorrupt | | Winners: Corrupt | |
|---|---|---|---|---|
| | $\theta_0$ | $\theta_1$ | $\theta_0$ | $\theta_1$ |
| Log(capital) | 0.0012 | 0.0004 | -0.0074*** | -0.0014 |
| | (0.0025) | (0.0009) | (0.0023) | (0.0008) |
| Log(formal employee) | -0.0132*** | -0.0052*** | -0.0124*** | -0.0047*** |
| | (0.0019) | (0.0006) | (0.0013) | (0.0004) |
| Log(#members) | -0.0023 | -0.0008 | 0.0014 | -0.0003 |
| | (0.0018) | (0.0006) | (0.0017) | (0.0004) |
| State-owned | 0.0095 | 0.0131 | -0.0104 | 0.0015 |
| | (0.0252) | (0.0090) | (0.0210) | (0.0109) |
| Log(budget) | 0.0093** | -0.0045*** | 0.0082** | -0.0056*** |
| | (0.0040) | (0.0013) | (0.0037) | (0.0007) |
| Auction Characteristics | Y | Y | Y | Y |
| $N$ | 30651 | 30651 | 67533 | 67533 |

**Notes**: The first two columns present the estimated $\theta_0$ and $\theta_1$ for winners in non-corrupt auctions, while the third and fourth columns display the estimated $\theta_0$ and $\theta_1$ for winners in corrupt auctions. All regressions control for category fixed effects, number of bidders fixed effects, price weights fixed effects, fixed quality weights fixed effects, procurement year, and firm entry year. Standard errors are clustered at the procurement category level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

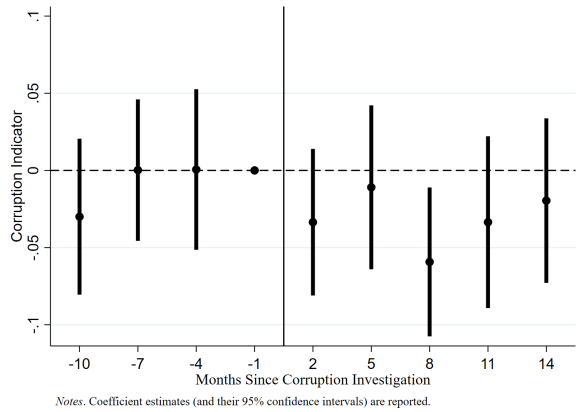
Figure A7: The Effect of Corruption Investigation



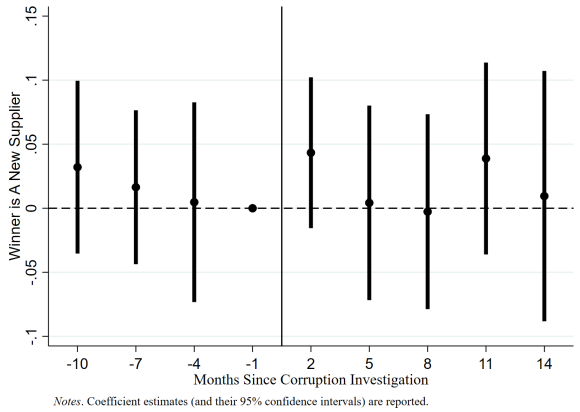

(a) Open Auction



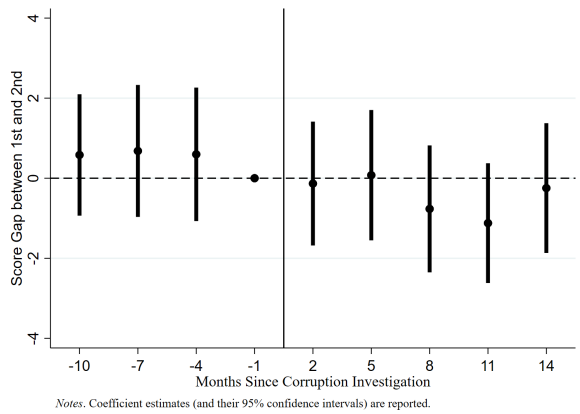

(b) Number of Bidders



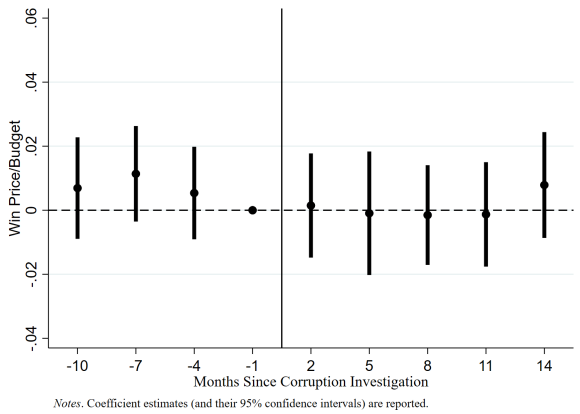

(c) Corruption Indicator




(d) New Winning Firm



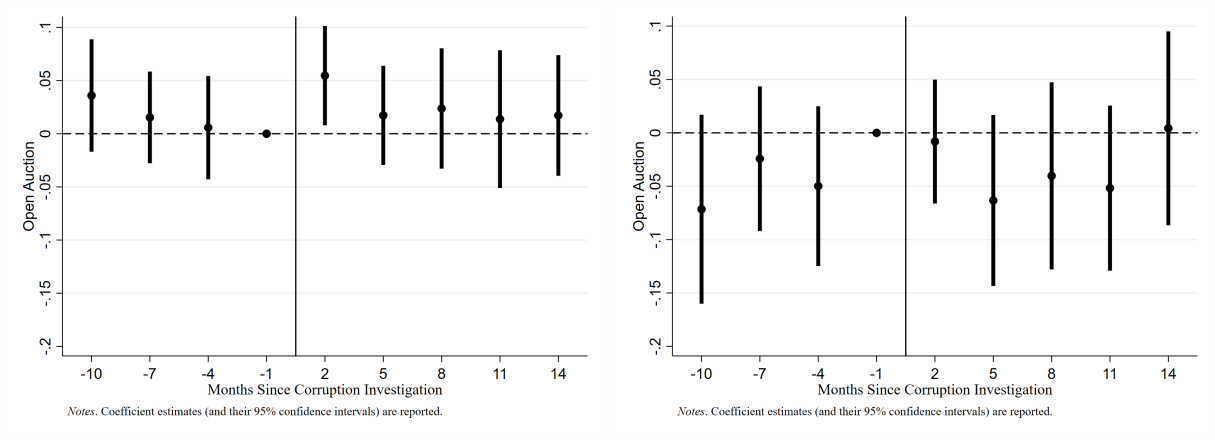

(e) Score Gaps




(f) Final Price/Budget

**Notes**: The regression model incorporates time-fixed effects and procurement government department-fixed effects. Standard errors are clustered at the department level.

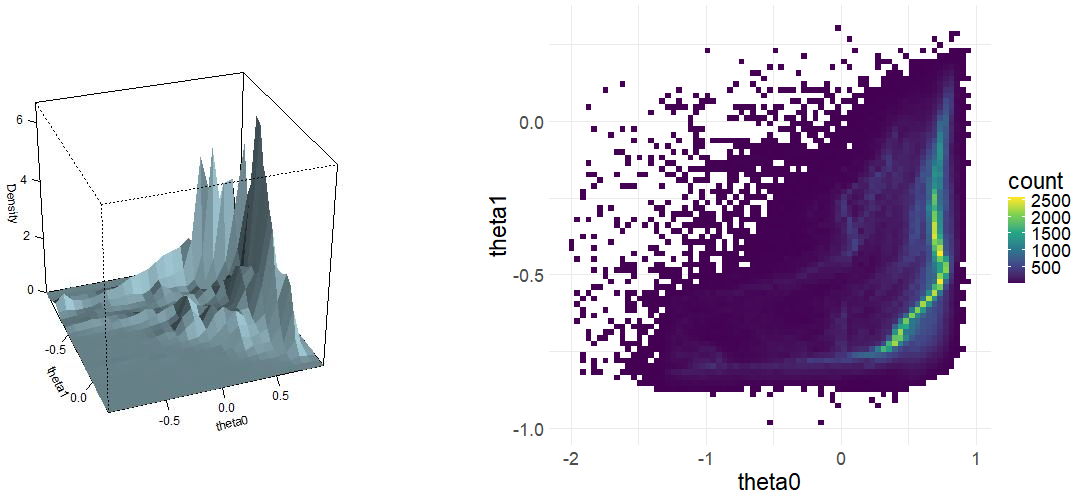
Figure A8: The Effect of Corruption Investigation by Officials with Different Levels




(a) Open Auction: High-level Officials (b) Open Auction: Low-level Officials

**Notes**: The regression model incorporates time-fixed effects and procurement government department fixed effects. Standard errors are clustered at the department level.

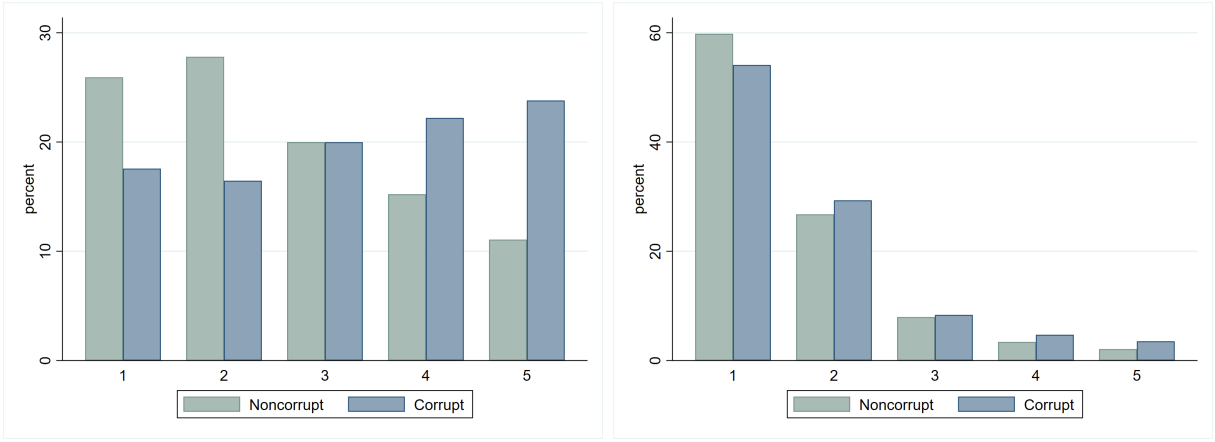
Figure A9: $\boldsymbol{\theta}$ Distribution




(a) 3D Density distribution of $\theta_0$ and $\theta_1$ (b) 2D Density distribution of $\theta_0$ and $\theta_1$

**Notes**: (a) shows the 3D density distribution of $(\theta_0, \theta_1)$. (b) shows the 2D density distribution. Lighter colors are associated with larger density.

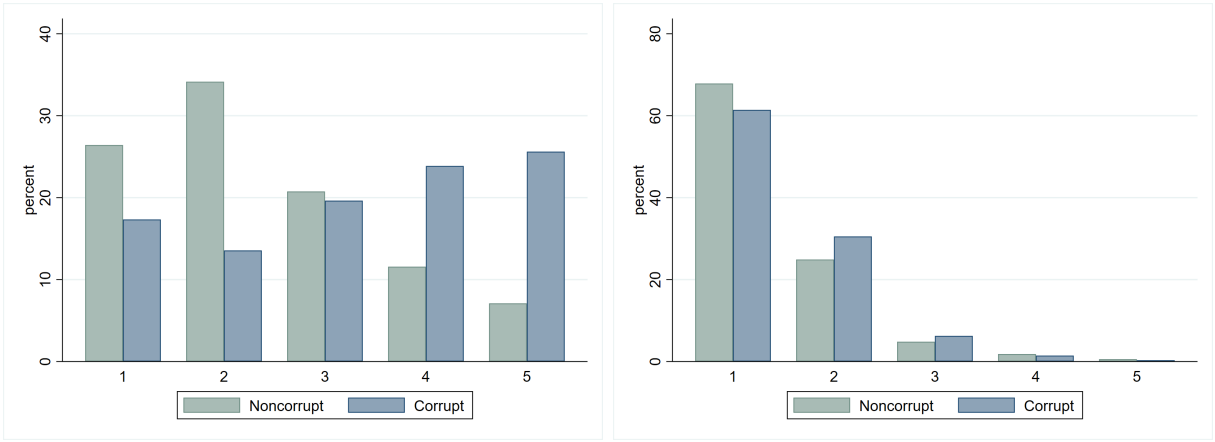
Figure A10: $\theta_0$ Quintile Distribution




(a) All Bidders (b) Winners

**Notes**: I rank the estimated $\theta_0$ within different groups based on procurement categories, number of bidders, price weights, and fixed quality weights. Lower ranking numbers indicate higher efficiency. (a) shows the ranking distribution of all bidders, categorized by corrupt and non-corrupt cases. (b) displays the ranking distribution of winners in procurement auctions.

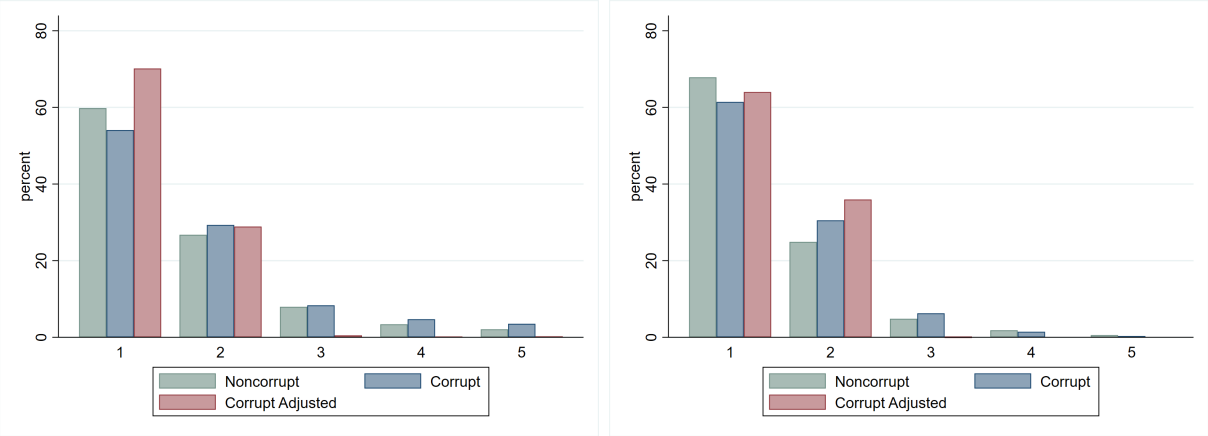
Figure A11: $\theta_1$ Quintile Distribution




(a) All Bidders (b) Winners

**Notes**: I rank the estimated $\theta_1$ within different groups based on procurement categories, number of bidders, price weights, and fixed quality weights. Lower ranking numbers indicate higher efficiency. (a) shows the ranking distribution of all bidders, categorized by corrupt and non-corrupt cases. (b) displays the ranking distribution of winners in procurement auctions.

