A Field Experiment on Antitrust Compliance

Kei Kawai
University of California, Berkeley
& University of Tokyo & NBER

Jun Nakabayashi *†
Kyoto University

April 4, 2024

Abstract

We study the effectiveness of firms' compliance programs by conducting a field experiment in which we disclose to a subset of Japanese firms that the firm is potentially engaging in illegal bid-rigging. We find that the information that we disclose affects the bidding behavior of the treated firms: our test of bid-rigging is less able to reject the null of competition when applied to the bidding data of the treated firms after the intervention. We find evidence that this change is not the result of firms ceasing to collude, however. We find evidence suggesting that firms continue to collude even after our intervention and that the change in the bidding behavior we document is the result of active concealment of evidence by cartelizing firms.

KEYWORDS: Regulatory Compliance, Scoring Auctions, Collusion.

*Contact information, Kawai: kawaixkei@gmail.com, Nakabayashi: nakabayashi.jun.8x@kyoto-u.ac.jp
†We thank Ben Handel, David Weinstein, and seminar audiences at various venues for thoughtful comments. The usual disclaimer applies.

1 Introduction

In many policy areas, there is an increasing trend towards delegating the day-to-day monitoring and enforcement of regulations to the regulated firms themselves (Sigler and Murphy, 1988; Ayres and Braithwaite, 1992). This trend reflects the rapid growth in the scope, scale and complexity of government regulations that have not been matched by commensurate increases in regulatory resources.[1] In many policy areas, regulatory enforcement takes the form of a hierarchical structure in which the firm's compliance function handles much of the routine regulatory violations, while the regulatory agencies are primarily charged with overseeing the compliance function of the firms.[2] According to one estimate, regulatory compliance now accounts for about 1.34% to 3.33% of the total wage bill of U.S. firms (Trebbi and Zhang, 2023).

While the trend towards delegating monitoring and enforcement functions to firms have been especially salient for policy areas such as investor protection[3] and environmental protection,[4] this trend is also starting to affect areas such as antitrust which have traditionally had a strong emphasis on direct regulatory enforcement.[5] For example, implementation of antitrust compliance programs and designation of antitrust compliance officers have become key components of many consent decrees.[6] Moreover, in some recent antitrust violation

[1]See e.g., Davis (2017) for a discussion of the rapid growth in government regulations in the U.S.

[2]The regulatory hierarchy is described in Ayres and Braithwaite (1995) as follows: "achievement of regulatory objectives is more likely when agencies display both a hierarchy of sanctions and a hierarchy of regulatory strategies of varying degrees of interventionism. The regulatory design requirement we describe is for agencies to display two enforcement pyramids with a range of interventions of ever-increasing intrusiveness (matched by ever-decreasing frequency of use). Regulators will do best by indicating a willingness to escalate intervention up those pyramids or to deregulate down the pyramids in response to the industry's performance in securing regulatory objectives."

[3]For example, the Bank Secrecy Act and the Dodd-Frank Act require financial institutions to designate a compliance officer who is responsible for implementing compliance programs within the firm. The compliance officer is required to certify to the regulators that the firm is in compliance with all regulations.

[4]See, e.g., the EPA's audit policy, formally titled "Incentives for Self-Policing: Discovery, Disclosure, Correction and Prevention of Violations". According to the EPA, the audit policy provides "several major incentives for regulated entities to voluntarily discover, self-report and correct violations of federal environmental laws and regulations", "making formal EPA investigations and enforcement actions unnecessary."

[5]See, e.g., Sokol (2017, 2012). Sokol (2017) discusses the literature on antitrust compliance including both work that emphasizes the role of compliance in antitrust as well as those that advocate for a more limited role.

[6]See, e.g., US v. DirecTV Group Holdings and AT&T (2017), US v. Charleston Area Medical Center

cases, the U.S. DOJ has started "seeking court-supervised probation as a means of assuring that the company devises and implements an effective compliance program"(Assistant Attorney General Baer, 2014).[7] The DOJ's emphasis on effective compliance, treating it as a goal in and of itself, departs from the traditional regulatory model that focuses on ex-post punishment of violations.

Although compliance functions within firms have become an important part of antitrust enforcement, and of regulation of firms more generally, there is still relatively limited evidence on the effectiveness of within-firm compliance functions. In this paper, we provide empirical evidence on one aspect of regulatory compliance, specifically, the extent to which firms can take remedial action when confronted with evidence of illegal activity. How firms respond to evidence of regulatory violations—whether firms take steps to end wrongdoing, ignore evidence and continue to engage in wrongdoing, or actively seek to conceal incriminating evidence—can shed light on how best to incorporate within-firm compliance into the regulatory environment. It can also shed light on the effectiveness of various policy tools that seek to guide firm behavior without being legally binding, such as voluntary programs and regulatory guidelines. These policy levers require at least some level of voluntary compliance from firms to be effective.

In order to study how firms respond to evidence of illegal activity, we conduct a field experiment in which we disclose to a set of construction firms in Japan evidence suggesting that they are potentially engaging in illegal bid-rigging.[8] We first develop a statistical test

---

and St. Mary's Medical Center (2016), etc.

[7]See, for example, the following remarks by Bill Baer, the Assistant Attorney General of the Antitrust Division on Sept 10, 2014.

> We also expect companies to take compliance seriously once they have pleaded guilty or have been convicted. Taking compliance seriously includes making an institutional commitment to change the culture of the company. Companies should be fostering a corporate culture that encourages ethical conduct and a commitment to compliance with the law.
>
> In such cases, the division will consider seeking court-supervised probation as a means of assuring that the company devises and implements an effective compliance program. We reserve the right to insist on probation, including the use of monitors, if doing so is necessary to ensure an effective compliance program and to prevent recidivism.

[8]Bid-rigging in procurement auctions is illegal in Japan and firms face fines of up to 10 % of all relevant sales in the affected market, in addition to the direct monetary gains that results from collusion. Individuals

of bid-rigging for scoring auctions which we apply to bidding data from auctions let by the Ministry of Land Infrastructure and Transportation in Japan. We identify 240 firms whose bidding behavior is inconsistent with competitive bidding. We then subject a random subset of these firms to a treatment in which we send out a letter that explains the statistical test we developed and the outcome of the test for the firm receiving the letter (i.e., competition is rejected). We compare the subsequent bidding behavior of the treated and the control firms to identify how firms respond to evidence implicating them of bid-rigging.

Our first finding is that the treated firms change their bidding behavior in such a way that our statistical test of bid-rigging becomes less able to reject the null of competition. Using Fisher's randomization test, we find evidence against the strong null hypothesis that the treatment induces no change in firm bidding behavior with respect to the ability of our statistical test to detect collusion.

Our second finding is that the change in the treated firms' bidding behavior is likely to be the result of an adaptive response by the firms to evade detection without stopping collusion. First, we do not find significant decreases in the level of bids or increases in the quality of the proposals—changes typically associated with a shift from collusion to competition. Moreover, we find a significant increase in the frequency with which losing bidders submit invalid bids, i.e., bids that are higher than the reserve price. Because we study procurement auctions, bids above the reserve price have no chance of winning. These bidding patterns reflect, if anything, a more blatant form of collusion in which designated losers effectively stop participating.

Overall, the results of our experiment suggest that, at least for the subset of construction companies participating in procurement auctions in Japan, existing levels of compliance capacity within colluding firms are unlikely to complement formal regulatory actions in achieving regulatory compliance.[9] These findings are somewhat different from previous

involved in bid rigging can face up to 3 years in prison.

[9]It should be noted that in Japan, there is no equivalent of *qui tam* law suits in which private parties are compensated for assisting governments recover damages from illegal activities. For example, in the U.S., the False Claims Act allows private parties to bring suit on behalf of the government and receive a portion of the damages recovered. See, e.g., Kovacic (2001), and Engstrom (2013) for analysis on *qui tam* law suits. On the topic of private enforcement of laws more generally, see, e.g., Landes and Posner (1975) and Polinsky

studies that document firms stopping illegal behavior soon after allegations of wrongdoing are made (See, e.g., Christie et al., 1994 for the NASDAQ collusion case, and Monticini and Thornton, 2013, for the LIBOR manipulation case). Our findings suggest that, in the absence of widespread publicity or impending regulatory action, internal compliance functions may be ineffective at changing firm behavior.

The results of the experiment also shed some light on the value of using behavioral screens when firms adapt. In our experiment, the cartels seemed to adapt by instructing designated losers to submit invalid bids. A bidding pattern in which qualified bidders refrain from participating, however, is a well-known indicator of collusion.[10] Hence, multiple screens of collusion can be complementary. Moreover, we argue below that adapting to our screen necessitated the cartel to either reduce the winning bid or have designated losers bid above the reserve price. Our findings thus support the point made in previous work (e.g., Porter, 2005; Harrington, 2008a; Marshall and Marx, 2012) that screens can make cartels less profitable and more costly to maintain even when firms can adapt, especially if multiple screens are applied together.

Related literature. There is a large literature in law and organizational behavior that analyzes how within-firm compliance functions complement the work of regulatory agencies, for example, Braithwaite (1985), Ayres and Braithwaite (1995), Parker (2002), and Sokol (2012). There is also a small theoretical literature on regulatory compliance in economics that studies the trade-off between regulatory capture and efficient use of private information, e.g., Gehrig and Jost (1995) and Grajzl and Murrell (2007). In these models, firms are better informed about the business environment than the regulators making firms better positioned to identify the types of regulations that are efficient. This informational asymmetry makes it potentially more efficient to delegate rule-setting to firms. However, given the obvious conflict of interest, the regulator may not want to delegate rule-setting and enforcement

(1980).

[10] See, e.g., the document published by the U.S. DOJ titled "Preventing And Detecting Bid Rigging, Price Fixing, And Market Allocation In Post-Disaster Rebuilding Projects".

entirely to the firms.[11]

Our paper is also related to those that study firm adaptation in environments where regulatory agencies use screens to select the set of firms that receive additional scrutiny.[12] For example, Wollmann (2019) and Cunningham et al. (2021) document evidence that firms adapt to the merger notification threshold set by the Hart-Scott-Rodino Act. The possibility of firm adaptation implies that the design of screens should account for the firms' equilibrium response. The importance of taking an equilibrium view of regulatory screening has been made previously by Cyrenne (1999), LaCasse (1995), Harrington (2004), Ortner et al. (2022), etc.

Previous work examining how firms react to evidence of incriminating evidence include Christie et al. (1994) and Monticini and Thornton (2013). Christie et al. (1994) document immediate changes in the quotes offered by market makers in NASDAQ after newspapers reported potential collusion by dealers. Monticini and Thornton (2013) find evidence of banks stopping underreporting of LIBOR rates after a newspaper reported potential manipulation of the rates. For both the NASDAQ collusion case and the LIBOR manipulation case, there was substantial publicity generated by the news reports. In contrast, the firms in our study were not subject to media exposure.

Lastly, our paper is related to the literature on detecting cartels in auctions. Early seminal work includes Hendricks and Porter (1988), Baldwin et al. (1997) and Porter and Zona (1993, 1999). More recent work includes Bajari and Ye (2003), Abrantes-Metz et al. (2006), Athey et al. (2011), Conley and Decarolis (2016), Schurter (2017), Kawai and Nakabayashi (2022), Chassang et al. (2022), Martin et al. (2022), Kawai et al. (2023), Seibel and Škoda (2023) and Baránek et al. (2023).[13] The strategy for cartel detection we adopt in this paper extends

[11]Other related work include Innes (1999), who studies remediation activities that are offered voluntarily by violators and Kaplow and Shavell (1994), who study self-reports of violations by perpetrating firms. There is also a literature that studies the efficacy of leniency programs and the incentives they create for firms to report their involvement in collusion, e.g., Motta and Polo (2003), Aubert et al. (2006), Spagnolo (2005), Chen and Harrington (2007), Harrington (2008b) and Miller (2009).

[12]Relatedly, Blundell et al. (2020) study a dynamic environment in which regulatory violations lead to heightened scrutiny.

[13]Other related work includes Pesendorfer (2000), who studies bidding rings with and without side-payments, and Asker (2010), who studies knockout auctions among cartel members. Ohashi (2009) and

ideas proposed in Kawai et al. (2023) to scoring auctions.[14]

2 Institutional Background and Auction Format

Our paper analyzes the bidding behavior of firms that participate in auctions let by the Ministry of Land Infrastructure and Transportation (MLIT) in Japan. This section provides a brief description of the institutional background and the auction format used by the MLIT.

MLIT is the largest procurement buyer in Japan, letting in each year about 9,000 auctions for construction projects, worth a combined total of about 1.7 trillion yen (about $17 billion USD). The range of projects let by the MLIT includes road paving, building and repairing bridges, installation of electrical equipment and other machinery as well as civil engineering work.

Since around 2006-7, almost all MLIT auctions are let using scoring auctions. In a scoring auction, each bidder, in addition to the price, submits a proposal which is converted into a scalar quality measure, q. Allocation is determined by each bidder's score, s, an index that combines both price (p) and quality (q) of the bid. The MLIT transforms (p, q) into a score according to the following rule:

$$s = q/p.$$

Bidders submit sealed bids simultaneously and the project is allocated to the bidder with the highest score, subject to the secret reserve price.

