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

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

[1]*INSPER, São Paulo, Brazil*

Abstract

This online appendix accompanies *Cheap Signals, Costly Proof: Award-Layer Triage for Cartel Enforcement* and contains: (A) the cover-bidding framework, assumptions A1–A6, and proofs of Lemma 1 and Propositions 2–4 invoked in the main text; (B) core supporting tables and figures; (C) identification audits (overlap-restricted estimates, sign-reversal decomposition, Oster bounds, and the anti-leakage audit); (D) threshold and specification robustness tables; (E) operational metrics and detector comparison; and (F) the staggered-design estimators that return null at the available bandwidths. References to sections, equations, and tables in the main text use the same labels as the published manuscript.

Appendix A. Cover-Bidding Framework and Proofs

This appendix formalizes the reduced-form separating-equilibrium framework the paper reads against. We deliberately keep the framework spare: the empirical contribution is screening under incomplete observability, and the role of the framework is to identify the participation primitive a cartel-deployed cover bidder generates in award-record data and to discipline the interpretation of the empirical objects, not to provide a fully-specified mechanism-design treatment of cartel formation.[1]

*Corresponding author.

Email address: darciogm1@insper.edu.br (Darcio Genicolo-Martins)

[1]The cartel's allocation rule, side-payment scheme, and within-cartel sanctions are micro-founded in Marshall and Marx (2012) and Asker (2010); we take their characterization as given and parameterize the cartel's reduced-form deployment problem.

Primitives.. A cartel of n firms is active in a procurement market k with $|k|$ auctions per period. The cartel allocates the designated winner to one cartel member submitting bid b^* in each cartel-targeted auction and chooses the per-period deployment count $m \geq 0$, defined as the number of auctions per period in which the cartel deploys at least one cover bidder.[2] A cover bidder in a cartel-targeted auction submits a bid $b_\ell > b^*$. Per-deployment cost is $c_1 > 0$. Market-level detection probability is $\theta_k \in (0, 1)$, declining in oversight capacity. Let $R(m, \theta_k)$ denote gross expected cartel rents from deploying m auctions in market k per period; rents are net of the cartel's distributive sharing rule (whose details are not material to the empirical claims). The cartel's expected net surplus per period is $\pi(m, \theta_k, c_1) = R(m, \theta_k) - (c_1 + \theta_k \phi_0)\, m$, where $\phi_0 > 0$ is the cartel-wide marginal detection penalty per cover-deployment; we treat detection penalty as cartel-level (not pegged to individual cover-bidder identity), consistent with oversight that detects the *pattern* of cover bidding rather than individual outcomes.

Assumptions.. We impose:

A1 (Rent positivity at the margin) $\partial R/\partial m|_{m=0} > c_1 + \theta_k \phi_0$ at the observed θ_k, ensuring the cartel optimally deploys at least one cover bidder.

A2 (Diminishing rents) $\partial^2 R/\partial m^2 < 0$, which by linearity of the cost terms in m yields $\partial^2 \pi/\partial m^2 < 0$.

A3 (Cartel-allocation incentive compatibility) The within-cartel allocation rule is incentive-compatible in the bidding subgame: for any non-designated cartel firm, rents $< \kappa$, where rents is the per-auction profit from deviating to a winning bid and $\kappa > 0$ is the internal cartel sanction or capacity-revelation cost triggered by the deviation. We treat A3 as a primitive equilibrium condition the cartel's optimal contract design is required to satisfy; Marshall and Marx (2012) and Asker (2010) derive sufficient conditions on side-payment and trigger-strategy schemes.

[2]For tractability we treat each cartel-targeted auction as covered by a single cover bidder; the framework is unchanged if multiple cover bidders are deployed per auction provided the per-deployment cost c_1 is re-interpreted accordingly.

A4 (Type partition) Each non-cartel firm i is of type C (cover bidder, deployed by the cartel in some fraction of auctions) or type G (genuine entrant). Hybrid or intermediate types are absorbed into the empirical type-G class.[3]

A5 (Stationary deployment within market) The cartel's per-period deployment m is constant within market k over the analysis window; deployment events are conditionally independent across auctions given m and $|k|$.

A6 (Selection on the wins-zero subset) Conditional on $\{\text{wins}_i = 0\}$, type-G firms are permanent competitive losers (low-intensity participants, in the market at low rate) while type-C firms are cartel-deployed participants. This selection effect implies $\lambda_G < \lambda_C$ on the wins-zero subset, where λ_C, λ_G are per-period auction-participation rates of type C and type G firms respectively. A6 is the model-completing assumption that disciplines the type-discrimination claim of Proposition 2; without A6 the wins-zero subset is uninformative about type.

Empirical observables.. Award-record data report a participation count $T_i = \text{tenders\_count}_i$, a win indicator $\text{wins}_i \in \{0, 1\}$, and a tenure τ_i (years active) for each firm i.

Roadmap.. Lemma 1 identifies the wins $= 0$ filter as the equilibrium choice of type C. Proposition 2 establishes $\log(1 + T_i)$ as a Bayesian-monotone ranking statistic for type C on the conditioning set $\{\text{wins} = 0\}$. Proposition 3 characterizes the cartel's optimal deployment m^* and its comparative statics. Proposition 4 discusses why the broad-sample association β and the overlap-restricted β^{ov} are different empirical objects under the framework's deployment problem.

Lemma 1 (Wins-zero separating equilibrium). *Consider the bidding subgame in which a cartel allocates the designated winner to one cartel member submitting bid b^* and deploys m cover bidders in a cartel-targeted auction, each cover bidder ℓ choosing a bid b_ℓ. Under A3, every separating equilibrium of this subgame has $b_\ell > b^*$ for every*

[3]The empirical implementation does not require the binary partition to be the unique data-generating process; hybrid types whose deployment intensity is intermediate are absorbed by the rank-based screening statistic of Proposition 2.

cover bidder ℓ, so $\Pr(win \mid C) = 0$*: the bright-line filter wins* $= 0$ *identifies the type-C subset of bidders. Genuine entrants of type G face no κ-sanction constraint and bid $b_G < b^*$ whenever expected payoff from winning is positive. The cover-bidder strategy correspondence is $\{b : b > b^*\}$; the lemma pins down the equilibrium win probability and the wins-zero filter, not a unique numerical bid.*