Figure A12: $\theta$ Quintile Distribution for Winners after the Adjustment




(a) $\theta_0$ (b) $\theta_1$

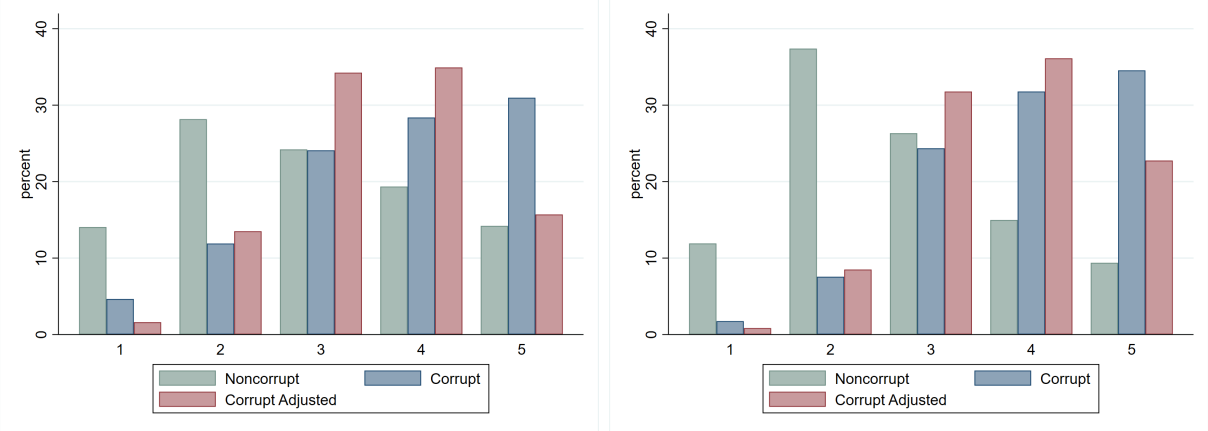
**Notes**: I rank the estimated $\theta_0$ and $\theta_1$ within different groups based on procurement categories, number of bidders, price weights, and fixed quality weights. Lower ranking numbers indicate higher efficiency. (a) shows the ranking of $\theta_0$. (b) shows the ranking of $\theta_1$.

Table A6: Firm Characteristics and $\theta$ of Bidders

| | Losers: Noncorrupt | | Losers: Corrupt | |
|---|---|---|---|---|
| | $\theta_0$ | $\theta_1$ | $\theta_0$ | $\theta_1$ |
| Log(capital) | -0.0053*** | -0.0082*** | -0.0027** | -0.0033* |
| | (0.0017) | (0.0016) | (0.0010) | (0.0017) |
| Log(formal employee) | -0.0074*** | -0.0070*** | -0.0023*** | -0.0053*** |
| | (0.0017) | (0.0014) | (0.0008) | (0.0014) |
| Log(#members) | -0.0040*** | -0.0039*** | -0.0032*** | -0.0048*** |
| | (0.0013) | (0.0013) | (0.0009) | (0.0011) |
| State-owned | -0.0217 | 0.0178* | -0.0130 | -0.0083 |
| | (0.0230) | (0.0100) | (0.0112) | (0.0086) |
| Log(budget) | 0.0075*** | 0.0020 | 0.0037* | 0.0039 |
| | (0.0020) | (0.0021) | (0.0022) | (0.0031) |
| Auction Characteristics | Y | Y | Y | Y |
| $N$ | 88736 | 88736 | 195349 | 195349 |

**Notes**: The first two columns present the estimated $\theta_0$ and $\theta_1$ for losers in non-corrupt auctions, while the third and fourth columns display the estimated $\theta_0$ and $\theta_1$ for losers in corrupt auctions. All regressions control for category fixed effects, number of bidders fixed effects, price weights fixed effects, fixed quality weights fixed effects, procurement year, and firm entry year. Standard errors are clustered at the procurement category level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A13: $\theta$ Quintile Distribution for Losers




(a) $\theta_0$ (b) $\theta_1$

Table A7: Implementing Anonymous Evaluation on RFP

| Counterfactual | ΔPrice | ΔQuality | ΔScore(Welfare) |
|---|---|---|---|
| **Product Procurement** | | | |
| Competitive Bid | -10.3% | 3.81% | 6.4% |
| | [-11.94%, -8.67%] | [1.01%, 6.62%] | [4.47%, 8.34%] |
| Competitive Bid+Entry | -15.31% | 6.35% | 11.23% |
| | [-17.66%, -12.96%] | [1.1%, 11.59%] | [7.24%, 15.22%] |
| **Construction and Service Procurement** | | | |
| Competitive Bid | -14.11% | 4.7% | 7.14% |
| | [-16.5%, -11.73%] | [1.86%, 7.54%] | [4.84%, 9.45%] |
| Competitive Bid+Entry | -4.25% | 7.39% | 8.6% |
| | [-6.91%, -1.58%] | [2.17%, 12.61%] | [3.98%, 13.23%] |

**Notes**: In this exercise, the adjusted estimated $\theta$ values are used in the competition model to estimate how predetermined winners bid in a real competitive environment. In the entry exercise, there are two more bidders, the number is estimated from Figure 8. 95% confidence intervals are included in the brackets.

Table A8: Increasing Price Weights by 20pp

| Counterfactual | ΔPrice | ΔQuality | ΔScore(Welfare) |
|---|---|---|---|
| Noncorrupt | -4.63% | -12.33% | -7.29% |
| | [-6.58%, -2.68%] | [-12.71%, -11.96%] | [-7.69%, -6.89%] |
| Corrupt | -9.79% | -7.38% | -2.61% |
| | [-11.17%, -8.41%] | [-9.35%, -5.42%] | [-4.05%, -1.16%] |
| Overall | -8.28% | -9.11% | -4.14% |
| | [-9.45%, -7.12%] | [-10.44%, -7.78%] | [-5.14%, -3.14%] |

**Notes**: In this exercise, I examine the change in bid behaviors by increasing the price weights while holding other conditions constant. For predetermined winners in corrupt cases, I utilize the adjusted $\theta$ and conditions from the corruption model, assuming all previous bidders were non-fake bidders. The results are presented with 95% confidence intervals in brackets.

# Appendix B: Data

## B1 Estimation of TFP

The estimation of total factor productivity (TFP) for firms is based on administrative enterprise income tax records obtained from the Chinese State Administration of Tax (SAT), which is responsible for tax collection and auditing in China. The SAT maintains firm-level records of tax payments and other financial statement information used in tax-related calculations. In this study, I followed the codes and methodology provided by Chen et al. (2021), with the time period extended to 2007-2016 instead of 2008-2011.

They use the structure in the model of constant elasticity demand to write firm value-added $r_{it}$ as:

$$\ln r_{it} = \frac{\theta - 1}{\theta}[\kappa \ln k_{it} + (1-\kappa)\ln l_{it} + \phi_{it}]$$

where $k_{it}$ is the capital input, $l_{it}$ is labor input, and $\phi_{it}$ is firm residual productivity. To get the $\phi_{it}$, I need to know the parameters $\kappa$ and $\theta$. Set the $\theta = 5$ as the benchmark, I have to estimate the $\kappa$. The first-order condition of cost minimization gives the relation as follows:

$$\ln \frac{w l_{it}}{r_{it}} = \ln[(1-\kappa)\frac{\theta}{\theta - 1}] + v_{it}$$

where $w$ is the price of labor, and $v_{it} \sim iid$ and $E[v_{it}] = 0$ is the measurement error in factor prices. $\frac{wl_{it}}{r_{it}}$ is the ratio of labor input to revenue. Then I can get the industry average $\kappa$ from the above equation. Finally, I can calculate the firm TFP:

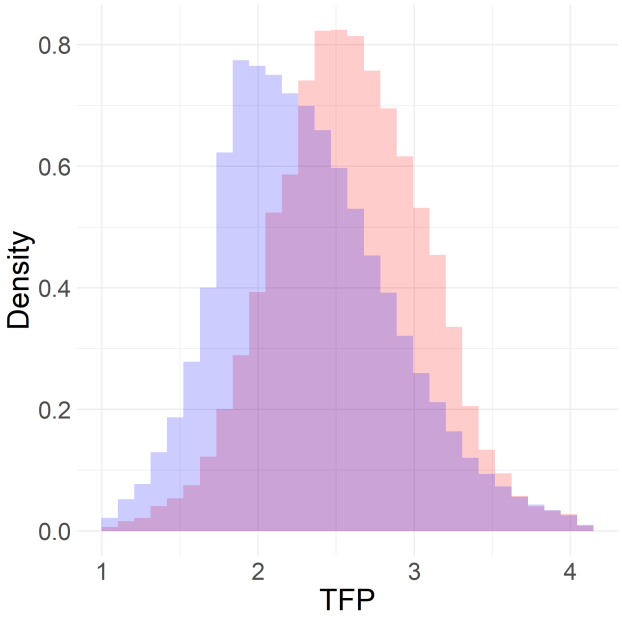
$$\hat{\phi_{it}} = \frac{\theta}{\theta - 1}\ln r_{it} - \kappa \ln k_{it} - (1-\kappa)\ln l_{it}$$

I present the distribution of estimated Total Factor Productivity (TFP) for two sets of firms: the entire dataset from the State Administration of Tax (SAT) and firms that participated in public procurement. Figure B1 displays the distributions, with the blue distribution representing all firms in all available years, and the red distribution representing firms that have engaged in the procurement competition. The comparison of these distributions reveals that firms involved in public procurement tend to exhibit higher levels of TFP, indicating a general superiority in terms of productivity.

## B2 Large Performance Variance in Public Procurement Data

In the public procurement data, many firms perform substantially differently across different procurement auctions. This variability often defies straightforward explanation through differences in procurement requirements or scoring methodologies. While I assess the requisites

Figure B1: The Estimated TFP Distribution




**Notes**: The blue distribution represents the estimated total factor productivity (TFP) using all available data from the SAT dataset across multiple years. On the other hand, the red distribution represents the estimated TFP using only the firms that have matched records in the public procurement data.

stipulated in call-for-tender documents, these divergences don't manifest to the extent seen in bid outcomes.

For instance, consider the case of Wonders Information Corporation, a public sector IT software and service provider with a registered capital of 0.2 billion USD and over 1,300 copyrighted software products. Its expertise in medical information procurement stands out. Foresight Industry Research Institute's study on the medical information system market underscores its low concentration and vibrant competition dynamics, characterized by a "big industry, small enterprises" pattern. In 2020, Wonders Information Corporation held a 2.39% market share, ranking third in the medical information industry.

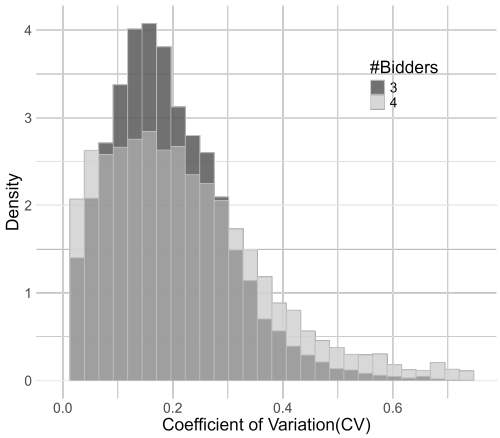
To illustrate, I present three procurement cases involving medical information systems in which Wonders Information Co. participated in Figure B2 (a), revealing its divergent performances. In the first case, Wonders emerges as a dominant winner, notably sharing an identical price bid with Company F. In the second case, Wonders experiences a dramatic loss, with Company I securing a resounding win; both losing entities have similar scores. It's important to note that both these cases align with the large winning margin pattern mentioned earlier. In the final case, Wonders seemingly regains its customary form, placing a reasonable bid despite stiffer competition from Company J, which boasts a higher quality score. Intriguingly, the winner J is a state-owned company, while Company L, the dominant

market player, got a very low score.

Figure B2: Performance Variance of the Same Firm

| Company | price score | quality score | total score | rank |
|---|---|---|---|---|
| Dominate Winner | | | | |
| Wonders | 18.51 | 77.63 | 96.14 | 1 |
| F | 18.51 | 27.07 | 45.58 | 2 |
| G | 20 | 25.32 | 45.32 | 3 |
| Absolute Loser | | | | |
| Wonders | 28.12 | 7.4 | 35.52 | 2 |
| H | 30 | 4.6 | 34.6 | 3 |
| I | 28.52 | 47.3 | 75.82 | 1 |
| Competitor | | | | |
| Wonders | 6.62 | 81.77 | 88.39 | 2 |
| J | 6.64 | 88.7 | 95.34 | 1 |
| K | 10 | 57.13 | 67.13 | 3 |
| L | 6.75 | 52.8 | 59.55 | 4 |

(a) A Case Study: Wonders Co.




(b) Distribution of CV

**Notes**: (a) shows a case study of the firm Wonders Co. participates in different procurement projects. (b) calculates the score standard deviation for each firm that participates in more than one procurement auction.

To quantitatively assess the varying bidding scores of firms across multiple auctions, I calculate the coefficient of variation of scores for each firm. The resulting distributions are presented in Figure B2 (b), differentiating between auctions featuring three bidders and those with four bidders. As depicted, the average coefficient of variation is substantial. A significant number of firms have a coefficient of variation exceeding 0.5, implying that their standard deviation is half the size of the mean score. When comparing the coefficient of variation between three-bidder and four-bidder auctions, it becomes evident that the former tends to exhibit a larger coefficient of variation, signifying greater variability in bidding scores within the auctions featuring three bidders.

The substantial variation in performance cannot be fully explained by the characteristics of the procurement projects. This stylized fact provides compelling evidence that firms may not consistently act as competitors across different auctions. Instead, they appear to assume varying roles, which encompass predetermined corrupt winners, genuine competitors, and even zombie bidders.

### B3 Expert Survey

Here I list the translated survey question design.

1. Based on your experience, what is the normal markup compared to the budget: (please put an estimated number)

2. Are there any unreasonable requirements on the entry threshold?
   **A.** Yes (please highlight them) **B.** No

3. Is the price weight reasonable?
   **A.** Yes **B.** No, too high **C.** No, too low

4. Are there any brand preferences you can tell from the technology score part?
   **A.** Yes (please highlight them) **B.** No

5. Are there any other unnecessary score points set for the technology section?
   **A.** Yes (please highlight them) **B.** No

6. Are there any unnecessary requirements for certificates in the business section?
   **A.** Yes (please highlight them) **B.** No

7. Are there any unnecessary requirements for workers in the business section?
   **A.** Yes (please highlight them) **B.** No

8. Are there any unnecessary requirements for previous experiences in the business section?
   **A.** Yes (please highlight them) **B.** No

9. Are there any other unnecessary requirements in the business section?
   **A.** Yes (please highlight them) **B.** No

10. Are there any other unnecessary subjective scoring rules?
    **A.** Yes (please highlight them) **B.** No

11. What is the level of requirement compared to the scoring rules? (5 is just right, 0 is the lowest, 10 is the highest) **0-10**

12. How many companies are competitive under the scoring rules?
    **A.** Less than 5 firms **B.** 5-10 firms **C.** 10-20 firms **D.** 20-50 firms **E.** 50+ firms

13. What is the possibility that there is a predetermined winner? (0 is the lowest, 10 is the highest) **0-10**

# Appendix C: Model and Proof

To show how to obtain the equilibrium bidding strategy, without loss of generality, I can rewrite the firm's choice of the bid combination $(p, \boldsymbol{q})$ as choosing a score $s$ and quality $\boldsymbol{q}$. Therefore the price bid is $p(s, \boldsymbol{q})$. To examine the competitive equilibrium of the scoring auction, I follows Asker and Cantillon (2008), Hanazono et al. (2013), and Hanazano et al. (2020) to decompose the multidimensional bidding process into two backward steps. First, upon winning at a score $s$, the firm selects a profit-maximizing price $p$ and quality $\boldsymbol{q}$ combination. Then the firm selects the optimal score $s$.

