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

Appendix A. Guide to the Online Appendix

This appendix is organized around the evidentiary burden of the main text. The manuscript argues that award records can allocate forensic attention before bid-level proof is assembled. The appendix therefore supplies three kinds of support: the formal ranking logic, the construction and timing of the labels, and the audits that discipline alternative interpretations. Works cited in this appendix appear in the reference list of the main paper.

Table A.1: How the Appendix Supports the Main Claims

Main-text claim	Appendix support	Location
Persistent zero-win participation is a triage statistic	Reduced-form ranking result and legal-scope separation	Appendix B
CADE cobidders are a valid validation target	Exact CADE–BEC matching, direct-defendant anchors, and cobidder definitions	Appendix C
The ranking is not only participation volume, opportunity-set exposure, or leakage	Participation-stratified permutation, exposure-adjusted audit, leakage, direct-defendant, and temporal audits	Appendix D
Price evidence is scope evidence	Overlap reversal and high-value segment decomposition	Appendix E
Static thresholds require adaptive implementation	Adaptation simulations and implementation implications	Appendix F
Award-layer and bid-layer evidence should be sequenced	Gatekeeping algorithm and distinction from joint scoring	Appendix G

Notes: The appendix is not a second empirical paper. It documents the assumptions, construction choices, and validation exercises needed to interpret the main text as an evidence-allocation design.

Appendix B. Evidence-Allocation Framework and Legal Scope

The model used in the main text is intentionally spare. Its job is to justify an ordering rule for scarce forensic attention, not to model the full internal organization of a cartel. Issues such as side payments, discipline, and leniency incentives are treated in the cartel literature (Asker, 2010; Marshall and Marx, 2012). The narrower object here is the pre-proof enforcement problem: why a firm that repeatedly loses, while continuing to appear in procurement records, can be a useful target for bid-level follow-up.

Appendix B.1. Primitives

Consider a procurement environment with repeated tender-items. In a cartel-targeted tender, one firm is assigned the winning role and submits the designated bid b^*. A losing-role participant submits a bid $b_\ell > b^*$ and is not meant to win. The losing role has a cost $c_1 > 0$. Winning by mistake or by deviation creates an internal cartel sanction $\kappa > 0$, interpreted broadly as the loss of cartel discipline, future rents, or capacity-allocation rights. Let r denote the private gain from deviating into the winner role in a tender where the firm was assigned to lose.

Firms outside the winning role are of two reduced-form types. Type C firms are cartel-deployed losing-role participants. Type G firms are ordinary losing participants: they may be inefficient, capacity constrained, inexperienced, or simply unlucky. The type notation is a device for stating the maintained ranking condition; it is not an empirical classification rule. The empirical screen conditions on the observed event $W_i = 0$, where W_i is the firm's number of wins, and ranks firms by participation count T_i.

Lemma 1 (Zero-win losing role). *If the internal sanction from winning in a losing role exceeds the private deviation gain, $\kappa > r$, a losing-role firm's equilibrium bid lies above the designated winning bid: $b_\ell > b^*$. Hence $\Pr(W_i > 0 \mid C) = 0$ over tenders in which the firm is deployed in the losing role.*

Proof. If the losing-role participant submits $b_\ell > b^*$, it loses and receives the deployment payoff net of the role cost. If it deviates to a winning bid, it gains r but pays the sanction κ. The deviation is unprofitable when $\kappa > r$. The result fixes the win probability of the losing role, not the exact numerical bid. Any bid above the designated winner's bid is consistent with the reduced-form role. □

Lemma 1 explains why the zero-win conditioning set is a coherent place to begin. The conditioning set is intentionally broad: many genuine firms also never win. The evidentiary content comes from the second observable, participation intensity.