Proof of Lemma 1. At the bidding stage, the cartel's deployment m and market detection probability θ_k are taken as given; the cartel-wide detection penalty $\theta_k \phi_0 m$ is therefore sunk from the perspective of any individual cover bidder choosing b_ℓ. Write the cover bidder's bidding-stage payoff conditional on choosing b as $u_C(b) = -c_1 - \tilde{\phi} - \kappa \cdot \Pr(\text{win} \mid b)$, where c_1 is the per-deployment cost, the sunk term $\tilde{\phi} = \theta_k \phi_0$ is the cover bidder's share of the cartel-wide detection cost (independent of own bid), and κ is the personal cartel-internal sanction triggered *only* by an off-equilibrium winning deviation. Compare the equilibrium-candidate strategy $b > b^*$ against the single deviation $b' < b^*$ that becomes the winning bid (i.e., b' undercuts both b^* and any genuine-entrant bid). Equilibrium-candidate payoff: $\Pr(\text{win} \mid b) = 0$ and $u_C(b) = -c_1 - \tilde{\phi}$. Deviation payoff: $\Pr(\text{win} \mid b') = 1$ and $u_C(b') = -c_1 - \tilde{\phi} - \kappa +$ rents. Under A3, rents $< \kappa$, so $u_C(b') < u_C(b)$ and no profitable deviation exists. Therefore in any separating equilibrium, every cover bidder chooses some $b > b^*$. By the auction's lowest-bid rule and the assumption that b^* is the designated winning bid, $\Pr(\text{win} \mid C) = 0$ in every such equilibrium. Type G faces no κ penalty and prefers $b_G < b^*$ whenever the post-cost rent from winning is positive. The argument pins down the equilibrium win probability and the wins-zero filter; it does not pin down a unique cover bidder bid, since any $b > b^*$ delivers the same payoff $-c_1 - \tilde{\phi}$. The bid correspondence is $\{b : b > b^*\}$, and the empirical filter wins $= 0$ identifies the type-C subset across this entire correspondence. □ □

Proposition 2 (Linearized cumulative exposure). *Define the per-period auction-participation rate of a type-C firm in market k as $\lambda_C(\theta_k, c_1) \equiv m^*/|k|$, where m^* is the cartel's optimal per-period deployment count (Proposition 3) and $|k|$ is the number of auctions per period in market k; λ_C is therefore the fraction of auctions per period in which a typical cover bidder participates. Under A5 (stationary deployment with conditional independence) and the Poisson approximation $T_i \sim \text{Poisson}(\lambda \tau_i)$, the expected participation count over tenure τ_i years is $\mathbb{E}[T_i \mid type = C, \tau_i] = \lambda_C \tau_i$ for type C and $\mathbb{E}[T_i \mid type = G, \tau_i] = \lambda_G \tau_i$ for type G, with $\lambda_G < \lambda_C$ on the conditioning subset $\{wins_i = 0\}$ by A6. The log-statistic $\log(1 + T_i)$ is a monotone, rank-equivalent transformation of*

T_i. Under the monotone-likelihood-ratio property of the Poisson family with $\lambda_C > \lambda_G$, the posterior $\Pr(type = C \mid T_i = t, wins_i = 0)$ is non-decreasing in t; equivalently, $\log(1 + T_i)$ is a Bayesian-monotone ranking statistic for type C on the conditioning set $\{wins = 0\}$. Combined with Lemma 1, the joint statistic $(\log(1 + T_i), \mathbf{1}[wins_i = 0])$ is the sufficient pair of award-record observables for type-C discrimination under A4–A6; the binary frequent-loser rule $\mathbf{1}[T_i > k]$ is one information-coarsening of this posterior, with the empirical choice $k = \text{med} + 1.5 \times \text{IQR}$ a convention motivated by the empirical participation distribution rather than by the framework.

Proof of Proposition 2. The expectation result $\mathbb{E}[T_i \mid \text{type}, \tau_i] = \lambda\tau_i$ is the mean of a Poisson process with rate λ over horizon τ_i (A5). The monotonicity claim follows from Bayes' rule on the wins-zero subset: $\Pr(C \mid t, \text{wins} = 0) = \Pr(t \mid C, \text{wins} = 0)\,\pi'_C / [\Pr(t \mid C, \text{wins} = 0)\,\pi'_C + \Pr(t \mid G, \text{wins} = 0)\,(1 - \pi'_C)]$, where π'_C is the prior conditional on wins=0. This posterior is non-decreasing in t if and only if the likelihood ratio $\Pr(t \mid C, \text{wins} = 0) / \Pr(t \mid G, \text{wins} = 0)$ is non-decreasing in t. For Poisson likelihoods with rates $\lambda_C > \lambda_G$ (the inequality holds on $\{\text{wins} = 0\}$ by A6), the monotone-likelihood-ratio property of the Poisson family is standard (Karlin and Rubin, 1956). Hence the posterior is non-decreasing in T_i, and the log transform $\log(1 + T_i)$ inherits the ranking. Coarsening to a binary indicator $\mathbf{1}[T_i > k]$ preserves rank-monotonicity above and below k but is no longer Bayesian-sufficient; this information-loss is the price of operational simplicity. □ □

Proposition 3 (Optimal cover-bidder deployment). *Under A1–A2 and the net-surplus definition $\pi(m, \theta_k, c_1) = R(m, \theta_k) - (c_1 + \theta_k\phi_0)\,m$, the cartel's optimal cover-bidder deployment rate m^* satisfies the first-order condition $\partial R/\partial m = c_1 + \theta_k\phi_0$. By A2 the solution is unique and interior. Implicit differentiation of the FOC yields the comparative statics $\partial m^*/\partial\theta_k = \phi_0/(\partial^2 R/\partial m^2) < 0$ and $\partial m^*/\partial c_1 = 1/(\partial^2 R/\partial m^2) < 0$; both ratios are negative because A2 ensures the denominator $\partial^2 R/\partial m^2 = \partial^2\pi/\partial m^2$ is negative. Cover-bidder deployment is more aggressive where the principal cost of detection is lower (§7.3).*

Proof of Proposition 3. Differentiating $\pi(m, \theta_k, c_1) = R(m, \theta_k) - (c_1 + \theta_k\phi_0)m$ with respect to m and setting $\partial\pi/\partial m = 0$ yields $\partial R/\partial m = c_1 + \theta_k\phi_0$. A1 ensures the optimum is interior ($m^* > 0$); A2 ensures the solution is unique. Define $F(m, \theta_k, c_1) = \partial R/\partial m - c_1 - \theta_k\phi_0$; the FOC is $F = 0$. By the implicit function theorem, $\partial m^*/\partial\theta_k =$

$-F_{\theta_k}/F_m = \phi_0/(\partial^2 R/\partial m^2)$ and $\partial m^*/\partial c_1 = -F_{c_1}/F_m = 1/(\partial^2 R/\partial m^2)$. Both are negative because A2 gives $\partial^2 R/\partial m^2 < 0$. □ □

Propositions 2–3 together imply that frequent-loser participation is not distributed uniformly over items: under A1–A2, the cartel deploys cover bidders preferentially in items where its expected surplus is highest, conditional on the deployment cost c_1 and the local detection probability θ_k. Items the cartel selects into therefore differ from items it does not on a vector of unobserved characteristics $\mathbf{u}$ that governs the cartel's expected surplus—chiefly, the rents available from the designated winner's bid relative to the competitive equilibrium price. The next proposition formalizes the implication for the two empirical objects β and β^{ov}.