In the first step, if the firm $i$ wins the auction at score $s_i$, its profit is given by $\pi(s_i|s_i > s_{-i}) = P(s_i, \boldsymbol{q_i}) - C(\boldsymbol{q_i}, \boldsymbol{\theta_i})$. The firm chooses the optimal $\boldsymbol{q_i}$ to solve the profit maximization problem:

$$q(s_i, \boldsymbol{\theta_i}) = arg \max_{\boldsymbol{q}} \pi(s_i|s_i > s_{-i}) = arg \max_{\boldsymbol{q}} P(s_i, \boldsymbol{q_i}) - C(\boldsymbol{q_i}, \boldsymbol{\theta_i})$$

Plug in the optimal $\boldsymbol{q}$ back to the profit function:

$$\pi(s_i, \boldsymbol{\theta_i}) = P(s_i, q(s_i, \boldsymbol{\theta_i})) - C(q(s_i, \boldsymbol{\theta_i}), \boldsymbol{\theta_i}) \tag{C1}$$

$\pi(s_i, \boldsymbol{\theta_i})$ is analogue to the profit-maximizing quantity in the general production decision problem. Applying the envelope theorem, the profit function $\pi(s_i, \boldsymbol{\theta_i})$ satisfies two first-order conditions with respect to $\boldsymbol{q}$ and $s$:

$$P_{q^l}(s_i, q(s_i, \boldsymbol{\theta_i})) = C_{q^l}(q(s_i, \boldsymbol{\theta_i}), \boldsymbol{\theta_i}) \tag{C2}$$

$$\pi_s(s_i, \boldsymbol{\theta_i}) = P_s(s_i, q(s_i, \boldsymbol{\theta_i})) \tag{C3}$$

The two equations stipulate that, conditional upon winning, the marginal cost is equal to the marginal revenue, and the marginal profit corresponds to the marginal price with respect to the price bid. $\pi_s$ is strictly negative by Assumption 1. The quality is chosen endogenously, and the target score $s$ is a sufficient statistic for the optimal bid in the profit-maximizing problem (C1).

Next, in the second step, the firm chooses the target score $s = \sigma(\boldsymbol{\theta})$ to maximize expected profit:

$$\sigma(\boldsymbol{\theta}) = arg\ max_s [P(s_i, q(s_i, \boldsymbol{\theta_i})) - C(q(s_i, \boldsymbol{\theta_i}), \boldsymbol{\theta_i})] Pr\{s_i > s_{-i}\} \tag{C4}$$

**Theorem 1.** The competitive equilibrium bidding strategy $(p(\boldsymbol{\theta}), q(\boldsymbol{\theta}))$ is given by

$$P(\sigma(\boldsymbol{\theta}), \boldsymbol{q}(\boldsymbol{\theta})) = C(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{\theta}) - P_s(\sigma(\boldsymbol{\theta}), \boldsymbol{q}(\boldsymbol{\theta})) \frac{G(\sigma(\boldsymbol{\theta}))}{(n-1) g(\sigma(\boldsymbol{\theta}))}$$

$$\Delta C_{\boldsymbol{q}}(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{\theta}) = \Delta P_{\boldsymbol{q}}(\sigma(\boldsymbol{\theta}), \boldsymbol{q}(\boldsymbol{\theta}))$$

*Proof.* If a bidder $i$ with $\boldsymbol{\theta_i}$ bids $(p_i, \boldsymbol{q_i})$ and wins at $s_i = w_p \frac{\underline{p}}{p_i} + w_q \boldsymbol{q_i}$. The winning probability is $Prob\{s > s_{-i}\}$. Rewrite the decision problem (C4) as

$$max_s \pi(s, \boldsymbol{\theta}) G(s)^{n-1} \tag{C5}$$

Differentiating the above equation with respect to $s$, I have the first-order condition:

$$\pi_s(s, \boldsymbol{\theta}) G(s)^{n-1} + \pi(s, \boldsymbol{\theta})(n-1) G(s)^{n-2} g(s) = 0$$

Rearrange the above equation,

$$\pi(s, \boldsymbol{\theta}) = -\pi_s(s, \boldsymbol{\theta}) \frac{G(s)}{(n-1) g(s)} \tag{C6}$$

Replace the $\pi_g(s, \boldsymbol{\theta})$ with equation (C3) and $\pi(s, \boldsymbol{\theta})$ with equation (C1) and $s = \sigma(\boldsymbol{\theta})$, the first-order condition can also be written as

$$P(\sigma(\boldsymbol{\theta}), \boldsymbol{q}(\boldsymbol{\theta})) = C(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{\theta}) - P_s(\sigma(\boldsymbol{\theta}), \boldsymbol{q}(\boldsymbol{\theta})) \frac{G(\sigma(\boldsymbol{\theta}))}{(n-1) g(\sigma(\boldsymbol{\theta}))}$$

At the equilibrium, the bid satisfies the profit-maximizing condition, marginal revenue equals to the marginal cost $\Delta C_{\boldsymbol{q}} = \Delta P_{\boldsymbol{q}}$. □

**Corollary C.1.** The pseudotype $k$ is monotone in $\boldsymbol{\theta}$ and monotone in $s$.

*Proof.* From equation (5), I take derivatives of $k(.)$ in respective to $\theta^m$:

$$\frac{dk(\sigma(\boldsymbol{\theta}), \boldsymbol{\theta})}{d\theta^m} = \sigma_{\theta^m}(\boldsymbol{\theta}) - \frac{[\pi_s \sigma_{\theta^m}(\boldsymbol{\theta}) + \pi_{\theta^m}]\pi_s - \pi[\pi_{ss}\sigma_{\theta^m}(\boldsymbol{\theta}) + \pi_{s\theta^m}]}{\pi_s^2}$$

$$= -\frac{\pi_s \pi_{\theta^m} - \pi[\pi_{ss}\sigma_{\theta^m}(\boldsymbol{\theta}) + \pi_{s\theta^m}]}{\pi_s^2}$$

$$= \frac{\pi \pi_{s\theta^m} - \pi_s \pi_{\theta^m}}{\pi_s^2} + \frac{\pi \pi_{ss} \sigma_{\theta^m}(\boldsymbol{\theta})}{\pi_s^2}$$

Assumption 1 gives the first part as negative. [51] $\sigma_{\theta^m}$ is negative. The final sign depends on $\pi_{ss}$.

$$\pi_{ss}(s, \boldsymbol{\theta}) = P_{ss}(s, q(s, \boldsymbol{\theta})) + \sum_{l=1}^{L} P_{sq^l}(s, q(s, \boldsymbol{\theta})) q_s^l(s, \boldsymbol{\theta})$$

Given the form of the scoring rule, $P_{ss}$ is positive, and $P_{sq^l}$ is negative. The equation (**??**)

[51] Assumption 1 makes bidder's profit function satisfy the log-submodularity condition $\frac{\partial^2 log\ \pi(s, \boldsymbol{\theta})}{\partial s \partial \theta^m} < 0$. The log-submodularity condition gives $\pi \pi_{s\theta^m} - \pi_s \pi_{\theta^m} < 0$. Please see the proof in Hanazano et al. (2020).

in Theorem 1 gives

$$C_{q^l}(q(s,\boldsymbol{\theta}),\boldsymbol{\theta}) = P_{q^l}(s, q(s,\boldsymbol{\theta}))$$

Taking the derivative with respect to $s$:

$$P_{q^l s}(s, q(s,\boldsymbol{\theta})) = [C_{q^l q^l}(q(s,\boldsymbol{\theta}),\boldsymbol{\theta}) - P_{q^l q^l}(s, q(s,\boldsymbol{\theta}))]q^l_s(s,\boldsymbol{\theta})$$

Since Assumption 2 gives $P_{q^l q^l - C_{q^l q^l}} < 0$, therefore I have $q_s(s,\boldsymbol{\theta}) < 0$ and then $\pi_{ss} > 0$. Combined all together, $k$ in monotone in $\boldsymbol{\theta}$. Since $s = \sigma(\boldsymbol{\theta})$ is monotone equilibrium and the pseudotype $k$ is monotone in firm types, pseudotype $k$ is monotone in score $s$. □

**Corollary 1.1** In the competitive equilibrium, the winning scoring bid is the expectation of the strongest rival's pseudotype $k$:

$$s^{win} = E[k^{rival(1)}|s^{win}]$$

*Proof.* As $k$ is monotone in $s$, I rewrite the decision problem (C5) with respect to pseudotype $k$:

$$max_s \pi(s(k),\boldsymbol{\theta})G(s(k))^{n-1} = max_s \pi(s(k),\boldsymbol{\theta})F_k(s^{-1}(s(k)))^{n-1}$$

The first-order condition is

$$\pi_s(s,\boldsymbol{\theta})F_k(k)^{n-1} + \pi(s,\boldsymbol{\theta})(n-1)F_k(k)^{n-2}f_k(k)\frac{1}{s'(k)} = 0$$

Since $\pi_s = P_s < 0$, I divide both sides by $\pi_1$ gives

$$F_k(k)^{n-1} + (s-k)(n-1)F_k(k)^{n-2}f_k(k)\frac{1}{s'(k)} = 0$$

Solving the differential equation, I get

$$s(k) = \int_{\underline{k}}^{k} \frac{\tau(n-1)f_k(\tau)F_k(\tau)^{n-2}}{[F_k(k)]^{n-1}}\,d\tau \tag{C7}$$

Meanwhile, the expectation of the strongest rival's pseudotype $k$ can be written as:

$$E[k^{rival(1)}|s^{win}] = E[k^{rival(1)}|k^{rival(1)} < k^{win}] = \int_{\underline{k}}^{k^{win}} \frac{\tau(n-1)f_k(\tau)F_k(\tau)^{n-2}}{[F_k(k^{win})]^{n-1}}\,d\tau \tag{C8}$$

Equation (C7) and (C8) give:

$$s^{win} = E[k^{rival(1)}|s^{win}]$$

□

**Corollary 1.2** With corruptly unrelated quality requirement $\boldsymbol{w}_{q^U} > \mathbf{0}$:

$$E[k^{rival}|s^{win}] > E[k^{rival}_{corrupt}|s^{win}]$$

and the equation in Corollary 1.1 doesn't hold, instead

$$s^{win} > E[k^{rival(1)}_{corrupt}|s^{win}]$$

*Proof.* Based on Corollary C.1 and Corollary 1.1, $k$ is monotone with respect to $\boldsymbol{\theta}$. Similarly, $s$ is also monotone with respect to $\boldsymbol{\theta}$. Consequently, this implies that $k$ will be monotone in $s$. In instances of corruption, the customized scoring rules downgrade the scores of the competitors, such that $\tilde{k}^{rival} < k^{rival}$.

□

**Corollary 1.3** When there are many corrupt bids involving weak zombie bidders, the estimated $k$ for genuinely strong bidders in competitive auctions is upwardly biased, while the estimate for zombie bidders is downwardly biased.

*Proof.* The pseudotype is estimated from:

$$\hat{k}_i = s_i + \frac{\hat{G}(s_i)}{(n-1)\hat{g}(s_i)}$$

For strong competitors in competitive auctions, their estimated $\hat{G}(s_i)$ for mixing of corrupt and competitive auctions is larger than the $\hat{G}(s_i)$ estimated from all competitive auctions due to the presence of many weak "zombie" bids. Additionally, $\hat{g}(s_i)$ is smaller because these weak zombie bids tend to cluster in the lower tail. Therefore, $\hat{k}_i$ is upwardly biased.

For weak zombie bids, representing less efficient firms with worse $\boldsymbol{\theta}$, based on the monotonicity of $k$, they experience a downward bias. □

# Appendix D: Empirical Test

## D1 Monte Carlo Simulation

Before applying the proposed estimation and test to real scoring auction data, I conduct experiments on simulated datasets with known underlying data-generating processes (DGP). The objective is twofold: to demonstrate the effectiveness and accuracy of the proposed iterative estimation method and to explore the trade-off between type I and type II errors, ultimately identifying the most suitable approach, i.e. unadjusted p-values or adjusted p-values, for identifying corruptly manipulated auctions.

I begin by generating 1,000 competitive scoring auctions. Each auction involves three draws,[52] representing three bidders, with scores sampled from a truncated normal distribution characterized by a mean of 80, a standard deviation of 15, and a range spanning from 0 to 100. This setup closely mimics the score distribution observed in real data. Figure D1 (a) illustrates the distribution of winning margins, revealing no missing mass around 0. Subsequently, I generate an additional 1,000 scoring auctions. In these auctions, the scores are drawn from the same distribution as in the competitive scenario, but with a modification: one extra point is added for predetermined winners.[53]. Furthermore, the scores of losers are downgraded by a random value $m$ drawn from a uniform distribution within the range $m \in [10, 20]$ Figure D1 (b) displays the distribution of the winning margin, with no observations approaching 0.

Next, I introduce different corruption ratios to create datasets comprising 1,000 auctions, blending both corrupt and competitive scenarios. Figure D1 (c) and (d) depict the distribution of winning margins for corruption ratios of 30% and 70%, respectively. These visual representations reveal that when the proportion of corrupted cases is relatively small, there is no discernible missing mass pattern around 0. However, when the corrupt cases dominate, the observed pattern aligns with that observed in three-bidder auctions in real auction data, as illustrated in Figure 4 (a).

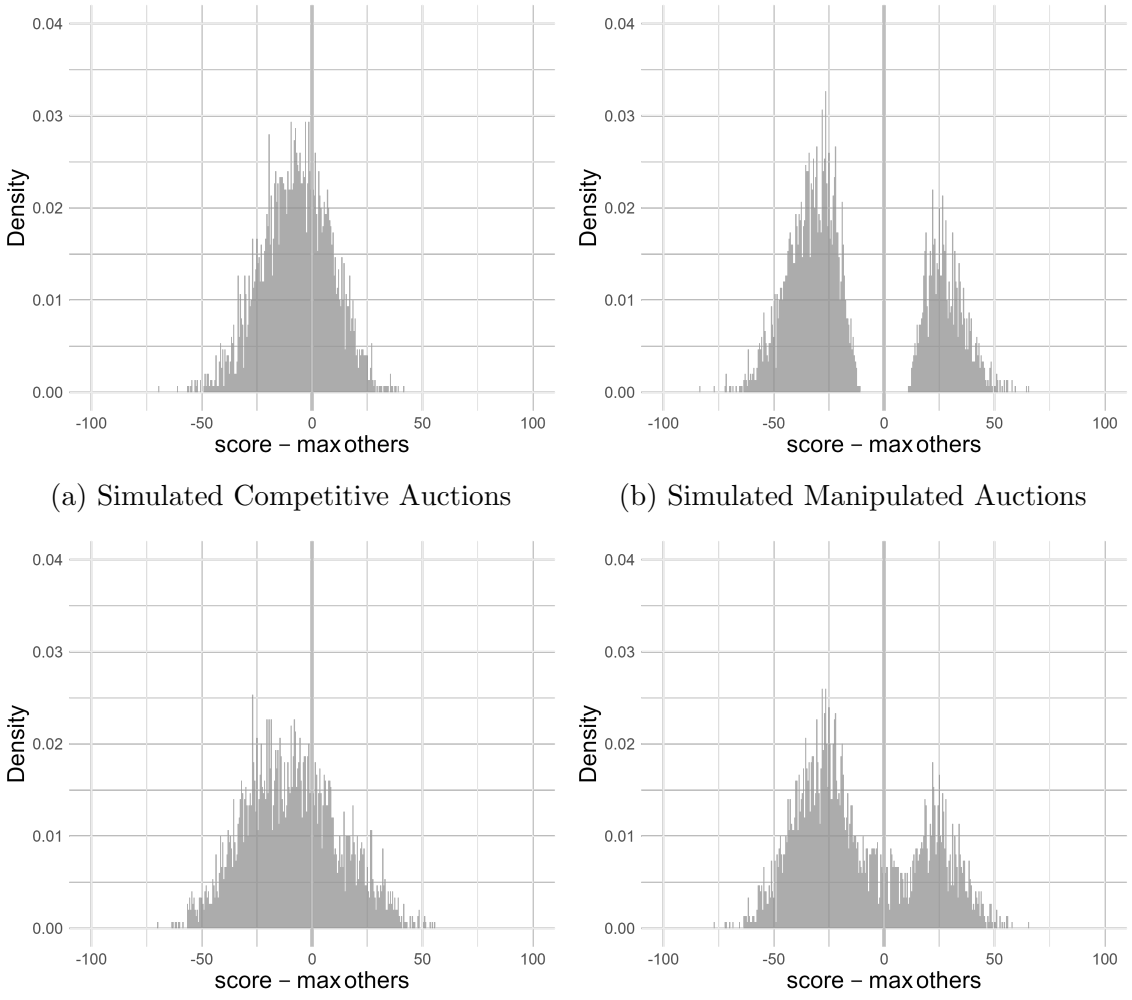
## D2 Corruption Proportion Estimation: Monte Carlo Results

With the simulated auctions, I estimate the NPMLE and the proportion of corrupt auctions under varied corruption ratios and contrast the estimation with the DGP. Figure D2 shows an example where the corruption ratio is 50%. The NPMLE yields two distributions: one with a negative mean and another with a positive mean. The estimated proportions are 51.2% and 48.8%, respectively, which closely align with the actual corruption ratio of 50%.