Appendix B.2. Participation as a Ranking Statistic

Within the zero-win set, let participation counts follow a Poisson approximation:

$$T_i \mid \tau_i, q \sim \text{Poisson}(\lambda_q \tau_i), \qquad q \in \{C, G\},$$

where τ_i is the length of time firm i is observed and λ_q is the type-specific participation rate. The screening assumption is

$$\lambda_C > \lambda_G \quad \text{on the conditioning set } W_i = 0.$$

This is the substantive assumption behind the frequent-loser ranking. Cartel-deployed losing-role participants, if present, are repeatedly deployed to provide losing participation; ordinary zero-win firms participate less often. The condition $\lambda_C > \lambda_G$ is *maintained*, not derived: it is the behavioral assumption under which repeated zero-win participation becomes a monotone ranking statistic for loser-side exposure. The empirical sections then ask whether the ranking implied by this assumption has validation content against adjudication-adjacent labels. The type notation is a device for stating the maintained ranking condition; it is not an empirical classification rule.

Proposition 2 (Monotone loser-side suspicion). *Under the Poisson approximation and $\lambda_C > \lambda_G$, the posterior probability $\Pr(C \mid T_i = t, W_i = 0, \tau_i)$ is non-decreasing in t. Consequently, T_i and $\log(1 + T_i)$ are rank-equivalent screening statistics for loser-side cartel-adjacency exposure within the zero-win stratum.*

Proof. By Bayes' rule, the posterior odds of type C against type G are proportional to the likelihood ratio

$$\frac{\Pr(T_i = t \mid C, \tau_i)}{\Pr(T_i = t \mid G, \tau_i)} = \exp[-(\lambda_C - \lambda_G)\tau_i] \left(\frac{\lambda_C}{\lambda_G}\right)^t.$$

Because $\lambda_C > \lambda_G$, the likelihood ratio is increasing in t. This is the monotone-likelihood-ratio property of the Poisson family (Karlin and Rubin, 1956).[1] The posterior increases with participation count, and the log transformation $\log(1 + T_i)$ preserves the ranking. □

The empirical binary rule used in the main text is an operational coarsening of Proposition 2. It flags firms above the median plus 1.5 times the interquartile range of zero-win participation. The threshold is not a structural parameter. It is an auditable way to turn a monotone ranking into an operational list. Agencies could instead use the continuous score, a capacity-constrained top-k list, or a threshold calibrated to their investigative budget. The posterior language belongs to the reduced-form screening model. The empirical implementation uses the statistic as a ranking device for forensic priority, not as an estimated probability of legal membership.

Appendix B.3. Evidentiary Scope

The proposition delivers a ranking result. Even if the model is correct, a high value of $\log(1 + T_i)$ raises the posterior probability of loser-side cartel-adjacency exposure only relative to other zero-win firms. It does not identify an agreement, assign cartel membership, prove that any specific bid was cover, or establish a damages base. Those claims require evidence from the richer record: bid values, communications, leniency material, repeated winner rotation, or other case-specific facts.

This boundary is the legal reason for the staged architecture in the main text. The award layer can allocate attention because it ranks firms on a statistic that is cheap, observable, and monotone under a transparent behavioral assumption. The bid layer and case record then do the evidentiary work that liability requires, because liability turns on conduct and agreement, not on suspicious participation alone.

The framework does not imply that all zero-win firms occupy a cartel losing role, that frequent-loser status is legal membership, or that any specific bid is a cover bid. It supplies only the ranking logic for triage.

Proposition 3 (Scope separation). *Under Lemma 1 and Proposition 2, the frequent-loser statistic is sufficient to order expected investigative value within the zero-win stratum. It is not sufficient for liability or damages.*

[1] Karlin, S. and H. Rubin (1956). The theory of decision procedures for distributions with monotone likelihood ratio. *Annals of Mathematical Statistics* 27(2), 272–299.

Proof. Triage requires an ordering of expected investigative value. Proposition 2 supplies that ordering within the zero-win stratum. Liability and damages require different predicates: proof of agreement, conduct in particular tenders, and an economic measure of price injury. The statistic contains participation and win information only. It omits the facts needed for those predicates. The same statistic can be useful for allocating forensic attention and insufficient for legal adjudication. □

Appendix C. Data Construction and Validation Labels

This appendix records the construction choices behind the award-layer screen and the adjudication-based labels. The aim is to make the empirical object reproducible and to keep each source in its proper evidentiary role. The screen is built from procurement records. The validation labels come from CADE adjudications and their BEC co-bidding environments. The design therefore tests whether a cheap award-layer statistic ranks firms near legally anchored cartel cases, while leaving legal membership to the case record.

