Cheap Signals, Costly Proof: Award-Layer Evidence Triage in Cartel Enforcement

Darcio Genicolo-Martins[1,*], Paulo Furquim de Azevedo[1]

[1] *INSPER, São Paulo, Brazil*

Abstract

Cartel enforcement begins before legal proof exists. Agencies often observe contract-award records long before they recover the bid-level evidence needed to prove coordination. This paper asks whether that cheap award layer can allocate costly forensic attention. Using São Paulo's BEC procurement platform (2009–2019) and CADE adjudications, we construct a *frequent-loser* screen from award records alone: among zero-win firms, persistent participation ranks loser-side cartel-adjacency risk. The screen targets forensic priority, not cartel membership. We validate the ranking against adjudication-anchored cobidder labels, discipline it with participation and opportunity-set exposure checks, timing and leakage audits, and direct-defendant scope checks, and compare it with bid-distribution forensics on the same target. The exposure checks ask whether the ranking survives among more similar procurement opportunity sets, rather than merely reflecting more chances to appear near CADE defendants. The award-layer signal adds non-redundant information to bid-layer forensics. Used sequentially, the award-layer ranking reduces the bid-microdata pool for the forensic stage by 83%while recovering 131 of 193 adjudicated cobidders. The results support an enforcement-design claim: cheap administrative records can discipline where proof-producing investigation should begin while liability remains in the richer evidentiary record.

Keywords: cartel screening, public procurement, award-layer data, enforcement triage, incomplete observability, bid rigging

*Corresponding author.
Email address: darciogm1@insper.edu.br (Darcio Genicolo-Martins)
We thank seminar participants at INSPER for comments; all remaining errors are our own.

JEL No.: D44, D73, H57, K21, L41

1. Introduction

Cartel enforcement begins before legal proof exists. An agency deciding where to investigate rarely starts with the communications, bid histories, leniency testimony, or tender-level conduct evidence that would support liability. It starts with thinner information: administrative records, complaints, procurement anomalies, repeated interactions, and case leads that make some firms and markets worth a closer look. The first legal-economic problem is therefore not proof itself. It is the allocation of proof-producing effort.

That distinction matters in procurement. The empirical cartel-screening literature usually starts from the bid layer: submitted bid amounts, bid ranks, within-tender dispersion, bid histories, and sometimes bidder networks. Classic procurement screens compare bidding patterns under competition and collusion (Porter and Zona, 1993; Bajari and Ye, 2003); variance and descriptive screens formalize suspicious bid distributions (Abrantes-Metz et al., 2006; Harrington, 2008) and field-test them in procurement settings (Imhof et al., 2018; Imhof, 2019). Recent work extends the approach through bidder groups (Conley and Decarolis, 2016), machine learning (Huber and Imhof, 2019; Wallimann et al., 2023), and robust screen design (Chassang et al., 2022). Those objects are close to the evidentiary stage, but they are not always cheap to obtain at scale. Many procurement systems make contract-award records routinely visible while keeping per-bidder bid files in operational systems that require case-specific recovery. An enforcement agency then faces a sequencing problem: before opening the costly bid layer, can it use the cheap award layer to decide where bid-layer forensics should begin?

This paper studies that question on São Paulo's BEC procurement platform, covering 1,654,401 tender-items from 2009 to 2019. BEC records a routine award layer—participants, winners, item codes, negotiated prices, procuring units, years, and modalities—and also preserves a richer bid layer recoverable through the LANCES export. CADE procurement-cartel adjudications provide external legal anchors. This institutional combination lets us study an earlier stage than most bid-distribution screens: the point at which an agency observes award records but has not yet recovered the bid-level evidence needed for forensic evaluation.

The award-layer statistic is deliberately simple. Among firms with zero wins, we rank participation intensity using $\log(1 + \text{tenders\_count})$ and implement an auditable binary rule that flags *frequent losers*. The economic logic is monotone rather than classificatory. If coordinated bidders use loser-side participation to create the appearance of competition while avoiding the win, persistent zero-win participation can rank loser-side exposure under a transparent behavioral condition. The statistic does not identify a firm type or assign cartel membership. In the paper's staged design, the rank allocates forensic attention; the richer bid layer remains where tender-level conduct and agreement are evaluated.

That legal scope determines the validation target. Direct CADE defendants are legal membership anchors, but they are not the natural object of a zero-win loser-side screen. Procurement cartels often allocate winning roles, rotate designated winners, and use other firms on the losing side of the tender (Pesendorfer, 2000; Asker, 2010). The same role-allocation logic is central in cartel theory and dynamic case evidence (Marshall and Marx, 2012; Clark et al., 2021). Large-scale procurement evidence shows how such arrangements can appear in administrative bidding records (Kawai and Nakabayashi, 2022). We therefore validate the screen against always-loser firms that bid alongside direct CADE defendants in adjudicated cartel environments. These cobidders define an adjudication-anchored exposure target. Direct defendants anchor legally adjudicated environments; cobidders are the loser-side footprint the award layer can plausibly rank before bid-level proof is recovered.

The empirical design is organized around the threats that would make such a screen uninformative for enforcement. A raw co-bidding association could reflect participation volume or opportunity-set exposure rather than signal: high-volume firms, or firms active in the same product, buyer, year, and modality environments as CADE defendants, have more chances to appear near those legal anchors. We therefore compare observed concentration with participation-disciplined placebos and exposure-adjusted audits that compare firms within more similar opportunity sets for meeting CADE defendants. Appendix D reports the exact exposure, leakage, and timing audit construction. A retrospective score could reuse information from the target period, so we report timing exercises that form the score before the evaluation window and distinguish relaxed from strictly ex ante definitions. A loser-side statistic could be overread as a cartel-membership detector, so we test whether it also classifies direct defendants. It should not.

The results support a limited but operationally useful claim. The frequent-loser ranking concentrates adjudication-anchored cobidders inside the always-loser stratum, including when the score is formed before the target window. It performs poorly against direct CADE defendants, consistent with its loser-side scope rather than with a generic cartel-membership interpretation. Cobidders inside the frequent-loser stratum also look different from other frequent losers: they are deployed more broadly, appear closer to legal cartel anchors, operate in more concentrated product portfolios, and display bid-level patterns consistent with, but not diagnostic of, credible losing roles. Those patterns give the adjacency target economic content and make it harder to reduce the result to ordinary high-volume losing alone.

The main institutional result is a cost-of-evidence result from the sequencing exercise. On the same adjudication-anchored target, the award-layer score adds information to an Imhof–Wallimann-style bid-distribution benchmark. The two layers therefore enter at different stages: award records prioritize where to look, and bid records evaluate what is found. Used as a gatekeeper, the award-layer ranking cuts the bid-microdata pool for the forensic stage by 83% while recovering 131 of 193 adjudicated cobidders. Joint scoring with award and bid features is the full-observability upper bound; sequential gatekeeping is the architecture relevant for agencies that must decide where to spend costly bid recovery.

This paper makes three contributions. First, it reframes procurement-cartel screening as evidence allocation under costly observability. Existing bid-distribution screens show how to evaluate suspicious bidding once bid-level data have been recovered; we study the prior institutional stage, where agencies observe award records but must decide where recovery should start. This turns screening toward enforcement triage, not toward treating a statistical flag as proof (Harrington, 2008; Sánchez Graells, 2019); robust-screen design formalizes the same distinction between flags and evidentiary conclusions (Chassang et al., 2022). Second, the paper shows that a minimal award-layer statistic—persistent zero-win participation—ranks a legally limited loser-side adjacency target anchored in CADE adjudications. Third, it evaluates a sequential architecture in which cheap award-layer triage gates costly bid-layer forensics, reducing the evidentiary footprint without moving liability out of the richer case record.

The boundaries are part of the design. The award layer ranks forensic priority; the validation target is adjudication-anchored exposure; and the bid layer remains the proof-producing stage. Price evidence is scope information, not a damages exercise. Strategic response also shapes implementation: a static cutoff invites gaming, while refreshed thresholds, continuous ranks or capacity-constrained queues, bunching checks, and bid-layer follow-up fit the staged enforcement architecture.

The broader lesson is institutional. A legal system that waits for full proof before ranking investigative priorities will spend scarce forensic capacity poorly; a legal system that treats cheap suspicion as proof will overreach. The useful middle ground is a disciplined triage rule: one that is cheap enough to run before proof exists, informative enough to order forensic attention, and legally limited enough to leave liability where it belongs—in the richer evidentiary record.

2. Costly Proof and Observable Awards

The paper's empirical object is an institutional sequence, not a stand-alone classifier. An enforcement agency first observes a thin administrative record, then decides whether to spend resources recovering richer bid-level evidence, and only later can that evidence contribute to legal proof. This section defines that sequence for BEC procurement: the legal role of a screen, the information layers available to the agency, and the adjudication-anchored labels used to validate a loser-side ranking. Subsection 2.1 states the screening-before-proof problem. Subsection 2.2 separates the award layer from the forensic-recoverable bid layer. Subsection 2.3 defines the CADE anchors and the loser-side cartel-adjacency target used for validation.

2.1. Screening Before Proof

Antitrust agencies do not choose investigation under full information. They choose it while the facts needed for liability are still costly to obtain. In a procurement-cartel case, proof requires evidence of coordinated conduct: communications, allocation rules, bid rotation, cover bidding, or tender-level patterns inconsistent with independent rivalry. Brazilian competition law treats bid rigging as an infraction under Lei 12.529/2011; comparable competition-law frameworks likewise require evidence beyond suspicious administrative patterns (Hovenkamp, 2005). A pattern in award records does not contain those legal predicates.

The institutional question is therefore earlier than proof and narrower than liability. The Becker–Stigler tradition treats enforcement intensity, investigation, and sanctions as costly instruments (Becker, 1968; Stigler, 1970). Cartel-screening work makes the same allocation logic operational: when investigation is costly and misconduct is only partially observable, agencies should use cheap screens to decide where expensive inquiry has the highest expected return (Baker, 2003; Harrington, 2008; OECD, 2022). Leniency policy illustrates the same enforcement trade-off in cartel cases (Wils, 2007, 2016). For procurement cartels, the first-stage problem is not whether a cheap statistic can decide liability. It is whether a cheap statistic can order the queue for recovering proof-producing evidence.