Proposition 4 (Screening-value vs. treatment-effect objects). *Let $D_i \in \{0,1\}$ indicate frequent-loser presence in item i, $\mathbf{x}_i$ the vector of observables absorbed by fixed effects and matching covariates, and $\mathbf{u}_i$ the vector of unobservables governing the cartel's deployment optimum. Suppose the cartel's optimal deployment satisfies $\Pr(D_i = 1 \mid \mathbf{x}_i, \mathbf{u}_i) = g(\mathbf{x}_i, \mathbf{u}_i)$ with g strictly increasing in the rent component of $\mathbf{u}_i$. Then (1) the broad-sample $\beta = \mathbb{E}[y_i \mid D_i = 1, \mathbf{x}_i] - \mathbb{E}[y_i \mid D_i = 0, \mathbf{x}_i]$ integrates over conditional distributions of $\mathbf{u}_i$ that differ systematically in the rent component (Bayes' rule on the deployment density $g(\mathbf{x}, \mathbf{u})f(\mathbf{u} \mid \mathbf{x})$); (2) within the common-support cell Ω, defined by propensity scores $\Pr(D_i = 1 \mid \mathbf{x}_i)$ bounded away from 0 and 1, the conditional support of $\mathbf{u}_i$ overlaps strictly across $D_i = 1$ and $D_i = 0$ at each $\mathbf{x}_i \in \Omega$, so within-cell comparisons between treated and untreated items are well-defined; the proposition makes no claim that the conditional distributions of $\mathbf{u}_i$ within Ω are equal or that the conditional means are bounded, only that their supports overlap and that ATT or PS-trimmed reweighting is mathematically applicable; and (3) under the auxiliary premise that the rent component shifts the realized log price in the same direction it shifts the deployment probability, the gap $\beta - \beta^{ov}$ inherits the sign of the rent co-movement, and an observed sign reversal $\beta > 0$, $\beta^{ov} < 0$ is consistent with the deployment problem rather than refuting it. The proposition does not derive the sign reversal; it identifies the rent-component-co-movement premise as a sufficient condition under which the observed sign reversal is rationalizable within the framework. The proposition is therefore better read as a structural remark than as an identification result; we retain the proposition form for cross-reference convenience.*

Proof of Proposition 4. Claim 1 follows from the law of iterated expectations: the broad-sample β averages $\mathbb{E}[y_i \mid D_i, \mathbf{x}_i, \mathbf{u}_i]$ over the conditional distributions of $\mathbf{u}_i$ given D_i and $\mathbf{x}_i$, and Proposition 3 combined with Bayes' rule applied to $f(\mathbf{u} \mid D = 1, \mathbf{x}) \propto g(\mathbf{x}, \mathbf{u})f(\mathbf{u} \mid \mathbf{x})$ implies $\mathbf{u}_i \mid D_i = 1$ first-order stochastically dominates $\mathbf{u}_i \mid D_i = 0$ in the rent component

at the same $\mathbf{x}_i$. Claim 2 follows from the definition of common support: cells in Ω have propensity score $\Pr(D_i = 1 \mid \mathbf{x}_i) \in (a, b)$ for some $0 < a < b < 1$. By Bayes' rule, the conditional density $f(\mathbf{u} \mid D_i = d, \mathbf{x}_i)$ is positive at any $\mathbf{u}$ where $f(\mathbf{u} \mid \mathbf{x}_i)$ is positive and $g(\mathbf{x}_i, \mathbf{u}) \in (0, 1)$; the latter holds generically when g is strictly monotone and the propensity score is interior. Therefore the conditional supports of $\mathbf{u}$ across $D = 1$ and $D = 0$ within Ω overlap strictly. The proposition asserts only this support-overlap condition, not that the conditional distributions of $\mathbf{u}$ within Ω are equal or that conditional means are bounded; common support is a necessary condition for ATT or PS-trimmed estimators to be mathematically defined, not for the unobservable distributions across treatment status to match. Claim 3 is the comparative statement: under the rent-component co-movement premise, the broad-sample average inherits a positive contribution from the high-rent cells, while β^{ov} averages over a sub-population in which the high-rent contribution is mechanically attenuated. The gap $\beta - \beta^{\text{ov}}$ therefore inherits the sign of the rent co-movement under the premise; an observed sign reversal is consistent with this gap exceeding $|\beta^{\text{ov}}|$, but the proposition does not derive sign-reversal as a unique prediction. If the premise fails, the proposition is silent and the observed sign reversal must be explained by some channel other than the deployment problem. □ □

The proposition is the formal counterpart to the screening-value reading in §7.2: under the rent-component co-movement premise, the observed sign reversal is consistent with the deployment problem rather than refuting it. The proposition does not assert that the deployment problem *predicts* the sign reversal; it identifies a sufficient condition (the rent-component co-movement) under which the sign reversal is rationalizable within the framework. The model does not by itself adjudicate whether the premise holds; the empirical evidence of §6, §8.1, and the segment-level decomposition of §7.2 together discipline the interpretation.

The framework is silent on the bid-level shape of b_C beyond the equilibrium constraint $b_C > b^*$. The early empirical cover-bidding literature (Porter and Zona, 1993, 1999; Bajari and Ye, 2003) reads the constraint loosely: cover bidders submit bids *visibly* above the winning bid because the cartel's purpose is to manufacture the appearance

of competition, not to compete credibly (the *textbook* reading, R1). A more recent strand reads it tighter: cover bids must remain plausible enough that the procurement officer does not flag the bid distribution as anomalous, and the cover bidder cycles through cover roles across tenders generating cross-bid dispersion within firm (the *credible-cover-bidding* reading, R2; Marshall and Marx 2012; Asker 2010). The framework above is consistent with either reading. The within-stratum bridge of §6.2 discriminates between them: cobidders bid systematically closer to the eventual winner than FL non-cobidders ($d = -0.281$, $p < 10^{-6}$) with elevated within-firm dispersion ($d = +0.147$); a multivariate logit holding $\log(1 + \text{tenders\_count})$ constant confirms both signs at $p < 10^{-3}$. The directions are opposite to R1 and consistent with R2; we adopt R2 as the operational reading. The framework's predictions about firm-level participation intensity (Proposition 2), deployment proximity to direct CADE defendants, and portfolio specialization are unaffected by the R1/R2 distinction; only the bid-level reading of b_C is.