Appendix C.1. Award-Record Universe

The BEC procurement universe covers 2009–2019 and contains 1,654,447 tender-item records before the final price-regression restriction. The main price sample contains 1,654,401 observations. For the screening exercise, the unit is the firm. A firm enters the outer screening universe if it appears as a participant in at least one BEC tender during the sample window. This gives 41,444 participating firms.

For each firm i, we construct two award-layer variables:

$$T_i = \text{number of BEC tender-items in which firm } i \text{ participates},$$

and

$$W_i = \text{number of BEC tender-items won by firm } i.$$

The always-loser stratum is the set $\{i : W_i = 0\}$ and contains 16,843 firms. The continuous screening score is $s_i = \log(1 + T_i)$ within that stratum. The discrete frequent-loser flag is

$$FL_i = \mathbf{1}\{W_i = 0,\ T_i \geq 14\}.$$

The threshold comes from the always-loser participation distribution: median plus 1.5 times the interquartile range, with statistical cutoff 13.5 and integer rule $T_i \geq 14$. This flags 2,735 firms, or 16.2% of the always-loser stratum.

Table C.1: Screen Construction

Object	Value	Construction role
BEC participants	41,444	Outer firm universe for award-record construction
Always-losers	16,843	Zero-win stratum in which the screen ranks firms
Participation threshold	14	Integer rule from median +1.5×IQR
Frequent losers	2,735	Discrete first-stage flag within always-losers
Continuous score	$\log(1 + T_i)$	Rank used when a threshold is not required

The threshold is not estimated from CADE labels. It is fixed from the participation distribution alone. That matters because the validation exercise then asks a genuine question: whether a pre-specified award-layer rule concentrates firms later observed near adjudicated cartel defendants.

Appendix C.2. Matching CADE Defendants to BEC Firms

CADE adjudications identify firms by legal name and tax identifier (CNPJ). The starting CADE portfolio contains 65 firm-defendants across 12 adjudicated procurement-cartel cases linked to BEC subject matter. We match CADE defendants to BEC participant records on the full 14-digit CNPJ string, zero-padded to preserve leading zeros that would otherwise be truncated by numeric parsing. The match is implemented as exact string equality on the firm-identifier field (*códigofornecedor*) common to the CADE and BEC extracts; no fuzzy or name-based matching is used, and no manual overrides enter the direct-defendant set. Of the 65 firm-defendants, 47 have at least one BEC tender-item record during the sample window and define the direct-defendant set used in validation. The remaining defendants either operate outside BEC or do not appear under matchable identifiers in the BEC participation extract; they are omitted from the direct-defendant label.

Appendix C.3. CADE Anchors and Cobidder Labels

CADE adjudications provide the legal anchors. A direct CADE defendant is a firm named in an adjudicated procurement-cartel case and matched to BEC participation records as described in Appendix C.2. The linked BEC portfolio contains 47 direct defendants. Direct defendants are not the main target of the frequent-loser screen. They are used to locate cartel-affected bidding environments.

The validation target is the always-loser cobidder. A firm is a cobidder if it belongs to the always-loser stratum and appears in at least one BEC tender-item with a direct CADE defendant. This definition yields 193 cobidders. The label is cartel-adjacency, not legal membership. It says the firm appears on the loser side of tenders that include legally adjudicated cartel anchors. It does not say the firm was a defendant, signed an agreement, or submitted a cover bid in a particular tender.

For the conservative validation benchmark, we restrict CADE cases to adjudications available before the end of 2020. This leaves 4 cases, 30 BEC-active direct defendants, and 210 always-loser cobidders. Among those cobidders, 108 are frequent losers. This benchmark sacrifices sample size to avoid using later case resolution to define the validation target.