This is the sense in which the paper studies triage rather than classification. A model that uses submitted bid amounts, bid ranks, and within-tender bid moments assumes the forensic layer has already been opened. That is the right information set for evaluating suspicious bidding, but not for deciding where the first costly bid-file request should be sent. An award-layer screen has value only at this earlier point. Its output is forensic priority, not an accusation.

The legal-economic object is *forensic priority*. A useful triage statistic must be computable from records already available before case-specific bid recovery; it must rank firms or environments where richer evidence of cartel conduct is more likely to be worth recovering; and it must remain legally limited, so that administrative suspicion does not substitute for proof. The BEC setting supplies the two data layers and the external adjudication anchors needed to evaluate whether an award-layer statistic can play that role.

2.2. Award and Bid Layers in BEC

São Paulo's BEC platform is useful because it preserves the two information layers that an enforcement sequence requires. The sample contains 1,654,401 tender-items from 2009 to 2019. Two procurement modalities dominate. *Convite* is a sealed-bid invitation procedure under Lei 8.666/93 with a statutory three-bidder quorum and a low value cap. *Pregão* is a reverse electronic auction under Lei 10.520/2002, with progressively lower real-time offers and no minimum-bidder rule. The procedures differ, but both generate a routine award record before any case-specific bid recovery occurs. This matters because procurement rules, auction format, and bureaucratic implementation can shape both bidding behavior and observed procurement outcomes. Auction format affects participation and bidding incentives (Athey et al., 2011; Decarolis, 2014); public procurement outcomes also depend on bureaucratic incentives, competence, discretion, and rules (Coviello and Gagliarducci, 2017; Decarolis et al., 2020); rules and discretion create an additional institutional margin (Decarolis et al., 2025).

The *award layer* is that routine administrative envelope. It records participant identity, winner identity, item code, negotiated price, procuring unit, year, and modality. These fields show who appeared, who won, what was bought, and where a firm repeatedly participated without winning. They do not reveal the bids submitted by each participant, the timing of offers, bid revisions, or within-tender dispersion. The award layer is therefore cheap enough for broad monitoring, but too thin to supply legal proof of coordination.

The *bid layer* is the forensic-recoverable layer. BEC preserves per-bidder offers, timestamps, and bid-revision sequences in the LANCES export. Those records support bid-distribution forensics: within-tender dispersion, skewness, spread, distance from the winning bid, and related Imhof–Wallimann-style features (Imhof et al., 2018; Imhof, 2019). Machine-learning variants combine these bid moments at scale (Huber and Imhof, 2019; Wallimann et al., 2023). But these features require the bid file to be recovered, cleaned, linked, and inspected. They are not available at the same cost or timing as an award-record query.

The empirical design treats that cost difference as substantive rather than incidental. If bid microdata were already available for every firm at no meaningful cost, the agency could run forensic screens directly. The institutional problem arises because the bid layer is available but costly: feasible for case work and validation, but not the natural first object for system-wide triage. The issue is therefore cost, timing, and processing burden rather than data non-existence.

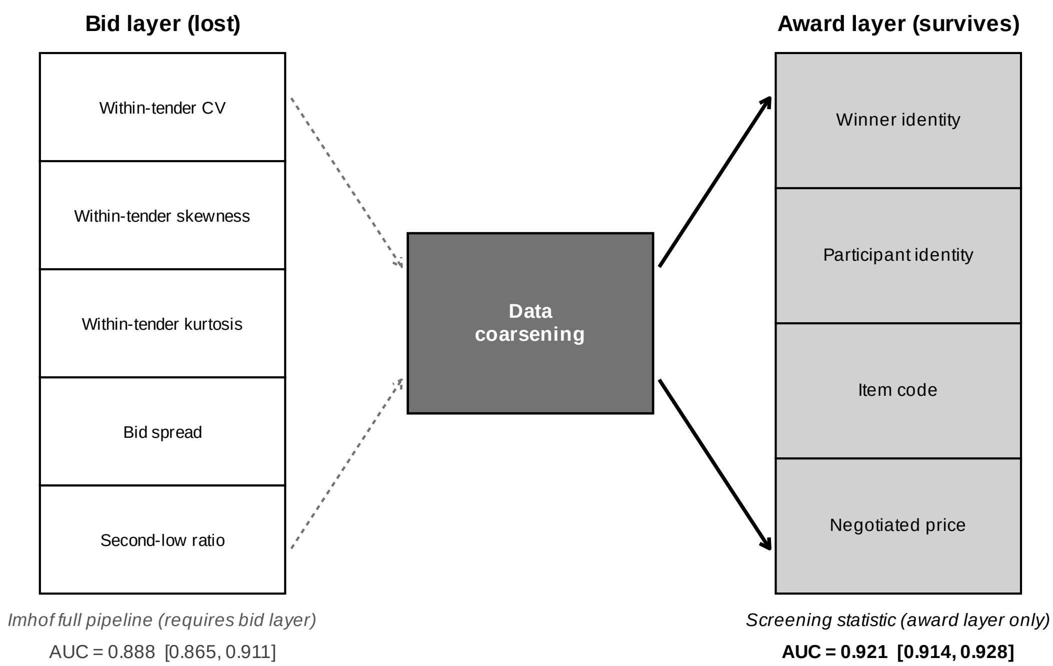

Figure 1: Information layers in procurement-cartel enforcement. Bid-distribution forensics requires per-bidder bid amounts and within-tender bid moments; the award-layer screen uses participant identity, winner identity, item code, and negotiated price. The question is whether the cheap layer can rank where the costly bid layer should be opened first.

Figure 1 summarizes the distinction. It is an information-cost diagram, not a detector horse race. The bid layer is not absent; it is forensic-recoverable through LANCES and therefore costly to open at scale. Bid-distribution methods remain the natural forensic tools once bid microdata have been recovered. Award-layer triage asks whether the thinner routine record can rank where that recovery should start.

2.3. CADE Anchors and Cartel-Adjacency Labels

CADE adjudications supply the legal anchors, not the screen's labels mechanically. The linked BEC portfolio contains 12 adjudicated procurement-cartel cases with 65 firm-defendants, of which 47 are active in BEC during the sample window. Final decisions span 2015–2025, with long process-to-ruling lags. That timing is central: if a cheap screen is useful, it is because it can move investigative attention before the completed adjudication record exists.

The adjudication record creates two objects with different roles. Direct CADE defendants are legal membership anchors: firms named in adjudicated procurement-cartel cases and matched to BEC participant records. They locate environments where legal process established cartel conduct. They are not, however, the natural target of a zero-win loser-side statistic. Direct defendants may occupy winner-side, allocation, or rotating roles in procurement cartels. A screen built from repeated losing should not be judged by whether it classifies those firms as if it were a generic cartel-membership detector.

The validation target is narrower. A cobidder is an always-loser firm that appears in at least one BEC tender-item with a direct CADE defendant. This definition yields 193 cartel-adjacent loser-side firms in the main validation target. Cobidder status is not legal membership and is not treated as a latent truth about agreement. It is an adjudication-anchored exposure label: the firm is observed on the losing side near a legally established cartel anchor. The label is therefore appropriate for evaluating forensic priority, not liability.

This distinction avoids two errors. The first would be to collapse defendants and cobidders into a single cartel class, overstating what award records can establish. The second would be to require a loser-side screen to recover direct defendants, misstating what the statistic is designed to rank. The relevant question is instead whether persistent zero-win participation concentrates the loser-side firms most plausibly worth bid-layer follow-up around adjudicated cartel environments. Because adjacency can also arise from opportunity—firms active in the same items, years, modalities, or buyers as CADE anchors have more chances to meet them—the validation design later disciplines participation volume and opportunity-set exposure separately from the ranking signal.

Table 1 fixes the roles used throughout the paper. The outer BEC population defines where procurement participation is observed. The always-loser stratum defines the side of the bidding pool the award layer can rank. Frequent losers are the operational screen. Direct defendants are legal anchors. Cobidders are the loser-side adjacency target for validation.

Table 1: Data Layers, Populations, and Legal-Economic Roles

Object	Size / layer	Role in the analysis
Award layer	Routine record	Participants, winners, item codes, prices, procuring units, years, and modalities; the cheap triage layer.
Bid layer	LANCES export	Per-bidder offers, timestamps, revisions, and bid moments; the costly forensic layer.
All BEC participants	41,444 firms	Outer firm universe for award-record construction.
Always-losers	16,843 firms	Zero-win stratum in which the award-layer screen ranks loser-side exposure.
Frequent losers	2,735 firms	Operational first-stage flag within always-losers; not a legal category.
Direct CADE defendants in BEC	47 firms	Legal anchors from adjudicated procurement-cartel cases.
Always-loser cobidders	193 firms	Firms losing alongside direct defendants; cartel-adjacent validation target, not legal membership.

Notes: Direct CADE defendants are legal anchors. Always-loser cobidders are adjudication-anchored exposure labels. Frequent losers are an operational screen, not a legal category. The categories separate legal membership from screening exposure.

The rest of the paper keeps these objects separate. The screen operates inside the always-loser stratum. The validation target is cobidder adjacency. Direct defendants anchor the environments but do not define the screen's success. Bid-layer forensics remains the proof-producing stage. This separation lets the paper evaluate a legally limited triage architecture without turning suspicion into adjudication.

3. Award-Layer Triage

The award layer observes outcomes before it observes conduct. A losing bid may be sincere, poorly priced, capacity constrained, exploratory, strategic, or coordinated. The screen therefore uses what the award layer can measure at the triage stage: which firms appeared, which firms won, and which firms kept appearing without ever winning. The object is a ranking for forensic priority, not proof of conduct or a classification of cartel membership.