Appendix B. Core Supporting Tables and Figures

The sample-construction tables (modal mix, conditional descriptive statistics) and the heterogeneity tables (detection-regime quartiles, modal contrast, McCrary density, robustness to exclusion of CADE-involved items) referenced in the body are collected here, together with the always-loser participation distribution that motivates the median + 1.5×IQR cutoff (Figures B.1–B.2) and the detection-regime gradient (Figure B.3).

Appendix C. Identification Audits

Table C.1 reports the Cinelli and Hazlett (2020) robustness value and Oster (2019) bounds for the broad-sample association; the leakage decomposition and the permutation null behind the conservative benchmark are reported in Tables 8 and 7 of §6 (kept there as the discrimination centerpiece rather than as appendix audits). The leave-one-out instrumental-variable placebo (Table C.2) confirms the participation primitive operates through the threshold-crossing population the construct identifies; random subsamples

Table B.1: Tender Counts by Procurement Modality and Year

Year	Pregão	Convite	Total
2009	137,424	232,765	370,189
2010	134,181	253,196	387,377
2011	72,894	140,305	213,199
2012	129,664	305,808	435,472
2013	125,324	328,616	453,940
2014	118,274	310,947	429,221
2015	92,324	226,503	318,827
2016	57,804	122,841	180,645
2017	120,194	261,205	381,399
2018	60,518	121,698	182,216
2019	11,857	26,976	38,833

Notes: Winner-level observations from BEC (São Paulo) 2009–2019. Pregão = electronic reverse auction; Convite = sealed-bid invitation.

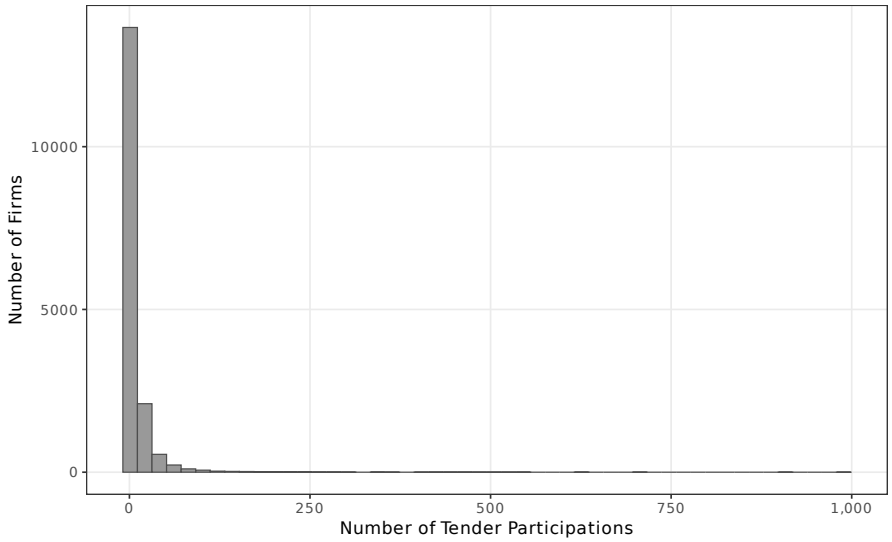

Figure B.1: Distribution of tender-item participation counts across always-loser firms in BEC, 2009–2019.

Table B.2: Conditional Descriptive Statistics: Residualized Prices by FL Status

	FL-Present (losers = 1)	FL-Absent (losers = 0)
Panel A: Unconditional		
Mean log(price)	4.9241	2.4938
Difference	2.4303	
Panel B: Conditional on Item + Year + PBU FE		
Mean residualized price	0.0472	-0.0024
SD residualized price	1.5734	0.8704
Difference (gap)	0.0495	
Overlap coefficient	0.978	
t-statistic	8.81	
p-value	0.000000	
Cohen's d	0.0390	
Observations	79,452	1,574,949

Notes: Residuals from regressing log(negotiated price) on item, year, and purchasing unit (PBU) fixed effects. The gap represents the mean difference in residualized prices between FL-present and FL-absent tenders, after absorbing common item-level, temporal, and PBU-level variation. Overlap coefficient = $2\Phi(-|\Delta\mu|/\sqrt{\sigma_1^2 + \sigma_0^2})$ measures the proportion of the two residual distributions that overlap; values near 1 indicate near-identical distributions. Cohen's d measures the standardized effect size.

Table B.3: Frequent-Loser Price Coefficient by Procuring-Unit Size and Procedure

Sub-sample	Coefficient	SE	N
By procuring-unit size quartile:			
Q1 (smallest buyers)	0.2138	(0.1670)	7,503
Q2	0.0111	(0.0513)	87,571
Q3	0.0660**	(0.0304)	302,258
Q4 (largest buyers)	0.0165	(0.0160)	1,257,069
By procedure type:			
Pregão (reverse electronic auction)	0.0933***	(0.0255)	546,549
Convite (sealed-bid invitation)	0.0382**	(0.0186)	1,107,852

Notes: Quartile gradient consistent with variation in the screening signal across detection regimes (§7.3); we do not identify a causal channel. Item and year FE (PBU FE for procedure split). SE clustered at item level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

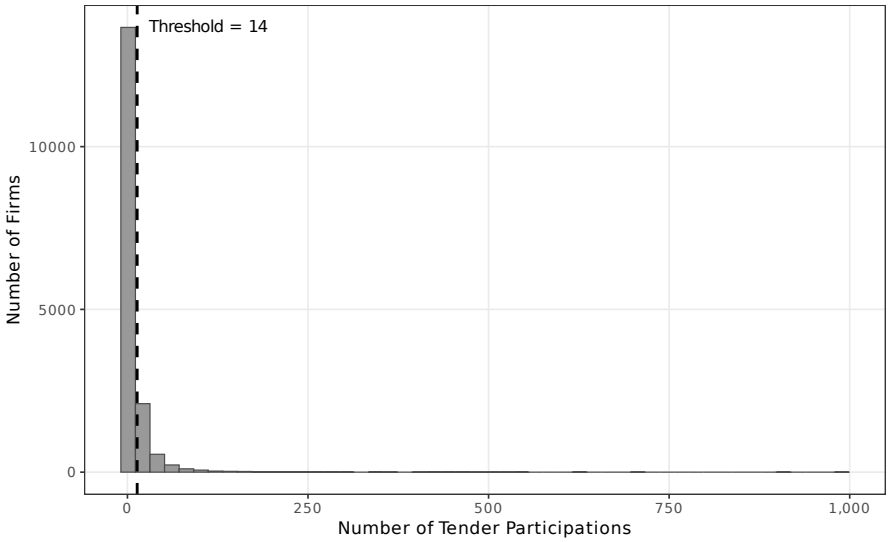

Figure B.2: Identification of the frequent-loser threshold via median + 1.5× IQR on the always-loser participation distribution; the cutoff yields 14 tender-items (13.5 statistical threshold) and flags 2,735 firms.