[52] The choice of three bidders aligns with the prevalent scenario in real auction data

[53] Predetermined winners typically attain higher scores compared to competitive winners and are closer to a perfect score of 100.

Figure D1: Distribution of Winning Margins




(a) Simulated Competitive Auctions

(b) Simulated Manipulated Auctions

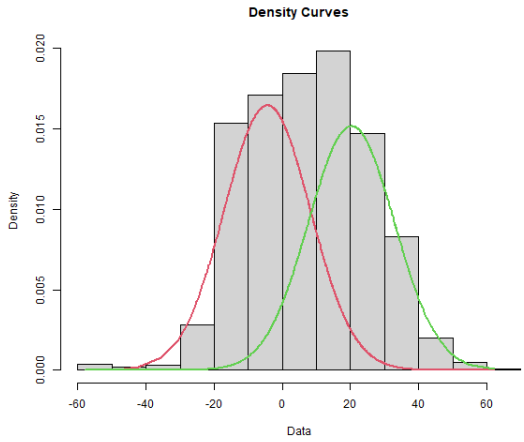
(c) Corruption Ratio 30%

(d) Corruption Ratio 70%

**Notes**: The X-axis represents the score gap between a bidder's own score and the highest score among all bidders, excluding the bidder's own score.

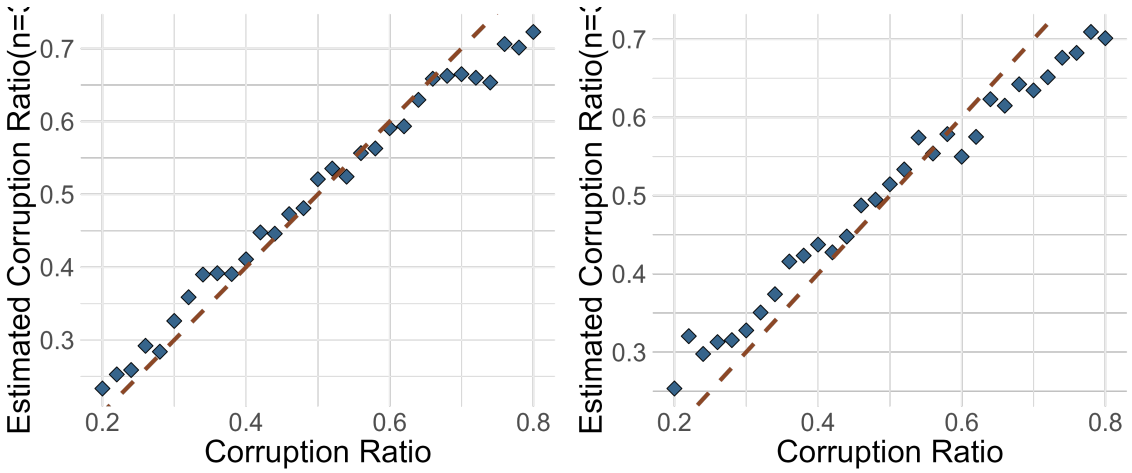
Figure D3 illustrates the estimation in comparison with the DGP. The x-axis represents the corruption ratio in the DGP, and the y-axis depicts the estimated proportion of corruption. A closer mass of dots to the dashed red 45-degree line implies higher estimation accuracy. As demonstrated in the graphs, across differing corruption ratios, the estimations approximate the actual corruption ratio, typically rendering a slightly inflated corruption proportion when the actual corruption ratio is minimal and slightly deflated when it is substantial. Overall, the results provided by the NPMLE are closely aligned.

Figure D2: An Example of NPMLE on Mixture Analysis




**Notes**: The example is mixing 500 competitive auctions with 500 corrupt auctions. By estimating the pseudotype together, I plot the distribution of $s_{win} - k^{rival(1)}$.

Figure D3: The Comparison between Estimation and DGP




(a) Auctions with Three Bidders (b) Auctions with Four Bidders

**Notes**: 1000 auctions are simulated with the different number of bidders as Section 6. The red dashed line is a 45-degree line to show the comparison between the estimation and the real DGP.

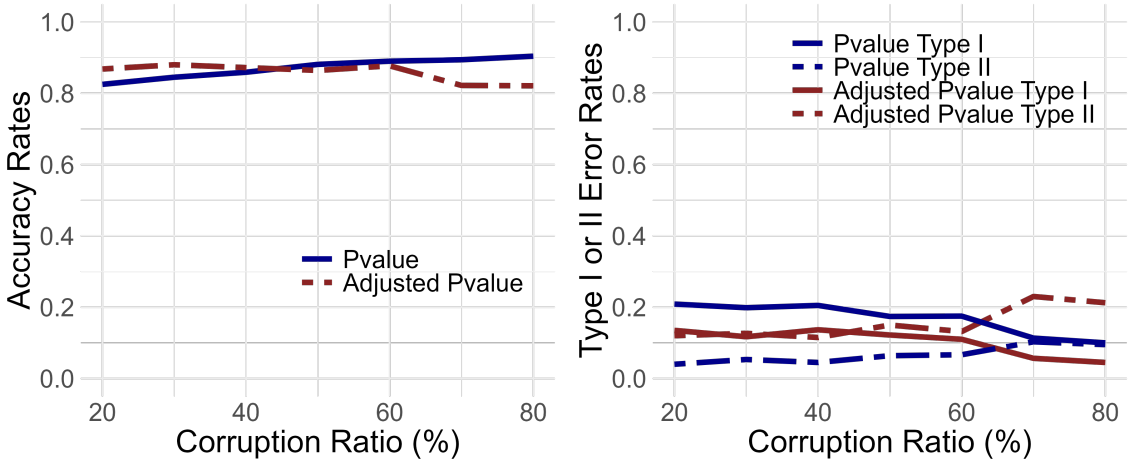
## D3 Auction-level Test: Monte Carlo Results

For the Monte Carlo study of the auction-level test, I first discuss the choice of statistical test types for evaluating systematic one-directional bias. Researchers commonly employ two categories of methods: parametric approaches like the t-test, and non-parametric methods exemplified by the Wilcoxon-Mann-Whitney test. Numerous studies (Hodges Jr and Lehmann, 1956; Fay and Proschan, 2010; De Winter, 2019) have extensively examined the

test power and type I error comparisons between these two approaches. Given that most auctions involve only three bidders and the underlying distribution of score is unknown, I employ a strategy of conducting multiple tests using both the t-test and Wilcoxon-test for each auction, thus enhancing test power. The resulting outcomes encompass both unadjusted p-values and adjusted p-values, which account for multiple testing concerns.

Then for each corruption ratio, I implemented the test steps, separately employing both the Wilcoxon test and the t-test. To enhance test power, I set the significance level at 0.1, considering that each auction involved three bidders.[54] One way to classify auctions is to reject the null hypothesis of competitiveness if either a p-value from the Wilcoxon test or the t-test was smaller than 0.1. However, to control the overall false positive rate arising from multiple tests, I also calculated adjusted p-values. Therefore, the second and more conservative classification method considered auctions as non-competitive if either an adjusted p-value from the Wilcoxon test or the t-test was smaller than 0.1. Utilizing the known underlying data-generating process, I calculated accuracy rates, type I error rates, and type II error rates for both classification methods under varying corruption ratio scenarios.

Figure D4: Comparison between Unadjusted and adjusted P-value (3 bidders)




(a) Test Accuracy Rate

(b) Type I and II Error Rate

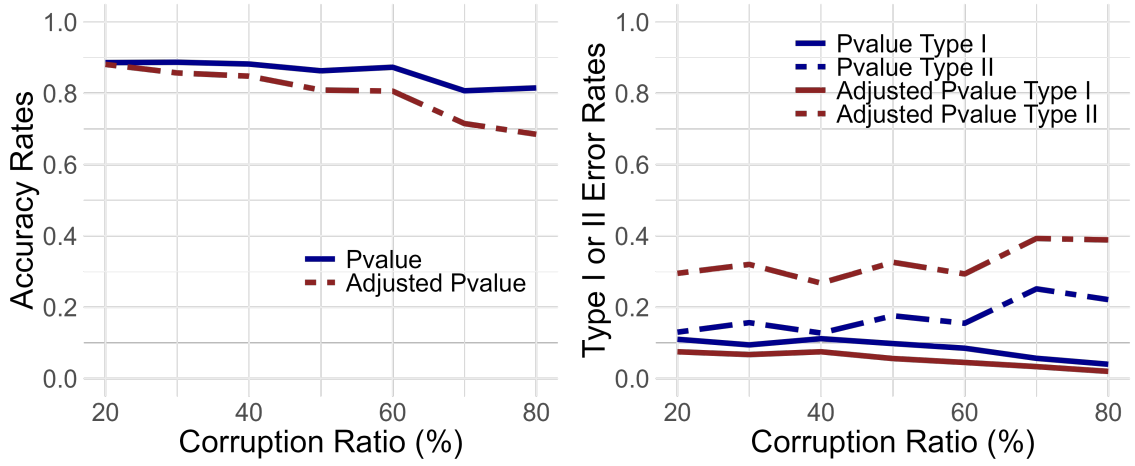
**Notes**: (1) The accuracy rate is computed by adding the number of auctions from a competitive DGP failing to reject the null and the number of auctions from a corrupt DGP rejecting the null. The sum is then divided by 1000, the total number of auctions. (2) The Type I error rate is calculated by dividing the number of auctions from a competitive DGP but still rejecting the null hypothesis by the total number of competitive auctions. (3) The Type II error rate is calculated by dividing the number of auctions from a corrupt DGP but failing to reject the null hypothesis by the total number of corrupt auctions.

The summary of results is presented in Figure D4 (a) and (b). As the corruption ratio increases, the accuracy rate of using unadjusted p-values rises from 81% to 87%, surpassing

[54]For auctions with more bidders, such as six or more, a significance level of 0.05 could be considered.

that of using adjusted p-values when the corruption rate is around 45%. Conversely, the accuracy rate of using adjusted p-values declines from 86% to 77% as they tend to be overly conservative in cases of dominant corruption. Figure D4 (b) illustrates a trade-off between type I error and type II error, with a decrease in type I error leading to an increase in type II error and a reduction in test power. By setting the significance level at 0.1, the type I error using unadjusted p-values remains restrained below 0.1 when the corruption ratio exceeds 0.5, whereas, in all scenarios deploying adjusted p-values, it is maintained below 0.1 and even converges to 0.05 when corruption is pervasive.

Figure D5: Comparison between Different Significance Levels




(a) Accuracy Rate (3 bidders with $\alpha = 0.05$) (b) Type I and II (3 bidders with $\alpha = 0.05$)

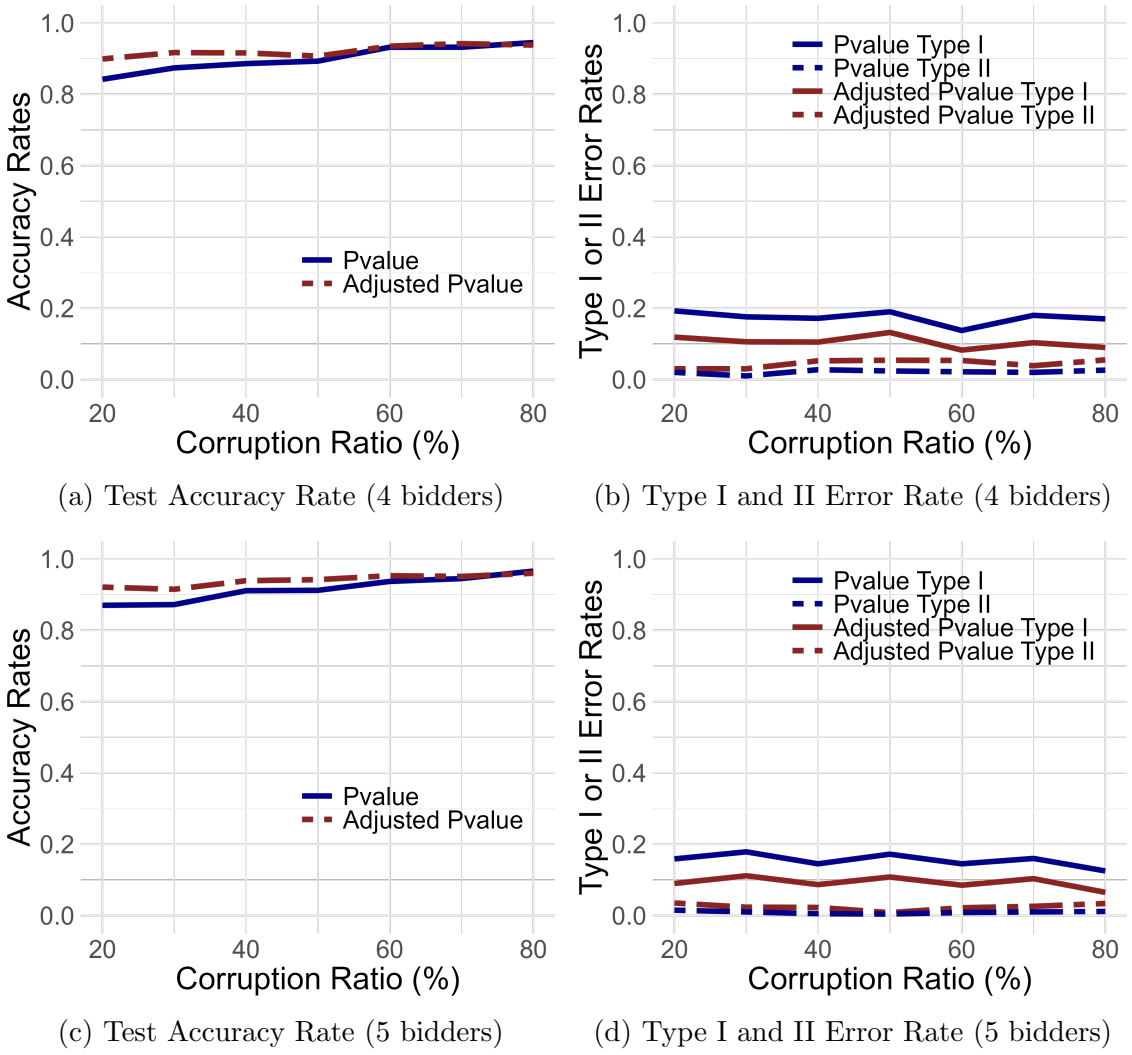
**Notes**: (1) The accuracy rate is computed by adding the number of auctions from a competitive DGP failing to reject the null and the number of auctions from a corrupt DGP rejecting the null. The sum is then divided by 1000, the total number of auctions. (2) The Type I error rate is calculated by dividing the number of auctions from a competitive DGP but still rejecting the null hypothesis by the total number of competitive auctions. (3) The Type II error rate is calculated by dividing the number of auctions from a corrupt DGP but failing to reject the null hypothesis by the total number of corrupt auctions.