This section defines that ranking and the empirical implications that discipline it. Subsection 3.1 introduces persistent zero-win participation as the award-layer primitive. Subsection 3.2 explains why the continuous ranking can be converted into an operational queue or binary frequent-loser flag. Subsection 3.3 states the tests that the rest of the paper uses to evaluate whether the ranking is informative for enforcement.

3.1. Persistent Zero-Win Participation

Let T_i denote the number of BEC tender-items in which firm i participates, and let W_i denote the number it wins. The award-layer screen operates inside the zero-win stratum, $\{i : W_i = 0\}$, which contains 16,843 firms. Within that stratum, the continuous score is

$$s_i = \log(1 + T_i).$$

The log transformation stabilizes the scale but does not change the ordering. The economic object is the rank: among firms the award layer observes only as losers, which ones appear persistently enough to warrant bid-layer follow-up?

The maintained condition is monotonic rather than classificatory. If cartel-deployed losing participants are present, they should appear more persistently in the zero-win stratum than ordinary losing firms, because such a role is useful only through repeated deployment in cartel-targeted tenders. The condition is maintained, not derived. Appendix B formalizes it as a monotone-likelihood-ratio ranking result. The main text uses it more modestly: persistent zero-win participation is a cheap statistic that can rank loser-side exposure under a transparent behavioral assumption and direct attention to the bid layer where conduct can be evaluated. It does not assign a firm type or cartel membership.

The zero-win restriction is therefore a scope restriction. Many firms never win for ordinary competitive reasons: poor fit, high costs, capacity constraints, market exploration, subcontracting positioning, or simple bad luck. The screen's content comes from persistence inside that heterogeneous loser-side set. If the ranking has enforcement value, it should concentrate adjudication-anchored loser-side adjacency beyond what generic participation volume, product-market overlap, and opportunity to meet CADE anchors would mechanically produce.

3.2. From Ranking to an Operational Queue

The continuous score is the central object because agencies can translate a rank into different investigative protocols. A team with fixed capacity could inspect the top k firms. A monitoring unit could refresh the rank each quarter and examine bunching below old cutoffs. A case team could use the score to request LANCES bid files for a survivor pool before applying bid-distribution forensics. These implementations differ, but all preserve the same sequencing logic: award records first, costly bid recovery second.

For an auditable binary implementation, we also report a frequent-loser flag. The rule marks zero-win firms whose participation count exceeds the median plus 1.5 times the interquartile range of the always-loser participation distribution. In BEC, the statistical cutoff is 13.5, implemented as $T_i \geq 14$, which flags 2,735 firms, or 16.2% of the always-loser stratum. The cutoff is fixed from the participation distribution alone, before any CADE label enters the analysis.

The cutoff is an auditable administrative implementation of the rank, not a structural or legal threshold. The paper does not rest on the number 14; it rests on the continuous loser-side participation ranking. Two analysts applying the same rule to the same award records obtain the same first-stage list. Above the cutoff does not mean cartel member; below it does not mean clean. Because a static threshold can be gamed, the enforcement architecture treats the binary flag as one implementation among several: agencies can also use continuous ranks, top-k queues, refreshed thresholds, or budget-calibrated survivor pools. The strategic-adaptation analysis later returns to this point.

3.3. Testable Implications

The screen is useful if it survives the threats implied by its limited information. Table 2 organizes the validation map. The first implication is the core validation claim: persistent zero-win participation should rank adjudication-anchored loser-side cobidders within the always-loser stratum. The main threat is mechanical exposure: frequent participants, and firms operating in the same procurement environments as CADE defendants, have more opportunities to appear near them even under ordinary competition.

The second implication is a scope check. A loser-side statistic should perform worse against direct CADE defendants than against always-loser cobidders, because it ranks a different role in the procurement environment. The third implication asks whether the cobidder target has economic content: among frequent losers, cobidders should differ from other frequent losers in directions consistent with loser-side exposure, proximity to defendants, market concentration, and bid-level losing behavior. These profiles give the validation target economic content; they do not prove conduct or assign legal status.

The final implications evaluate the enforcement architecture. If award records and bid records carry different information, the award-layer score should add non-redundant signal to bid-distribution forensics. If the screen has institutional value, using it before bid recovery should reduce the bid-microdata footprint while preserving much of the adjudication-anchored recall. And because firms can respond to published administrative rules, the binary threshold should be treated as vulnerable to strategic adaptation rather than as a fixed rule immune to gaming.

Together, these tests evaluate the screen at the stage for which it is designed: the allocation of costly forensic attention before liability evidence is assembled. The BEC cutoff is an implementation detail; the portable object is the staged architecture that uses cheap administrative records to discipline where costly cartel forensics should begin.

4. Validating the Loser-Side Ranking

This section validates the award-layer ranking, not the legal status of ranked firms. The ranking is useful for enforcement only if it orders the loser-side firms for which bid-layer follow-up is most valuable. The validation target is therefore the one defined in Section 2.3: always-loser firms observed alongside direct CADE defendants in adjudicated procurement-cartel environments. These cobidders are adjudication-anchored exposure labels. They are not treated as legal members of a cartel.

Table 2: Testable implications and validation threats

Implication	Prediction	Test	Main threat
Loser-side ranking	Persistent zero-win participation ranks always-loser cobidders.	Conservative benchmark, timing audit, exposure-adjusted validation.	Participation volume, market overlap, and opportunity to meet CADE anchors.
Scope asymmetry	Performance is weaker for direct defendants than for cobidders.	Direct-defendant AUC and temporal direct-defendant check.	Misreading the screen as generic cartel-membership detection.
Economic profile	Cobidders differ from other frequent losers in loser-side exposure and bid patterns.	Firm-level and bid-level profile comparisons.	Product-market composition and ordinary high-volume losing.
Complementarity	Award-layer signal adds information to bid-layer forensics.	Same-target award, bid, and combined AUCs.	Full-observability benchmark confused with operational sequencing.
Gatekeeping	Award-first sequencing reduces bid-microdata recovery with limited recall loss.	Sequential survivor-pool exercise and cost-recall frontier.	Survivor-pool choice and arbitrary budget choice.
Strategic response	Static cutoffs are more fragile than ranks or refreshed queues.	Adaptation diagnostics.	Simulation is not an equilibrium model.

Notes: Predictions evaluate whether the award-layer ranking is useful for evidence allocation. They do not define legal membership or proof.

The validation exercise has to do more than report a high AUC. The main empirical threat is exposure. Firms that participate often, or that operate in the same procurement environments as CADE defendants, have more opportunities to appear near those anchors even under ordinary competition. Procurement-screen designs therefore have to separate suspicious conduct from market structure, bidder heterogeneity, and ordinary repeated participation (Porter and Zona, 1993; Bajari and Ye, 2003). Network and large-scale procurement screens sharpen the same concern (Conley and Decarolis, 2016; Kawai and Nakabayashi, 2022). A useful award-layer screen must therefore concentrate cartel-adjacent loser-side firms beyond what participation volume and opportunity-set exposure alone would imply. A second threat is timing: an enforcement screen should rank candidates before the target information is used. A third threat is legal overreach: a loser-side screen should not be read as a generic cartel-membership detector.

This section evaluates the ranking against those threats. Subsection 4.1 reports the baseline adjudication benchmarks. Subsection 4.2 disciplines the result with participation, exposure, leakage, and timing checks. Subsection 4.3 uses direct CADE defendants as a scope check. Subsection 4.4 summarizes what the validation establishes for the staged enforcement architecture.

4.1. Baseline Adjudication Benchmarks

The baseline asks a simple question before adding stricter discipline: do firms flagged by the award-layer rule appear disproportionately among always-loser firms located near adjudicated cartel anchors? The answer is yes. In the conservative adjudication benchmark, we restrict the CADE portfolio to cases resolved before the end of 2020. This leaves 4 cases, 30 BEC-active firm-defendants, and 210 always-loser cobidders. Among those cobidders, 108 are frequent losers.

This benchmark uses early legal anchors and asks whether the award-layer rank concentrates the associated loser-side footprint. Within the always-loser stratum, the observed frequent-loser share among conservative cobidders is 3.95%, compared with a participation-stratified placebo mean of 1.24%. The excess ratio is 3.2× with $p < 0.001$. The corresponding ROC AUC is 0.748 (95% DeLong CI [0.713, 0.783]) (DeLong et al., 1988). These numbers establish the baseline empirical fact: persistent zero-win participation is not random with respect to adjudication-anchored loser-side exposure.

The second benchmark adds timing discipline. The score is computed from 2009–2016 participation and evaluated against 2017–2019 cobidder labels. This is a temporal holdout, not a real-time field deployment: it asks whether earlier award records rank firms whose cartel adjacency is later observed in the adjudication-anchored record. The firm-level AUC is 0.864 (95% CI [0.858, 0.870]), below the full-sample AUC 0.924 but well separated from random ranking. Figure 2 reports the rolling-origin version of the same exercise.

The baseline results are deliberately read as ranking evidence. They show that the award-layer statistic is informative about where cartel-adjacent loser-side exposure later appears. The next subsection tightens that reading by asking how much of the signal remains after disciplining participation volume, market exposure, leakage, and strict timing.