Depending on the type of project, the quality component of the bid can be more or less predictable. For simple projects, the quality measure is a deterministic function of observable characteristics of the firm, such as how many similar projects the firm has completed in the past. For complicated projects, however, the quality of the proposals play a more important role in determining q. The range of quality measures a firm can obtain is typically between

Chassang and Ortner (2019) document how changes in auction design can affect the ability of bidders to sustain collusion. Clark et al. (2018) analyze the breakdown of a cartel and its implications on prices. For a survey of the literature, see Porter (2005) and Harrington (2008a).

[14] See e.g., Che (1993), Asker and Cantillon (2008, 2010), etc. for analysis of competitive equilibrium in scoring auctions.

$[100, 150]$. The lower bound of the quality measure is always fixed at $q = 100$, while the upper bound can be higher or lower depending on the complexity of the project.

One institutional detail that is worth mentioning at this point is that when the price exceeds the secret reserve price, the quality of the bid is either not recorded, or is assigned the lowest possible measure of 100. This fact explains certain features of the data pattern that we document in our analysis below.

3 Model and Test Statistic

This section specifies a model of scoring auctions and derives results that we use to construct tests for competitive bidding. Our model is similar to those of Che (1993), and Asker and Cantillon (2008, 2010). The results that we derive are similar to those in Kawai et al. (2023).

To preview the results, Corollary 1 states that, under the null of competition, conditional on the set of auctions in which two bidders almost tie for first place (i.e., have very similar scores), the price of the marginal winner is equal to the price of the marginal losers, on average. This result should be intuitive: for any bidder, conditional on winning or losing by a very small margin, the bidder cannot systematically more likely to have been on the winning side or on the losing side. The bidder should be a marginal winner with 50% probability and a marginal loser with 50% probability. In other words, winning or losing is as-if random conditional on being almost tied for first place. This implies that the average price (or quality, or any other characteristic) should be the same between marginal winners and losers, on average.[15]

Similarly, Corollary 1′ states that, conditional on the set of auctions in which two firms are almost tied for submitting the lowest price, the scores must also be similar between the bidder submitting the lowest bid and the bidder submitting the second lowest bid, on average. We use these results to test for competition: we test (1) whether or not the price of marginal losers are the same as that of the marginal winners, on average; and (2) whether or not the scores of bidders who marginally bid the lowest price are the same, on average as those of

[15] As we explain below, our results do not require that the firms be symmetric.

bidders who marginally bid above the lowest price. If there are systematic differences, it will be evidence against competition.

Game form. A buyer procures a single item from a finite set N of potential suppliers. The procurement contract is allocated through a sealed-bid auction with a secret reserve price r, which is drawn from a distribution F_r. Each potential bidder $i \in N$ decides whether or not to participate in each auction. Conditional on participation, a bidder incurs participation cost $k > 0$, and submits a bid $\mathbf{b}$ consisting of a price-quality pair $\mathbf{b} = (p, q)$. Profit from non-participation is normalized to 0. A bid is valid if the price is below the reserve price, i.e., $p \leq r$. The bidder who submits a valid bid with the highest score is allocated the project, where the score is computed according to the formula $s = q/p$. Ties are broken with uniform probability. We denote by $\vee \mathbf{s}_t$ the highest score among participating bidders in auction t, by $\mathbf{s}_{-i,t} \equiv (s_{j,t})_{j \neq i}$ the scores of firms other than i, and by $\vee \mathbf{s}_{-i,t} \equiv \max_{j \neq i} s_{j,t}$ the highest score among i's participating competitors. Let $\vee \mathbf{s}_{-i} \prec s_i$ denote the event that bidder i wins the contract, i.e. s_i is the highest score and possible ties are broken in favor of bidder i. Bids are publicly revealed at the end of each period.

The bidder's profit conditional on winning is given by $p - C(q, \theta)$, where θ is the privately-known cost type of the bidder. The function $C(\cdot, \theta)$ represents the cost of providing higher quality for each type θ. In practice, the cost function C depends on observable auction characteristics as well, but we suppress this dependence. We assume that $C(\cdot, \theta)$ is increasing, convex and continuously differentiable for each θ.

Information. Each bidder i privately observes a signal $z_i \in Z_i$. We do not impose any assumptions on the distribution of the signal profile $\mathbf{z}_t = (z_{i,t})_{i \in N} \in Z = \prod_{i \in N} Z_i$. Signals may be arbitrarily correlated. The distribution of z_i can be asymmetric. We denote by $F_Z(\cdot)$ the joint distribution of the signals.

Cost types $\theta = (\theta_i)_{i \in N} \in \mathbb{R}^N$ are drawn independently conditional on each private signal z_i, i.e.,

$$\theta_i | z_i \sim \theta_i | \mathbf{z}_t, \theta_{-i}.$$

Bidder i's cost type does not provide information about the cost type of other bidders beyond the information already provided in the private signal z_i.[16] This class of information structures nests asymmetric independent private values, correlated values, and complete information. We denote by $F_\theta(\cdot|\mathbf{z}_t)$ the conditional distribution of the profile of cost types θ given signals $\mathbf{z}$.

Indirect Profit Function Because any bid (p, q) that yields the same score s guarantees the bidder the same winning probability, optimal bidding behavior requires a bid (p, q) with $p/q = s$ to be the solution to the following maximization problem:

$$\pi_i(s, z_i) = \max_{p,q} \; p - \mathbb{E}[C(q, \theta_i)|z_i]$$
$$\text{s.t. } q/p = s,$$

where the expectation $\mathbb{E}$ is taken with respect to the distribution of θ_i conditional on z_i.[17] The objective function, $p - \mathbb{E}[C(q, \theta_i)|z_i]$, is the profit the firm obtains if it wins. We denote the solution of this problem as $\pi_i(s, z_i)$. The function $\pi_i(s, z_i)$ corresponds to the indirect profit function of the firm when it bids a score of s and its signal is z_i. Note that $\pi_i(s, z_i)$ is continuously differentiable in s, by the envelope theorem.[18]

Competitive Bidding We now derive implications of Bayes Nash equilibrium that we use to construct our test of competitive bidding. Our first result establishes that conditional on being a marginal winner or a marginal loser, any bidder believes that they win with probability greater than 50%.

Proposition 1. *For any bidder i and for any positive number $\eta > 0$, there exists ε such that*

$$\textit{prob}(i \textit{ wins} \mid z_i, \; |s_i - \vee \mathbf{s}_{-i}| < \varepsilon) \geq 1/2 - \eta.$$

[16] Because the signals are allowed to be correlated, $z_{i,t}$ helps bidder i predict the cost of other bidders.

[17] If bidder i's cost, θ_i, is part of the private signal, z_i, the expectation would be unnecessary.

[18] Note that $\pi_i'(s, z_i) < 0$ for any s.

Proof. Let $D_i(s_i|z_i)$ denote $\mathsf{prob}(s_i \succ \vee \mathbf{s}_{-i}|z_i)$, the probability of winning the auction conditional on signal z_i. Incentive compatibility of the bidder implies

$$D_i(s_i|z_i)\pi_i(s_i, z_i) \geq D(s_i + \varepsilon|z_i)\pi_i(s_i + \varepsilon, z_i)$$

and

$$D_i(s_i|z_i)\pi_i(s_i, z_i) \geq D(s_i - \varepsilon|z_i)\pi_i(s_i - \varepsilon, z_i).$$

Noting that $D_i(s_i|z_i)$ must be strictly positive and continuous in equilibrium,[19]

$$\begin{aligned}
\mathsf{prob}(i \text{ wins} \mid z_i,\ |s_i - \vee \mathbf{s}_{-i}| < \varepsilon) &= \frac{D_i(s_i|z_i) - D_i(s_i - \varepsilon|z_i)}{D_i(s_i + \varepsilon|z_i) - D_i(s_i - \varepsilon|z_i)} = \frac{1 - \frac{D_i(s_i - \varepsilon|z_i)}{D_i(s_i|z_i)}}{\frac{D_i(s_i+\varepsilon|z_i)}{D_i(s_i|z_i)} - \frac{D_i(s_i-\varepsilon|z_i)}{D_i(s_i|z_i)}} \\
&\geq \frac{1 - \frac{\pi_i(s_i,z)}{\pi_i(s_i-\varepsilon,z)}}{\frac{D_i(s_i+\varepsilon|z_i)}{D_i(s_i|z_i)} - \frac{\pi_i(s_i,z)}{\pi_i(s_i-\varepsilon,z_i)}} \geq \frac{1 - \frac{\pi_i(s_i,z_i)}{\pi_i(s_i-\varepsilon,z_i)}}{\frac{\pi_i(s_i,z_i)}{\pi_i(s_i+\varepsilon,z_i)} - \frac{\pi_i(s_i,z_i)}{\pi_i(s_i-\varepsilon,z_i)}} = \frac{\frac{1}{\pi_i(s_i,z_i)} - \frac{1}{\pi_i(s_i-\varepsilon,z_i)}}{\frac{1}{\pi_i(s_i+\varepsilon,z_i)} - \frac{1}{\pi_i(s_i-\varepsilon,z_i)}} \\
&= \frac{\frac{d}{ds}\left(\frac{1}{\pi_i(s_i,z_i)}\right) \times \varepsilon + o(\varepsilon)}{\frac{d}{ds}\left(\frac{1}{\pi_i(s_i,z_i)}\right) \times 2\varepsilon + o(\varepsilon)}
\end{aligned}$$

As $\pi_i(s_i, z_i)$ is continuously differentiable and $\pi'(s_i, z_i) < 0$, the last term converges to 0.5 as $\varepsilon \to 0$. □

Proposition 1 shows that the winning probability must not be much lower than 1/2 conditional on close bids.[20] As the next proposition shows, because at most one bidder can win, and because there are at least two close bidders conditional on the existence of close bids, it cannot be that firms' winning probability is strictly larger than 1/2. For any $\epsilon > 0$, let ϵ-close denote the event that the winning bid of the auction is within ϵ of the next highest score. For any $\varepsilon > 0$, let $\mathbb{E}[\cdot|\, \epsilon\text{-}\mathsf{close}]$ denote the expectation over $\mathbf{z}$ conditional on the event ϵ-close.

[19] Recall that participation costs k are strictly positive, which implies that $D_i(s_i)$ needs to be strictly positive. If $D_i(s_i)$ is not continuous at $s_i = s_0$, then there exists some bidder j who bids exactly s_0 with positive probability. In this case, all bidders other than j will have no incentive to bid in a small interval right below s_0, which in turn, makes bidding s_0 with positive probability suboptimal for bidder j. In other words, bidder j can gain strictly by bidding slightly below s_0. Hence, we have a contradiction.

[20] This proof is essentially the same as that of Proposition 1 of Kawai et al. (2023).

Proposition 2 (as-if random bids)**.** *For all $\eta > 0$ there exists $\epsilon > 0$ small enough such that*

$$\mathbb{E}\left[\left|\textsf{\textit{prob}}(i\ wins \mid z_i\ and\ |s_i - \vee \mathbf{s}_{-i}| < \epsilon) - \frac{1}{2}\right| \ \middle|\ \epsilon\textsf{\textit{-close}}\right] \leq \eta. \tag{1}$$

Proof. See Appendix. □

In words, Proposition 2 states that winning is as-if random conditional on close bids under competition. This result motivates our regression discontinuity test of competition. The following Corollary formalizes the link between equilibrium bidding and our regression discontinuity test.

Corollary 1. *Assume that bidders are playing a Bayes Nash Equilibrium. Then, for any smooth function $f : (p, q) \mapsto f(p, q) \in \mathbb{R}$,*

$$\lim_{\epsilon \searrow 0^+} |\mathbb{E}[f(p_i, q_i) \mid s_i - \vee \mathbf{s}_{-i} \in (0, \epsilon)] - \mathbb{E}[f(p_i, q_i) \mid s_i - \vee \mathbf{s}_{-i} \in (-\epsilon, 0)]| = 0.$$

In particular, if $f(p_i, q_i) = p_i$,

$$\lim_{\epsilon \searrow 0^+} |\mathbb{E}[p_i \mid s_i - \vee \mathbf{s}_{-i} \in (0, \epsilon)] - \mathbb{E}[p_i \mid s_i - \vee \mathbf{s}_{-i} \in (-\epsilon, 0)]| = 0. \tag{2}$$

.

Proof. See Appendix. □

The first part of Corollary 1 guarantees that, under the null of competition, the conditional expectation of $f(p_i, q_i)$, when bidder i is a marginal winner (i.e., $s_i - \vee \mathbf{s}_{-i} \in (0, \epsilon)$), is the same as the conditional expectation of $f(p_i, q_i)$ when bidder i is a marginal loser (i.e., $s_i - \vee \mathbf{s}_{-i} \in (-\epsilon, 0)$). In particular, setting $f(p_i, q_i) = p_i$, the second part of the Corollary 1 claims that the price of the marginal winner and the price of the marginal loser must be the same, in expectation. If, to the contrary, the conditional expectation of the marginal winner and the marginal loser is different, we can reject the null that bidder i is bidding competitively. We use expression (2) to test for competitive bidding.