Table B.4: Frequent-Loser Price Coefficient by Modality (Pregão vs. Convite)

Specification	Pregão-only	Convite-only	Full sample	Ratio (P/C)
Binary FL14	$+0.0959^{***}$	$+0.0392^{**}$	$+0.0653^{***}$	2.45×
	(0.0256)	(0.0188)	(0.0216)	
	$p < 10^{-3}$	$p = 0.037$	$p = 0.003$	
Continuous log(1+ max tc)	$+0.0262^{***}$	$+0.0124^{**}$	—	2.11×
	(0.0063)	(0.0049)	—	
	$p < 10^{-4}$	$p = 0.011$	—	
Joint specification: both regressors				
Binary FL14 (with continuous control)	-0.0652	-0.0828^{**}	—	—
	(0.0449)	(0.0350)	—	—
Continuous (with binary control)	$+0.0399^{***}$	$+0.0305^{***}$	—	—
	(0.0118)	(0.0101)	—	—
Items	543,752	1,105,852	1,653,658	—
Item, year, PBU FE	YES	YES	YES	—
Cluster (item)	YES	YES	YES	—
Minimum-bidder rule binds	NO	YES	MIXED	—

Notes: Outcome is $\log p_{\text{negotiated}}$. The pregão coefficient is roughly 2.45× the convite coefficient on the binary specification. Under the screening framing (§5.1), we read the modal contrast as variation in the observability regime within which the screening signal is recovered, not as a positive test of any specific institutional channel. SEs clustered at item level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B.5: McCrary Density Diagnostic at the Convite/Pregão Statutory Threshold

	Density at threshold			Modality
Threshold (Lei 8.666 art. 23)	Left	Right	Log-density disc.	Convite share gap
R$80,000 (pre-2018)	381	406	−0.063	0.098
R$176,000 (post-2017)	216	277	−0.250	0.092

Notes: Densities estimated by local linear smoothing of contract-record counts in ±50% windows around each statutory cap. The convite-share gap is the difference between the share of contracts using the convite modality just below versus just above the threshold. Both thresholds show small log-density discontinuities ($|\cdot| < 0.25$); neither shows the kind of bunching that would indicate strategic gaming of the modality assignment by buyers. The ~ 9.8 percentage-point convite-share drop across the cap is consistent with rule-driven assignment of value to modality, not with strategic selection.

Table B.6: Price Regressions Excluding CADE-Involved Markets

	(1) General+PBU	(2) Pregão
FL presence	0.0616***	0.0901***
	(0.0213)	(0.0256)
Observations	1,622,954	534,956
CADE tenders dropped	12,514	

Notes: All tenders involving any CADE-convicted firm are excluded. SE clustered at item level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

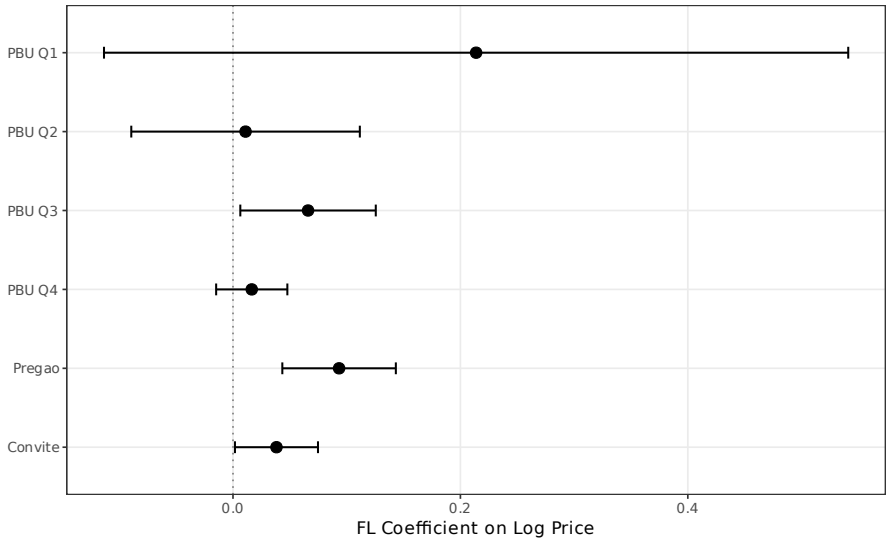

Figure B.3: Frequent-loser price coefficient by procuring-unit-size quartile; the Q1-vs-Q4 contrast (smallest to largest buyers) is 12.6×, intermediate quartiles imprecisely estimated. Companion to Table B.3.

of below-threshold always-losers do not produce comparable discrimination. Table C.3 reports the rolling-origin temporal-holdout AUC by test year (year-by-year decomposition of Figure 1 in the body). Table C.4 reports minimum-detectable-effect calculations for the failed RDD and DiD designs of §5.1: observed-to-MDE ratios well below one indicate the null findings are structurally underpowered, not refutations of the underlying mechanism. Tables C.5 and C.6 carry the overlap-restricted ATT estimates and the cell-dropping decomposition discussed in §7.2, kept here rather than in the body so the price-imprint section remains compact: the overlap-cell ATT and propensity-score-trimmed ATT formalize the within-support reweighting that produces the sign reversal, and the cell-dropping decomposition shows that only 1.06% of treated items lack a within-cell counterfactual and are strictly dropped.

Table C.1: Oster (2019) Coefficient Stability Bounds

	Full sample	Pregão	Convite
$\hat{\beta}_{\text{short}}$ (item + year FE)	0.0676	0.0954	0.0507
$\hat{\beta}_{\text{full}}$ (+ PBU FE)	0.0636	0.0933	0.0382
$\tilde{R}^2_{\text{short}}$	0.8790	0.8737	0.8857
$\tilde{R}^2_{\text{full}}$	0.8860	0.8821	0.8932
$R^2_{\max} = \min(1, 1.3\tilde{R}^2)$	1.0000	1.0000	1.0000
$\hat{\delta}$ (Oster bound)	261.6	634.00	43.62
$\beta^*(\delta = 1)$	-0.0005	0.0643	-0.1395
Observations	1,654,401	546,549	1,107,852

Notes: $\hat{\delta}$ is the degree of selection on unobservables (relative to observables) required to drive $\hat{\beta}$ to zero (Oster, 2019). $|\hat{\delta}| > 1$ indicates robustness: unobservables would need to be more important than observables to explain the coefficient. $\beta^*(\delta = 1)$ is the bias-adjusted coefficient assuming equal selection. $R^2_{\max} = 1.3\tilde{R}^2$ following Oster's recommended bound. Short model: item + year FE. Full model adds PBU FE (and convite control for full sample).