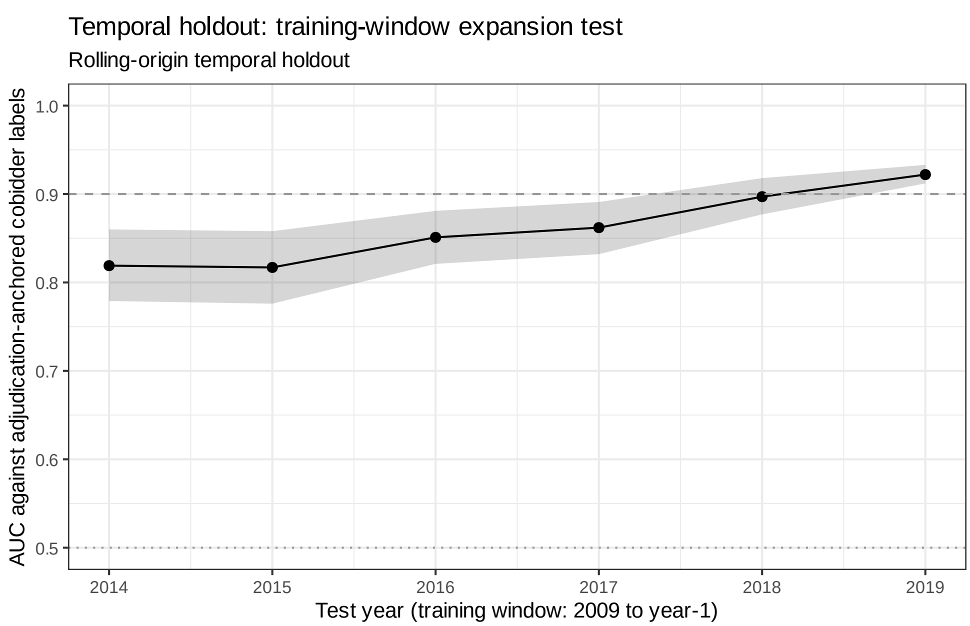

Figure 2: Temporal-holdout ROC curves. Scores are computed using participation observed before the test year; the test year supplies the out-of-time cobidder label. The exercise measures out-of-time discrimination against later-adjudicated cobidder labels, not field deployment.

4.2. Exposure and Timing Discipline

The baseline result becomes useful for enforcement when the main sources of mechanical exposure are made explicit. The first is participation volume. In the conservative benchmark, firms are grouped by participation quartiles within the always-loser stratum, cobidder status is permuted within quartiles 1,000 times, and the frequent-loser share among placebo cobidders is recomputed. This keeps the participation distribution fixed while breaking the particular adjacency to CADE defendants. The observed share, 3.95%, remains well above the placebo mean, 1.24%. This audit disciplines activity volume; it does not by itself make firms' opportunities to meet CADE anchors comparable.

A formal version of the same volume placebo permutes the frequent-loser label among always-losers while preserving the marginal participation distribution. Across $B = 2{,}000$ draws, the sham AUC distribution has mean 0.500, standard deviation 0.013, and 99th percentile 0.531. The observed FL14 AUC of 0.924 falls thirty-two standard deviations above the sham mean, with permutation $p < 1/2{,}000$. Volume alone does not reproduce the observed concentration. The price coefficient under the same sham distribution does not reject the null—observed +0.064 against sham mean +0.144, $p = 0.989$—which is consistent with the scope-not-damages reading of the price evidence in §7.2.

The second source is market opportunity. Firms active in the same procurement niches as CADE defendants have more chances to appear near those anchors even when conduct is ordinary. Exposure adjustment is therefore a required audit, not optional robustness. The audit asks whether the ranking survives after comparing firms within more similar opportunity sets—defined by procurement environments in which a firm could have appeared with CADE anchors, using item, modality, year, and buyer information where support permits. Its purpose is to ask whether the award-layer rank continues to concentrate adjudication-anchored cobidders after the markets in which firms could meet CADE anchors are made more comparable. The audit is not a legal-membership test and does not eliminate all opportunity-set concerns. Appendix D.2 reports the exact construction and retained support.

Timing is disciplined in two layers. The temporal holdout above forms the participation score before the evaluation window and tests later labels. A stricter version applies the same ex ante logic to the screen definition itself: zero-win status, the participation threshold, the frequent-loser flag, and the continuous score are all formed using only pre-window records. In that strict train-period audit, the threshold becomes 7 and the firm-level frequent-loser AUC is 0.767 (95% CI [0.734, 0.800]). The continuous score gives AUC 0.750 (95% CI [0.706, 0.795]). The decline relative to the relaxed holdout is informative: real-time screening has less information than a full-window reconstruction, but the ranking remains useful under a stricter field-like information set. It remains a timing audit, not causal prediction.

Two structural facts make the timing audit harder to dismiss as a same-firm artifact. First, firm persistence between the 2009–2016 and 2017–2019 windows is 8.7%: roughly nine in ten always-loser firms in the test window are firms not seen in the training window. The temporal holdout therefore evaluates a substantially fresh firm population rather than the same firms in different years. Second, when the training window is shortened further (scores frozen on 2009–2015, evaluated against cobidders linked to CADE adjudications closed after 2019—a strictly out-of-time target), the firm-level AUC is 0.791 for the frequent-loser flag and 0.854 for the continuous score. The ranking remains informative when the classifier is trained at one horizon, the cobidder label is defined by adjudications at a later horizon, and the bulk of test-period firms were not in the training set. Table 3 reports the full three-classifier battery.

Table 3: Three-Classifier Timing Battery

	vs cobid_all		vs cobid_post2019	
Classifier	FL flag	Continuous	FL flag	Continuous
clf_2015 (train 09–15)	0.791	0.851	0.786	0.854
clf_2017 (train 09–17)	0.856	0.897	0.844	0.894
clf_2019_full (in-sample)	0.924	0.939	—	—

Notes: AUC against two cobidder targets across three classifiers. cobid_all uses cobidders linked to any CADE adjudication in the panel. cobid_post2019 uses cobidders linked only to adjudications closed after 2019, a strictly out-of-time target that cannot be in the clf_2015 or clf_2017 training data.

The leakage audit addresses a different concern: item-level validation can mechanically reuse the same firm histories that construct the score. We therefore treat the in-sample item-level AUC, 0.995, as a mechanical reference point rather than as the main validation estimate. Holding cobidder firms out of score construction within cross-validation folds leaves AUC 0.891 (95% CI [0.887, 0.894]). The surviving component is the signal relevant for triage: the award-layer rank continues to order cartel-adjacent loser-side exposure when the most direct reuse of firm labels in score construction is weakened.

4.3. Direct-Defendant Scope Check

The direct-defendant exercise aligns the validation target with the screen's legal-economic role. Direct CADE defendants are the legal anchors that locate adjudicated cartel environments. They need not be persistent zero-win firms: in procurement cartels, defendants may occupy winner-side, allocation, or rotating roles. The award-layer statistic, by construction, ranks the loser side of the procurement record. The empirical prediction is therefore asymmetry: the rank should be informative for cobidders and much less informative for direct defendants.

The data follow that prediction. Against direct defendants in the broader BEC firm universe, AUC is 0.49, close to random classification. This is the scope check doing its work. If the statistic were a generic cartel-membership detector, it should rank direct defendants as well as cobidders. It does not, which is the expected result for a zero-win loser-side ranking. Read together with the cobidder results, this pattern supports the staged interpretation. Direct defendants anchor cartel-affected environments; the award layer ranks the loser-side exposure around those anchors; the bid layer then supplies the conduct evidence for case-specific evaluation.

A complementary universe-anchored scope matrix systematically varies the universe (always-loser firms vs full BEC firm registry) and the positive class (cobidders vs direct defendants) to map where the ranking targets succeed and where they fail. The matrix produces a stronger result than a stand-alone null: raw participation count on the full BEC universe against direct CADE defendants returns AUC 0.383, which is *below* random. A high-participation BEC firm is less likely, not more likely, to be a direct CADE defendant under the loser-side scoring logic. The score does not simply fail to identify winner-heavy defendants; the relevant direction is reversed, consistent with the role separation observed in D4 (14.9% of direct defendants are always-losers; median win rate 0.261 versus 0.086 for cobidders). Table 4 reports the full eight-row scope matrix.

Table 4: Universe-Anchored Stratum Scope Matrix

Score	Universe	Positive class	AUC	Reading
Binary FL	Always-losers	Cobidders	0.924	Within-stratum prioritization
Continuous log_tc	Always-losers	Cobidders	0.939	Best within-stratum ranking
Binary FL	All BEC firms	Direct CADE	0.491	At chance; no membership detection
Participation count	All BEC firms	Direct CADE	0.383	**Below chance: actively repels winner-heavy**
Binary FL (frozen 09–16)	Always-losers as of 09–16	Cobidders	0.767	Weaker than pooled within-stratum
Continuous (train 09–16)	Items 17–19	Cobidder items	0.770	Some prospective generalization
Binary FL (frozen 09–16)	Items 17–19	Cobidder items	0.565	Weak binary item-level holdout
Continuous (train 09–16)	Items 17–19	Direct-defendant items	0.511	Direct-defendant null survives timing

Notes: AUC across an eight-row matrix that varies the score, the universe of firms or items being ranked, and the positive class. Row 4 is below chance because the loser-side score discriminates *against* high-participation winner-heavy defendants. The matrix structurally rules out a generic-detector reading.

4.4. Validation Summary

Table 5 is a threat map rather than a proof table. Each row corresponds to a threat identified before the tests are read: legal anchoring, participation volume, market exposure, timing, leakage, and scope asymmetry. The resulting claim is staged: the award-layer ranking supports a priority queue for the bid-layer stage, where tender-level conduct can be examined with richer evidence.

Taken together, the benchmarks support the screen at the stage for which it is designed. Persistent zero-win participation orders adjudication-anchored loser-side exposure well enough to allocate forensic attention, not to validate the legal status of ranked firms. The next step is economic interpretation: among firms already selected by the frequent-loser rule, do cobidders display a distinct loser-side profile rather than a mechanical collection of high-volume losers?

Table 5: Validation Architecture for the Award-Layer Ranking

Benchmark	Question	Result
Legal anchoring	Are frequent losers overrepresented among pre-2020 adjudication-anchored cobidders?	3.2× excess; AUC 0.748
Participation-volume discipline	Does concentration survive volume-preserving placebo assignment?	Placebo mean 1.24% versus observed 3.95%
Exposure discipline	Does the ranking remain informative under more comparable opportunity sets for meeting CADE anchors?	Exposure-adjusted audit over comparable opportunity sets; see Appendix D.2
Temporal holdout	Does pre-window participation rank later cobidder labels?	AUC 0.864
Leakage audit	Does signal survive when cobidders are held out of score construction?	Out-of-fold AUC 0.891
Direct-defendant scope check	Does the ranking preserve the distinction between legal anchors and loser-side exposure?	Direct-defendant AUC 0.49

Notes: Rows address distinct threats to interpreting the ranking as forensic priority. The benchmarks validate a legally limited ranking, not cartel membership, liability, or proof.