Changing the significance level from 0.1 to 0.05, Figure D5 shows the accuracy rates alongside type I and II errors. Generally, a high level of accuracy is still sustained as presented in Figure D5 (a). However, when contrasted with Figure D4, the test power is substantially reduced, particularly when the corruption ratio is exceedingly high, such as above 80%.

I also replicate the Monte Carlo Study for cases with four and five bidders and present the summary of test accuracy rates, type I and II error rates in Figure D6. In comparison to Figure D4 (a), Figure D6 (a) and (c) indicate that auctions with more bidders exhibit higher overall accuracy rates. One consistent trend is that when the underlying real corruption ratio exceeds 50%, unadjusted p-values outperform adjusted p-values.

In conclusion, in terms of accuracy rates, the test steps perform reasonably well in this

Figure D6: Comparison between Unadjusted and adjusted P-value (4 or 5 bidders)




(a) Test Accuracy Rate (4 bidders)

(b) Type I and II Error Rate (4 bidders)

(c) Test Accuracy Rate (5 bidders)

(d) Type I and II Error Rate (5 bidders)

**Notes**: (1) The accuracy rate is computed by adding the number of auctions from a competitive DGP failing to reject the null and the number of auctions from a corrupt DGP rejecting the null. The sum is then divided by 1000, the total number of auctions. (2) The Type I error rate is calculated by dividing the number of auctions from a competitive DGP but still rejecting the null hypothesis by the total number of competitive auctions. (3) The Type II error rate is calculated by dividing the number of auctions from a corrupt DGP but failing to reject the null hypothesis by the total number of corrupt auctions.

Monte Carlo study. The choice between using the unadjusted p-value or adjusted p-value depends on the specific objectives and the underlying situation. If the goal is to maximize the overall accuracy rate, especially when corruption is prevalent, as in this setting, the unadjusted p-value proves more powerful. Therefore, the unadjusted p-value is chosen as the primary result, while the adjusted p-value results are also presented as a conservative measure. However, if the agency prioritizes minimizing type I errors, if the estimated corruption level is low, or if there is a high average number of bidders, then the adjusted p-value may

be a more suitable choice.

### D4 Auction-level Test: More

I apply the auction-level test to the real public procurement auction data.[55] To provide a concrete example of how the statistical tests work, I used an auction described as the "for-sure corrupt case" in Section 4, Table 1, as an illustration.

This auction involved three bidders with scores of 98.3, 51.2, and 46.7. Given the winning score of 98.3 and the auction characteristics, including budget, price/variable quality weight, procurement category, and number of bidders, there is a distribution of $\hat{k}$ for bidders with scores lower than 98.3, as shown in Figure D7 (a). The expected $\hat{k}$ of rivals is 90.72, inferred from the winning score of 98.3 and represented by the red dashed vertical line. However, the realized rivals' $\hat{k}$ are 54.79 and 53.03, drawn by the solid blue lines and located in the low tail of the distribution. The p-value of the Wilcoxon test is 0.002, leading to the rejection of the null hypothesis. Therefore, the auction is successfully identified as corrupt using my method.

The same is true for the four-bidder procurement auction described in Section 4, Table A1. Given the winning score of 98.25 and the auction characteristics, the distribution of $\hat{k}$ is shown in Figure D7 (b). The expected $\hat{k}$ is 84.89, while the realized rivals' $\hat{k}$ are 66, 64, and 60, drawn by the solid blue lines and located in the low tail of the distribution. The p-value of the Wilcoxon test is 0.008, resulting in the rejection of the null hypothesis and labeling the auction as corrupt.

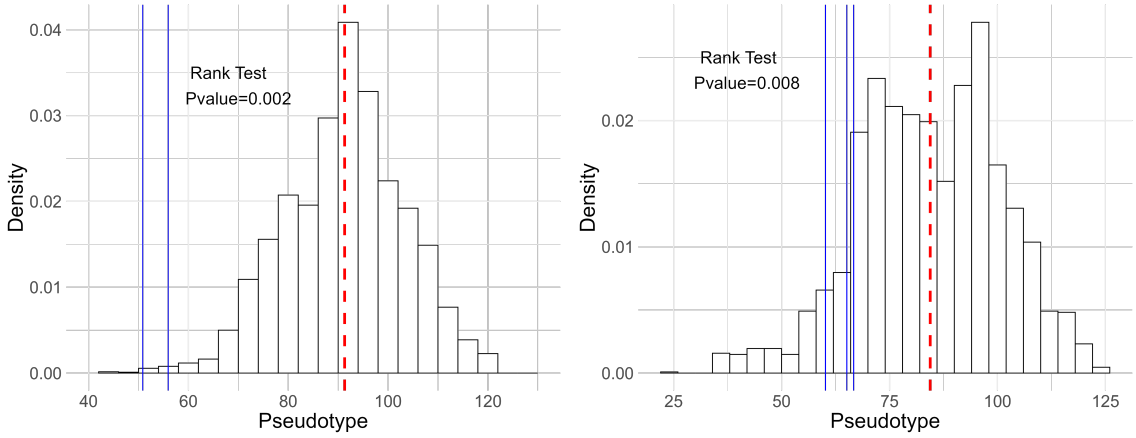
Furthermore, I plot the distributions of the density of winning margins separately for the procurement auctions that reject and do not reject the null hypothesis in Figure D8, grouped by 4-bidders and 5-bidders. Figure D8 (a) and (b) show the distributions of winning margins for auctions that fail to reject the null. The distributions are concentrated around the threshold 0, with a thin tail of low scores. Figure D8 (c) and (d) show the distributions of winning margins for auctions that reject the null. There are missing mass areas around 0, and the low-score tail is heavy, indicating a significant number of exceptionally low-quality bids. The missing gap is more pronounced when there are fewer bidders. The comparison of graphs shows that the screening process works as expected.

### D5 Auction-level k-means Clustering

Instead of estimating an unbiased pseudotype through the iterative approach proposed in Section 6, I can also use the biased pseudotype and clustering method to identify the

[55] The desired significance level $\alpha = 0.1$ is achieved after five iterations for the t-test and six iterations for the Wilcoxon test.

Figure D7: Test Examples




(a) Test for Table 1 (b) Test for Table A1

**Notes**: The distributions of pseudotype are conditional on the winning scores and take procurement project characteristics into consideration.

corrupt auctions. In the following sections, I provide a detailed implementation of this clustering approach and a comparison with the main method in Section 6. Overall, the main method from Section 6 consistently outperforms the clustering method in terms of prediction accuracy, as evidenced by both Monte Carlo simulations and real data validated through expert surveys.

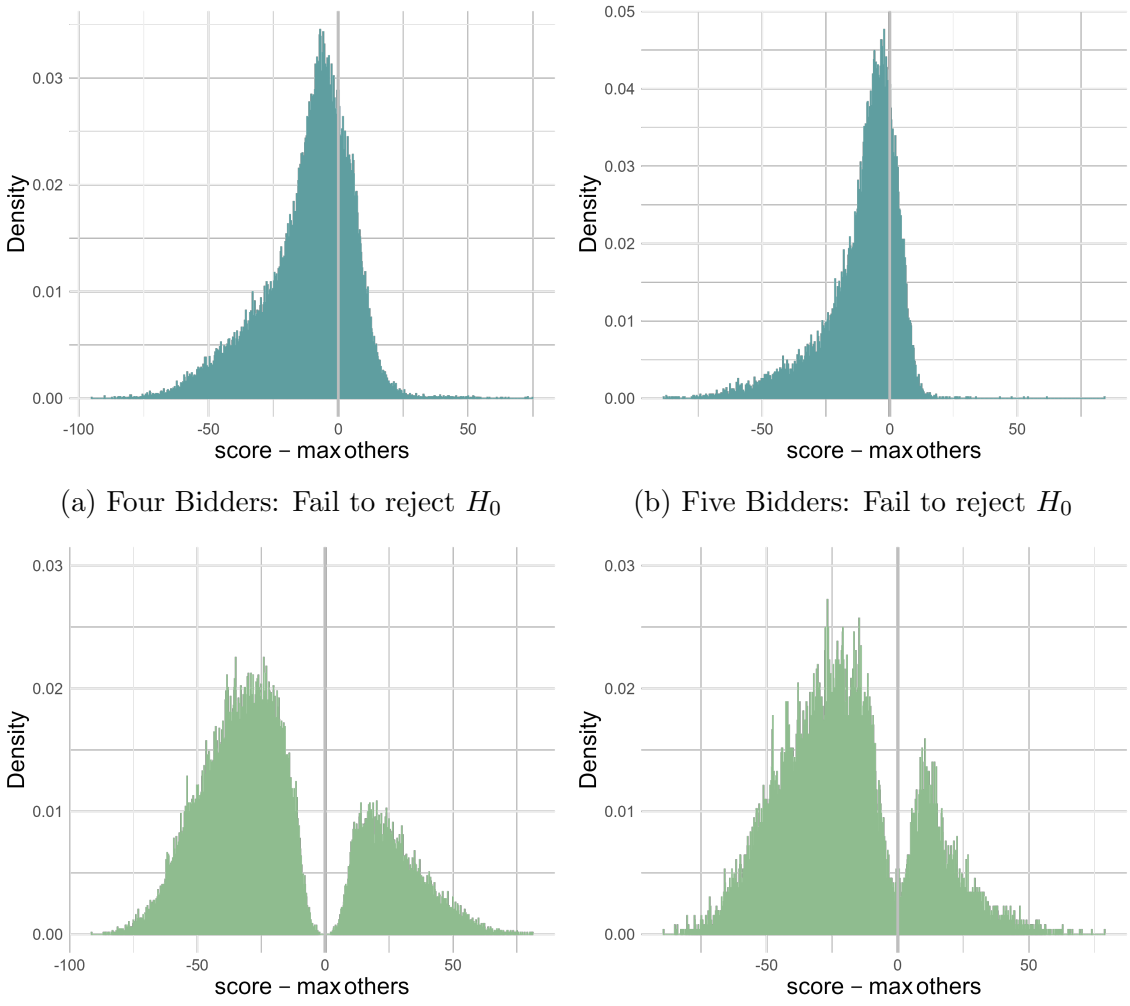
In Section 5, it's demonstrated that in the presence of scoring rule manipulation, both scores and pseudotypes $k$ are downgraded, such that $E[k^{rival}|s^{win}] > E[k_a^{rival}|s^{win}]$. As a result:

$$E[k^{rival} - s^{win}] > E[k_{corrupt}^{rival} - s^{win}]$$

This distinction provides a rule for employing the clustering method to categorize auctions as either corrupt or noncorrupt.

For each auction $t$ with $n$ bidders, I construct an observation vector composed of the demeaned pseudotypes $k^{rival} - s^{win}$ of all rivals. This vector, of dimension $n-1$, takes the form $[k^{rival(1)} - s^{win}, k^{rival(2)} - s^{win}, ..., k^{rival(n-1)} - s^{win}]$. To classify this multi-dimensional data, I use the k-means clustering method, a popular unsupervised learning clustering method in computer science. To show how the vectors of noncorrupt auctions and corrupt auctions behave differently, I use the 50% corruption ratio with three bidders in the Monte Carlo study as an example. In such a scenario, given three bidders, there are two rival bids, resulting in a 1x2-dimensional vector. This can be visualized on a 2D plot. Figure D9 (a) presents the corruption labels based on the data-generating process (DGP). Noncorrupt auctions concentrated on the upper position of the 45-degree line while corrupt auctions are mainly

Figure D8: Distribution of Winning Margins




(a) Four Bidders: Fail to reject $H_0$

(b) Five Bidders: Fail to reject $H_0$

(c) Four Bidders: Reject $H_0$

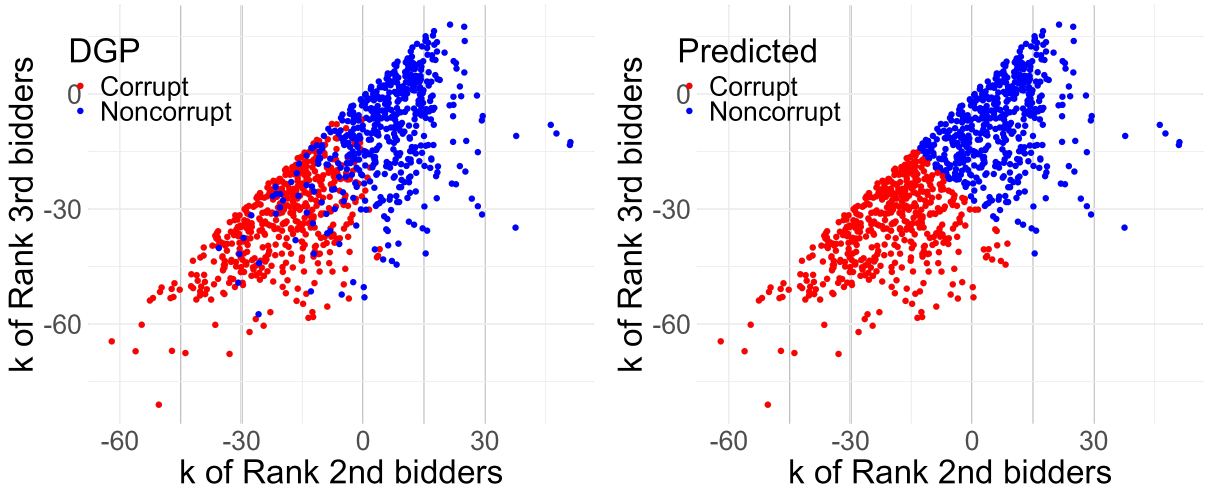
(d) Five Bidders: Reject $H_0$

**Notes**: The X-axis represents the score gap between a bidder's own score and the highest score among all bidders, excluding the bidder's own score.

located in the lower area. Despite some overlap in the central region, a clear separation is evident. Implementing the k-means clustering to identify different clusters, the predicted results, shown in Figure D9 (b), perform pretty well, with some discrepancies observed in the overlapping middle area.

I compare the accuracy rates of the main test results with the k-means predictions across varying corruption ratios, as shown in Figure D10 (a). Overall, for auctions with three bidders, the main test surpasses the k-means method in performance under most corruption ratio conditions. Figure D10 (b) shows the auctions with four bidders, and in all scenarios, the main test outperforms the K-means.

Figure D9: k-means Clustering: Observed V.S. Predicted

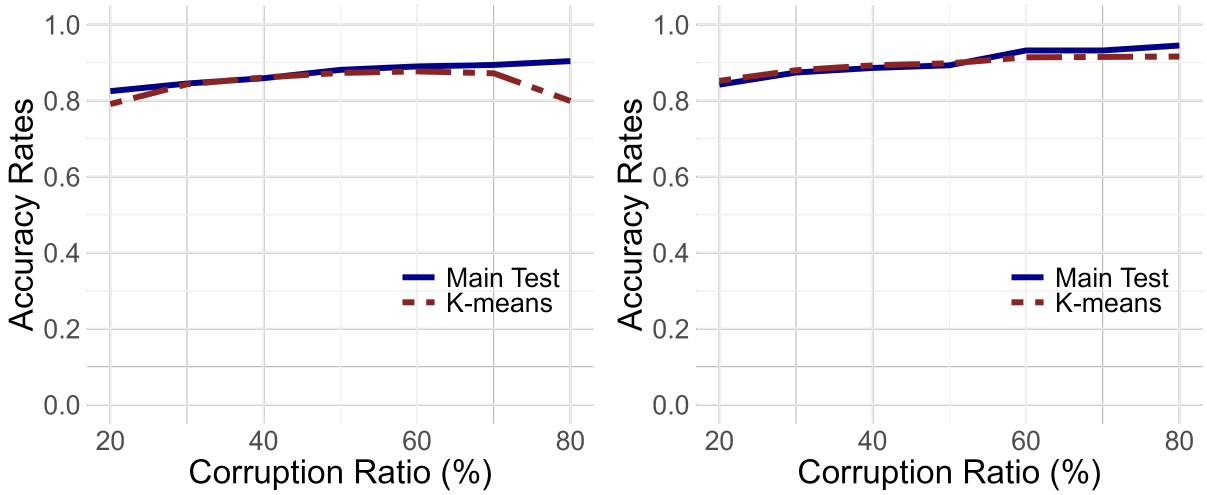



(a) DGP observed corruption labels

(b) K-means predicted corruption labels

**Notes**: (a) plots the real corrupt auctions in red v.s. noncorrupt auctions in blue. (b) plots the predicted corrupt auctions in red v.s. predicted noncorrupt auctions in blue.