Appendix D. Threshold and Specification Robustness

The IQR-multiplier sensitivity (Table D.1) places the headline price coefficient as positive and significant under five multipliers ranging from 1.0× to 3.0× IQR, declining

Table C.2: Placebo IV: Sub-Threshold Always-Losers as Instrument

	First Stage	2SLS
	FL presence	Log Price
Sub-threshold supply (LOO)	0.000040***	
	(0.000007)	
F-statistic	33.0	
FL presence (instrumented)		5.2445***
		(1.3200)

Notes: Placebo instrument uses supply of sub-threshold always-losers (below the IQR cutoff). The sub-threshold first stage ($F = 33$) exceeds the Stock and Yogo (2005) 10% maximal IV size critical value of 16.38 for a single instrument. The implausibly large 2SLS coefficient (5.24) is therefore not attributable to weak-instrument bias. Rather, it reflects the absence of a genuine cover-bidding mechanism: sub-threshold always-losers do not drive procurement price anomalies, confirming that the main instrument's price effect operates through the cover-bidding channel rather than through correlated demand shocks. SE clustered at item level.

Table C.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 trains the construct on years preceding the test year (rolling-origin scheme starting in 2009) and evaluates AUC against the cobidder ground truth realized in the test year. Bootstrap 95% CI from 1,000 resamples. Source: `scripts/17_temporal_holdout_roc.R`.

Table C.4: Minimum Detectable Effects for Failed RDD and DiD Designs

Test	Observed coef	SE	MDE @ 80%	MDE @ 50%	N	Obs/MDE
RDD R$80k pre-Decreto, FL prevalence	−0.0038	0.0033	0.0092	0.0065	7,836	0.41
RDD R$176k post-Decreto, FL prevalence	+0.0054	0.0170	0.0476	0.0333	1,053	0.11
RDD R$80k pre-Decreto, convite share (1st-stage)	+0.0046	0.0198	0.0553	0.0387	5,324	0.08
RDD R$176k post-Decreto, convite share (1st-stage)	+0.1556	0.0759	0.2126	0.1488	472	0.73
DiD Decreto, FL prevalence	−0.0059	0.0105	0.0294	0.0206	3,676	0.20
DiD Decreto, log(value) — placebo	+0.0616	0.0574	0.1608	0.1125	3,676	0.38
DiD Decreto, n_firms	−0.0022	0.0016	0.0044	0.0031	3,676	0.49
DiD Decreto, convite share (1st-stage)	+0.0234	0.0356	0.0998	0.0698	3,676	0.23

Notes: Power calculations for the design-based identification strategies that returned null. The MDE columns report the smallest effect detectable at 80% and 50% power for the realized N. Observed-to-MDE ratios well below one indicate the null is structurally underpowered, not a refutation of the underlying mechanism. Source: `scripts/23_mde_calculations.R`.

Table C.5: Item-Level Scope Checks: Conditional Association Under Stricter Overlap

Specification	Coefficient	SE	p-value	N
baseline_fe	+0.064	0.021	0.003	1,653,658
overlap_cell_att	−0.097	0.015	< 0.001	1,517,066
overlap_ref_att	−0.097	0.015	< 0.001	1,434,636
ps_att_trimmed	−0.307	0.020	< 0.001	400,687

Notes: Reports β (`baseline_fe`) and the overlap-restricted β^{ov} (`overlap_cell_att`, `overlap_ref_att`, `ps_att_trimmed`). Under the screening framing of §5.1 these are different empirical objects rather than alternative estimates of the same parameter; the sign reversal is the diagnostic that locates the construct as a screening-value object.

Table C.6: Sign-Reversal Decomposition: Where Does $\widehat{\beta}^{ov} < 0$ Live, and Which Items Get Dropped?

	Coef. / Mean	SE	p-value	N
Panel A. Headline specifications on common sample.				
Broad sample $\widehat{\beta}$	+0.064	0.021	0.003	1,654,401
Overlap-cell ATT $\widehat{\beta}^{ov}$	−0.097	0.015	< 0.001	1,517,868
Panel B. $\widehat{\beta}^{ov}$ within overlap-cell sample, by subgroup.				
$\widehat{\beta}^{ov}$ in non-direct-CADE items	−0.097	0.016	< 0.001	1,487,055
$\widehat{\beta}^{ov}$ in direct-CADE items	−0.061	0.080	0.445	30,813
$\widehat{\beta}^{ov}$ in pregão items	−0.099	0.020	< 0.001	475,923
$\widehat{\beta}^{ov}$ in convite items	−0.098	0.013	< 0.001	1,041,945
Panel C. Treated items in dropped vs. surviving cells (means).				
	In dropped	In surviving	Difference	t-test p
Convite share among treated items	0.496	0.638	−0.142	< 0.001
Direct-CADE-item share among treated items	0.019	0.024	−0.005	0.333
Cobidder-item share among treated items	0.139	0.078	+0.061	< 0.001
Mean log reference price among treated items	6.691	5.490	+1.201	< 0.001
Mean number of bidders among treated items	10.533	9.047	+1.486	< 0.001

Notes: Panel A reports the broad-sample $\widehat{\beta}$ and the overlap-cell ATT-weighted $\widehat{\beta}^{ov}$, both estimated on items with non-missing log negotiated price and within-item, year, and PBU fixed effects. Panel B re-estimates $\widehat{\beta}^{ov}$ on the same overlap-cell ATT design within four subgroups (pregão vs. convite items; non-direct-CADE vs. direct-CADE items). Panel C compares treated items (FL present) that fall in cells dropped by the overlap restriction (no untreated counterfactual) versus treated items in surviving cells; differences in convite share, direct-CADE-item share, cobidder-item share, log reference price and number of bidders test the screening-value reading directly. The screening framework predicts that items dropped by overlap restriction should disproportionately carry markers of deployment value (higher convite, direct-CADE, cobidder-item rates).
Source: `scripts/59_sign_reversal_decomp.R`.

monotonically as the cutoff tightens—the pattern the framework's continuous primitive predicts (§8.1). Standard-error clustering (Table D.2) under item-level (baseline), procuring-unit-level, and two-way Cameron et al. (2011) clustering keeps the coefficient significant at $p < 0.01$ under all three regimes.