5. The Economic Profile of Cartel-Adjacent Losers

The validation results show that the award-layer ranking concentrates adjudication-anchored loser-side exposure. This section asks whether that exposure target has economic content. The comparison is intentionally demanding: among firms already inside the frequent-loser stratum, do cobidders differ from other frequent losers in dimensions relevant for loser-side forensic prioritization?

The answer matters for interpretation. If cobidders look like other frequent losers once both groups cross the same participation threshold, the validation target would be easier to read as mechanical high-volume losing. If they differ in deployment breadth, proximity to legal anchors, market specialization, and bid-level losing behavior, the target becomes more useful for forensic prioritization. This profile-based reading is consistent with evidence that auction collusion can operate through bidder roles, repeated relationships, and cartel organization rather than through a single universal price statistic (Baldwin et al., 1997; Porter and Zona, 1999; Pesendorfer, 2000; Asker, 2010; Clark et al., 2021). The evidence below gives the target economic content, without turning adjacency into legal membership or proof.

5.1. Comparing Cobidders to Other Frequent Losers

The firm-level comparison holds fixed the award-layer selection margin. Both groups are always-losers, and both cross the frequent-loser threshold. Cobidders are frequent losers observed with direct CADE defendants; the comparison group is frequent losers without that adjacency. The question is whether the adjudication-anchored target marks a different loser-side profile among firms already selected by persistent losing. The comparison does not by itself remove all market-composition concerns; its narrower role is to ask whether the target is indistinguishable from other firms already selected by persistent losing.

The differences are consistent with the triage interpretation. Cobidders are deployed more broadly: they participate in 136.5 tender-items, compared with 76.7 for other frequent losers ($d = +0.67$). They also face more unique winners, 24.8 versus 13.5 ($d = +1.00$). These are deployment measures, not legal conclusions: they show that cobidders occupy a larger loser-side footprint within the already-selected frequent-loser pool.

They are also closer to the legal anchors used in validation. Among cobidders, 1.5% of loss-bids face a direct CADE defendant, compared with 0.2% for other frequent losers ($d = +0.46$). Their portfolios are more concentrated in product space as well, with item-group HHI 0.380 against 0.288 ($d = +0.39$). The combination of broader deployment and narrower product concentration is the kind of profile that can make the target useful for forensic prioritization.

The profile is not one-dimensional. Cobidders have a lower share of own loss-bids in repeat winner-pair clusters than other frequent losers, 21.6% versus 33.4% ($d = -0.38$). That wrong-signed measure is informative: the cobidder label is not a mechanical proxy for a single winner-pair pattern. This is useful discipline: the profile is not a clean mechanism test. The stronger regularities are deployment breadth, proximity to CADE anchors, and product specialization within the frequent-loser stratum.

5.2. Bid-Level Behavior

The LANCES export lets us ask whether the distinction also appears inside submitted bids. This evidence uses the forensic layer; it is deliberately kept outside the award-layer screen. For firms with at least five usable losing bids in convite or pregão, we compute the gap between each losing bid and the winning bid in the same tender-item,

$$\frac{\text{bid} - \text{winning bid}}{\text{winning bid}},$$

winsorized at $[-0.99, 10]$ before firm-level aggregation. The usable sample contains 182 cobidders and 2,369 frequent-loser non-cobidders.

Two patterns matter. First, cobidders bid closer to winners: their median bid gap is 0.582, compared with 0.809 for other frequent losers ($d = -0.281$). Under a credible-losing-role interpretation, losing bids need not be visibly extreme; plausible losing bids can preserve the appearance of rivalry. This pattern is not uniquely diagnostic, and it does not show that any particular bid was a cover bid. Second, cobidders show higher within-firm dispersion in bid gaps, 1.207 versus 1.099 ($d = +0.147$), consistent with firms appearing in more varied losing positions across tenders.

The bid-level distinction is not just participation intensity in another form. In a logit of cobidder status within the frequent-loser stratum, controlling for $\log(1 + \text{tenders\_count})$, both signs remain: the median-gap coefficient is -0.708 ($z = -4.14$), and the dispersion coefficient is $+0.442$ ($z = +3.62$). The forensic layer therefore shows that the adjudication-anchored loser-side target is not reducible to participation intensity alone.

5.3. Interpretation

The profile gives economic content to the validation target. Within the frequent-loser stratum, cobidders are more broadly deployed, closer to legal cartel anchors, more concentrated in product space, closer to winners in bid gaps, and more dispersed across their own losing bids. These are dimensions on which a forensic queue should plausibly differ if the award-layer rank is capturing more than generic high-volume losing. The profile evidence supports forensic prioritization by making the cobidder target harder

Table 6: Cobidder Profile Within the Frequent-Loser Stratum

Dimension	Cobidders	FL non-cobidders	Difference
Total participations	136.5	76.7	$d = +0.67$
Unique winners faced	24.8	13.5	$d = +1.00$
Loss-bids facing direct CADE defendant	1.5%	0.2%	$d = +0.46$
Portfolio item-group HHI	0.380	0.288	$d = +0.39$
Repeat winner-pair share	21.6%	33.4%	$d = -0.38$
Median bid gap to winner	0.582	0.809	$d = -0.281$
Within-firm SD of bid gap	1.207	1.099	$d = +0.147$

Notes: Descriptive within-stratum comparisons. They characterize the cartel-adjacent loser-side profile; they do not establish legal membership, liability, or cover bidding in any particular tender.

to reduce to generic high-volume losing; it does not identify cover bidders or prove agreement.

The boundary is honest. Most of the within-stratum quadrant heterogeneity does not survive propensity-score matching on participation history. The largest cell (low-HHI, low-pairs) has an unmatched coefficient of +0.081 ($p = 0.072$) that shrinks under matching to +0.062 ($p = 0.235$); neither estimate reaches conventional significance at the 5% level, and the shrinkage under matching is what we expected if the cell-level heterogeneity is driven by participation volume rather than by mechanism. What survives is the bid-level signature—cobidders' bid gap to winner is closer ($d = -0.28$) and more dispersed than non-cobidder FLs, a moment-level pattern that is harder to attenuate by matching on firm-level covariates. The descriptive distinctness is robust; the causal/mechanistic distinctness is narrower, concentrated in bid-level behavior rather than in quadrant-level heterogeneity.

This evidence also clarifies the division of labor between data layers. The award layer ranks a priority set using information available before bid recovery. The LANCES bid layer then provides additional evidence that the priority set has loser-side behavior worth forensic attention. The profile evidence does not establish agreement, membership, liability, damages, or cover bidding in any specific tender. It fits the paper's staged logic: cheap records order attention, and richer records evaluate what the screen has prioritized.

The next section turns from interpretation to implementation. It asks whether the award-layer rank can be placed before bid-distribution forensics in an operational sequence that saves bid-microdata recovery while preserving adjudication-anchored recall.

6. From Award-Layer Triage to Bid-Layer Forensics

The preceding sections validate the award-layer ranking on the margin for which it is designed. Section 4 shows that persistent zero-win participation orders adjudication-anchored loser-side exposure; Section 5 shows that the exposed firms have an economically distinctive profile rather than merely high procurement volume. The remaining question is how an enforcement agency should use such a signal. The answer is not to stop at the award layer, and it is not to replace bid-layer forensics. A thin administrative record can rank priorities, but it cannot by itself evaluate the bid conduct that richer cartel evidence requires.

This section therefore turns from validation to sequencing. The screen is useful because it sits before costly proof-producing work. It uses records the agency can observe broadly: who participated, who won, and who kept appearing without winning. Bid-layer forensics uses submitted offers, within-tender bid moments, and distance from the winning bid. Those data are more probative but also more expensive to recover, clean, harmonize, and inspect at scale. The issue is not whether a cheap signal should replace a costly one; it is whether the cheap signal can discipline where the costly one should be opened first.

The section evaluates that architecture in three steps. First, it places the award-layer ranking and an Imhof–Wallimann-style bid-distribution benchmark on the same adjudication-anchored target, so performance can be read against information cost. Second, it asks whether the two layers are complements: the award layer measures repeated loser-side exposure across tenders, while the bid layer measures conduct inside tenders after the bid file is opened. Third, it implements the operational gatekeeper: award records create a survivor pool, and bid forensics rank firms inside that pool. The object is a cost-recall frontier for forensic attention, not a horse race under full observability.

This sequencing is the paper's central enforcement implication: administrative records rank where proof-producing forensics should begin, while bid-level evidence remains the stage at which conduct is evaluated. The value is measured by how much forensic scope falls while adjudication-anchored signal is preserved.

6.1. Bid-Layer Benchmark Under Costly Observability

The natural forensic comparator is the bid-distribution tradition. Imhof–Wallimann-style screens use within-tender bid moments—dispersion, skewness, spread, distance from the winning bid, and related features—to flag patterns consistent with coordinated bidding (Abrantes-Metz et al., 2006; Imhof et al., 2018; Imhof, 2019; Huber and Imhof, 2019; Wallimann et al., 2023). These features are closer to conduct than the award-layer screen is. They use bids themselves rather than the administrative trace of participation and awards. They are also more costly: the agency must recover usable bid histories, harmonize offers, construct within-tender moments, and map those moments to firms.

The relevant comparison is therefore not sophistication. The bid-layer benchmark is more forensic by construction. The question is how much ranking information is already available before the bid file is opened. Table 7 places the award-layer measures and the bid-layer benchmark on the same adjudication-anchored target: CADE-cobidder labels within the always-loser pool for which bid-distribution features can be computed. Holding the target fixed is important. It means that the table varies the information envelope rather than the legal-economic object being ranked. Table 7 is an information-cost table before it is a classification table.