Remark on Corollary 1 Corollary 1 claims that the average price of marginal winners is the same as those of the marginal losers. This result does not rely on the bidders being symmetric. The result holds even with asymmetric bidders because, for every bidder, there is an equal probability of being the marginal winner and being the marginal loser (expression (1)). When computing the average price of the marginal winner and the marginal loser, each bidder contributes equally to the former and the latter.

Instead of comparing marginal winners and marginal losers (as does expression (2)), we can also compare bidders who submit the lowest bid ($\wedge\mathbf{p}$) and bidders who bid just above it ($\wedge\mathbf{p}+\epsilon$)). The following Corollary states that, with the additional restriction that the distribution of $\{p_i\}_{i\in N}$ is smooth, we can replace the conditioning set $s_i - \vee\mathbf{s}_{-i} \in (-\epsilon, 0)$ and $s_i - \vee\mathbf{s}_{-i} \in (0,\epsilon)$ in expression (2) with $p_i - \wedge\mathbf{p}_{-i} \in (0,\epsilon)$ and $p_i - \wedge\mathbf{p}_{-i} \in (-\epsilon, 0)$.

Corollary 1′. *Suppose that bidders are playing a Bayes Nash Equilibrium. Suppose the distribution of* $\min_{j\neq i} p_j | z_i$ *admits a continuous density function. Then, for any smooth function* $f : (p, q) \mapsto f(p, q) \in \mathbb{R}$,

$$\lim_{\epsilon \searrow 0^+} \left| \mathbb{E}\left[f(p_i, q_i) \,|\, p_i - \wedge\mathbf{p}_{-i} \in (0,\epsilon) \right] - \mathbb{E}\left[f(p_i, q_i) \,|\, p_i - \wedge\mathbf{p}_{-i} \in (-\epsilon, 0) \right] \right| = 0.$$

In particular, if $f(p_i, q_i) = p_i/q_i$,

$$\lim_{\epsilon \searrow 0^+} \left| \mathbb{E}\left[s_i \,|\, p_i - \wedge\mathbf{p}_{-i} \in (0,\epsilon) \right] - \mathbb{E}\left[s_i \,|\, p_i - \wedge\mathbf{p}_{-i} \in (-\epsilon, 0) \right] \right| = 0. \tag{2′}$$

.

Proof. See Appendix. □

Expression (2′) states that the score of the bidder who submits the lowest bid must be the same as the score of the bidder who bids just above the lowest bid, on average. In our empirical exercise, we use both (2) and (2′) to test for competition.

4 Identifying Firms That Do Not Bid Competitively

Illustration: Case of Nippo Corporation We now illustrate how to use the results obtained in the previous section to test for competitive bidding. We do so by focusing on the bidding pattern of Nippo Corporation, a major paving company which participated in a bid-rigging scheme that was discovered and prosecuted by the Japanese competition authority (JFTC).[21] We illustrate the usefulness of our test by partitioning the sample of auctions to those let during the period in which the bidders were likely to have been colluding (Fiscal year 2012) and to those that were let in subsequent years (Fiscal year 2013-2015), and applying our test separately.[22]

Our first test of competition is based on whether or not there exists a significant price difference between marginal winners and marginal losers (Corollary 1). In order to implement the test, for each auction t in which Nippo Corporation participated, let $\Delta^s_{i,t}$ denote the margin of defeat for each non-winner i. Specifically, we define $\Delta^s_{i,t} = \frac{s_{i,t}-s_{i^*,t}}{s_{i^*,t}}$, where i^* denotes the winner of the auction, and $s_{i,t}$ and $s_{i^*,t}$ denote the scores of firms i and i^*.[23] We define $\Delta^s_{i,t}$ for each non-winner $i(\neq i^*)$, and it corresponds to the margin of defeat, measured as a percentage of the winner's score.[24] It is negative by construction, and is close to 0 for marginal losers. We also define $\Delta^p_{i,t} = \frac{p_{i,t}-p_{i^*,t}}{p_{i^*,t}}$, the price difference between bidder i and the winner, measured as a fraction of the winner's price. Corollary 1 implies that, as $s_i - s_{i*} \to 0$, we have $\mathbb{E}[p_i - p_{i^*}] \to 0$. We test for competition by testing whether or not $\mathbb{E}[\Delta^p_{i,t}]$ converges to 0, as $\Delta^s_{i,t}$ converges to 0.

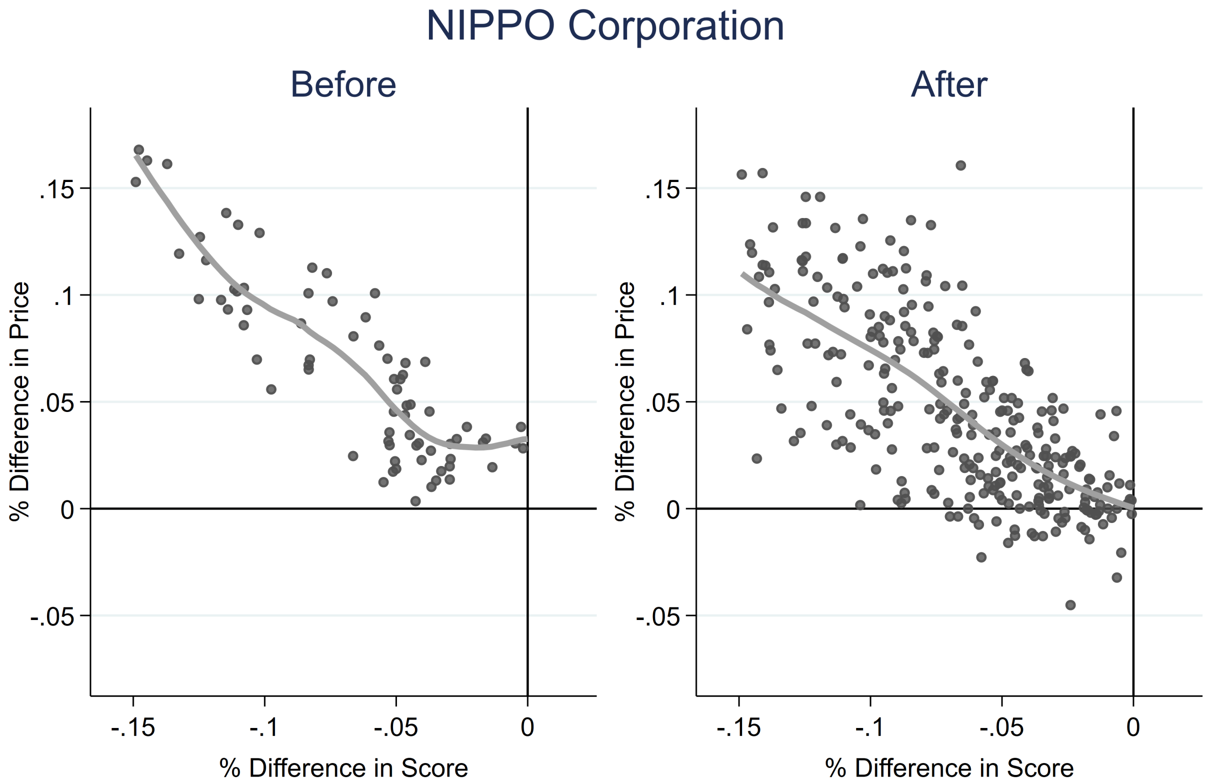
Figure 1 is a binned scatter plot of $(\Delta^s_{i,t}, \Delta^p_{i,t})$ for the collusive period (left panel) and the post-period (right panel). Figure 1 also plots the global fourth order mean regression of $\Delta^p_{i,t}$

[21] The firms were prosecuted for colluding on paving auctions let by NEXCO-East, a publicly owned operator of highways in eastern Japan. Although the bidders were not prosecuted for auctions let by the MLIT, there were news reports suggesting that they may have been colluding on MLIT auctions as well (for example, see Sankei Shimbun Jan. 30, 2015.)

[22] The partitioning is based on an article in Shukan Kinyobi dated November 30, 2012 in which the JFTC is reported as launching an investigation into paving firms in the Tohoku region for collusion (Shukan Kinyobi 20 (46), November 30, 2012, pages 32-33).

[23] There were no ties for the highest score in the sample.

[24] We normalize the margin by the winner's score in order to make comparable auctions with low and high maximum scores.



Left panel corresponds to the auctions in which NIPPO participated before being investigated for collusion. Right panel corresponds to auctions after the investigation.

Figure 1: Binned scatter plot of $\Delta^s_{i,t}$ and $\Delta^p_{i,t}$.

on $\Delta^s_{i,t}$. Notice that, in the left panel, the mean regression does not pass through the origin, as we have $\lim_{\epsilon \searrow 0^+} \mathbb{E}\left[\Delta^p_{i,t} | \Delta^s_{i,t} \in (-\epsilon, 0)\right] > 0$. In other words, $\Delta^p_{i,t}$ does not converge to 0 as $\Delta^s_{i,t}$ converges to 0. On the other hand, in the right panel of Figure 1, the mean regression goes through the origin.

Panel (A) of Table 1 reports our local linear estimates of $\mathbb{E}\left[\Delta^p_{i,t} | \Delta^s_{i,t}\right]$ at $\Delta^s_{i,t} = 0$ for the pre-period (first column) and the post period (second column).[25] We find that, for the sample of auctions before the investigation, the intercept of $\mathbb{E}\left[\Delta^p_{i,t} | \Delta^s_{i,t}\right]$ at $\Delta^s_{i,t} = 0$ is estimated to be 0.037, and statistically different from 0 at the 95% confidence level. Hence, we can reject

[25] Our estimates are obtained using a local linear regression with a coverage error rate optimal bandwidth and a triangular kernel with a bias correction procedure as proposed in Calonico et al. (2014). We obtain standard errors by clustering at the auction level.

the null hypothesis that $\lim_{\epsilon \searrow 0^+} \mathbb{E}\left[\Delta^p_{i,t} | \Delta^s_{i,t} \in (-\epsilon, 0)\right] = 0$ with 95% confidence. On the other hand, for the sample of auctions after the investigation, our estimate of the intercept is 0.002, and statistically indistinguishable from 0.

	(1) Before	(2) After
Panel (A) : $\mathbb{E}\left[\Delta^p_i | \Delta^s_i = 0\right]$		
$\hat{\beta}$	0.037^{***}	0.002
	(0.006)	(0.004)
Bandwidth $\hat{h}$	0.019	0.025
Obs.	73	324
Panel (B) : $\mathbb{E}\left[\tilde{\Delta}^s_i | \tilde{\Delta}^p_i = 0\right]$		
$\hat{\beta}$	-0.042^{***}	-0.014
	(0.007)	(0.015)
Bandwidth $\hat{h}$	0.027	0.039
Obs.	73	276

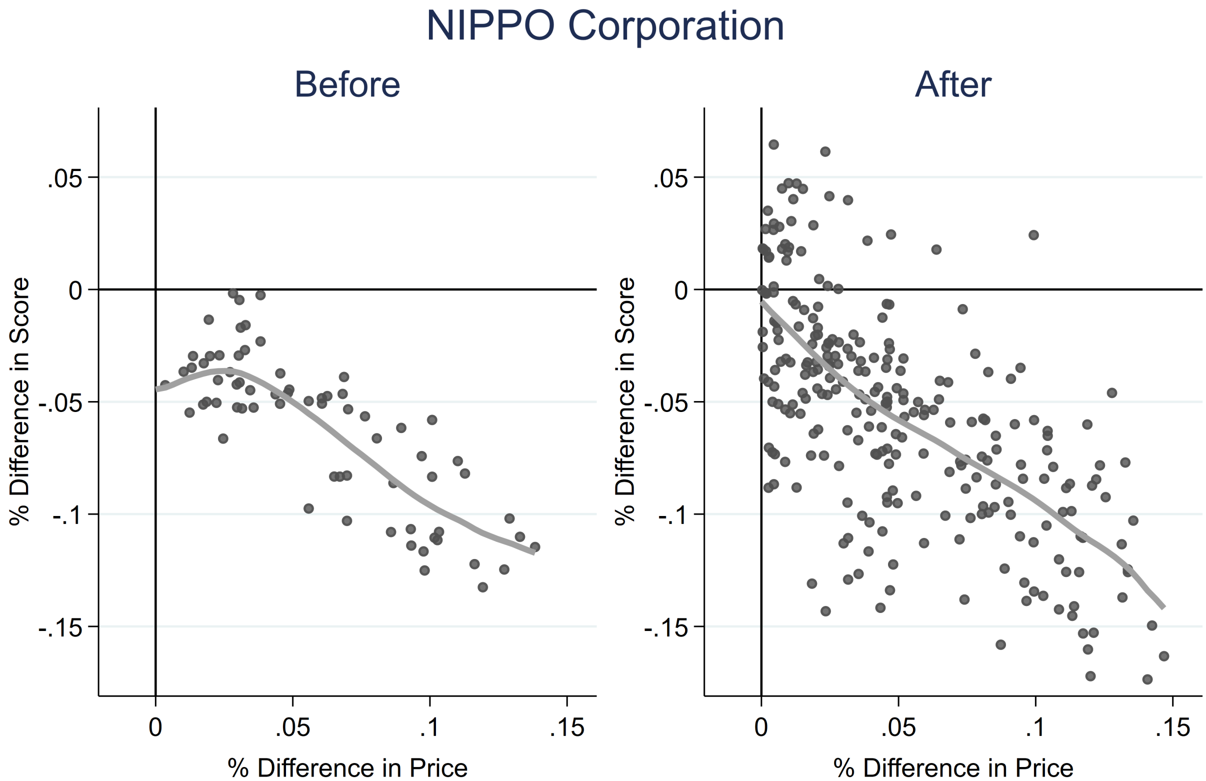
*, **, and *** respectively denote significance at the 10%, 5%, and 1% levels.