Table C.2: Validation Labels

Label	Count	Interpretation
Direct CADE defendants in BEC	47	Legal anchors from adjudicated procurement-cartel cases
Always-loser cobidders	193	Zero-win firms appearing with direct defendants; main cartel-adjacency target
Conservative cases	4	Pre-2020 adjudication benchmark
Conservative direct defendants	30	BEC-active defendants in the conservative benchmark
Conservative cobidders	210	Always-loser cobidders in the conservative benchmark
Conservative frequent-loser cobidders	108	Frequent losers inside the conservative cobidder target

The direct-defendant label is deliberately kept separate from the cobidder label. The former is a legal membership anchor. The latter is an exposure label tailored to the loser-side ranking. That separation is central to the paper's interpretation: direct defendants identify legally anchored environments, while cobidders provide the validation target for a statistic designed to find persistent losing participation.

Appendix C.4. Timing

The temporal-holdout exercise computes participation scores from 2009–2016 and evaluates discrimination against labels realized in 2017–2019. This split follows the enforcement logic rather than a purely statistical convention. At the moment an agency uses the award layer, it does not know the completed case file. The relevant question is whether earlier participation helps rank firms whose cartel adjacency is later revealed through adjudicated anchors.

The timing split is conservative in one respect and imperfect in another. It is conservative because the score is computed before the target period. It is imperfect because CADE adjudications often occur after the conduct and can cover conduct spanning multiple years. For that reason, the temporal design is an out-of-time discrimination exercise, not causal prediction or a field deployment. It asks whether the screen ranks later-adjudicated cartel-adjacent losers before the later labels are used to construct the score.

Appendix C.5. What Is Excluded From the Screen

The award-layer screen does not use bid amounts, bid ranks, within-tender dispersion, reference prices, LANCES timestamps, CADE outcomes, or price-regression residuals. Those objects enter later as validation, forensic benchmarks, or descriptive economic corroboration. Keeping them outside the screen is essential. It ensures that the first-stage statistic can be computed in award-only environments and that the bid layer remains available as an independent forensic stage.

Appendix D. Validation Audits

The validation design has to address distinct mechanical channels. First, frequent participants have more chances to appear near any defendant, so raw co-bidding rates are not enough. Second, firms may operate in the same procurement environments as CADE anchors, giving them more opportunities to meet those anchors even under ordinary competition. Third, item-level labels can mechanically reuse the same participation history that constructs the score. The main text reports the headline results; this appendix records the audit logic and the interpretation each audit licenses.

Appendix D.1. Participation-Stratified Permutation

The conservative CADE benchmark fixes the direct-defendant list using cases adjudicated before the end of 2020. Within the always-loser stratum, we then compare the observed frequent-loser share among cobidders with a permutation distribution that preserves participation volume. Specifically, firms are grouped by participation-volume quartile defined within the always-loser stratum, cobidder status is randomly permuted within those quartiles, and the frequent-loser share among placebo cobidders is recomputed 1,000 times. The number of cobidders per quartile is preserved across permutations by construction; the within- quartile shuffle reassigns the cobidder marker over the same set of firms, leaving the per-quartile marginal unchanged. The p-value reports the share of permutations whose placebo frequent-loser share is at least as extreme as the observed share.

The observed share is 3.95%, compared with a permutation mean of 1.24%. The ratio is 3.2×, with $p < 0.001$, and the conservative ROC AUC is 0.748 (95% CI [0.713, 0.783], computed via the DeLong method). The permutation answers the participation-volume concern: the frequent-loser concentration near CADE defendants is larger than activity volume alone predicts. It does not by itself address opportunity-set exposure, which is the purpose of the next audit.

Appendix D.2. Exposure-Adjusted Audit

The participation permutation does not by itself address market opportunity. Firms can appear near CADE anchors because they buy or sell in the same procurement environments, not because their loser-side behavior is informative for triage. The exposure-adjusted audit addresses that threat at the firm level: it asks whether the frequent-loser ranking remains informative after comparing firms within more similar opportunity sets for meeting direct CADE defendants.

The opportunity set is defined from the award layer, using the procurement environments in which a firm could have appeared with a CADE anchor. The construction uses item, modality, year, and buyer/procuring-unit information where the data provide common support. Observations without the information needed to place a firm in a comparable opportunity set are not used for that comparison. The retained support is therefore the subset in which firms can be compared to others facing similar chances to encounter direct CADE defendants in routine procurement records. The audit is reported as a diagnostic construction rather than as a single new headline statistic.