The award-layer signal performs strongly despite its thinner information set. Within this common pool ($N = 16{,}779$), the binary frequent-loser flag reaches AUC 0.921 (95% CI [0.914, 0.928]), and the continuous participation score reaches AUC 0.884 ([0.860, 0.908]). The full bid-layer benchmark reaches AUC 0.888 ([0.865, 0.911]). When the two layers are combined, discrimination rises to 0.962 ([0.954, 0.969]).

These results should be read as an information-cost result. The award-layer statistic does not make bid-distribution forensics unnecessary, and the paper does not ask it to. Rather, a cheap record-layer statistic discriminates the same adjudication-linked loser-side target in the range of a richer bid-layer benchmark. That is enough to justify the sequencing question: if the award layer contains useful information before bid recovery, and if it adds information to the bid layer rather than merely duplicating it, then it can be used to decide where bid-layer forensics should begin.

Table 7: Discrimination at Different Information Costs

Model	Information required	AUC	95% CI
Frequent-loser flag	Award records: participation count and zero wins	0.921	[0.914, 0.928]
Participation intensity	Award records: log(1 + tenders_count) and zero wins	0.884	[0.860, 0.908]
Bid-layer benchmark	Per-bid offer values and within-tender bid moments	0.888	[0.865, 0.911]
Award layer + bid layer	Participation intensity plus bid-distribution moments	0.962	[0.954, 0.969]

Notes: Five-fold cross-validated AUC against CADE-cobidder labels within the always-loser pool for which the bid-distribution benchmark can be computed ($N = 16{,}779$). The target is fixed; models differ in required information. The comparison is about information cost, not substitution. The combined specification is a full-observability diagnostic, not the operating rule.

The AUC levels above use the full 16,779-firm cross-validation pool. The DeLong incremental test reported in the text (+0.035 for the frequent-loser flag, $p = 0.014$) is computed on the stricter same-sample subset of firms with complete bid-layer features ($N = 11{,}676$), where the bid-layer benchmark reaches AUC 0.846 and the frequent-loser flag 0.881. DeLong requires paired observations on identical firms, so this same-sample increment (+0.035) differs slightly from the across-pool level gap.

6.2. Complementarity Across Information Layers

Similar discrimination does not by itself establish a useful two-stage architecture. If persistent zero-win participation were only a noisy proxy for bid-distribution moments, the award layer would matter mainly when bid microdata are missing. The stronger test is incremental: once bid-layer moments are observed, does the award layer still contribute information about the adjudication-anchored loser-side target?

It does. Adding the award-layer statistic to the bid-layer benchmark raises discrimination to 0.962 ([0.954, 0.969]). In the identical same-sample audit, where all models are evaluated only on firms with complete bid-layer features, the frequent-loser flag adds +0.035 AUC over the full Imhof–Wallimann benchmark alone (DeLong $p = 0.014$) (DeLong et al., 1988). The combined specifications gain +0.096 to + 0.098 AUC over the bid-layer benchmark ($p < 0.001$). The increment is not an outperformance claim; it shows that the two layers are not measuring the same thing.

The distinction is substantive. Bid moments describe conduct inside tenders after the bid file has been opened: spread, dispersion, rank distance, and related within-tender patterns. Participation intensity describes repeated loser-side exposure across tenders before that file is opened. The award layer is therefore closer to investigative priority; the bid layer is closer to forensic evaluation. An enforcement architecture should preserve that division of labor. The award layer creates a reason to look first in some places rather than others. The bid layer then helps evaluate what is found after looking.

The combined model in Table 7 is thus a complementarity diagnostic, not the operational architecture. Running the combined model for every firm assumes full bid-layer observability before triage has occurred, which is precisely the resource problem the paper studies. The enforcement use of non-redundancy is sequential: use broad, cheap award records to create a survivor pool, then apply costly bid-distribution forensics within that pool. The next subsection implements that gatekeeper directly.

6.3. Gatekeeping the Forensic Stage

The operational question is whether the cheap layer can reduce the forensic workload while preserving much of the adjudication-anchored signal. This is a different estimand from stand-alone classification performance. A rule that requires bid-level features for every firm may score well, but it has already paid the full cost of forensic observability. A gatekeeper is useful only if it changes the size of the set for which that cost must be paid. Table 8 is therefore a cost-of-evidence table, not only a classification table.

The sequential rule imposes that institutional constraint directly. First, firms are ranked using only the award layer, and the top $K_1 = 2{,}000$ firms form the survivor pool. Second, bid microdata are recovered or processed only for those survivors. The bid-distribution benchmark is then applied within that smaller pool to select the final top-k priorities. The joint model, by contrast, gives the classifier both award-layer and bid-layer features for every firm before triage. It is an upper-bound benchmark under full observability, not an implementable first-stage policy when bid recovery is costly.

Table 8 reports the resulting cost-recall frontier. The evaluation pool contains 11,676 always-loser firms with computable bid-distribution features and 193 CADE-cobidder positives. At the top 500 flags, award-layer screening alone finds 66 positives with precision 0.132. The bid-layer benchmark alone finds 73 positives with precision 0.146, but it requires bid microdata for the entire pool. Joint scoring finds 112 positives with precision 0.224, again under full bid-layer observability. The sequential rule finds 96 positives with precision 0.192 while opening bid microdata for only 2,000 firms.

The same trade-off appears at the top 1,000 cutoff. Joint scoring recovers 142 positives, or recall 0.736, after observing bid-layer features for all 11,676 firms. The sequential rule recovers 131 positives, or recall 0.679, while requiring bid microdata for only 2,000 firms. Relative to joint scoring, the sequential architecture reduces the bid-microdata footprint by 83% at a recall cost of 8%. Put differently, at the top-1,000 cutoff the sequential rule recovers 131 of 193 cobidders while requiring bid microdata for only 2,000 of 11,676 firms. This is the cost-of-evidence result: much of the signal is retained after the forensic bid layer is opened for a much smaller survivor pool.

Table 8: Gatekeeping the Bid-Layer Forensic Stage

Rule	k	TP	Precision	Recall	Bid-microdata firms
Award-layer only	500	66	0.132	0.342	0
Bid-layer only	500	73	0.146	0.378	11,676
Joint scoring, award + bid	500	112	0.224	0.580	11,676
Sequential: award $\rightarrow$ bid, $K_1 =$ 2,000	500	96	0.192	0.497	2,000
Joint scoring, award + bid	1,000	142	0.142	0.736	11,676
Sequential: award $\rightarrow$ bid, $K_1 =$ 2,000	1,000	131	0.131	0.679	2,000

Notes: Evaluation pool: 11,676 always-loser firms with computable bid-distribution features; positives are 193 CADE-cobidders. The bid-microdata column reports how many firms require per-bid offer values before the rule can be applied. Joint scoring is a full-observability upper bound; the sequential rule is the operational gatekeeper under costly bid recovery.

The sequential rule is not designed to dominate the full-information joint model. Its value is that it approaches much of the full-information signal without requiring full-information forensic scope. That is the enforcement margin on which the paper's contribution rests. In an agency setting, scarce analyst time, data cleaning, subpoena capacity, and case-development attention cannot be applied with equal intensity to every participant. A useful screen does not decide liability; it decides which files deserve the next, more expensive form of evidentiary attention.

6.4. Timing the Gatekeeper

The gatekeeper exercise also needs a timing discipline. A ranking that uses the full sample may perform well because the feature history is long and close to the adjudication target. A stricter audit asks whether the same sequencing logic remains useful when the scores are formed before the target window. This is not a field deployment in real time; the labels remain adjudication-based. The restriction is on the information used to rank firms before bid-layer forensics is allocated.

The temporal audit constructs award-layer and bid-layer features using the 2009–2016 window and evaluates against later adjudication-anchored cobidder exposure. The holdout pool contains 8,257 always-loser firms with complete train-window bid features and 142 positives. This design asks whether the staged architecture survives when the ranking layer is separated in time from the validation target.

It does. At the top 500 cutoff, the award-only rule loses 47% of its in-sample true positives, joint scoring loses 24%, and the sequential rule loses only 9%. In levels, the sequential rule recovers 87 positives at top 500, compared with 85 under joint scoring. At top 1,000, it recovers 114 positives, compared with 111 under joint scoring. The corresponding precisions are 0.174 and 0.114 for the sequential rule.

The temporal result reinforces the sequencing interpretation rather than expanding it. It should not be read as proof that sequential gatekeeping is superior to joint scoring, that a fixed cutoff is the right rule in every setting, or that an agency can infer agreement from the award layer alone. It shows that conditioning bid-layer forensics on an earlier award-layer screen remains competitive under timing discipline while preserving the smaller forensic footprint. The conclusion of Section 6 is therefore operational: cheap administrative records can structure the order in which expensive proof-producing evidence is assembled, while the richer bid layer remains the place where conduct is evaluated.

7. Scope, Limits, and Price Corroboration

Section 6 carries the paper's central enforcement-design result. Price evidence enters this section in a narrower, but now substantive, role. The screen is not designed to estimate cartel markups, damages, or welfare losses. Overcharge measurement has its own evidentiary tradition (Bryant and Eckard, 1991; Connor, 2007), distinct from screening (Marshall and Marx, 2012; Harrington, 2008). What the price record can do, however, is reveal the empirical signature of the cover-bidding mechanism that motivates the screen. The naive FL-price association is positive (++0.064); under overlap-cell ATT weighting it reverses to negative (-0.097, $p < 0.001$). This sign reversal is not a specification artifact. It is the predicted decomposition into a positive selection component (cartels concentrate in cells with structurally higher underlying prices) and a negative within-cell mechanism component (cover bidders depress the observed winner price relative to the reference price by manufacturing the appearance of competition). Both components are documented empirically in this section.