Table 1: Intercept of Partial Linear Regression, NIPPO Corporation.

As we discussed at the end of Section 3, we also consider a second test of competition in which we compare the score of the bidder who bid the lowest price and that of the bidder who bid marginally above the lowest price (Corollary 1′). Under the null of competition (i.e., static Nash equilibrium), the average scores should be equal to each other.

In order to implement this second test, let $i^\dagger$ denote the bidder who submits the lowest price in each auction, and for each $i \neq i^\dagger$, define $\tilde{\Delta}^p_{i,t} = \frac{p_{i,t} - p_{i^\dagger,t}}{p_{i^\dagger,t}}$. The variable $\tilde{\Delta}^p_{i,t}$ is the price difference between i's bid and the lowest bid, as a percentage of the lowest bid. By construction, $\tilde{\Delta}^p_{i,t}$ is positive. We also define $\tilde{\Delta}^s_{i,t} = \frac{s_{i,t} - s_{i^\dagger,t}}{s_{i^\dagger,t}}$, the score difference between bidder i and $i^\dagger$, measured as a percentage of the score of $i^\dagger$.

Figure 2 is a binned scatter plot of $(\tilde{\Delta}^p_{i,t}, \tilde{\Delta}^s_{i,t})$ for all bidders $i \neq i^\dagger$ for the pre-period (left panel) and the post-period (right panel). In the left panel, we find that bidders who bid



Left panel corresponds to the auctions in which NIPPO participated before being investigated for collusion. Right panel corresponds to auctions after the investigation.

Figure 2: Binned scatter plot of $\tilde{\Delta}^p_{i,t}$ and $\tilde{\Delta}^s_{i,t}$.

marginally higher than $i^\dagger$ have much lower scores, on average, than $i^\dagger$, the bidder who bid the lowest price. This also means that these bidders have much lower quality than $i^\dagger$. The right panel corresponds to the set of auctions after the investigation. In this panel, we find that bidders who bid marginally higher than $i^\dagger$ have, on average, the same score as $i^\dagger$. The curves in the figure correspond to the global fourth-order mean regression of $\tilde{\Delta}^s_{i,t}$ on $\tilde{\Delta}^p_{i,t}$.

The bottom panel of Table 1 reports the local linear estimate of $\mathbb{E}\left[\tilde{\Delta}^s_{i,t}|\tilde{\Delta}^p_{i,t}\right]$ at $\tilde{\Delta}^p_{i,t} = 0$. We find that the estimate is negative and statistically significant for the sample of auctions before the investigation, at -0.042. The estimate for the post period is -0.014 and it is statistically indistinguishable from 0.

Why collusive bidding may fail the test Corollary 1 and 1′ show that the regressions $\mathbb{E}\left[\Delta^p_{i,t}|\Delta^s_{i,t}\right]$ and $\mathbb{E}\left[\tilde{\Delta}^s_{i,t}|\tilde{\Delta}^p_{i,t}\right]$ should converge to 0 as $\Delta^s_{i,t}$ and $\tilde{\Delta}^p_{i,t}$ converge to 0. The Corollaries do not, however, show why these conditions might fail under collusion. Here, we discuss why our test has power against certain types of collusive behavior.

Under most bid rigging arrangements, the ring preallocates projects so that one of its members is a designated winner and other members are designated losers. Often there is heterogeneity among the ring members in terms of their ability to submit high quality proposals. The quality of the proposals depends on the technology and machinery available to each firm. When the designated winner is a low quality firm, the designated losers will have to submit a price that is substantially higher than that of the designated winner. For these auctions, the designated losers often come out as narrow losers (as intended), but with a much higher price and a much higher quality measure than the winner. The designated winner, on the other hand, comes out as the narrow winner, but with a much lower price and a much lower quality measure than the marginal losers. This bidding pattern is consistent with the data pattern illustrated in the left panel of Figure 1.

When the cartel designates a high quality bidder to be the winner of the auction, the low quality bidders can bid lower than the designated winner in terms of prices and still ensure that the auction is allocated to the intended bidder. In practice, however, many low quality bidders bid slightly higher than the designated winner, perhaps out of caution. When this is the case, bidders who bid marginally above the low price bidder are often low quality designated losers whose overall scores are much lower than the designated winner. These types of bidding behavior results in bidding patterns illustrated in the left panel of Figure 2.

Screening for non-competitive bidders In order to screen firms that are likely to be colluding, we apply the two tests we discussed above to each firm that participated in at least 5 MLIT auctions between April 2015 and March 2017. Specifically, for each firm during the sample period, we test the following two hypotheses:

$$H_0 : \mathbb{E}\left[\Delta^p_{i,t} | \Delta^s_{i,t} = 0^-\right] = 0 \quad \text{v.s.} \quad H_1 : \mathbb{E}\left[\Delta^p_{i,t} | \Delta^s_{i,t} = 0^-\right] > 0$$

and

$$H_0 : \mathbb{E}\left[\tilde{\Delta}^s_{i,t} | \tilde{\Delta}^p_{i,t} = 0^+\right] = 0 \quad \text{v.s.} \quad H_1 : \mathbb{E}\left[\tilde{\Delta}^s_{i,t} | \tilde{\Delta}^p_{i,t} = 0^+\right] < 0.$$

Out of a little more than 4,000 firms that we tested, we initially select 1,143 firms that fail either of the tests at 5% significance.[26] We then further screen these firms by visually inspecting the binned scatter plot and determine whether or not the intercept seems clearly different from zero. The reason for adopting this two-step procedure instead of mechanically selecting the firms with the highest t-statistics is because there is occasionally a disconnect between the magnitude of the t-statistic and how strong the evidence appears based on the binned scatter plot. Hence, we choose a threshold of 5% confidence in terms of the t-statistic to initially preselect the set of potentially non-competitive firms and complement it with a follow-up visual inspection of the binned scatter plot of each firm to avoid making type I errors. Korting et al. (2023) discusses the benefits of combining a formal econometric test with a visual examination. Ultimately, we end up with 240 firms which we feel confident classifying as noncompetitive.[27]

Because we conduct tests for all firms in the data that participate in at least 5 auctions, the set of firms that we classify as noncompetitive is likely to include both true positives and false positives. In order to get a sense of the *false discovery rate* (FDR), or the proportion of competitive firms that we wrongly identify as uncompetitive, we plot the empirical CDF of the estimated t-statistic from our first test (i.e., test of $\mathbb{E}\left[\Delta^p_{i,t} | \Delta^s_{i,t} = 0^-\right] = 0$) along with the CDF of the standard normal in Figure 3.[28] Note that under the null that every single firm is bidding competitively, our test statistic follows the standard normal distribution. Figure

[26]For the first test, we select firms with t-statistics of above 1.65. For the second test, we select those with t-statistics of below -1.65. Of the 1,143 firms that we initially select, 635 firms failed the first test, 601 firms failed the second test, and 93 failed both.

[27]We initially select 242 firms, but we end up dropping 2 firms after performing the clustering procedure we discuss in the next section. See Online Appendix Section OA for details.

[28]The Online Appendix contains a corresponding figure for our second test.

3 shows clearly, however, that the empirical CDF exhibits large excess mass at the right tail of the distribution relative to the standard normal: for example, 15% of the estimated t-statistics are higher than 1.65 as opposed to 5%. This implies that the (positive) FDR at critical value of 1.65 is at most 0.05/0.15$\approx$ 33%.[29] Using empirical Bayes shrinkage to shrink the distribution of the t-statistic only changes the fraction of t-statistic exceeding 1.65 to 14%.

As we discussed above, we use a cutoff of 1.65 in terms of the t-statistics to determine the initial set of noncompetitive firms which we then follow up with a visual inspection of the binned scatter plots. If, instead of visual inspection, we were to mechanically pick the same number of firms as our two-step procedure, but based on the value of the t-statistics alone, the corresponding FDR, computed using Benjamini et al. (1995), would be about 2%. To the extent that our visual inspection is no worse at screening out type I errors than mechanically choosing the firms with the highest t-values, the FDR among the chosen 240 firms is lower than 2%.[30]

Finally, we note that while the selection of the 240 firms are based on somewhat subjective criteria, all of the subsequent statistical analysis are based on randomization inference which gives us exact p-values.

5 Experimental Design

Assignment of Treatment The 240 firms that we identify as noncompetitive in the previous section often bid on the same auctions. In order to deal with correlated shocks across firms that bid on the same auctions, we partition the 240 firms by grouping those

[29]The original definition of the FDR proposed by Benjamini and Hochberg (1995) is $\mathbb{E}[V/R|R>0]\Pr(R>0)$, where R is the number of total rejections and V is the number of Type 1 errors. The positive FDR is defined to be $\mathbb{E}[V/R|R>0]$ (See, e.g., Storey, 2002).

[30]Korting et al. (2023) conclude that visual inference achieves a type I error rate that is slightly lower than the Imbens and Kalyanaraman (2012) and Calonico et al. (2014) procedures. Moreover, they find that visual and econometric inferences are complementary.

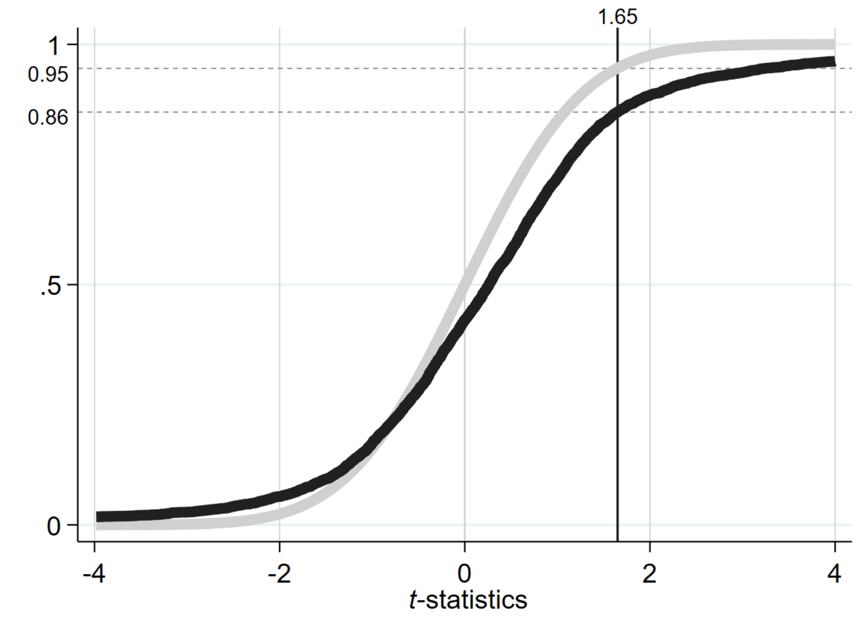


Figure 3: Histogram of t-statistic for $\mathbb{E}[\Delta^p_{i,t}|\Delta^s_{i,t} = 0^-]$.

that frequently bid together using a clustering algorithm.[31] The resulting partition has the property that firms within each group bid on the same auction frequently while firms in different groups rarely do. We end up with 26 groups of firms. We then construct 13 matched pairs based on the group's geographical location, type of work (e.g., landscaping, paving, etc.), and the number of firms in the group.

Finally, we assign treatment status with rerandomization (see, e.g., Morgan and Rubin, 2012) to achieve balance between the treatment and the control with regards to the mean winning bid, mean of the t-statistics of the two tests, and the average number of auctions firms participate.[32] We take the effect of rerandomization into account when conducting our statistical tests below.

Treatment We send (physical) letters to 13 groups, or a total of 107 firms on Feb 12, 2019. We send the letters to the addresses of the firms recorded in the MLIT's registry. In the letter, we first explain that we have developed a screen for bid rigging, and that we

[31]In particular, we use a hierarchical agglomerative linkage procedure. We provide the details in Online Appendix Setion OA.

[32]We explain in more detail the rerandomization procedure in Online Appendix Section OA.

are exploring its usefulness and applicability. In the letter, we explain the mechanics of our test by walking through Figures 1 and 2 for Nippo Corporation that we discussed in Section 4. We then include corresponding figures for the firm in question and discuss the similarities between the firm's own bidding patterns with those of Nippo Corporation before the JFTC investigation, i.e., the regression lines do not pass through the origin. We also include in the letter a list of auctions that we use in our analysis so as to emphasize that our analysis is specific to each firm. Note that there is a lag between the period of analysis contained in the letter (April 2015 to March 2017) and the date the letters are sent out (Feb 12, 2019). Finally, we ask the firm whether the firm is aware of various screening methods to detect noncompetitive bidding, and whether such screens can help the firm improve antitrust compliance.[33] We include a return envelope for the firm to send back its reply, asking them to do so by March 15, 2019.[34] A copy of the letter sent to the firms in the treatment group is in the Online Appendix.