The result is reported as an exposure-discipline check rather than as a new liability measure. Its interpretation is limited: it reduces the concern that the cobidder concentration is only a mechanical byproduct of participation in defendant markets. It does not prove agreement, identify cartel members, estimate damages, or establish that all market-opportunity confounding has been removed.

Table D.1: Validation Audits and Threats Addressed

Audit	Threat addressed	Unit	What is held fixed or withheld	Main result or location	Interpretation
Participation-stratified permutation	Participation volume alone creates cobidder concentration	Firm	Participation-volume quartile within always-losers	3.2× excess; AUC 0.748	Concentration exceeds volume-preserving placebo assignment
Exposure-adjusted audit	Market opportunity to meet CADE anchors drives adjacency	Firm	Comparable opportunity sets using item, modality, year, and buyer/procuring-unit support	Reported in text; see Appendix D.2	Disciplines, but does not eliminate, opportunity-set exposure
Leakage audit	Score construction reuses firms that define item-level labels	Item/firm	Cobidder firms withheld from score construction within folds	Out-of-fold AUC 0.891	Signal survives the most direct label-reuse channel
Temporal holdout	Retrospective score uses later information	Firm/year	Score formed before the test year	AUC 0.864; Table D.3	Ranking remains separated from random classification out of time
Direct-defendant scope check	Loser-side ranking is misread as generic cartel-membership detection	Firm/item	Target changed from cobidders to direct CADE defendants	Direct-defendant AUC 0.49	Near-random direct-defendant classification is expected under loser-side scope

Notes: The audits validate forensic priority, not cartel membership, liability, damages, or proof. "Held fixed" and "withheld" describe the audit discipline each exercise adds relative to the baseline adjudication benchmark.

Appendix D.3. Leakage Decomposition

The in-sample item-level AUC is intentionally not treated as the main validation number. It is 0.995, and part of that performance is mechanical: the same firms can help define both the participation score and the item-level cobidder label. The audit decomposes that headline into a structural component and a leakage component.

The surviving component is large: AUC remains in the 0.86–0.89 range under the holdout audits. The lost component, roughly 0.10–0.13 AUC, is disclosed as mechanical. That disclosure is important because the main claim is an evidence-allocation claim. A cheap record-layer statistic retains meaningful ranking power after the most obvious mechanical channels are removed.

Table D.2: Leakage Audit

Audit	What changes	AUC	95% CI
In-sample reference	Full-sample score predicts any-cobidder item label	0.995	[0.995, 0.995]
Direct-defendant scope check	Same score predicts direct-CADE item label	0.506	[0.505, 0.507]
Cobidder-firm holdout	Held-out cobidders are removed from score construction within folds	0.891	[0.887, 0.894]
Temporal holdout	2009–2016 score predicts 2017–2019 cobidder labels	0.864	[0.858, 0.870]
Temporal direct-defendant check	Same temporal score predicts direct-CADE labels	0.511	[0.510, 0.513]

Notes: The reference AUC is partly tautological. The firm-holdout and temporal rows isolate the component that survives when labels are withheld from score construction or placed out of time. Direct-defendant rows check whether the loser-side statistic accidentally behaves like a general membership detector. Direct-defendant AUCs differ slightly across firm-level, item-level, and temporal audit definitions, but remain near random in each specification.

Appendix D.4. Direct-Defendant Scope Check

The direct-defendant result is a falsification of overreach. If the statistic were a generic cartel-membership detector, it should classify direct CADE defendants as well as cobidders. It does not. The direct-defendant AUC is 0.49, close to random classification. The item-level direct-defendant rows in Table D.2 use different samples and labels, but they are likewise near random. That pattern is informative: the statistic behaves like a loser-side ranking, while direct defendants are legal anchors that often appear on the winner side of the arrangement.

Appendix D.5. Year-by-Year Holdout

The rolling-origin exercise trains the score on years preceding each test year and evaluates cobidder discrimination in the test year. Table D.3 reports the year-by-year decomposition behind the temporal-holdout figure in the main text.

The year-by-year pattern is not a claim that the environment becomes progressively easier to police. The samples and adjudication links change across years. The relevant point is that the score remains separated from random classification in each rolling-origin test year.