Figure D10: Comparison between Main Test and K-means in Monte Carlo



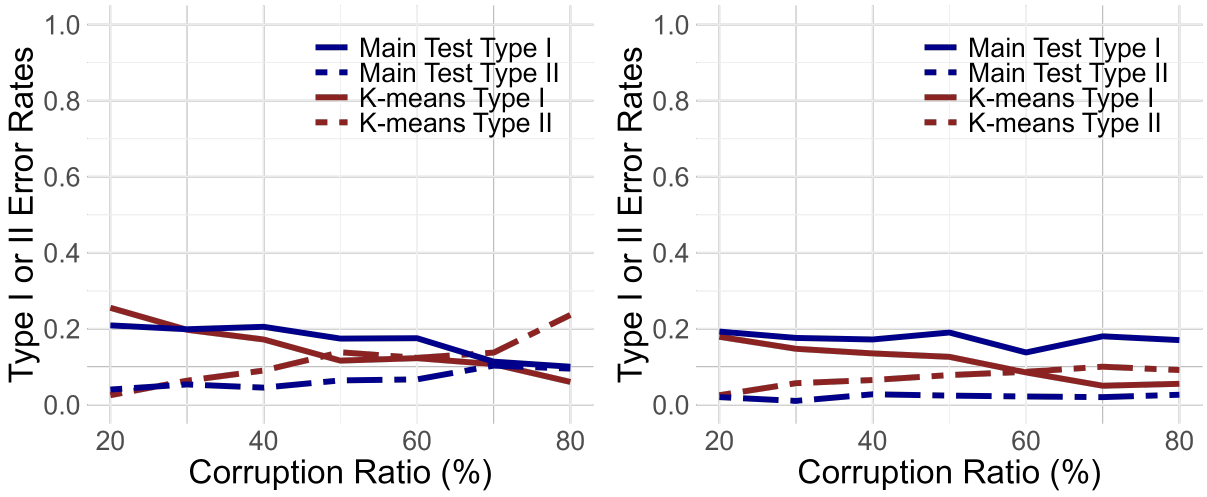

(a) Accuracy Rates with 3 Bidders

(b) Accuracy Rates with 4 Bidders

**Notes**: (a) plots the accuracy rates under different corruption ratios, with the blue solid line representing the main test results and the red dashed line representing the k-means cluster prediction. (b) plots the accuracy rates for the auctions with four bidders.

Figure D11 (a) and (b) show the type I and II errors for auctions consisting of three or four bidders. Evidently, in both scenarios, the test power of the main method surpasses that of the k-means approach.

Figure D11: Comparison between Main Test and K-means in Monte Carlo




(a) Errors with 3 Bidders

(b) Errors with 4 Bidders

**Notes**: (a) plots the Type I and II errors under different corruption ratios in auctions with 3 bidders. (b) plots the Type I and II errors under different corruption ratios in auctions with 4 bidders.

Then I apply the k-means method to all the public procurement data. The results are presented in Table D1. The proportions of auctions labeled as corrupt by the K-means method align closely with the primary results that employ Unadjusted P-values in cases of 3 and 4 bidders. For instances with 5 bidders, the proportion approximates the adjusted P-values. The comparison further provides evidence for the assertion in Section 6 that with a greater number of bidders, the adjusted p-value is the preferable choice. The overall prediction consistency rate between the unadjusted p-value in the main results and the K-means method is 82.83%. Comparing the results with the expert audit study, the K-means results reach a consistency (accuracy) rate of 79.59%, less than the main results of 81.4%.

Table D1: K-means Estimation V.S. Main Results

| | All | 3 bidders | 4 bidders | 5 bidders |
|---|---|---|---|---|
| Unadjusted P-value $\alpha = 0.1$ | 64.71% | 69.74% | 51.18% | 47.74% |
| Adjusted P-value $\alpha = 0.1$ | 56.83% | 62.36% | 43.39% | 38.32% |
| K-means Estimation | 66.04% | 72.48% | 55.8% | 48.36% |
| Same as Unadjusted P | 83.29% | 86.82% | 77.77% | 73.45% |

**Notes**: The corruption proportions are calculated from the main test method and the K-means clustering method, conditional on auction characteristics.

# Appendix E: Participation Test

I employ the participation test proposed by Conley and Decarolis (2016) to examine if there are specific groups of bidders who participate abnormally together in procurement auctions. The participation test compares the participation patterns of a suspected group of firms, denoted as $g$, in the auctions of interest with the participation patterns in reference groups, denoted as $H$. The objective is to determine whether the statistic reflecting the participation patterns of group $g$ is an extreme event compared to the reference distribution, which follows a uniform distribution with support points represented by analogous statistics for all groups in $H$. Let $T$ denote the total number of auctions, and the indicator variable $d_{it}$ takes a value of 1 if firm $i$ participates in auction $t$. For a group $g$ with a size of $N^g$, the fraction of auctions participated in by $K \leq N^g$ members of group $g$ is calculated as:

$$f_K^g = \sum_{t=1}^{T} \mathbb{1}\{K = \sum_{i \in g} d_{it}\}$$

Similarly, the count for firms in reference group $h \in H$ is defined as:

$$f_K^h = \sum_{t=1}^{T} \mathbb{1}\{K = \sum_{i \in h} d_{it}\}$$

I test the hypothesis that firms in group $g$ do not exhibit unusually coordinated participation by assessing whether $f_K^g$ is drawn from the distribution of $f_K^h$.

The test relies on the selection of the suspected group $g$ and the reference group $H$. While theoretically applicable to any group, the computational demands of analyzing the extensive auction data make it impractical to consider all possible groups. To address this, I adopt the approach proposed by Conley and Decarolis (2016), which estimates the links between firms to predict the probability of belonging to the same cartel. The estimation involves three key steps. First, I compile the data on firm associations, including common ownership, temporary bidding consortia, subcontract exchanges, and location proxies. Next, I identify all pairs of firms that participated in the same procurement auction. Then, I employ the cartel membership probit model estimation from Conley and Decarolis (2016) to predict the probabilities of cooperative group membership for each pair of firms. To reduce the computational burden, I focus on the top 10% firms that participate in auctions most frequently. The complements of these predicted probabilities are interpreted as a dissimilarity array. In the third step, I construct the dissimilarity array using the hierarchical clustering algorithm (Gordon, 1999) to classify firms into different clusters. Finally, I conduct the statistical test by comparing the selected clusters with the remaining non-clustered firms.

# Appendix F: Model Extension with Bidder Asymmetry

In this Appendix, I extend the basic model to incorporate some bidder asymmetry using firm characteristics $z$. For example, local and non-local firms have different information and have different bidding strategies even when they have the same firm type $\boldsymbol{\theta}$.

## F1 Model

In a real competition, $n$ potential ex-ante asymmetric and risk-neutral firms participate in the auction, denoted by $i = 1, 2, ..., n$. The number of competitors is known. Each firm draws its firm type $(\boldsymbol{\theta}|z)$ independently, conditional on firm characteristics $z$ from a publicly known absolutely continuous conditional distribution function $F(\boldsymbol{\theta}|z)$. The type $\boldsymbol{\theta}$ and firm characteristics $z$ are private information for each firm. The distribution of $z$, $F_z$, is also publicly known. Firms with different types have different costs $C(\boldsymbol{q}, \boldsymbol{\theta}|z)$, which depends on the quality bid $\boldsymbol{q}$ they submit and the firm type combination $(\boldsymbol{\theta}|z)$. Then Theorem 1 and Theorem 2 become the following.

**Theorem 1.** *The asymmetric, which is also conditional symmetric, Bayes–Nash equilibrium with interior-solution bids equilibrium is given by*

$$P(s, q(s, \boldsymbol{\theta}|z)) = \underbrace{C(q(s, \boldsymbol{\theta}|z), \boldsymbol{\theta}|z)}_{cost} - \underbrace{P_s(s, q(s, \boldsymbol{\theta}|z)) \frac{\int_z G(s|z) f_z(z)\, dz}{(n-1) \int_z g(s|z) f_z(z)\, dz}}_{\textit{Markup: relative advantage of MC}} \tag{D1}$$

$$C_{q^l}(q(s, \boldsymbol{\theta}|z), \boldsymbol{\theta}|z) = P_{q^l}(s, q(s, \boldsymbol{\theta}|z)) \tag{D2}$$

*with the condition that $P_{q^l q^l} - C_{q^l q^l} < 0$ to ensure the profit maximizing.*

*Proof.* If a bidder with $\boldsymbol{\theta}_i$ bids $(p_i, \boldsymbol{q_i})$ and wins at $s_i = w_p \frac{\underline{p}}{p_i} + w_q \boldsymbol{q_i}$. The winning probability is $Prob\{s(\boldsymbol{\theta}_j|z_j) < s_i, j \neq i\}$. Rewrite the decision problem (C4) as

$$max_{s_i} \pi(s_i, \boldsymbol{\theta}_i|z_i) \int_{z_{-i}} \prod_{j \neq i} G(s_i|z_j) f_z(z_j) dz_{-i} = max_{s_i} \pi(s_i, \boldsymbol{\theta}_i|z_i) (\int_z G(s_i|z) f_z(z)\, dz)^{n-1}$$

Differentiating the above equation with respect to $s$, I have the first-order condition:

$$\pi_s(s_i, \boldsymbol{\theta}_i|z_i)(\int_z G(s_i|z) f_z(z)\, dz)^{n-1} + \pi(s_i, \boldsymbol{\theta}_i|z_i)(n-1)(\int_z G(s_i|z) f_z(z)\, dz)^{n-2} \int_z g(s_i|z) f_z(z)\, dz = 0$$

Rearrange the above equation,

$$\pi(s_i, \boldsymbol{\theta}_i|z_i) = -\pi_1(s_i, \boldsymbol{\theta}_i|z_i) \frac{\int_z G(s_i|z) f_z(z)\, dz}{(n-1) \int_z g(s_i|z) f_z(z)\, dz} \tag{D3}$$

Replace the $\pi_s(s_i, \boldsymbol{\theta}_i|z_i)$ with equation (C3) and $\pi(s_i, \boldsymbol{\theta}_i|z_i)$ with equation (C1), the first-order condition can also be written as

$$P(s_i, q(s_i, \boldsymbol{\theta_i}|z_i)) = C(q(s_i, \boldsymbol{\theta}_i|z_i), \boldsymbol{\theta}_i|z_i) - P_s(s_i, q(s_i, \boldsymbol{\theta}_i|z_i)) \frac{\int_z G(s_i|z) f_z(z)\, dz}{(n-1)\int_z g(s_i|z) f_z(z)\, dz}$$

At the equilibrium, the bid satisfies the profit-maximizing condition $P_{q^l} = C_{q^l}$. From the equation (C2) and Assumption 2, $P_q = \frac{w_q p^2}{w_p \underline{p}}$. I can get the equation (D2). □

**Corollary 1.1.** *In the competitive equilibrium, the winning scoring bid is the expectation of the strongest rival's pseudotype k:*

$$s^{win} = E[k^{rival(1)}|s^{win}] \tag{D4}$$

*Proof.* Rewrite the decision problem (C4) with type $\boldsymbol{\theta}$ distribution:

$$max_{s_i} \pi(s_i, \boldsymbol{\theta}_i|z_i)(\int_z G(s_i|z) f_z(z)\, dz)^{n-1} = max_{s_i} \pi(s_i, \boldsymbol{\theta}_i|z_i)(\int_z F(s^{-1}(s_i|z)) f_z(z)\, dz)^{n-1}$$

The first-order condition is

$$\pi_s(s, \boldsymbol{\theta}|z)(\int_z F(\boldsymbol{\theta}|z) f_z(z)\, dz)^{n-1} + \pi(s, \boldsymbol{\theta}|z)(n-1)(\int_z F(\boldsymbol{\theta}|z) f_z(z)\, dz)^{n-2} \int_z f(\boldsymbol{\theta}|z) f_z(z)\, dz \frac{1}{s'(\boldsymbol{\theta}|z)} = 0$$

Since $\pi_s = P_s < 0$, I divide both sides by $\pi_s$ gives

$$(\int_z F(\boldsymbol{\theta}|z) f_z(z)\, dz)^{n-1} + (s(\boldsymbol{\theta}|z) - k)(n-1)(\int_z F(\boldsymbol{\theta}|z) f_z(z)\, dz)^{n-2} \int_z f(\boldsymbol{\theta}|z) f_z(z)\, dz \frac{1}{s'(\boldsymbol{\theta}|z)} = 0$$

Solving the differential equation, I get

$$s(\boldsymbol{\theta}|z) = \int_{\underline{\boldsymbol{\theta}}}^{\boldsymbol{\theta}} \frac{(n-1)\int_z f(\tau|z) f_z(z)\, dz (\int_z F(\tau|z) f_z(z)\, dz)^{n-2}}{(\int_z F(\boldsymbol{\theta}|z) f_z(z)\, dz)^{n-1}} k\, d\tau = E[k^{rival(1)}|s^{win}, z]$$

Then take expectations on both sides with respect to $z$:

$$s^{win} = E_z[s(\boldsymbol{\theta}|z)] = E_z[E[k^{rival(1)}|s^{win}, z]] = E[k^{rival(1)}|s^{win}]$$

□

### F2 Estimation and Test

The estimates of the distribution of scores also need to take the asymmetry of bidders into consideration. I have $A$ independent scoring auctions, and for each auction $a$, there

are $n_a$ firms. I observe the bids $(p_{ia}, q_{ia})$, the score $s_{ia}$ for each bidder, and some firm characteristics $\boldsymbol{z}_i$, as well as partial auction-specific covariates $\boldsymbol{x_a}$. The CDF and PDF of equilibrium scores can be non-parametrically estimated:

$$\hat{G}_s(s, n, \boldsymbol{x}, \boldsymbol{z}) = \frac{1}{T h_{g_n} h_{g_x} h_{g_z}} \sum_{a=1}^{A} \frac{1}{n_a} \sum_{i=1}^{n_a} \mathbb{1}(s_{ia} \leq s) K_G\left(\frac{n - n_a}{h_{g_n}}, \frac{\boldsymbol{x} - \boldsymbol{x_a}}{h_{g_x}}, \frac{\boldsymbol{z} - \boldsymbol{z_i}}{h_{g_z}}\right)$$

$$\hat{g}_s(s, n, \boldsymbol{x}, \boldsymbol{z}) = \frac{1}{T h_s h_{g_n} h_{g_x} h_{g_z}} \sum_{a=1}^{A} \frac{1}{n_a} \sum_{i=1}^{n_a} K_g\left(\frac{s - s_{it}}{h_s}, \frac{n - n_a}{h_{g_n}}, \frac{\boldsymbol{x} - \boldsymbol{x_a}}{h_{g_x}}, \frac{\boldsymbol{z} - \boldsymbol{z_i}}{h_{g_z}}\right)$$

where $\mathbb{1}(.)$ is the indicator function; $K_G$ and $K_g$ are kernels; and $h_{g_n}, h_{g_x}, h_{g_z}, h_s$ are bandwidths.