This distinction matters for the paper's legal-economic position. The decomposition does not deliver a damages estimate because the observed price under cover bidding is a depressed artifact of the theater, not an undistorted measurement of the cartel's true overcharge. The decomposition does, however, establish that the FL screen identifies environments where the underlying cover-bidding mechanism is operative. The price record therefore disciplines the legal-economic interpretation in two directions: it rules out a naive damages reading (the positive baseline coefficient is selection, not effect), and it rules out the dismissal of price evidence as uninformative (the within-cell sign reversal is consistent with the mechanism the screen targets).

The section reports the broad descriptive price imprint (§7.1), the formal decomposition into selection and mechanism components (§7.2), the legal-economic implications (§7.3), and the adaptive deployment alternatives that follow from these findings (§7.4). The institutional implication remains: the award-layer signal is a trigger for richer evidence collection, not a stand-alone measure of price injury.

7.1. Price Evidence as Scope

The price regressions are reported to assess economic non-neutrality, not to estimate price injury. If the award-layer screen ranks environments where cartel-adjacent loser-side participation is more likely, it is informative to ask whether those environments also differ in negotiated unit prices. The answer can only play a corroborative role. The estimates below are descriptive associations, not cartel markups, causal effects of frequent-loser presence, or inputs into a damages calculation.

In the broad item-level regressions, frequent-loser presence is positively associated with log negotiated unit price. The coefficient is 0.0677^{***} with item and year fixed effects and 0.0636^{***} after adding procuring-unit fixed effects. The association is larger in pregão, 0.0933^{***}, than in convite, 0.0382^{**}. Across the main descriptive specifications, the implied price range is $+3.6\% - +7.7\%$.

The modality split is useful as scope discipline, not mechanism identification. If frequent losers were only mechanical quorum fillers under the three-bidder rule, the price imprint should be stronger in convite. Instead, the larger association appears in pregão, the electronic-auction environment without the same formal invitation structure. This contrast weighs against a simple procedural-artifact interpretation, but it still does not identify a causal procurement-price effect.

Table 9: Descriptive Price Imprint

Exercise	Estimate	Standard error	Use in the paper
Broad sample, item and year FE	0.0677^{***}	0.0230	Economic non-neutrality
Broad sample, item, year, PBU FE	0.0636^{***}	0.0215	Positive after buyer controls
Pregão subsample	0.0933^{***}	0.0255	Stronger in electronic auctions
Convite subsample	0.0382^{**}	0.0186	Smaller under quorum-prone procedure

Notes: Dependent variable is log negotiated unit price. Estimates are descriptive associations in the BEC item-level sample and are not interpreted as causal price effects, cartel markups, overcharges, or damages estimates.

The price imprint therefore corroborates the economic relevance of the priority set, but it does not carry the paper's central claim. The central claim remains about allocating the next unit of proof-producing effort. The next subsection uses the overlap and segment results to make that boundary explicit.

7.2. Decomposing the Sign Reversal

The sign reversal between the broad-sample coefficient (++0.064) and the overlap-cell ATT coefficient (−0.097) is not a specification puzzle to explain away or a boundary to be conservative around. The two coefficients measure two different objects, and the gap between them admits a clean economic decomposition. The broad sample averages across cells; cartels with cover bidders are not randomly distributed across cells but systematically concentrate where the underlying price level is structurally higher. The overlap-cell ATT removes that across-cell variation and isolates the within-cell effect; within cells, the cover-bidding theater mechanically depresses the observed winner price relative to the reference price. The naive baseline mixes the two opposing forces; the sign-reversed ATT spec isolates the within-cell mechanism.

We document both components empirically. The decomposition follows the standard separation between descriptive comparisons and causal estimands (Imbens and Rubin, 2015); coefficient stability and sensitivity logic reinforce the same caution (Oster, 2019; Cinelli and Hazlett, 2020).

Selection: cartels in high-price cells.. Among non-treated items (items without frequent-loser presence), mean log negotiated price rises monotonically across cell FL-share quintiles: 1.35 at Q1 (cells where ∼1.4% of items are FL-present) and 6.93 at Q5 (cells where ∼41% of items are FL-present). The gap is $\Delta = 5.58$ log-points across cells, a structural ordering of which products and procurement environments cartels choose to operate in. An item-level OLS regression of non-treated log price on cell FL-share with all five cell dimensions as marginal fixed effects yields a coefficient of +3.55 (SE 0.23, $N = 1{,}439{,}255$). The selection component is positive, large, and survives partialling out each cell-dimension marginal effect.

Mechanism: cover-bidding theater within cell.. Within overlap cells, frequent-loser presence depresses the observed winner-to-reference price ratio by −0.048 log-points (SE 0.004, $p < 10^{-30}$). The same regression with log(n_firms) as control returns +0.008 (SE 0.004, not significant): the within-cell price effect operates entirely through the bidder-count channel. Frequent-loser presence is associated with +0.507 log-bidders ($\approx$ 66%

more bidders) within the same cell type (SE 0.006, $p < 10^{-30}$); those additional bidders mechanically compress the winning bid toward the reference price. The cover-bidding theater is empirically detectable both in the bidder-count inflation and in its downstream price depression.

The bidder-count boundary.. The two opposing forces exchange dominance at the tender-density threshold. In sparse tenders (1–3 or 4–6 bidders), frequent-loser presence is associated with higher prices ($\Delta = +0.092$ and $+0.096$ in the two quartiles respectively): selection dominates because cover-bidding theater requires a critical mass of bidders to be visible. In dense tenders (11+ bidders), frequent-loser presence is associated with *lower* prices ($\Delta = -0.015$): the mechanism dominates because cover-bidder inflation is operative. The 7–10 bidder bucket is the transition ($\Delta = +0.052$). The selection-mechanism boundary is therefore observable at the tender-density threshold itself.

Together, the two components yield the empirical signature of cover-bidding theory: positive selection across cells, negative mechanism within cell, and a sharp bidder-count transition separating the two regimes.

Under overlap-cell ATT reweighting, the aggregate coefficient reverses to -0.097; under propensity-score trimming it becomes -0.307. Only 1.06% of treated items lack a within-cell counterfactual, so the reversal is not mainly a dropped-observation result. It is a weighting result. The broad price imprint and the overlap estimate put weight on different parts of the procurement distribution, and those parts carry different economic content: the broad spec averages selection-into-high-price-cells together with the cover-bidding theater that depresses observed prices, while the ATT spec isolates the latter.

Table 10 summarizes the decomposition.

The progression is monotone across the three specifications: broad-sample $+ + 0.064$ (item, year, PBU fixed effects; $p = 0.003$); overlap-cell unweighted $+ + 0.044$ ($p = 0.035$); overlap-cell ATT -0.097 (within-cell ATT weighting, $p < 10^{-9}$). The sign change occurs between the unweighted overlap restriction and ATT reweighting, not at the sample restriction itself. Within the overlap-ATT specification the negative coefficient is not driven by a single heterogeneous subset: it survives across both procurement modalities

Table 10: Selection + Mechanism Decomposition of the FL-Price Sign Reversal

Component	Estimate	SE / p	Interpretation
Broad-sample baseline	$++0.064$	$p = 0.003$	Joint effect: selection + mechanism
Across-cell selection (non-treated price on cell FL-share)			
Raw OLS	+22.65	SE 0.03	Cells where cartels concentrate are structurally high-priced
+ item-group + year FE	+24.04	SE 0.56	Selection beyond marginal cell dimensions
+ all five marginal cell-dim FE	+3.55	SE 0.23	Selection coefficient after partialling out each cell dim
Within-cell mechanism (winner-to-reference ratio)			
Within cell, no controls	−0.048	SE 0.004, $p < 10^{-30}$	Cover bidders compress winner-to-reference ratio
+ log(n_firms) control	+0.008	SE 0.004, n.s.	Effect operates through bidder-count channel
M1: log-bidder inflation	+0.507	SE 0.006, $p < 10^{-30}$	≈ 66% more bidders in FL-present tenders
Sign-reversal boundary by bidder count			
1–3 bidders (sparse)	+0.092	—	Selection regime: FL = higher price
7–10 bidders	+0.052	—	Transition
11+ bidders (dense)	−0.015	—	Mechanism regime: FL = lower price

Notes: Selection component: log negotiated price among non-treated items only, regressed on cell FL-share (across-cell variation). Mechanism component: log winner-to- reference ratio regressed on FL presence with overlap-cell fixed effects (within-cell variation). All standard errors clustered by overlap cell. The bidder-count rows report Δ(FL-present − FL-absent) on winner-to- reference within each bidder-count bucket. See Appendix C and AN-039, AN-040 in the supplementary site for the underlying regressions.

(Convite -0.099, $p < 10^{-5}$; Pregão -0.098, $p < 10^{-13}$), every PBU-size quartile, and tender-value quartiles Q1 through Q3. The only systematic positive cell within overlap-ATT is tender-value Q4, the segment already discussed. The subgroup decomposition therefore strengthens the scope reading rather than threatening it: the negative reading appears in routine large-buyer environments, while the high-value tail is where price evidence stops aligning with the loser-side scope.

The segment decomposition makes the boundary sharper. In the lowest three tender-value quartiles, the overlap estimates are negative. In the highest tender-value quartile, the estimate remains positive, $+0.041$ ($p = 0.045$). That is the segment where a cartel interpretation is most plausible ex ante because fixed coordination costs are spread over larger contracts. Even there, the result is scope evidence, not a markup estimate. It tells the reader where the price imprint is economically concentrated; it does not establish price injury.

Table 11: Where Price Evidence Stops

Exercise	Estimate	p / SE	Interpretation
Broad sample, item+year+PBU FE	$+0.064$	$p = 0.003$	Positive baseline before scope discipline
Overlap-cell unweighted	$+0.044$	$p = 0.035$	Attenuates under common-support restriction
Overlap-cell ATT	-0.097	$p < 0.001$	**Sign reversal under within-cell ATT weighting**
PS-trimmed ATT	-0.307	0.020	Stronger negative estimate after trimming
Tender-value Q4, overlap ATT	$+0.041$	$p = 0.045$	Positive only in high-value segment
Direct-CADE overlap items	-0.061	$p = 0.45$	No separate price association for legal anchors

Notes: Dependent variable is log negotiated unit price. The overlap rows restrict or reweight comparisons toward cells with treated and untreated support. Estimates are used to discipline the scope of price interpretation, not to estimate causal effects, overcharges, markups, or damages.