Table D.3: Temporal-Holdout AUC by Test Year

Test year	AUC	95% CI	N firms	N CADE
2014	0.819	[0.779, 0.860]	8,631	88
2015	0.817	[0.776, 0.858]	10,122	102
2016	0.851	[0.821, 0.881]	11,699	127
2017	0.862	[0.832, 0.891]	13,057	148
2018	0.897	[0.877, 0.918]	14,391	161
2019	0.922	[0.912, 0.933]	15,597	181

Notes: Each row computes the screening score using participation observed before the test year and evaluates AUC against the cobidder label in that test year. The exercise is out-of-time discrimination, not a field deployment.

Appendix E. Scope of Price Evidence

The main text treats price evidence as corroboration rather than causal identification. This appendix records the price decomposition that keeps that interpretation disciplined: the overlap sign reversal and the high-value-segment positive result. The strategic-adaptation simulation is reported separately in Appendix F.

Appendix E.1. Overlap and Segment Scope

The broad price association is positive, with the headline range $+3.6\% - +7.7\%$. Under overlap-cell ATT reweighting, the aggregate coefficient reverses to -0.097; under propensity-score trimming it becomes -0.307. This does not refute the screening interpretation. It shows that the broad estimate and the overlap estimate average different parts of the procurement distribution.

The segment decomposition is the useful object. The lowest three tender-value quartiles have negative coefficients (-0.065, -0.057, and -0.040 in the broad specification). The highest tender-value quartile is positive in both the broad and overlap designs: $+0.046$ and $+0.041$ ($p = 0.045$). Trimming heavily weighted overlap cells does not eliminate the negative aggregate estimate; it moves from -0.097 with no trim to -0.118 after dropping the top decile of cell weights and -0.133 after dropping the top half.

The legal-economic consequence is scope, not damages. The price results show that flagged environments are economically non-neutral, especially in the high-value segment where cartel rents are most plausible. They do not supply a causal price effect, a markup, or a damages calculation.

Appendix F. Adaptive Deployment Diagnostics

The award-layer screen is an operational rule, so its performance must be evaluated under plausible adaptive responses. This appendix records a diagnostic simulation that holds the empirical environment fixed and asks how the participation score behaves under simple evasive changes. The scenarios are intentionally simple. Their purpose is to show which adaptations weaken a static binary rule more than others and to inform the operational recommendations in the main text.

Appendix F.1. Simulation Scenarios

Five candidate adaptation strategies are tested against the baseline frequent-loser rule. Rotation of losing-role firms distributes losing participation across a larger pool of identities. Occasional wins introduce a small positive win-share so that targeted firms do not satisfy the strict $W_i = 0$ filter. CNPJ splitting fragments a single firm's participation across multiple legal identifiers. Threshold-aware capping holds participation just below the median-plus-1.5×IQR cutoff so that the firm avoids the flag. The final scenario combines capping with rotation.

Rotation leaves AUC essentially unchanged (0.931 versus baseline 0.928). Occasional wins also leave AUC nearly unchanged (0.929), because the strict zero-win rule removes those firms from the flagged set rather than mis-ranking them. CNPJ splitting lowers AUC to 0.891 by breaking the participation history that the persistence signal depends on. Threshold-aware capping is more damaging, lowering AUC to 0.801. Combining capping with rotation lowers AUC to 0.632.

Table F.1: Strategic-Adaptation Summary

Scenario	AUC	Change vs baseline
Baseline	0.928	—
Rotation	0.931	0.3 pp
Occasional wins	0.929	0.2 pp
CNPJ splitting	0.891	−3.7 pp
Threshold capping	0.801	−12.7 pp
Capping plus rotation	0.632	−29.5 pp

Notes: Diagnostic simulation holding the empirical environment fixed. The scenarios show vulnerabilities of a static binary cutoff; they are not a full equilibrium model of evasion or evidence that the rule is immune to manipulation.