Table D.1: Price Coefficient Across IQR Multiplier Thresholds

Multiplier	Threshold	FL Firms	Coef.	SE	Observations
1.0 × IQR	10	3,442	0.0791***	(0.0227)	1,654,401
1.5 × IQR	14	2,735	0.0636***	(0.0215)	1,654,401
2.0 × IQR	17	2,093	0.0598***	(0.0221)	1,654,401
2.5 × IQR	20	1,778	0.0543**	(0.0228)	1,654,401
3.0 × IQR	24	1,456	0.0501**	(0.0237)	1,654,401

Notes: DV: log negotiated price. Each row re-estimates the baseline specification (item + year + PBU fixed effects) using a different IQR multiplier to define the FL threshold. The threshold equals median $+k\times$ IQR of the tenders-count distribution among always-losers. The baseline (1.5× IQR) is highlighted. SE clustered at item level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table D.2: Alternative Clustering of Standard Errors

Clustering level	Coef.	SE	95% CI
Item level (baseline)	0.0636***	(0.0215)	[0.0214, 0.1058]
PBU level	0.0636***	(0.0144)	[0.0353, 0.0919]
Two-way (item + PBU)	0.0636***	(0.0240)	[0.0165, 0.1107]

Notes: DV: log negotiated price. All rows report the same baseline specification (item + year + PBU fixed effects) with different clustering of standard errors. The point estimate is invariant to clustering; only the standard errors and confidence intervals change. The two-way cluster uses the Cameron et al. (2011) multi-way approach. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix E. Generalization Audit and Detector Comparison

The operational-metrics temporal-holdout audit (Table E.1) referenced in §9.3 and the same-sample incremental-AUC comparison (Table E.2) referenced in §10 are reported here, together with the cross-sector external-validity scope (Table E.3) that quantifies

the within-product-code identifying variation underwriting the headline price gap and the AUC decomposition (Table E.4) that reports the marginal contribution of each feature block to discrimination on the cobidder ground truth. The participation primitive $\log(1+\text{tenders\_count})$ alone reaches AUC near the full-feature ceiling; bid-based features contribute non-trivially and dropping the participation primitive collapses discrimination to bid-distribution-only performance, supporting the framework's identification of the participation primitive as the dominant signal at the award layer.

Table E.1: Operational Deployment Metrics: Precision, Recall, and Lift at Top-k. Headline column is the temporal-holdout block (left); the in-sample block (right) is reported for transparency and over-states precision by approximately 50% at top-500.

	Temporal holdout (score: 2009–2016) [headline]				In-sample (full 2009–2019) [transparency]			
Top-k	TP	Precision	Recall	Lift	TP	Precision	Recall	Lift
50	1	0.020	0.005	1.7×	15	0.300	0.078	26.2×
100	7	0.070	0.036	6.1×	17	0.170	0.088	14.8×
250	19	0.076	0.098	6.6×	40	0.160	0.207	14.0×
500	35	0.070	0.181	6.1×	66	0.132	0.342	11.5×
1,000	66	0.066	0.342	5.8×	97	0.097	0.503	8.5×
Reference quantities								
Always-loser pool size N				16,843				
Cobidder positives N_+				193				
Base cobidder rate				0.0115				
ROC AUC (log(1+tc))		Temporal holdout: 0.864				In-sample: 0.939		

Notes: Score: $\log(1+\text{tenders\_count})$, ranked descending across the always-loser pool. Lift is precision divided by the base cobidder rate. The temporal-holdout columns (left, headline) score firms using only 2009–2016 participation and evaluate on items in 2017–2019 against the 193 cobidder labels. The in-sample columns (right, transparency) score firms using 2009–2019 participation and are reported only to expose the ∼50% inflation at top-500 that arises because the score draws on participation occurring after CADE adjudications were already underway for some cartels. All operational claims in the paper are read off the temporal-holdout block. The screening statistic requires only contract-award records (winner identity, participant identity, item identifier); no bid microdata.

Appendix F. Adversarial-Adaptation Simulation

The screen's behavior under deliberate adversarial adaptation (§11.2) is bounded but not invariant. We examine four strategic adaptations. Each holds the screening statistic fixed and modifies the always-loser pool to mimic plausible evasive behavior, and we recompute the firm-level AUC of $\log(1+\text{tenders\_count})$ against the 193 CADE-cobidder labels under each (Table F.1). Two are observable adaptations the screen absorbs cleanly. Rotation (20% of cover bidders retire annually and are replaced by fresh CNPJs) and

Table E.2: Incremental Value of Loser-Side Concentration Relative to the Full Imhof Benchmark

Model	Data requirements	AUC	95% CI	Δ vs Imhof full	DeLong p
Imhof full pipeline	All bid values within each tender	0.846	$[0.819, 0.873]$	+0.000	—
Binary FL flag	Winner + participants only	0.881	$[0.871, 0.892]$	+0.035	0.014
Continuous participation count	Winner + participants only	0.877	$[0.857, 0.898]$	+0.031	0.077
Imhof full + binary FL	Bid microdata + award records	0.942	$[0.927, 0.957]$	+0.096	< 0.001
Imhof full + participation count	Bid microdata + award records	0.944	$[0.929, 0.958]$	+0.098	< 0.001

Notes: All models are evaluated on the exact same always-loser subsample for which the full bid-distribution feature set is available. The relevant comparison is therefore not "does FL beat Imhof everywhere," but whether participation-only information remains informative when rich bid microdata are absent and whether it adds non-redundant signal when those microdata are present. The combined specifications answer the second question directly.

Table E.3: External Validity Scope: Modality and Commodity-Service Boundaries

Dimension	Group	Metric	Value	95% CI
coverage	Commodity	share_items	0.887	—
coverage	Service	share_items	0.113	—
dominant_item_class	Commodity	auc_log_tc	0.905	$[0.879, 0.931]$
dominant_item_class	Service	auc_log_tc	0.966	$[0.961, 0.971]$
item_class_price	Commodity	coef_losers	0.045	$[0.002, 0.088]$
item_class_price	Service	coef_losers	0.128	$[0.024, 0.231]$
modal_primary_auc	convite_primary	auc_log_tc	0.816	$[0.758, 0.874]$
modal_primary_auc	pregao_primary	auc_log_tc	0.952	$[0.946, 0.958]$

Notes: Item classes use a transparent text rule on BEC item descriptions. Rows labelled `modal_primary_auc` reuse the always-loser firm split by primary modality; rows labelled `item_class_price` report the within-item price association separately for commodity-like and service-like items. The point is not to claim exportability to all procurement, but to document that the paper speaks most naturally to standardized-goods environments and selected simple services.
Source: `scripts/52_external_validity_scope.R`.