Baseline Summary Statistics Table 2 and 3 report the summary statistics of the auctions and the bidders for the sample of 240 firms that we select as noncompetitive. The summary statistics correspond to the sample of auctions we use to detect noncompetitive bidding, i.e., auctions let between April 2015 to March 2017. The average reserve price of the auctions is about 128 million yen, or about \$1.3 million. The winning bid is about 94% of the reserve price and the average number of bidders is 4.81. The fact that there are only very minor differences between the treatment and the control is by design. We assigned treatment by a matched pairs design and, moreover, we rerandomized the assignment to achieve balance for the winning bid.

Table 3 reports firm characteristics. We report the mean annual sales, accounting profits, the number of engineers employed by the firm, and the t-statistics corresponding to our intercept estimates. The first three variables are obtained from the registry maintained by the MLIT. Although the total number of employees is not recorded in the registry, the ratio

[33] We report the responses of the firms that we received in Online Appendix **??**.
[34] The response rate was 31.8%.

	Treatment	Control
Reserve	128.358	128.787
	(72.350)	(72.432)
WinBid / Reserve	0.939	0.939
	(0.034)	(0.034)
Bid / Reserve	0.942	0.942
	(0.039)	(0.039)
Quality	155.722	155.746
	(6.836)	(6.839)
Number of Bidders	4.794	4.794
	(2.792)	(2.799)
N	1,300	1,289

Note: The table shows the summary statistics of auctions let during fiscal years 2015 through 2017. Reserve is reported in millions of yen. Standard errors are in parenthesis.

Table 2: Summary Statistics (Auctions)

of engineers to the total number of employees is usually about 1:2 to 1:3, based on the subset of firms that report the number of employees on their web pages. Given that the average number of engineers is about 27 in our sample, this implies that the total number of employees is likely to be about 55 to 80, on average. The average t-statistics associated with the intercept estimates of each firm before the intervention is about 3.5 for the first test (i.e., test of $\mathbb{E}[\Delta^p_{i,t}|\Delta^s_{i,t} = 0^-] = 0$), and it is about -3.0 and -3.5 for the second test (i.e., test of $\mathbb{E}[\tilde{\Delta}^s_{i,t}|\tilde{\Delta}^p_{i,t} = 0^+] = 0$). By construction, for each firm, the t-statistics are higher than 1.65 for the first test or lower than -1.65 for the second test.

6 Results

In this section, we document changes in the firms' bidding behavior before and after the intervention.

	Treatment	Control
Annual Sales	2,087.12	2,122.89
	(2,278.15)	(3,623.24)
Annual Profits	146.52	137.95
	(192.54)	(297.86)
# Engineers	26.75	27.63
	(18.97)	(34.29)
t-stat ($\mathbb{E}\left[\Delta^p_{i,t})\right]$)	3.45	3.75
	(4.09)	(8.69)
t-stat ($\mathbb{E}\left[\tilde{\Delta}^s_{i,t})\right]$)	-3.00	-3.54
	(4.39)	(4.97)
N	107	133 (132)

Note: Sales and profits are reported in units of one million yen. There are 240 firms in our sample. We could not get data on annual sales, profits, and the number of engineers for one firm in the control group. The numbers for these variables reported in the table for the control group are based on the averages of 132 firms.

Table 3: Summary Statistics (Firms)

Changes in screening power We first document changes in the firms' bidding behavior in relation to the ability of our test to detect collusion.

Figure 4 is a binned scatter plot of (Δ^p, Δ^s) for the treated firms (top panels) and for the control firms (bottom panels). The sample of firms is the set of firms that we flagged using the first test. In each row, the left panel corresponds to the pre-intervention period, and the right panel corresponds to the post-intervention period.[35] Comparing the top left and right panels, we find that the intercept of the regression curve (depicted in gray) at $\Delta^s = 0$ is positive in the left while it is close to 0 in the right.[36] The local linear estimates of the intercept are 0.020 and 0.003, and the respective t-values are 3.54 and 0.78. This implies that marginal losers stop bidding substantially higher than the marginal winner in terms of prices after the intervention, suggesting that bidder behavior changed after the intervention

[35]The sample for the pre-intervention period consists of auctions let between April 1, 2015 and February 15, 2019, the date we sent out the letters. The sample for the post-intervention consists of auctions let after March 15, 2019, the date by which we asked the firms to respond to our survey.

[36]The curve in the figure is a (global) forth-order polynomial regression curve.

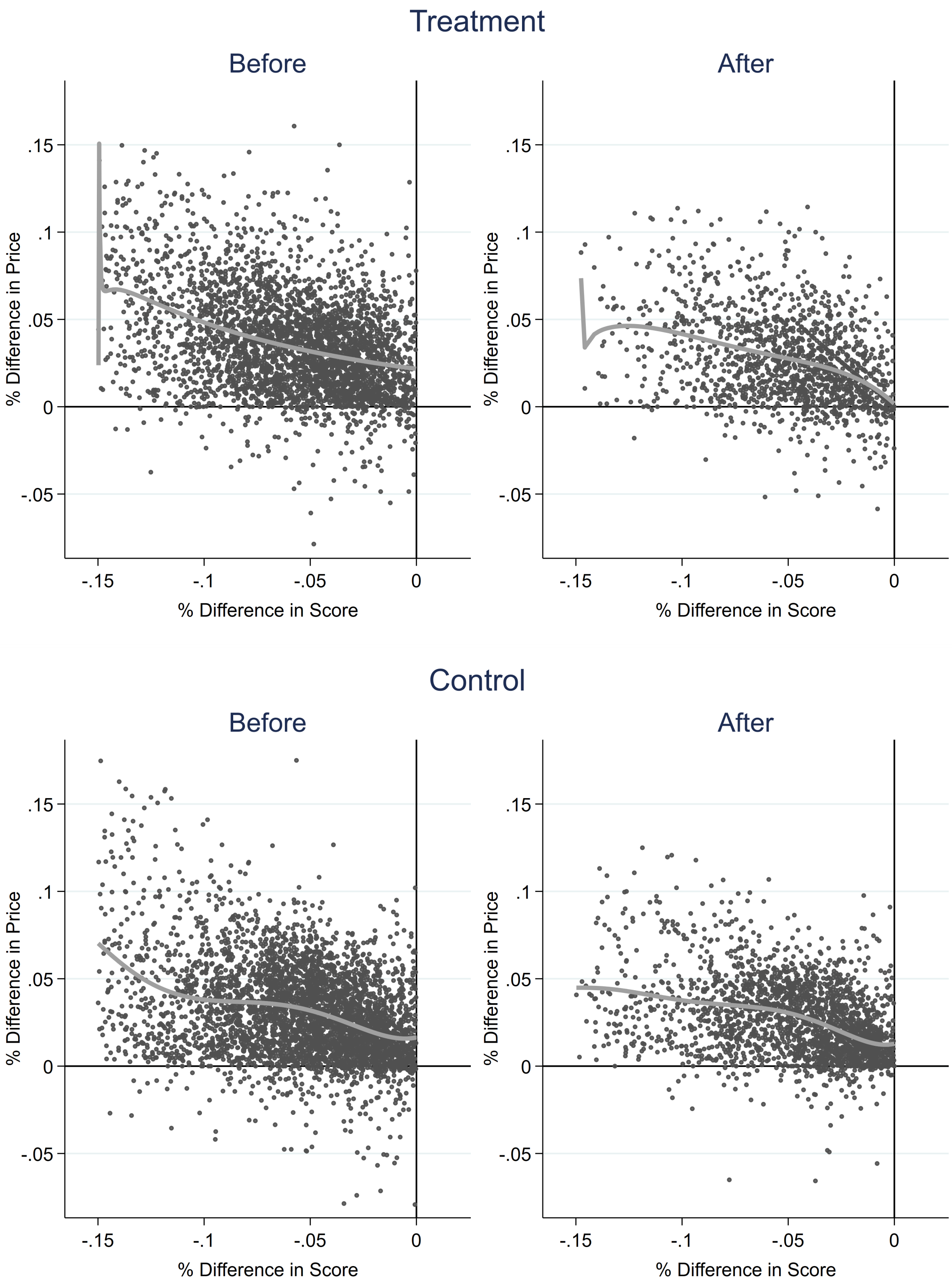


Figure 4: Binned scatter plots of (Δ^p, Δ^s).

for the treated firms. For the control firms, the regression curves seem to intersect the y-axis above the origin in both panels. The local linear estimates of the intercept are 0.019 and 0.014, and the respective t values are 3.78 and 3.77. The intercept estimates are statistically significant at the 5% level in both periods.

Similarly, Figure 5 is the binned scatter plots of $(\tilde{\Delta}^s, \tilde{\Delta}^p)$. The top two panels correspond to the treated firms and the bottom two correspond to the control firms. The sample is the set of firms that we flagged using the second test. For the treatment sample, we find that the estimated intercept of the the regression curve changes from -0.03 before the intervention to -0.001 after the intervention. The estimated intercept is statistically significant at the 5% level in the pre-intervention period while it is not statistically significant in the post period. For the control sample, the estimated intercept changes some, but much less drastically: The estimate is -0.024 before the intervention, and it is -0.015 after, although the estimate is only statistically significant in the pre-intervention period.

While Figures 4 and 5 are suggestive that the intervention induced changes in the bidding behavior of the treated firms, we wish to formally test whether or not the differences between the treatment and the control firms are statistically significant. In order to do so, we now run Fisher's test of the strong null hypothesis (See, e.g., Imbens and Rubin, 2015, for a textbook treatment).

Specifically, for each group $g \in \{1, \cdots, 26\}$, let $\beta_g^{\mathsf{Before}}$ (β_g^{After}) denote the estimated intercept, $\mathbb{E}[\Delta_{i,t}^p | \Delta_{i,t}^s = 0^-]$, computed using firms $i \in g$ and auctions in the pre-intervention (post-intervention) period.[37] Let Y_g denote the change in the estimated intercept before and after the intervention,

$$Y_g = \beta_g^{\mathsf{After}} - \beta_g^{\mathsf{Before}}.$$

Now, take a partition, G and G', of $\{1, \cdots, 26\}$ ($G \cap G' = \phi$, $G \cup G' = \{1, \cdots, 26\}$) such that one group from each of the 13 pairs belongs to G and the other belongs to G'. Consider

[37] For each g, we use only the subset of firms that we flagged using our first test.

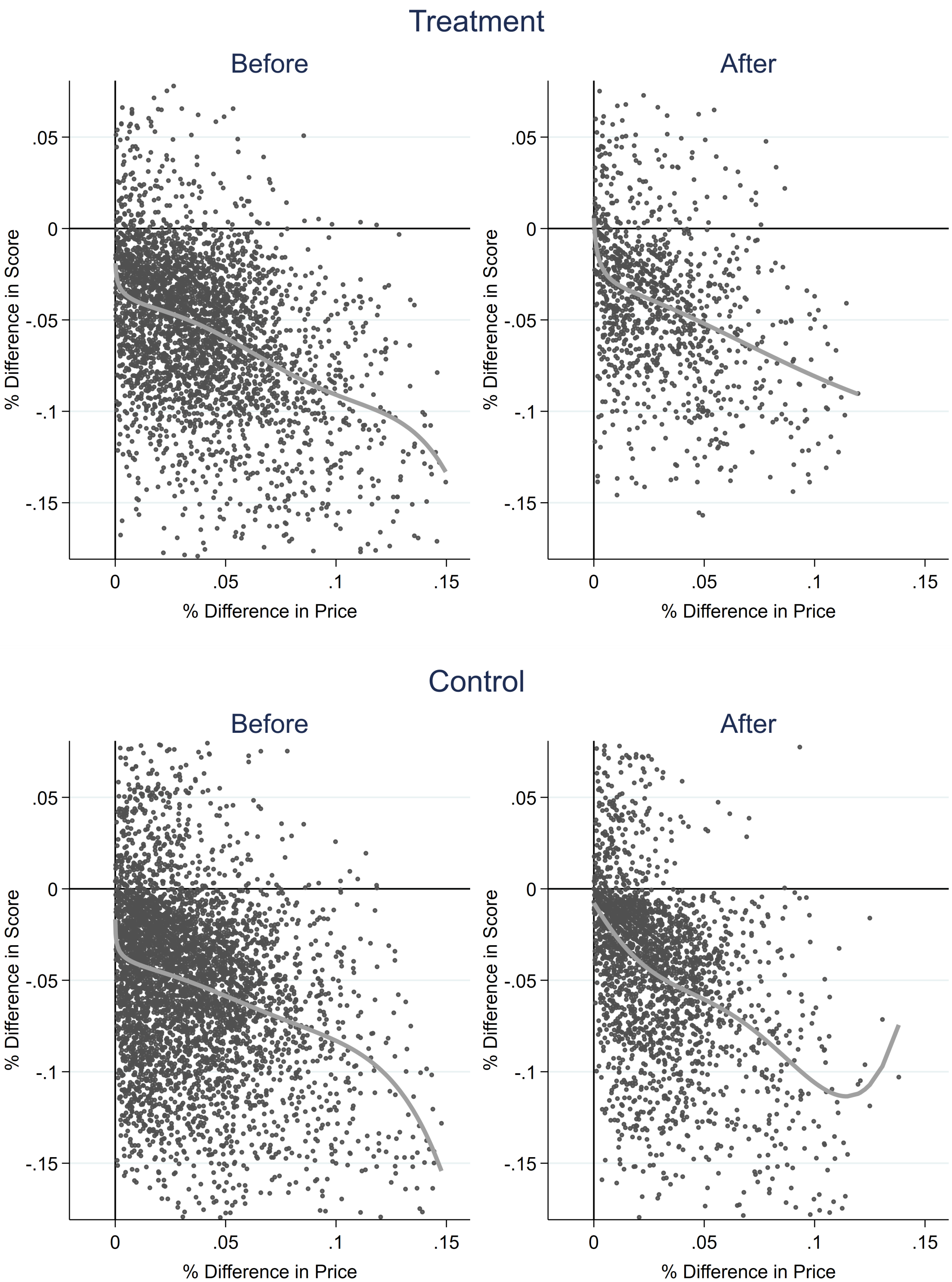


Figure 5: $\mathbb{E}[\tilde{\Delta}^p|\tilde{\Delta}^s]$

the difference in the average Y_g among groups $g \in G$ and among groups $g \in G'$:

$$\overline{Y}_{G-G'} = \frac{1}{\mid G \mid} \sum_{g \in G} Y_g - \frac{1}{\mid G' \mid} \sum_{g \in G'} Y_g. \tag{3}$$

The statistic $\overline{Y}_{G-G'}$ corresponds to a measure of the extra decline in the estimated intercept exhibited by groups in G relative to those in G'. In particular, if we set $G = G_T$ and $G' = G_C$, where G_T and G_C are the set of treated groups and control groups, $\overline{Y}_{G_T-G_C}$ corresponds to the extra decrease in the estimated intercept for the treatment groups relative to the control groups. Note, however, that we can compute $\overline{Y}_{G-G'}$ for an arbitrary partition G and G', not just for G_T and G_C.