I can rewrite the definition of pseudotype $k$ equation:

$$k(s(\boldsymbol{\theta}_i), \boldsymbol{\theta}_i | z_i) = s(\boldsymbol{\theta}_i | z_i) - \frac{\pi(s_i, \boldsymbol{\theta}_i | z_i)}{\pi_1(s_i, \boldsymbol{\theta}_i | z_i)} = s(\boldsymbol{\theta}_i | z_i) + \frac{\int_z G(s_i|z) f_z(z)\, dz}{(n-1) \int_z g(s_i|z) f_z(z)\, dz} \tag{D5}$$

Therefore the pseudotype defined above can be estimated using:

$$\hat{k}_{it} = s_{it} + \frac{\int_z \hat{G}(s_i|z) f_z(z)\, dz}{(n-1) \int_z \hat{g}(s_i|z) f_z(z)\, dz}$$

Using the new estimates I can redo the test. The test results are shown in Table F1: Compare to the test prediction from the basic model, the proportions of auctions failing to

Table F1: Test Results

| Proportion of Procurement Reject the Null | | | | |
|---|---|---|---|---|
| | All | 3 Bidders | 4 Bidders | 5 Bidders |
| Unadjusted P-value | 63.04% | 68.6% | 49.58% | 44.61% |
| Adjusted P-value | 55.37% | 61.23% | 42.13% | 36.87% |

**Notes**: The first row represents the results using unadjusted p-values and the second row represents the results using adjusted p-values.

reject the null increase as some pseudotype gaps are attributed to the asymmetry of bidders, the differences are small.

Then I compare the new predictions with the expert survey results in Table F2. The allover accuracy rate is 80%, which is still pretty high and captures fewer corrupt cases than the basic model.

Table F2: Expert Survey and Model Prediction

| Accuracy Rate=82.4% | Total=454 | Model Prediction | | |
|---|---|---|---|---|
| | | Corrupt (PP) | Not Corrupt (PN) | |
| Expert Survey | Corrupt (P) | 281 | 52 False N | Test Power 84.38% |
| | Not Corrupt (N) | 28 False P | 93 | Test Size 23.14% |

**Notes**: For the expert survey, if any evaluation survey reveals highlighted scoring rules accompanied by expert explanations and concerns, the procurement project is labeled as corrupt. The classification of model prediction uses unadjusted p-values with a significance level of 0.1.

# Appendix G: Model Extension with Evaluation Randomness

In a more realistic setting, the outcomes of procurement auctions can be influenced by the subjective evaluations of experts. This uncertainty can cause the target quality that bidders aim to maximize their expected profit to differ from the final quality score that is observed. However, the quality score calculation that takes the average of quality evaluations from experts can reduce the impact of uncertainty. To account for this randomness, I have developed an extended model in this appendix, which is compared to the basic model presented in Section 5. The results demonstrate that randomness does not have an important effect in my setting.

## G1 Model

I introduce the i.i.d. error $\epsilon_{iea}$ into the quality evaluation of each expert $e$ for the bidder $i$ in auction $a$.[56] Therefore the observed quality score that each expert gives is the sum of real quality $q_{ia}$ and error, $q^o_{iea} = q_{iea} + \epsilon_{iea}$. The final observed quality score that bidder $i$ in auction $t$ gets is $q^o_{ia} = \frac{\sum_{e=1}^{5} q_{iea}}{5} = \frac{\sum_{e=1}^{5} (q_{ia}+\epsilon_{iea})}{5} = q_{ia} + \epsilon_{ia}$. Then the observed score is the sum of real score plus the weighted error $s^o_{ia} = s_{ia} + w_q\epsilon_{ia}$

The new equilibrium under the extension now includes the expectation of the evaluation error. I rewrite the equilibrium conditions in Theorem 1 as follows:

$$P(s_{ia}, q(s_{ia}, \boldsymbol{\theta}_{ia})) = C(q(s_{ia}, \boldsymbol{\theta}_{ia}), \boldsymbol{\theta}_{ia}) - P_s(s_{ia}, q(s_{ia}, \boldsymbol{\theta}_{ia})) \frac{\int_\epsilon G(s_{ia} + w_q\epsilon_{ia})^{n-1} dF_\epsilon(\epsilon)}{\int_\epsilon (n-1) g(s_{ia} + w_q\epsilon_{ia}) G(s_{ia} + w_q\epsilon_{ia})^{n-2} dF_\epsilon(\epsilon)}$$

$$C_{q^l}(q(s_{ia}, \boldsymbol{\theta}_{ia}), \boldsymbol{\theta}_{ia}) = P_{q^l}$$

The new pseudotype is $k_{ia} = s_{ia} + \frac{\int_\epsilon G(s_{ia}+w_q\epsilon_{ia})^{n-1} dF_\epsilon(\epsilon)}{\int_\epsilon (n-1) g(s_{ia}+w_q\epsilon_{ia}) G(s_{ia}+w_q\epsilon_{ia})^{n-2} dF_\epsilon(\epsilon)}$. Theorem 2 still holds except the score is the real score $s_{ia}$ not the observed $s^o_{ia}$. Since now I don't know the real target score $s$ and I only know the observed score, for a given observed score, I can estimate the expected pseudotype:

$$E_s[k_{ia}|s^o_{ia}] = s^o_{ia} + \int_{s+\xi=s^o_{ia}} \frac{\int_\epsilon G(s_{ia} + w_q\epsilon_{ia})^{n-1} dF_\epsilon(\epsilon)}{\int_\epsilon (n-1) g(s_{ia} + w_q\epsilon_{ia}) G(s_{ia} + w_q\epsilon_{ia})^{n-2} dF_\epsilon(\epsilon)} dF_s(s)$$

$$= E_\epsilon[k_{ia}|s^o_{ia}] = s^o_{ia} + \int_\xi \frac{\int_\epsilon G(s^o_{ia} - w_q\xi + w_q\epsilon_{ia})^{n-1} dF_\epsilon(\epsilon)}{\int_\epsilon (n-1) g(s^o_{ia} - w_q\xi + w_q\epsilon_{ia}) G(s^o_{ia} - w_q\xi + w_q\epsilon_{ia})^{n-2} dF_\epsilon(\epsilon)} dF_\epsilon(\xi)$$

[56] The i.i.d. assumption is based on the random selection of experts for each auction and independent evaluation process.

The new test is

$$H_0 : E[E_\epsilon(k^{rival})|s_a^{win}] = E[E_\epsilon(k_a^{rival})|s_a^{win}]$$
$$\text{and } H_1 : E[E_\epsilon(k^{rival})|s_a^{win}] > E[E_\epsilon(k_a^{rival})|s_a^{win}]$$

Whether there are significant differences in the new test compared to the basic model test depends on how large the variance of the randomness is and whether there is systematic up or downward bias in the evaluation process. If the randomness is small, I will not see dramatically different test results.

## G2 Identification and Data

To conduct the test, I need to know the distribution of errors. Though I don't have a complete dataset including a quality evaluation of each expert for each bidder in every auction, I have a recent subsample from one city in Guangdong Province. The structure of the data is shown in Figure G1. For each bid, there are five randomly selected experts invited to evaluate the quality part of the bids. Each expert independently gives her or his own quality score. The final quality score is the average of the quality score given by the five experts.

Figure G1: An Example of Detailed Expert Evaluation Scores

| Supplier | Detailed Quality Scores | | | | | | | Price Score | Total Score | Order |
|---|---|---|---|---|---|---|---|---|---|---|
| | Name | Name | Name | Name | Name | Sum | Average | | | |
| | Expert 1 | Expert 2 | Expert 3 | Expert 4 | Expert 5 | | | | | |
| Dongjin Co. | 89.65 | 90 | 90 | 85 | 90 | 444.65 | 88.93 | 10 | 98.93 | 1 |
| Biaosheng Co. | 69 | 70.15 | 67.8 | 67.8 | 71.55 | 346.5 | 69.3 | 9.8672 | 79.1672 | 2 |
| Chuangmeng Co. | 53.8 | 59.95 | 53.1 | 53.1 | 56.85 | 272.8 | 54.56 | 9.886 | 64.446 | 3 |

The quality score $q_{iea}$ given by each expert $e$ for bidder $i$ and in auction $t$ is the sum of real quality $q_{ia}$ and the expert-level evaluation error $\epsilon_{iea}$. The identification of evaluation errors comes from a discrepancy in experts' evaluations:

$$q_{iea} - q_{ie'a} = \epsilon_{iea} - \epsilon_{ie'a}$$

The subtraction of quality evaluation from two different experts $e$ and $e'$ for the same bidder $i$ in auction $t$ helps me get rid of the real quality $q_{ia}$. The left is the subtraction between two i.i.d. variables. Therefore $\epsilon_{iea}$ is identified and further $\epsilon_{ia}$ is identified since $\epsilon_{ia} = \frac{\sum_e \epsilon_{iea}}{5}$ and $\epsilon_{iea}$ is i.i.d. The variance of $\epsilon_{ia}$ is much smaller than $\epsilon_{iea}$ because it is the average, $sd(\epsilon_{ia}) = \frac{sd(\epsilon_{iea})}{\sqrt{5}}$.

## G3 Estimation

I plot the histogram of $q_{iea} - q_{ie'a}$ for good, construct, and service procurement separately in Figure G2. The mean of $\hat{\epsilon}_{iea}$ is not different from 0. The standard deviations of $\epsilon_{ia}$ are 0.35 points for good procurement, 1.12 points for service procurement, and 2.29 points for construct procurement. The larger randomness in the evaluation of construction procurement may be due to a more subjective component in the evaluation of implementation plans. The demand and products in good procurement are more standardized compared to construction and service projects, resulting in lower variance.

I then re-estimate the model and conduct hypothesis testing based on the extended model. The results are shown in Table G1. The proportion of procurement auctions that reject the null hypothesis under the extended model is similar to the results in the basic model. For all cases, the results fall within the 95% bootstrapped confidence intervals of the results in the basic model, indicating that randomness does not play a significant role in my setting.

Table G1: Test with Randomness Results

| Proportion of Procurement Reject the Null | | | | |
|---|---|---|---|---|
| | All | 3 Bidders | 4 Bidders | 5 Bidders |
| Unadjusted P-value | 63.43% | 68.28% | 50.5% | 48.23% |
| Adjusted P-value | 57.03% | 62.71% | 44.4% | 38.74% |

**Notes**: The first row represents the results using unadjusted p-values and the second row represents the results using adjusted p-values.

Finally, I reconduct a validation test using expert surveys, and the results are shown in Table G2. The accuracy rate is 80.8%, which is slightly lower than the 81.4% accuracy rate in the basic model. However, the change is not significant.

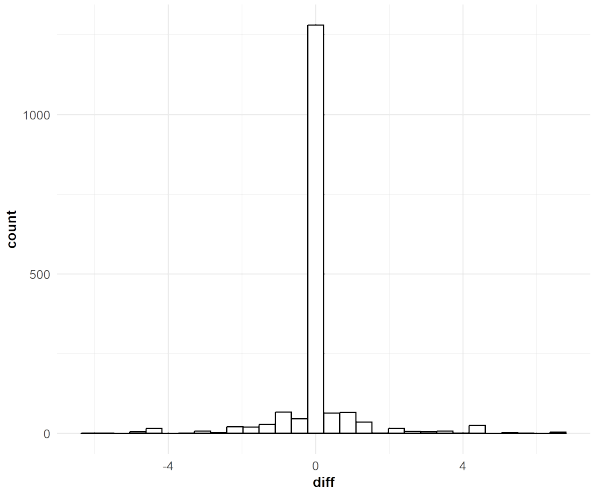
Table G2: Expert Survey and Model Prediction

| Accuracy Rate=75.4% | Total=454 | Model Prediction | | |
|---|---|---|---|---|
| | | Corrupt (PP) | Competitive (PN) | |
| Expert Survey | Corrupt (P) | 285 | 52 False N | Test Power 84.57% |
| | Competitive (N) | 26 False P | 92 | Test Size 22.03% |

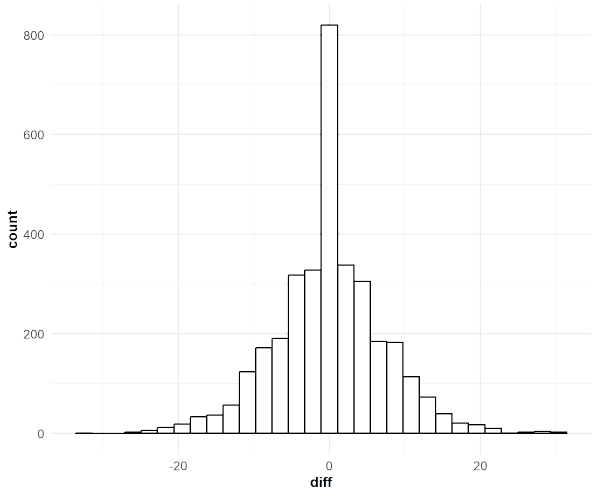
**Notes**: For the expert survey, if any evaluation survey reveals highlighted scoring rules accompanied by expert explanations and concerns, the procurement project is labeled as corrupt. The classification of model prediction uses unadjusted p-values with a significance level of 0.1.

Figure G2: Distribution of $q_{iea} - q_{ie'a}$ by Procurement Categories

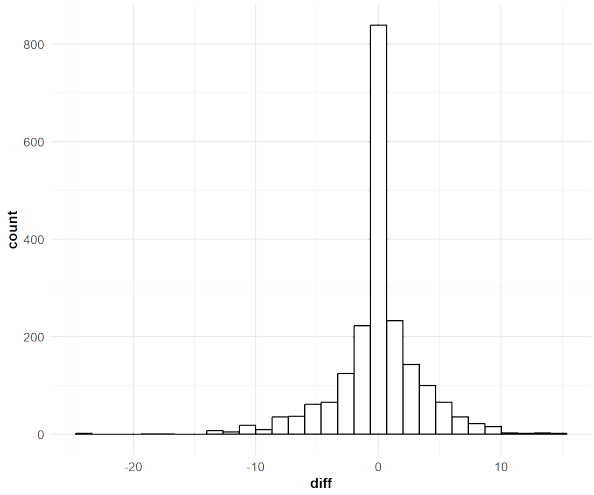
(a) Good Procurement




(b) Construct Procurement




(c) Service Procurement



# Appendix H: Which Firms Participate in Corruption

The impact of corruption on economic growth remains a topic of ongoing debate. One perspective posits that corruption hampers economic progress by hindering various processes and causing resource misallocation among companies (Hsieh and Klenow, 2009; Colonnelli and Prem, 2021; Brugués, Brugués, and Giambra, 2024). Conversely, an opposing viewpoint suggests that, in the presence of bureaucratic red tape, corruption might actually function as a form of "grease in the wheel," smoothing business processes and enhancing overall performance. In this subsection, I delve into what characteristics explain firms' participation in corrupt auctions aiming to provide evidence of resource misallocation due to the selection into corruption relationship.

To investigate this question, I use a Lasso probit model to identify firm and auction characteristics that have explanatory power on rejecting the null hypothesis in the previous test section. Specifically, I collect firm factors such as the location of firms, state ownership of firms, registered capital size, formal employee size, and firm productivity, as measured by Total Factor Productivity (TFP)[57]. Predetermined winners often tend to be located within the same region as the procurement department, and state-owned firms frequently have close relationships with local governments. The relationship between TFP and selection is not yet clear. Some studies suggest that highly efficient firms are more likely to seek corrupt relationships due to their profitability, while others find that connected firms may be less efficient, motivating them to engage in bribery as a means of compensating for their lack of competitiveness.