Table E.4: AUC Decomposition Across Feature Sets

Model	Features	AUC	95% CI	Marginal AUC vs. A
A	is_fl + imhof_cv + imhof_spread + tenders_count + n_bids	0.939	[0.928, 0.949]	0.939
B	imhof_cv + imhof_spread + tenders_count + n_bids	0.936	[0.924, 0.948]	0.003
C	is_fl	0.887	[0.868, 0.905]	0.052
D	imhof_cv + imhof_spread	0.785	[0.754, 0.816]	0.154

Notes: Random-forest classifiers trained on the cobidder ground truth, with five feature sets. Model A is the full set; subsequent models drop one block. The marginal contribution column reports the AUC delta from removing the listed feature(s) relative to A. Source: `scripts/20_auc_decomposition.R`.

occasional wins (5% of threshold-crossing firms win one tender) leave AUC essentially unchanged: the bright-line wins = 0 filter eliminates the small share of newly-winning cover bidders—a feature of the framework's separating equilibrium (Lemma 1), not a heuristic robustness—and rotation does not compress the $\log(1+\text{tenders\_count})$ ranking statistic across the always-loser pool.

Table F.1: Adversarial-Adaptation Simulation: AUC Degradation Under Strategic Cartel Adaptations

Scenario	AUC	Δ vs baseline	Retention
Baseline (no adaptation)	0.928	+0.000	100.0%
Rotation: 20% of cover bidders retire and are replaced	0.931	+0.003	100.3%
Occasional wins: 5% of FL14 firms get one win	0.929	+0.002	100.2%
CNPJ splitting: each FL14 firm splits into two CNPJs	0.891	−0.037	96.0%
Threshold cap: firms cap participation at 13 tender-items	0.801	−0.127	86.3%
Combined: threshold cap + 20% rotation	0.632	−0.295	68.2%

Notes: Each scenario applies a strategic adaptation to the BEC always-loser pool, holding the construct (log(1+tenders_count) ranking against 193 CADE cobidder labels) fixed. The construct is resilient to rotation and to occasional wins (the bright-line filter eliminates the small share of newly-winning cover bidders without compressing the ranking), and degrades modestly under CNPJ splitting. It is substantially vulnerable to threshold-aware capping and severely degraded when capping is combined with rotation: a sophisticated cartel that targets the threshold can compress AUC by roughly 30 points. Retention is the fraction of baseline AUC preserved.

Two are threshold-aware adaptations that degrade discrimination. CNPJ splitting (each frequent loser splits into two CNPJs carrying half the participation count) degrades AUC by 3.7 percentage points; threshold-aware capping (firms cap participation at 13 tender-items to stay below the cutoff) degrades AUC by 12.7 points; the combined attack (capping plus 20% rotation) drives AUC from 0.928 to 0.632, a 29.5-point compression. These adaptations target the binary cutoff specifically; the continuous primitive $\log(1+\text{tenders\_count})$ that Proposition 2 identifies as the sufficient ranking statistic is not

directly altered, but the binary discretization is. Threshold capping is the most exposed margin: a cartel that knows the cutoff can instruct cover bidders to remain below it.

The implication for the screening-vs-forensic-stage architecture of §10 is asymmetric. Under combined adaptation AUC drops to 0.632—still above random matching, still informative as a screening statistic, but substantially reduced. The forensic stage is unaffected by threshold-aware adaptation at the screening layer: bid-distribution moments operate on a different data envelope and are robust to the participation-pattern engineering that defeats the binary rule. A combined two-stage architecture in which a refreshable screening rule feeds a forensic stage on bid microdata is therefore more adaptation-robust than either component alone—the empirical evidence in §10 shows the same conclusion by distinct means. The screening statistic is not adaptation-proof; practical lifetime depends on whether the threshold can be periodically refreshed against new data and whether the screening stage is deployed alongside the bid-layer forensic stage where microdata are available.

Appendix G. Staggered Designs Attempted and Rejected

Two design-based strategies that would extract a causal estimate return null at the available bandwidths (§11.1). The sharp-RDD failure (gated by the McCrary density test) is documented in Table B.5 of Online Appendix B; this appendix reports the parallel diagnostic for the difference-in-differences strategy exploiting Decreto 9.412/2018's 2018 raise of the convite cap. The Sun and Abraham (2021) interaction-weighted estimator yields imprecise coefficients with non-monotone dynamic patterns incompatible with the smooth treatment-onset narrative the cap raise would imply (Table G.1); a stacked difference-in-differences design constructed around the cap-raise event window returns coefficients that flip sign across stacks and lack a clean pre-period (Table G.2). We treat both estimators as null at the available bandwidths. The interpretive consequences are limited to identification scope: the screening statistic does not require the cap raise to be a clean treatment, and the broad-sample β of §7.1 does not depend on the staggered design surviving. What the null does constrain is alternative framings of the empirical object that would require a causal identification of β; the screening framing trades that

claim for the screening-value interpretation of β and reports the failures here for full disclosure.

Table G.1: Staggered DiD: Callaway & Sant'Anna (Restricted Sample)

	C&S (2021)		TWFE	
Outcome	ATT	SE	Pre-trend	Post
Log Price	0.0143	(0.0396)	-0.0131	-0.0079
Log Firms (excl. FL)	-0.0220	(0.0333)	-0.0088	0.0113
Panel B: Statistical Power				
MDE (5%, two-sided)		0.0776		
MDE (80% power)		0.1109		
OLS benchmark		0.064		
Panel D: Sharp Entry Subsample				
C&S ATT	0.0143	(0.0389)		
Sharp entry markets		1,511 treated, 18,266 control		
Markets		19,777 (1,511 treated, 18,266 control)		
Sample restriction		$\geq$3 pre & $\geq$3 post observations		

Notes: Panel A: C&S = Callaway & Sant'Anna (2021) doubly-robust estimator. Panel B: MDE = minimum detectable effect at stated significance/power levels. Panel C: Rambachan & Roth (2023) sensitivity bounds under smoothness restriction $\bar{M}$ on maximum change in the slope of the trend. Panel D: restricts treated markets to those with zero FL presence before treatment year.

References

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

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

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *Quarterly Journal of Economics 134*(3), 1405–1454.

Table G.2: Stacked DiD vs. Callaway & Sant'Anna

	Stacked (Cengiz et al.)		C&S (2021)	
Outcome	ATT	SE	ATT	SE
Log Price	-0.0059	(0.0138)	0.0143	(0.0405)
Log Firms (excl. FL)	0.0162*	(0.0084)	-0.0220	(0.0339)
Stacked obs.		715,116		
Cohorts		6		
Event window		±5 years		
Cohort × Market FE	YES		—	
Cohort × Year FE	YES		—	
Clustering		Market level		

Notes: Stacked regression following Cengiz et al. (2019). Each cohort-year sub-experiment includes treated markets entering in year g and all never-treated markets, within a symmetric ±5-year event window. Cohort-specific market and year fixed effects absorb level differences across sub-experiments. SE clustered at market level. C&S = Callaway & Sant'Anna (2021) doubly-robust estimator from Table G.1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

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.

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.

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

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

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.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pp. 80–108. Cambridge: Cambridge University Press.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199.