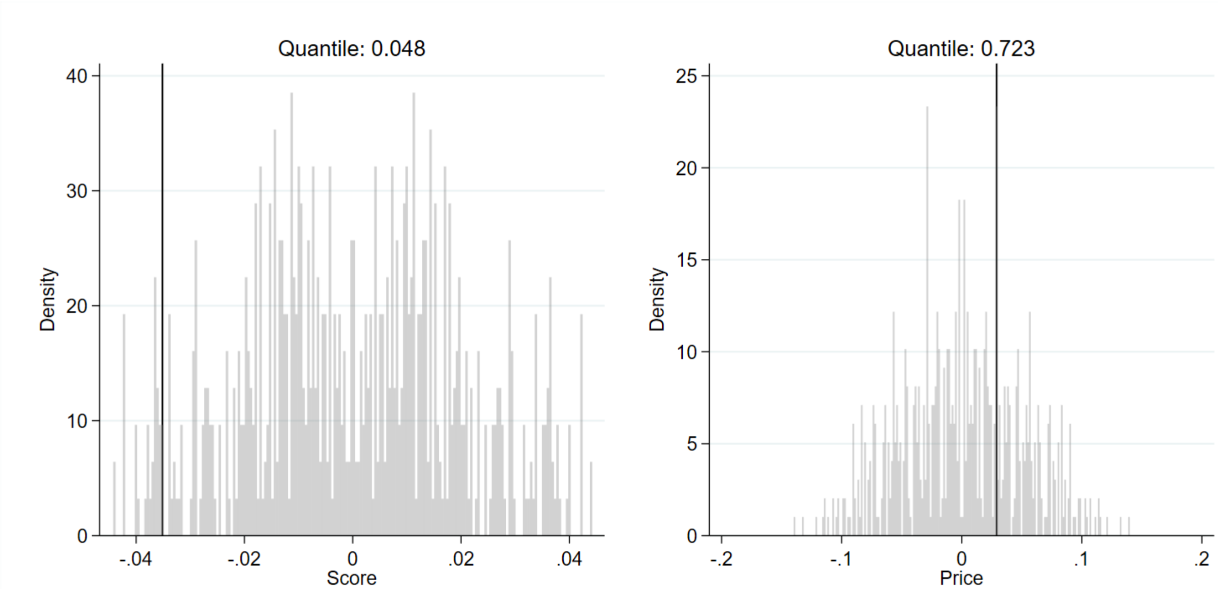
Under the strong null that our intervention had no effect on the bidding behavior of any of the treated groups whatsoever, the partition of the groups into the set of treated groups, G_T, and the control groups, G_C, is of no particular significance. In other words, under the strong null of no effect, the random variable, $\overline{Y}_{G_T-G_C}$, should have the same distribution as $\overline{Y}_{G-G'}$ for arbitrary partition G and G'. Fisher's randomization test compares the realized value of $\overline{Y}_{G_T-G_C}$ to the distribution of $\overline{Y}_{G-G'}$ for all possible partitions G and G'. If the value of $\overline{Y}_{G_T-G_C}$ is extreme relative the distribution of $\overline{Y}_{G-G'}$, we can reject the null that the treatment has no effect.

The left panel of Figure 6 reports the distribution of $\overline{Y}_{G-G'}$ for all possible partitions of $\{1, \cdots, 26\}$ that are consistent with our matched pair design and rerandomization criteria.[38] The value of $\overline{Y}_{G_T-G_C}$ is depicted as a vertical line in the figure. The realization of $\overline{Y}_{G_T-G_C}$ is quite extreme relative to the realizations of $\overline{Y}_{G-G'}$ for other possible partitions (4.8 percentile). The figure implies that the difference between the treatment and the control groups with regard to the average pre-post change in the estimate of the intercept ($\beta_g^{\mathsf{After}} - \beta_g^{\mathsf{Before}}$) is more extreme than 95.2% of possible partitions. This suggests that the change in the test statistic for the treatment group that we illustrate in Figure 4 is unlikely to have occurred

[38] Recall that we have a matched pair design in which one of the pair is treated and the other is not. The number of all possible treatment assignments is 2^{13}. Because we rerandomize to maintain balance between the treated and the control groups, we have a smaller number (704) of admissible assignments.

simply by chance.

We now repeat the above exercise with estimates of $\mathbb{E}[\tilde{\Delta}^s|\tilde{\Delta}^p = 0^+]$. For each group g, let $\tilde{\beta}_g^{\text{Before}}$ ($\tilde{\beta}_g^{\text{After}}$) denote the estimate of $\mathbb{E}[\tilde{\Delta}^s|\tilde{\Delta}^p = 0^+]$ for the sample of auctions let in the pre-intervention (post-interention) period.[39] For any partition G and G', we can construct $\overline{Y}_{G-G'}$ similarly as before, where $Y_g = \tilde{\beta}_g^{\text{After}} - \tilde{\beta}_g^{\text{Before}}$ and $\overline{Y}_{G-G'}$ is defined as in equation (3). The right panel of Figure 6 reports the distribution of $\overline{Y}_{G-G'}$ and the value of $\overline{Y}_{G_T-G_C}$. Note that for this test, we expect $\overline{Y}_{G_T-G_C}$ to be more positive than average. We find that this is the case (72.3 percentile) although it is not statistically significant at conventional levels.



Left panel corresponds to the distribution of $\overline{Y}_{G-G'}$ in which $Y_g = \beta_g^{\text{After}} - \beta_g^{\text{Before}}$, and the right panel corresponds to the distribution of $\overline{Y}_{G-G'}$ in which $Y_g = \tilde{\beta}_g^{\text{After}} - \tilde{\beta}_g^{\text{Before}}$.

Figure 6: Distribution of $\overline{Y}_{G-G'}$

In order to understand the joint significance of the two tests, Figure 7 plots the quantiles corresponding to the two tests for each of the possible partitions G and G'. The point in the figure that corresponds to the actual treatment and the control is $(0.048, 0.723)$. We find

[39]Similar as before, when we compute $\tilde{\beta}_g^{\text{Before}}$ and $\tilde{\beta}_g^{\text{After}}$, we condition on the set of firms that we detect with our second screen.

that the proportion of possible partitions that give more extreme outcomes, in the sense of lying to the top left of the 45 degree line that goes through $(0.048, 0.723)$, is 6.39%.

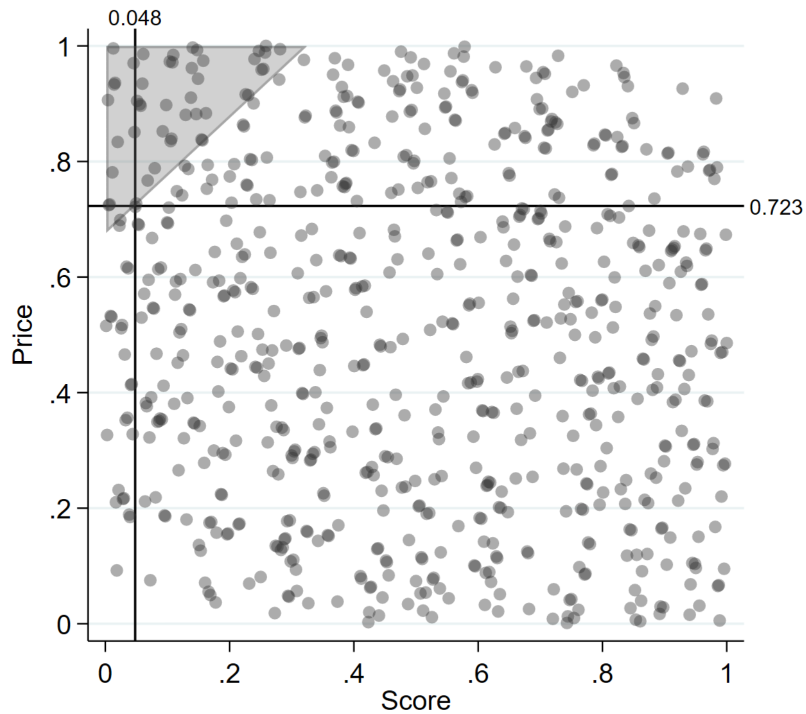
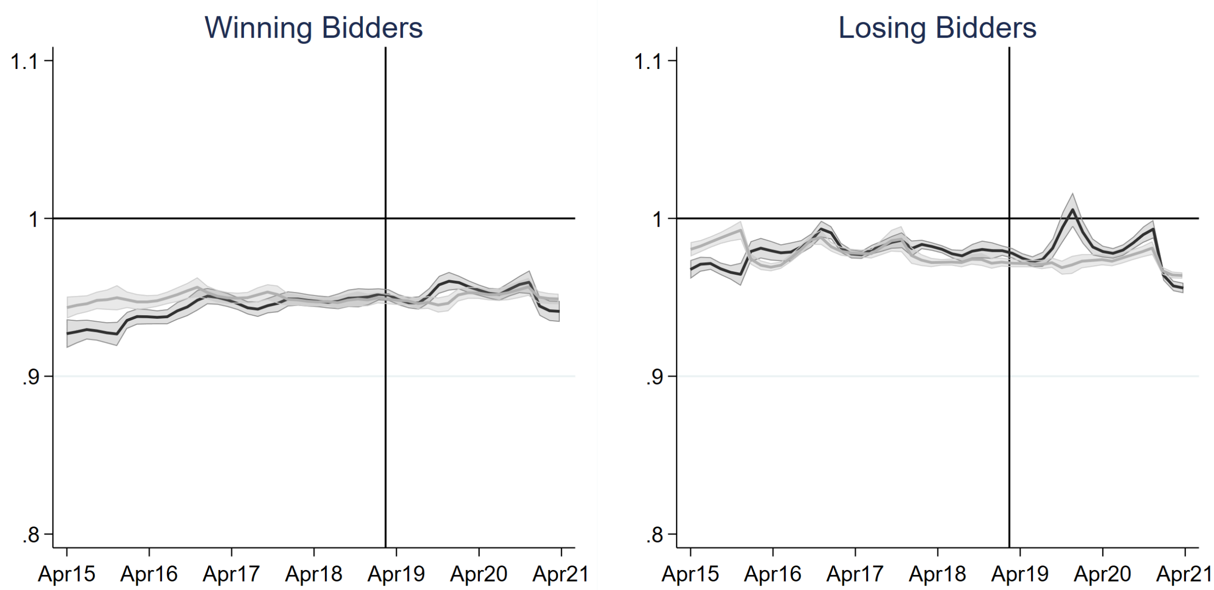


Figure 7: Scatter plot of the quantiles of the two tests.

Changes to other outcome variables We now explore the effect of the intervention on other auction outcomes to better understand how bidding patterns change, and in particular, to understand whether or not cartel firms stop colluding. Figure 8 plots the time series average of the normalized winning bid (left panel) and the average normalized losing bids (right panel) for both the treatment and the control groups. We normalize the bids by the reserve price when computing the averages to make them comparable across auctions. The light colored line corresponds to the control group, the dark colored line corresponds to the treatment group, and the solid vertical line corresponds to the date of the intervention. The shaded region around the lines correspond to the 95% confidence interval. We find that for both the treatment and the control groups, the bids are quite stable across time, at around 95% of the reserve price for the winning bid, and at around 97% for the losing bids. In particular, we do not see any breaks for the treatment group around the time of the intervention. Since breakdowns in cartels typically result in significant drops in prices, Figure 8 suggests that the firms in the treatment groups are unlikely to have stopped colluding after the intervention.