The direct-CADE overlap estimate is also null, -0.061 ($p = 0.45$), which reinforces the same interpretation. Price regressions do not recover the legal object of the case, and they do not supply a damages base. They show that the award-layer screen points to economically non-neutral environments in parts of the distribution, while the overlap reversal prevents price evidence from bearing more weight than it can support.

7.3. Implications for Legal-Economic Use

The legal-economic use of the screen follows from these boundaries. Price evidence may be consistent with economic relevance, but it does not transform an award-layer statistic into proof of agreement or price injury. The institutional use is staged: award records rank where costly bid-level and documentary evidence should be assembled; bid records, case documents, and legal process determine whether the conduct supports an antitrust case.

This distinction also clarifies portability. The architecture is more portable than the estimates: BEC estimates should not be exported as numerical external validity. What travels is the information problem. Many procurement systems observe participation and awards more cheaply than they can recover, clean, and evaluate bid histories. In such settings, a useful screen requires stable participant identities, observable winners and losers, repeated procurement, sufficiently comparable items, and a feasible path from the initial ranking to richer forensic follow-up. Without that path to proof, a cheap signal has little enforcement value. Another agency should not import the BEC cutoff or performance metrics mechanically; it should validate the ranking against its own enforcement anchors and observability constraints.

7.4. Adaptive Deployment

The screen should also be deployed as a moving queue rather than as a static legal trigger. Any administrative rule can generate strategic response once firms learn which observable patterns attract attention. In this setting, firms could rotate loser roles, engineer occasional wins, fragment identities, or avoid a known participation threshold. That possibility does not undermine the sequencing architecture, but it does rule out naive use of the binary cutoff as a permanent enforcement boundary. Strategic response weakens static cutoff use more than it invalidates the two-stage architecture. It is a standard concern for cartel enforcement and screening because firms may adapt to observable enforcement rules once those rules are understood (Green and Porter, 1984; Haltiwanger and Harrington, 1991). Screening rules face the same adaptation problem in enforcement settings (Harrington, 2008; Sánchez Graells, 2019).

The implementation implication is straightforward. Agencies should use continuous ranks or capacity-constrained top-k queues, refresh thresholds, monitor bunching below operational cutoffs, and combine award-layer screening with bid-layer forensics whenever bid files can be recovered. A firm that moves to avoid one observable pattern may create another: discontinuous participation, suspicious identity changes, or shifts in bid-level behavior. The architecture is therefore strongest when the award-layer queue is periodically updated and linked to forensic follow-up rather than treated as a one-time blacklist.

This is also the right legal posture. The screen is not a rule of liability and should not be used as one. It is a way to order scarce investigative attention when proof is costly and the first observable record is legally thin. Used adaptively, it gives an agency a disciplined place to begin without pretending that the beginning is the end of the evidentiary process.

8. Conclusion

Cartel enforcement often begins before the agency has the evidence needed to prove a case. Procurement records then pose a problem prior to liability: how should scarce investigative capacity be allocated when the cheap administrative record is broad but legally thin, and the richer bid-level record is costly to recover and evaluate? This paper treats procurement-cartel screening as evidence allocation under costly proof.

The empirical design separates the award layer from the bid layer. The award layer observes participation and outcomes before bid histories are opened. The bid layer observes submitted offers and within-tender conduct; it remains the natural stage for proof-producing forensic evaluation. The contribution is not to collapse these layers, but to sequence them: administrative records rank where investigation should begin, and bid-level evidence evaluates conduct inside the prioritized set.

The validation evidence supports that division of labor. Persistent zero-win participation orders adjudication-anchored loser-side exposure, survives participation, opportunity-set exposure, timing, and leakage discipline, and does not behave like a generic direct-defendant detector. The cobidder profile gives the target economic content without turning adjacency into membership or proof: cartel-adjacent firms are more broadly deployed, closer to legal anchors, more product-concentrated, and distinct in bid-level behavior. That makes the target harder to reduce to high-volume losing; it does not prove cover bidding or liability.

The central result is a sequencing result. Award-layer and bid-layer signals are complements, not substitutes. Used as a gatekeeper, the award-layer ranking reduces the bid-microdata pool by 83% while recovering 131 of 193 adjudicated cobidders. Joint scoring is the full-observability upper bound; sequential gatekeeping is the operational architecture for agencies deciding where to spend costly bid recovery.

The limits are part of the design. The screen creates forensic priority, not proof. The validation target is adjudication-anchored exposure, not cartel membership. Price evidence is scope evidence, not damages. The architecture is portable only where award records, repeated participation, comparable procurement categories, and recoverable bid files coexist. A static cutoff can be gamed, so implementation should use refreshed thresholds, continuous ranks or top-k queues, bunching checks, and bid-layer follow-up.

Cheap suspicion should neither be ignored nor treated as proof. Award-layer triage gives an agency a disciplined way to order scarce investigative attention before the richer evidentiary record is assembled. The statistic indicates where legal attention should start, not where legal responsibility ends.

References

Abrantes-Metz, R. M., L. M. Froeb, J. F. Geweke, and C. T. Taylor (2006). A variance screen for collusion. *International Journal of Industrial Organization 24*(3), 467–486.

Asker, J. (2010). A study of the internal organization of a bidding cartel. *American Economic Review 100*(3), 724–762.

Athey, S., J. Levin, and E. Seira (2011). Comparing open and sealed bid auctions: Evidence from timber auctions. *Quarterly Journal of Economics 126*(1), 207–257.

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

Baker, J. B. (2003). The case for antitrust enforcement. *Journal of Economic Perspectives 17*(4), 27–50.

Baldwin, L. H., R. C. Marshall, and J.-F. Richard (1997). Bidder collusion at forest service timber sales. *Journal of Political Economy 105*(4), 657–699.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*(2), 169–217.

Bryant, P. G. and E. W. Eckard (1991). A price-fixing probability index. *Review of Economics and Statistics 73*(4), 749–752.

Chassang, S., K. Kawai, J. Nakabayashi, and J. Ortner (2022). Robust screens for noncompetitive bidding in procurement auctions. *Econometrica 90*(1), 315–346.

Cinelli, C. and C. Hazlett (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society: Series B 82*(1), 39–67.

Clark, R., J.-F. Houde, and J. Kastl (2021). The dynamics of bidder collusion: Evidence from a long-lived cartel. *Journal of Political Economy 129*(8), 2353–2391.

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

Connor, J. M. (2007). Price-fixing overcharges: Legal and economic evidence. In *Research in Law and Economics*, Volume 22, pp. 59–153. Elsevier.

Coviello, D. and S. Gagliarducci (2017). Tenure in office and public procurement. *American Economic Journal: Economic Policy 9*(3), 59–105.

Decarolis, F. (2014). Awarding price, contract performance, and bids screening: Evidence from procurement auctions. *American Economic Journal: Applied Economics 6*(1), 108–132.

Decarolis, F., R. Fisman, P. Pinotti, and S. Vannutelli (2025). Rules, discretion, and corruption in procurement: Evidence from italian government contracting. *Journal of Political Economy Microeconomics 3*(2), 213–254.

Decarolis, F., L. M. Giuffrida, E. Iossa, V. Mollisi, and G. Spagnolo (2020). Bureaucratic competence and procurement outcomes. *Journal of Law, Economics, and Organization 36*(3), 537–597.

DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics 44*(3), 837–845.

Green, E. J. and R. H. Porter (1984). Noncooperative collusion under imperfect price information. *Econometrica 52*(1), 87–100.

Haltiwanger, J. and J. E. Harrington (1991). The impact of cyclical demand movements on collusive behavior. *RAND Journal of Economics 22*(1), 89–106.

Harrington, J. E. (2008). Detecting cartels. In P. Buccirossi (Ed.), *Handbook of Antitrust Economics*, pp. 213–258. MIT Press.

Hovenkamp, H. (2005). *Federal Antitrust Policy: The Law of Competition and Its Practice*. Thomson West.

Huber, M. and D. Imhof (2019). Machine learning with screens for detecting bid-rigging cartels. *International Journal of Industrial Organization 65*, 277–301.

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

Imhof, D. (2019). Detecting bid-rigging cartels with descriptive statistics. *Journal of Competition Law & Economics 15*(4), 427–467.

Imhof, D., Y. Karagök, and S. Rutz (2018). Screening for bid rigging—does it work? *Journal of Competition Law & Economics 14*(2), 235–261.

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

Marshall, R. C. and L. M. Marx (2012). *The Economics of Collusion: Cartels and Bidding Rings*. Cambridge, MA: MIT Press.

OECD (2022). Data screening tools for competition investigations. Technical Report 284, OECD Publishing, Paris.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

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

Porter, R. H. and J. D. Zona (1993). Detection of bid rigging in procurement auctions. *Journal of Political Economy 101*(3), 518–538.

Porter, R. H. and J. D. Zona (1999). Ohio school milk markets: An analysis of bidding. *RAND*

Journal of Economics 30(2), 263–288.

Sánchez Graells, A. (2019). “Screening for Cartels” in Public Procurement: Cheating at Solitaire to Sell Fool’s Gold? *Journal of European Competition Law & Practice 10*(4), 199–211.

Stigler, G. J. (1970). The optimum enforcement of laws. *Journal of Political Economy 78*(3), 526–536.

Wallimann, H., D. Imhof, and M. Huber (2023). A machine learning approach for flagging incomplete bid-rigging cartels. *Computational Economics 62*(4), 1669–1720.

Wils, W. P. J. (2007). Leniency in antitrust enforcement: Theory and practice. *World Competition 30*(1), 25–64.

Wils, W. P. J. (2016). The use of leniency in EU cartel enforcement: An assessment after twenty years. *World Competition 39*(3), 327–388.