I also include auction characteristics including price weights, budget, year, and procurement categories. Corrupt auction design usually puts lower price weights since higher price weights lead to more competition on price and reduce the markup. Project with large budget size is more vulnerable to corruption since the profit is higher.

I fit a Lasso-regularized probit model for predicting corrupt auctions as follows:

$$\min_{(\beta_0,\beta)\in\mathbb{R}^{p+1}} -[\frac{1}{N}\sum_{i=1}^{N} Corrupt_i \cdot (\beta_0 + x_i^T\beta) - \Phi(1+e^{\beta_0+x_i^T\beta})] + \lambda[(1-\alpha)||\beta||_2^2/2 + \alpha||\beta||_1]$$

where $Corrupt_i$ is from the test results in Section 6 and has value 1 if the test rejects the null hypothesis, and $x_i^T$ include auction characteristics and firm characteristics, consisting of both winners and losers.[58] The Lasso regularization helps identify words with the strongest predictive power while avoiding over-fitting. The results of this analysis are presented in

[57]The calculation of firm productivity TFP follows the method outlined in Chen et al. (2021) Additional details can be found in Appendix B, Section B2.

[58]I use 75% random sample to train the model and select an optimal $\lambda^*$.

Table H1.

Table H1: Characteristics Predictive of Corruption in Auctions

| | Winner | Loser | Auction Characteristics | |
|---|---|---|---|---|
| Variables | Coefs | Coefs | Variables | Coefs |
| Log(Capital Size) | -0.0316 | -0.0130 | Log(Budget Size) | 0.0647 |
| TFP | -0.0474 | -0.0919 | $w_p$ >Lowest bound | -0.199 |
| Stateowned | 0.223 | 0.0508 | | |
| Public Listed | 0.0867 | -0.0223 | | |
| Foreign Funded | 0.152 | -0.0616 | | |
| Distance to Gov Dept | -0.0145 | -0.0091 | | |
| log(#Group Members) | 0.0067 | -0.0186 | | |
| log(Formal Employee) | 0.0014 | -0.0117 | | |

**Notes**: The firm characteristics are sorted in descending order of coefficients in the winner column.

The second and third columns display the penalized coefficients for the characteristics of winning and losing firms respectively. The characteristics of the firms are arranged in descending order of coefficients in the column denoting winners. Winners in instances of corruption tend to have smaller capital sizes, lower Total Factor Productivity (TFP), and connections to state-owned entities. This suggests that firms that are less competitive are more likely to resort to corruption as a means to secure contracts. Additionally, winners in corrupt scenarios typically exhibit closer ties to local government procurement departments, indicative of local protectionism. Regarding the losers, firms that fail to win contracts in corrupt auctions generally exhibit significantly lower TFP and have smaller capital and workforce sizes. This underscores the presence of numerous non-competitive, or zombie firms in corrupt auctions, seemingly invited solely to satisfy the requirement for a minimum of three bidders. These findings collectively show the various dynamics and implications of corruption within procurement processes.

The final column presents the penalized coefficients corresponding to the characteristics of the auction. The outcomes align consistently with my hypothesis, indicating that projects with substantial budgets are more susceptible to corrupt practices, and the chosen price weight can, to a certain degree, serve as an indicator of corruption. Notably, auctions where the selected price weights exceed the lowest permissible bounds tend to exhibit lower likelihoods of corruption. This consistency reinforces the potential link between budget size,

price weight selection, and the propensity for corrupt practices in auctions.

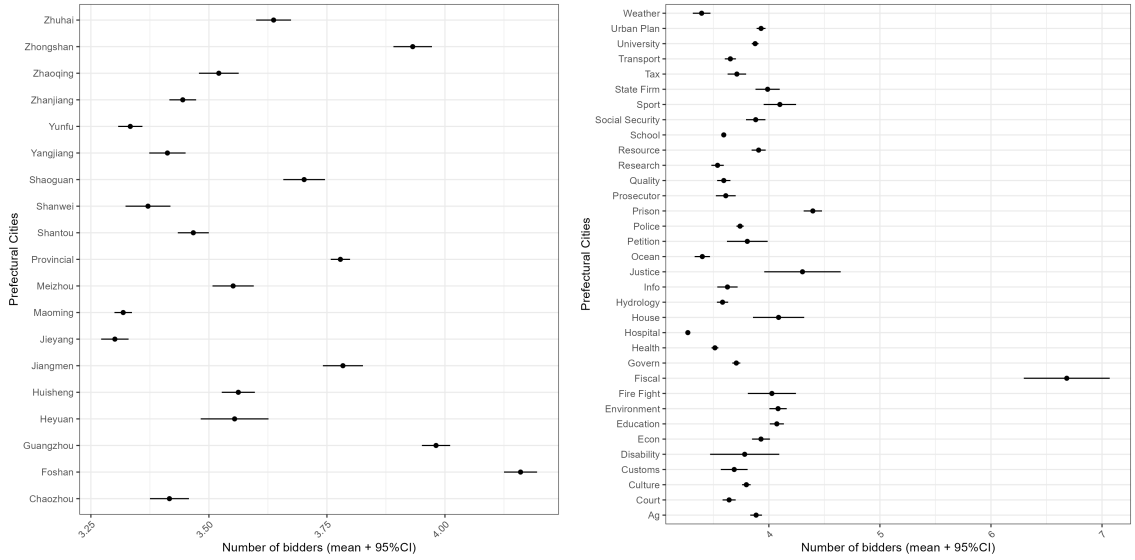
This analysis substantiates that authentic competition is integral in prioritizing efficient firms, whereas corruption induces secondary efficiency losses by disproportionately benefiting less proficient entities.

# Appendix I: City and Public Procurement Corruption

I conducted the statistical test for each public procurement auction and aggregated the auction-level predictions to specific department, county, or prefecture-city levels. This allows me to construct a corrupt index for each level, representing the proportion of public procurement auctions conducted by authorities rejecting the null hypothesis of a competitive auction.

In Figure I1, I plot the mean number of bidders with 95% confidence intervals by city and department. Cities such as Foshan, Zhongshan, and Guangzhou, the capital city of the province, have more bidders on average. Fiscal departments have the least number of bidders, which is consistent with the government's structure. The Fiscal Bureau monitors the procurement conducted by other departments, as it is responsible for local financial work, fiscal revenue organization, expenditure assurance, and the management of local financial funds.

Figure I1: Distribution of Number of Bidders




(a) Number of Bidders: by City

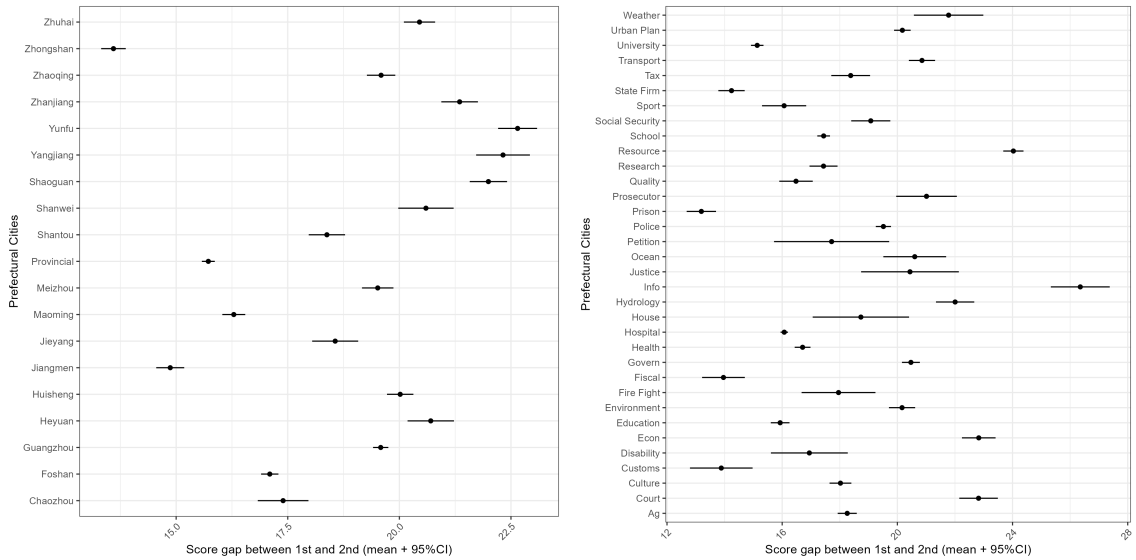
(b) Number of Bidders: by Department

**Notes**: Figure (a) is aggregating the procurement auction level data to the prefecture city level. In Figure (b), I relabeled the procurement departments by their functions, like the education department, the transportation department, the prison, etc.

Next, I plot the distribution of score gaps between winners and the strongest losers by city and department in Figure I2. Similar to the number of bidders, cities like Foshan, Zhongshan, and Guangzhou have smaller winning margins on average.

To measure the corruption level more directly, I construct the corrupt index by city and department in Figure I3. Zhongshan City shows the lowest corruption index, followed by

Figure I2: Distribution of Winning Margins




(a) Winning Margin: by City (b) Winning Margin: by Department

**Notes**: Figure (a) is aggregating the procurement auction level data to the prefecture city level. In Figure (b), I relabeled the procurement departments by their functions, like the education department, the transportation department, the prison, etc.

Jiangmen City and Foshan City. Interestingly, cities performing well in terms of corruption are mostly located around Guangzhou, except for Guangzhou itself.

To understand the city characteristics contributing to the corruption level, I collected city panel variables, including annual GDP, fiscal expenditure, and distance to the provincial capital government. I then ran city characteristics on city public procurement outcomes and the corruption index using the following specification at the city level:

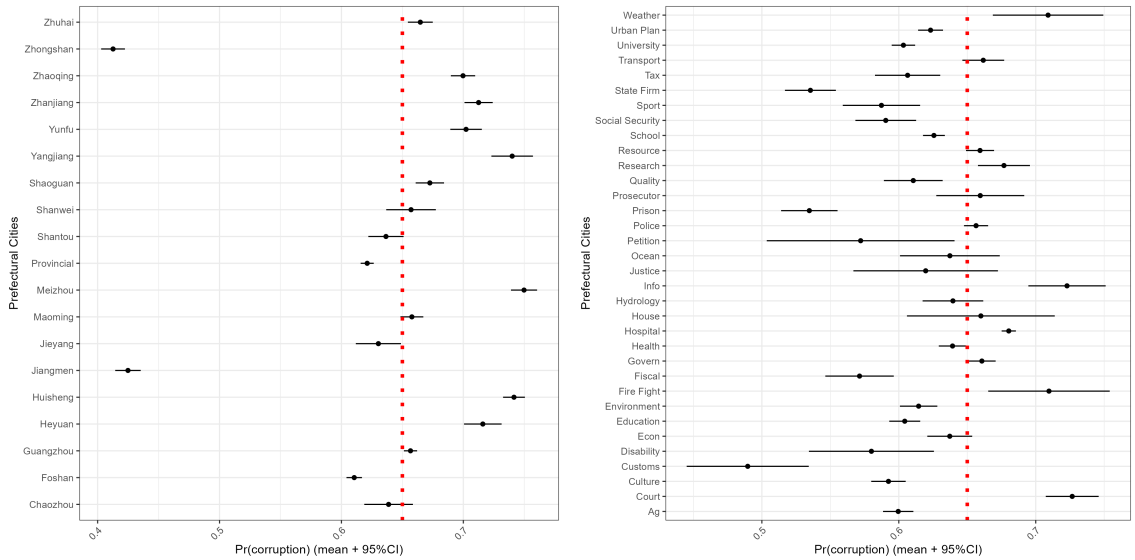
$$y_{ct} = \alpha_t + \beta_1 City\ Characteristics_{ct} + \epsilon_{ct}$$

where $y_{ct}$ represents aggregate public procurement dependent variables in city $c$ at year $t$, including the average number of bidders, winning margins, and the corruption index. I included city-level fiscal expenditure, the previous year's total GDP, GDP per capita, and the mean of public procurement budget size. Year fixed effects, $\alpha_t$, are also controlled. Additionally, I conducted procurement project-level regressions using the specification:

$$y_{act} = \alpha_t + \beta_1 City\ Characteristics_{ct} + \beta_2 Auction\ Characteristics + \epsilon_{ct}$$

where $y_{cat}$ represents the public procurement auction $a$ dependent variables in city $c$ at year $t$, including the number of bidders, score gaps between winners and the strongest losers, and

Figure I3: Distribution of Corruption Index




(a) Corruption Portion: by City (b) Corruption Portion: by Department

**Notes**: Figure (a) is aggregating the procurement auction level data to the prefecture city level. In Figure (b), I relabeled the procurement departments by their functions, like the education department, the transportation department, the prison, etc.

the corruption test results. Procurement auction characteristics, such as public procurement budget size and procurement category fixed effects, are added to control the variation in procurement projects.

The results in Table I1 show consistency between city-level and procurement-level regressions. The further the procurement is conducted away from the province's central government, the higher the corruption index and the higher the probability of rejecting the null hypothesis of real competition in procurement auctions. These results align with the Chinese proverb "The sky is high, and the emperor is far away," which describes a situation where local authorities or officials are not closely monitored or influenced by higher-level government authorities. Additionally, the results suggest that richer cities tend to have lower corruption levels, although this is not necessarily a causal relationship. Higher economic development could lead to more resources for regulation and increased transparency in public procurement. Conversely, lower corruption levels may make cities more attractive to investors and businesses, leading to further economic development.

Table I1: Firm Characteristics and $\theta$ of Bidders

| | City Level | | | Public Procurement Level | | |
|---|---|---|---|---|---|---|
| | Corruption | Score Gap | Bidders | Corruption | Score Gap | Bidders |
| log(Distance to Guangzhou) | 0.0121* | 0.3245* | -0.0424** | 0.0131** | 0.3752** | -0.0658*** |
| | (0.0062) | (0.1755) | (0.0163) | (0.0036) | (0.1432) | (0.0159) |
| Log(Procurement Budget) | 0.0101 | 0.0744 | 0.0800*** | 0.0090*** | -0.0257 | 0.1621*** |
| | (0.0098) | (0.3205) | (0.0297) | (0.0016) | (0.1099) | (0.0372) |
| Log(Fiscal Expenditure) | 0.0341* | 0.9740 | 0.0671 | 0.0135 | 0.2991 | 0.0803** |
| | (0.0201) | (0.6219) | (0.0506) | (0.0154) | (0.5233) | (0.0273) |
| Lag Log(GDP) | -0.0707** | -1.6135 | -0.1731 | -0.1139*** | -3.7184*** | 0.1693 |
| | (0.0318) | (1.2772) | (0.1223) | (0.0203) | (0.6995) | (0.0915) |
| Lag Log(GDP)$^2$ | 0.0021 | -0.0243 | 0.0097 | 0.0064*** | 0.1858*** | -0.0141* |
| | (0.0030) | (0.1100) | (0.0098) | (0.0008) | (0.0298) | (0.0060) |
| Lag Log(GDP per capita) | 0.0077 | 0.9234 | 0.2193*** | -0.0087 | -0.6099 | 0.1668*** |
| | (0.0281) | (0.8032) | (0.0649) | (0.0069) | (0.3904) | (0.0356) |
| Observations | 545 | 625 | 625 | 112981 | 113095 | 113095 |

**Notes**: The first three columns present the estimation for city panel data, while the second three columns display the estimation for public procurement auction data. All regressions control for year-fixed effects. The public procurement regressions also control for procurement category fixed effects and budget price. Standard errors are clustered at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$