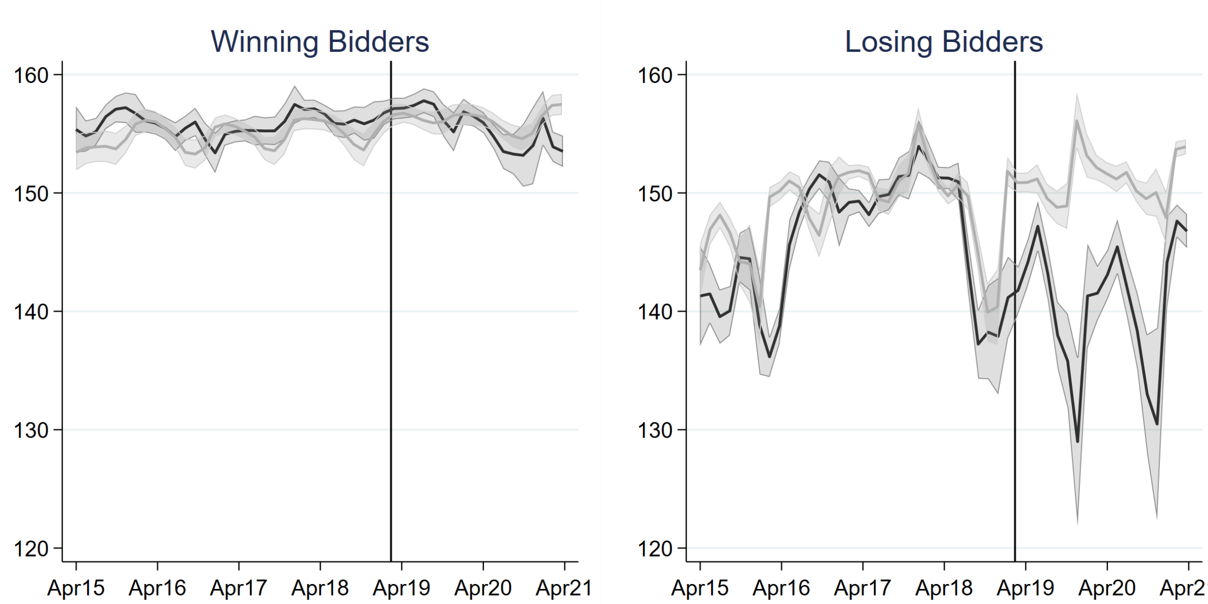


Left panel corresponds to the time series average of the winning bids (normalized by the reserve price) and the right panel corresponds to that of the losing bids (normalized by the reserve price). Light colored lines correspond to the control group and the dark lines correspond to the treatment group.

Figure 8: Time Series Plot of Bids, Normalized by the Reserve Price.

Figure 9 plots the time series average of the quality measure of the winning bidders (left panel) and the losing bidders (right panel). The dark lines correspond to the treatment and the light colored lines correspond to the control. Although there do not seem to be significant changes in the quality measure for the winning bidders, there is a substantial drop in the quality for the treatment group around the time of the intervention for the losing bidders.

As we mentioned in Section 2, firms that submit bids exceeding the secret reserve price are assigned a quality of 100 points, which is the lowest possible points attainable. In Figure 10 we plot the time series average of the probability that a losing bidder submits a bid that exceeds the reserve price.[40] The figure shows a substantial increase in the probability of submitting a bid above the reserve price for the treatment group after the intervention. The pattern in Figures 10 explain the significant drop in the quality measures of the losing bidders.

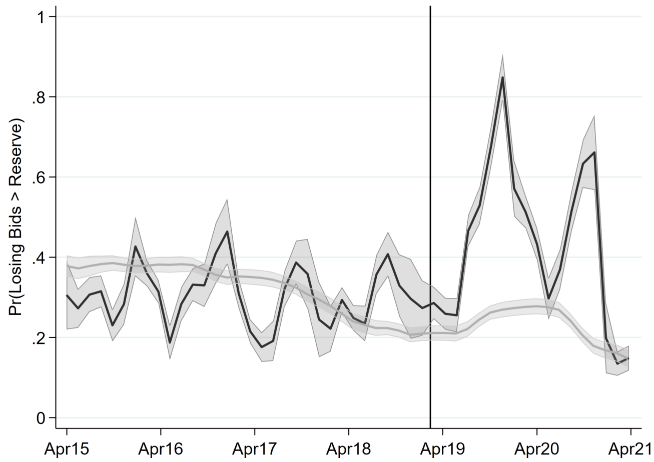


In the left panel, the dark line represents the raw quality of the winning bids in the treated auctions (i.e., auctions in which one or more treatment firms participates). The light colored line represents the raw quality of the winning bids for the control auctions (i.e., auctions in which one or more control firms participates). In the right panel, the dark line represents the raw quality of the losing bids for treated auctions. The light colored line represents the raw quality of the losing bids for control auctions.

Figure 9: Time Series Plot of Winner's Quality (Left) and Losers Quality (Right)

In order to assess the statistical significance of these findings, we again conduct Fisher's randomization test for each outcome. Specifically, let μ_g^{Before} and μ_g^{After} denote the pre- and post-intervention averages of a variable of interest (e.g., winning bid, losing bid, winner's quality, etc.) for group g. Let $Y_g \equiv \mu_g^{\mathsf{After}} - \mu_g^{\mathsf{Before}}$. For example, if μ_g^{Before} and μ_g^{After} denote pre- and post-intervention averages of the winning bid for group g, Y_g corresponds to the

[40]The probability that the winner's bid exceeds the secret reserve price is very close to zero, so we do not report the corresponding figure for the winner.



The dark line represents the probability that the losing bid is above the reserve price in treated auctions. The light colored line represents the corresponding probability for the control auctions.

Figure 10: Time Series Plot of Proportion of Invalid Bids

change in the average winning bid. Define $\overline{Y}_{G-G'}$ by expression (3), as before. We then test for the strong null of no effect by comparing $\overline{Y}_{G_T-G_C}$ to the distribution of $\overline{Y}_{G-G'}$.

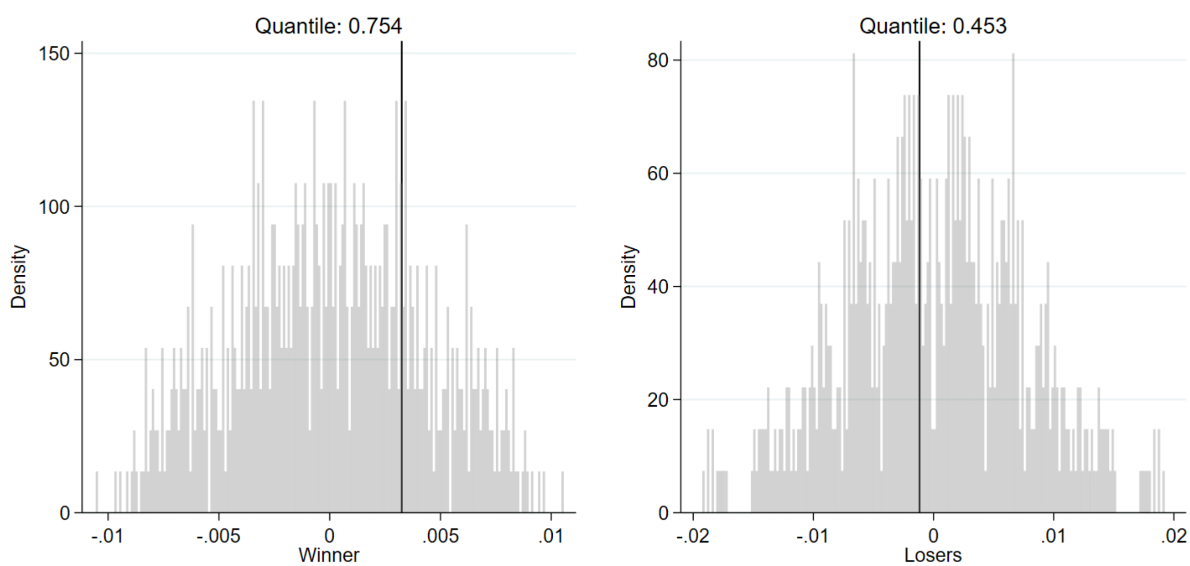


Figure 11: Histogram of $\overline{Y}_{G-G'}$ for Winning Bids (Left) and Losing Bids (Right).

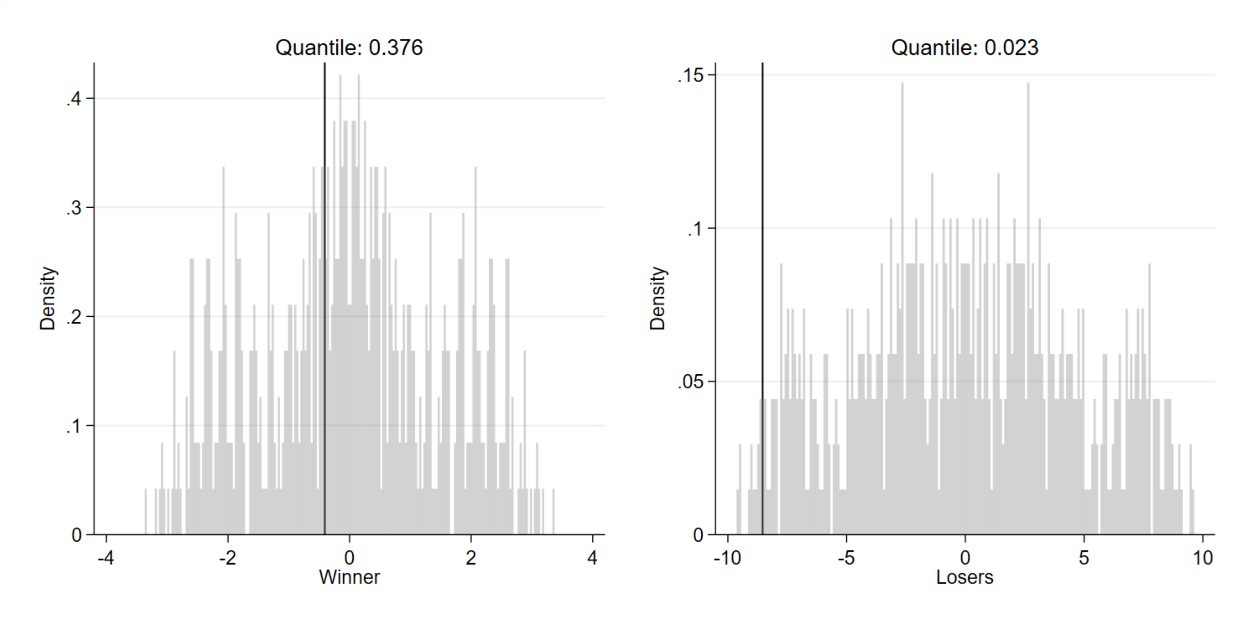


Figure 12: Histogram of $\overline{Y}_{G-G'}$ for Winner's Quality (Left) and Loser's Quality (Right).

The left panel of Figure 11 corresponds to the distribution of $\overline{Y}_{G-G'}$ for the winning bid and the right panel corresponds to that for the losing bids. We cannot reject the null that the intervention had no impact on the winning bid or the losing bids. These results are consistent with Figure 8 in which we find little changes between the pre- and post-intervention periods.

Figure 12 plots the histogram of $\overline{Y}_{G-G'}$ for the quality component of the bids. The left panel corresponds to the winner's quality and the right panel corresponds to the losers' quality. While we do not reject the null that the winner's quality is unaffected by the intervention, we find that $\overline{Y}_{G_T-G_C}$ is at the 2.3 percentile in the distribution. This suggests that the decrease in the quality of the losers' bids is unlikely to be the result of chance. Lastly, Figure 13 corresponds to the Fisher's test for the probability that one of the losers submits a bid above the reserve price.[41] Figure 13 shows that the test statistic is at the 96 percentile, suggesting that the higher proportion of invalid bids among the treated group is likely to be the result of our intervention.

[41]For this test, we construct an indicator variable for whether or not one of the losers submits a price above the reserve price. We then take averages of the indicator variable across all firms in g during the pre-intervention and post-intervention periods to construct μ_g^{Before} and μ_g^{After}.