Appendix F.2. Implementation Implications

The hierarchy is informative for implementation. The screen is not equally vulnerable to every adaptation strategy. Simple rotation and occasional wins are absorbed by the persistence signal; identity fragmentation begins to break the participation history; behavior targeted at the static cutoff is the most damaging. The implementation lessons are correspondingly specific. The rule should be refreshed periodically as new participation data arrive. Bunching just below the cutoff should be monitored, because threshold-aware capping leaves a visible pile-up. Continuous ranks or capacity-constrained top-k queues should be preferred where they are operationally feasible, since they are harder to game than a single published cutoff. And the award-layer screen should be paired with bid-layer forensics whenever bid files can be recovered, so that strategic adaptation against the cheap layer does not also succeed against the costly layer.

The simulation is a diagnostic rather than a full equilibrium model of evasion. Its implication is operational: strategic response exposes the vulnerability of a static binary rule more than it invalidates the two-stage architecture.

Appendix G. Forensic Sequencing Details

The sequential architecture in the main text separates two tasks that are often conflated. The award-layer stage ranks firms using fields that are available before bid recovery. The bid-layer stage then uses richer forensic features only for the firms that survive the first stage. This appendix states the sequence explicitly so that the empirical comparison is read as an evidence-allocation design rather than as a claim that thin records substitute for full bid files.

Appendix G.1. Stage 1: Award-Layer Ranking

Stage 1 uses the always-loser stratum and ranks firms by $\log(1 + T_i)$. An enforcement agency can translate the rank into a fixed threshold, a top-k list, or a budget-constrained request for further evidence. No bid values are needed at this stage. The output is a survivor pool, not a liability finding.

Appendix G.2. Stage 2: Bid-Layer Forensics

Stage 2 is run only inside the survivor pool. The bid-layer benchmark uses within-tender bid-distribution features: dispersion, skewness, kurtosis, spread, distance from the winning bid, and related moments. These features are closer to forensic evaluation because they describe how bids are placed inside tenders. They are also more costly because they require per-bid offer values and sufficient tender-level bid histories.

The comparison in the main text uses the bid-layer benchmark in two roles. First, it is an external benchmark for the award-layer screen: the cheap screen is informative because it reaches similar discrimination on a thinner data envelope. Second, it is the Stage 2 input in the sequential rule: the screen does not replace bid-distribution analysis; it decides where that analysis should begin.

Appendix G.3. Why Joint Scoring Is Not the Operational Baseline

A single joint model that uses both award-layer and bid-layer features for every firm is a useful upper-bound comparison, but it is not the relevant operational baseline for many agencies. Running the joint model requires bid microdata for the full evaluation pool before any triage has occurred. The sequential rule asks a different question: how much of the joint model's target capture can be preserved while recovering bid microdata for only a smaller first-stage pool?

That distinction is the enforcement contribution. Joint scoring is a prediction exercise under full observability. Sequencing is an evidence-allocation rule under partial observability. The second problem is the one agencies face when award records are routine but bid files require costly recovery.

Appendix G.4. Algorithm

The full gatekeeping procedure can be stated compactly.

Algorithm G.1: Award-Layer Gatekeeping

1. Define the firm universe from award records (participant identity, winner identity, item code, negotiated price, procuring unit, year, modality).
2. For each firm i, compute the win count W_i and the participation count T_i over the calibration window.
3. Restrict to the zero-win stratum $\{i : W_i = 0\}$.
4. Compute the award-layer score $s_i = \log(1 + T_i)$.
5. Select the survivor pool $\mathcal{S}$ using one of: a median-plus-1.5×IQR threshold, a top-K_1 rule, or a capacity-constrained budget rule.

6. Recover bid microdata only for firms in $\mathcal{S}$.
7. Compute the bid-layer forensic features inside $\mathcal{S}$.
8. Rank survivor firms using the bid-layer benchmark alone or a combined award-plus-bid rule, depending on the investigative protocol.
9. Deliver the top-k queue from that ranking for investigative review.
10. Treat the output as forensic priority, not as a liability finding.

Joint scoring requires bid microdata for the full pool before triage; sequential gatekeeping requires bid microdata only after the award-layer survivor pool is formed. The size of the survivor pool, the choice of threshold or top-K_1, and the recalibration cadence are all institutional choices that the algorithm leaves to the agency.