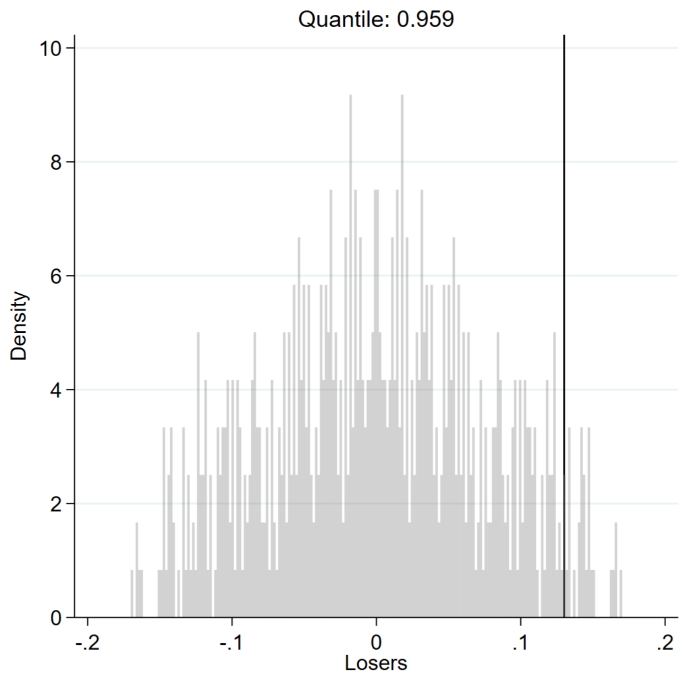


Figure 13: Histogram of $\overline{Y}_{G-G'}$ for Proportion of Invalid Bids.

Overall, the observed changes in the bidding pattern after the intervention do not suggest that bidders in the treated groups stopped colluding or that they began to compete. The bidding patterns that we document are, if anything, consistent with a more blatant form of collusion in which non-winners bid above the reserve price, guaranteeing that they lose. These findings suggest that the change in the screening power of our test for the treatment groups in the post-period documented earlier reflects firm adaptation while continuing to engage in collusion.

Bidding above the reserve price as an adaptive response We end this section with a discussion of why bidding above the reserve price is a natural response for a bidding ring that wishes to adapt to our screen. Recall that our first test compares the price of marginal winners and marginal losers. The test has power when a cartel allocates a project to a low quality firm and high quality designated losers bid substantially higher than the designated low quality winner. Often, the high quality bidder becomes the marginal loser, with much higher price than the marginal winner, failing our test. If the cartel wishes to adapt to the test without reducing profits through lower winning bids, the only available option, to the extent that manipulating the quality component is difficult in the short term, is to raise the bid of the designated losers. When the designated losers bid higher than the reserve price, the quality associated with the bid is either unrecorded or is assigned the lowest possible value of 100, guaranteeing that the differences in the score, Δ^s, becomes large and the bids to be "not close". This adaptive response is consistent with the treated firms beating our test, and with the bidding patterns documented in Figures 12 and 13.

We note, however, that bidding above the reserve price is akin to refraining from bidding, a common indicator of bid-rigging.[42] Our findings are consistent with the point made in previous work (e.g., Porter, 2005; Harrington, 2008a; Marshall and Marx, 2012), that adaptation to one screen can lead to violation of other screens of collusion.

7 Conclusion

As regulations have become more complex and broader in scope, regulators have come to rely more on within-firm compliance. This paper studies the effectiveness of firm compliance in taking remedial action when confronted with evidence of regulatory violation. We find that existing levels of compliance capacity among firms engaging in bid rigging are unlikely to complement formal regulatory enforcement. Our results differ somewhat from those in Christie et al. (1994) and Monticini and Thornton (2013) who find evidence that firms take prompt remedial action once regulatory violations become publicly known. Finally, our findings lend support to the view that, while firms can adapt to screens, adaptation to one screen can lead to violation of others, and that multiple screens when combined can make collusion harder and less profitable.

[42] See, e.g., the document published by the U.S. DOJ titled "Preventing And Detecting Bid Rigging, Price Fixing, And Market Allocation In Post-Disaster Rebuilding Projects".

References

ABRANTES-METZ, R., L. FROEB, J. GEWEKE, AND C. TAYLOR (2006): "A variance screen for collusion," *International Journal of Industrial Organization*, 24, 467–486.

ASKER, J. (2010): "A study of the internal organization of a bidding cartel," *The American Economic Review*, 724–762.

ASKER, J. AND E. CANTILLON (2008): "Properties of scoring auctions," *The RAND Journal of Economics*, 39, 69–85.

——— (2010): "Procurement when price and quality matter," *The RAND Journal of Economics*, 41, 1–34.

ATHEY, S., J. LEVIN, AND E. SEIRA (2011): "Comparing open and Sealed Bid Auctions: Evidence from Timber Auctions*," *The Quarterly Journal of Economics*, 126, 207–257.

AUBERT, C., P. REY, AND W. KOVACIC (2006): "The impact of leniency and whistle-blowing programs on cartels," *International Journal of Industrial Organization*, 24, 1241–1266, the Economics of Cartels, Cartel Policy, and Collusion.

AYRES, I. AND J. BRAITHWAITE (1992): *Responsive regulation: Transcending the deregulation debate*, Oxford University Press, USA.

——— (1995): *Responsive regulation: Transcending the deregulation debate*, Oxford University Press, USA.

BAER, B. (2014): "Prosecuting Antitrust Crimes," Remarks as Prepared for the Georgetown University Law Center Global Antitrust Enforcement Symposium.

BAJARI, P. AND L. YE (2003): "Deciding between competition and collusion," *Review of Economics and Statistics*, 85, 971–989.

BALDWIN, L. H., R. C. MARSHALL, AND J.-F. RICHARD (1997): "Bidder collusion at forest service timber sales," *Journal of Political Economy*, 105, 657–699.

Baránek, B., L. Musolff, and V. Titl (2023): "Detection of Collusive Networks in E-Procurement," *Working Paper*.

Benjamini, Y. and Y. Hochberg (1995): "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300.

Benjamini, Y., A. Krieger, and D. Yekurieli (1995): "Adaptive linear step-up procedures that control the false discovery rate," *Biometrika*, 93, 491–507.

Blundell, W., G. Gowrisankaran, and A. Langer (2020): "Escalation of Scrutiny: The Gains from Dynamic Enforcement of Environmental Regulations," *American Economic Review*, 110, 2558–85.

Braithwaite, J. (1985): *To Punish or Persuade*, State University of New York Press.

Calonico, S., M. Cattaneo, and R. Titiunik (2014): "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82, 2295–2326.

Chassang, S., K. Kawai, J. Nakabayashi, and J. Ortner (2022): "Robust Screens for Noncompetitive Bidding in Procurement Auctions," *Econometrica*, 90, 315–346.

Chassang, S. and J. Ortner (2019): "Collusion in Auctions with Constrained Bids: Theory and Evidence from Public Procurement," *Journal of Political Economy*, 127, 2269–2300.

Che, Y.-K. (1993): "Design competition through multidimensional auctions," *The RAND Journal of Economics*, 668–680.

Chen, J. and J. Harrington (2007): "Chapter 3 The Impact of the Corporate Leniency Program on Cartel Formation and the Cartel Price Path," in *The Political Economy of Antitrust*, ed. by V. Ghosal and J. Stennek, Elsevier, vol. 282 of *Contributions to Economic Analysis*, 59–80.

CHRISTIE, W., J. HARRIS, AND P. SCHULTZ (1994): "Why Did NASDAQ Market Makers Stop Avoiding Odd-Eighth Quotes?" *The Journal of Finance*, 49, 1841–1860.

CLARK, R., D. COVIELLO, J.-F. GAUTHIER, AND A. SHNEYEROV (2018): "Bid rigging and entry deterrence in public procurement: Evidence from an investigation into collusion and corruption in Quebec," *The Journal of Law, Economics, and Organization*, 34, 301–363.

CONLEY, T. G. AND F. DECAROLIS (2016): "Detecting bidders groups in collusive auctions," *American Economic Journal: Microeconomics*, 8, 1–38.

CUNNINGHAM, C., F. EDERER, AND S. MA (2021): "Killer acquisitions," *Journal of Political Economy*, 129, 649–702.

CYRENNE, P. (1999): "On antitrust enforcement and the deterrence of collusive behaviour," *Review of Industrial Organization*, 14, 257–272.

DAVIS, S. (2017): "Regulatory complexity and policy uncertainty: headwinds of our own making," *Working Paper*.

ENGSTROM, D. (2013): "Public Regulation of Private Enforcement: Empirical Analysis of DOJ Oversight of Qui Tam Litigation under the False Claims Act," *Northwestern University Law Review*, 107, 1689–1756.

GEHRIG, T. AND P. JOST (1995): "Quacks, lemons, and self regulation: A welfare analysis," *Journal of Regulatory Economics*, 7, 309–325.

GRAJZL, P. AND P. MURRELL (2007): "Allocating lawmaking powers: Self-regulation vs government regulation," *Journal of Comparative Economics*, 35, 520–545.

HARRINGTON, J. (2004): "Cartel pricing dynamics in the presence of an antitrust authority," *RAND Journal of Economics*, 651–673.

——— (2008a): "Detecting cartels," *Handbook of antitrust economics*, 213, 245.

——— (2008b): "Optimal Corporate Leniency Programs," *The Journal of Industrial Economics*, 56, 215–246.

Hendricks, K. and R. Porter (1988): "An Empirical Study of an Auction with Asymmetric Information," *American Economic Review*, 78, 865–883.

Imbens, G. and K. Kalyanaraman (2012): "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *The Review of Economic Studies*, 79, 933–959.

Imbens, G. and D. Rubin (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge University Press.

Innes, R. (1999): "Self-Policing and Optimal Law Enforcement When Violator Remediation is Valuable," *Journal of Political Economy*, 107, 1305–1325.

Kaplow, L. and S. Shavell (1994): "Optimal Law Enforcement with Self-Reporting of Behavior," *Journal of Political Economy*, 102, 583–606.

Kawai, K. and J. Nakabayashi (2022): "Detecting Large-Scale Collusion in Procurement Auctions," *Journal of Political Economy*, 130, 1364–1411.

Kawai, K., J. Nakabayashi, J. Ortner, and S. Chassang (2023): "Using Bid Rotation and Incumbency to Detect Collusion: A Regression Discontinuity Approach," *Review of Economic Studies*, 90, 376–403.

Korting, C., C. Lieberman, J. Matsudaira, Z. Pei, and Y. Shen (2023): "Visual Inference and Graphical Representation in Regression Discontinuity Designs," *The Quarterly Journal of Economics*, 138, 1977–2019.

Kovacic, W. (2001): "Private Monitoring and Antitrust Enforcement: Paying Informants to Reveal Cartels," *George Washington Law Review*, 69, 766–797.

LaCasse, C. (1995): "Bid rigging and the threat of government prosecution," *The RAND Journal of Economics*, 398–417.

Landes, W. and R. Posner (1975): "The Private Enforcement of Law," *The Journal of Legal Studies*, 4, 1–46.

Marshall, R. C. and L. M. Marx (2012): *The economics of collusion: Cartels and bidding rings*, Mit Press.

Martin, J., M. Pesendorfer, and J. Shannon (2022): "Underbidding for Oil and Gas Tracts," *Working Paper*.

Miller, N. (2009): "Strategic Leniency and Cartel Enforcement," *The American Economic Review*, 99, 750–768.

Monticini, A. and D. Thornton (2013): "The effect of underreporting on LIBOR rates," *Journal of Macroeconomics*, 37, 345–348.

Morgan, K. and D. Rubin (2012): "Rerandomization to improve covariate balance in experiments," *The Annals of Statistics*, 40, 1263 – 1282.

Motta, M. and M. Polo (2003): "Leniency programs and cartel prosecution," *International Journal of Industrial Organization*, 21, 347–379.

Ohashi, H. (2009): "Effects of transparency in procurement practices on government expenditure: A case study of municipal public works," *Review of Industrial Organization*, 34, 267–285.

Ortner, J., S. Chassang, K. Kawai, and J. Nakabayashi (2022): "Screening Adaptive Cartels," .

Parker, C. (2002): *The open corporation: Effective self-regulation and democracy*, Cambridge University Press.

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

Polinsky, M. (1980): "Private versus Public Enforcement of Fines," *The Journal of Legal Studies*, 9, 105–127.

Porter, R. (2005): "Detecting Collusion," *Review of Industrial Organization*, 26, 147–167.

Porter, R. H. and J. D. Zona (1993): "Detection of Bid Rigging in Procurement Auctions," *Journal of Political Economy*, 101, 518–38.

——— (1999): "Ohio School Milk Markets: An Analysis of Bidding," *RAND Journal of Economics*, 30, 263–288.

Schurter, K. (2017): "Identification and inference in first-price auctions with collusion," Tech. rep., working Paper, University of Chicago.

Seibel, R. and S. Škoda (2023): "Collusion by Exclusion in Public Procurement," *Working Paper.*

Sigler, J. and J. Murphy (1988): *Interactive Corporate Compliance: An Alternative to Regulatory Compultion*, Quorum Books.

Sokol, D. (2012): "Cartels, Corporate Compliance, and What Practitioners Really Think About Enforcement," *Antitrust Law Journal*, 78, 201–240.

Sokol, D. D. (2017): "Antitrust Compliance," in *The Oxford Handbook of Strategy Implementation*, Oxford University Press.

Spagnolo, G. (2005): "Divide Et Impera: Optimal Leniency Programs," *SSRN Working Paper.*

Storey, J. (2002): "A Direct Approach to False Discovery Rates," *Journal of the Royal Statistical Society. Series B*, 64, 479–498.

Trebbi, F. and M. B. Zhang (2023): "The Cost of Regulatory Compliance in the United States," *Working Paper.*

Wollmann, T. G. (2019): "Stealth consolidation: Evidence from an amendment to the Hart-Scott-Rodino Act," *American Economic Review: Insights*, 1, 77–94.