Transparency vs. Tacit Collusion in Algorithmic Ad Auctions: Optimal Disclosure with Budgets and Learning Bidders

Liz Lemma Future Detective

January 16, 2026

Abstract

Digital ad markets in 2026 are dominated by auto-bidding agents that learn from platform feedback under budget and performance constraints. This paper studies a platform's disclosure problem: what information about rivals' budgets and performance should be revealed through transparency APIs without enabling tacit collusion among learning algorithms. Motivated by AuctionNet-style environments (multi-agent, budgeted repeated auctions with configurable observation logs), we develop a clean repeated-auction model with a continuous disclosure knob controlling (i) bidders' ability to learn the market state (improving bidding efficiency and potentially raising revenue) and (ii) the quality of public monitoring (enabling stable bid suppression/market division among adaptive agents). In the model, expected revenue (and in some regions, welfare) is non-monotone in disclosure precision: moderate transparency improves competitive bidding by reducing learning frictions, while high transparency expands the set of sustainable collusive outcomes under public monitoring, lowering revenue and potentially distorting allocation. We complement the theory with an AuctionNet-based empirical design that implements multiple disclosure regimes by altering the observation space/logs, and we measure revenue, value, bid dispersion, spend volatility, and rotation/concentration statistics to detect tacit collusion. The results provide a mechanism-design and governance rationale for selective, noisy, or delayed transparency policies in modern ad auctions.

Table of Contents

budgets (baseline nonbinding), and a platform-chosen disclosure precision controlling both state learning and monitoring.

3. 3. Competitive benchmark (full-information): closed-form symmetric equilibrium multiplier and revenue/welfare formulas; interpretation as a pacing/shadow-price analog.
4. 4. Learning-efficiency channel (low transparency): how limited feedback biases multiplier choice downward; comparative statics of revenue and spend volatility with increasing signal precision.
5. 5. Monitoring-collusion channel (high transparency): perfect public equilibria with bid suppression/rotation; enforceability condition as a function of monitoring precision and patience.
6. 6. Main result: non-monotone platform objective and interior optimal disclosure; comparative statics in market concentration, patience, and heterogeneity; discussion of budgets/CPA as extensions.
7. 7. AuctionNet empirical design: implement disclosure regimes by modifying observation/log APIs; standardize bidder algorithms (e.g., pacing+bandit, IQL/BC/DT variants); define collusion and stability metrics.
8. 8. Experimental results: revenue–welfare curves by disclosure level; evidence of rotation/market division vs benign learning; robustness across auction formats (GSP vs FP) and slot structures.
9. 9. Implications for platform governance and antitrust: recommended disclosure designs (noise, delay, aggregation, anonymization) and auditing protocols.
10. 10. Limitations and extensions: richer bidder heterogeneity, endogenous entry, multi-slot externalities, privacy constraints; open problems.

1 Introduction and policy motivation

Digital advertising markets in 2026 are simultaneously *more transparent* and *more opaque* than they were a few years ago. They are more transparent because large platforms increasingly expose advertisers to "transparency APIs" and reporting dashboards that provide near-real-time information about auction outcomes (prices, ranks, impression shares, pacing diagnostics, and sometimes coarse signals of competitor behavior). Yet they are also more opaque because these disclosures are now filtered through privacy constraints, aggregation requirements, and noise-injection mechanisms that limit what can be inferred about any particular user, query, or rival. This tension is not merely a compliance detail. It changes the strategic environment faced by advertisers that bid through adaptive algorithms, and it changes the design problem faced by a platform that chooses what to reveal.

We motivate our analysis from a practical design question that platforms and regulators now confront explicitly: *How precise should public reporting be?* If the platform releases detailed, high-frequency auction statistics, advertisers can diagnose performance quickly, infer competitive conditions, and adjust bids or pacing rules. These benefits are often framed as improving "market efficiency" and reducing wasteful experimentation. However, the same stream of public statistics can also function as a coordination device. When bidders are algorithmic and repeated interaction is the norm, even relatively coarse public signals can support tacit coordination: algorithms can learn to avoid aggressive competition, punish deviations, and settle into stable bid-suppression patterns. In this sense, transparency is not innocuous. It can be an input into an equilibrium selection problem.

Two developments make this issue more salient in 2026. First, automated bidding has become the default for many large advertisers. Bids are no longer set by a human manager reading weekly reports; they are produced by learning systems that update continuously as new feedback arrives. Second, reporting policies have become objects of governance. Privacy laws and platform policies restrict user-level data access, while new market-design rules (including interoperability and transparency mandates in some jurisdictions) push in the opposite direction. The platform therefore faces a constrained information-design problem: disclosures must be sufficiently informative to support legitimate optimization and accountability, but not so informative that they facilitate collusion or exclusion.

Our central claim is that this is a *tradeoff*, not a one-sided argument for either maximal transparency or maximal opacity. On the one hand, low-precision disclosure can create a learning friction. When advertisers do not observe enough about auction outcomes, they may mis-calibrate competitive pressure and systematically underbid relative to the one-shot benchmark. This can depress revenue and, depending on the environment, can also reduce allocative efficiency by preventing high-value bidders from reliably winning. On the other hand, high-precision disclosure can create a monitoring technology. If bidders can publicly infer whether rivals deviated from a low-bid "understanding," then punishments become credible and collusion becomes sustainable in repeated interaction. The platform then risks moving from a competitive outcome with noisy learning to a coordinated outcome with suppressed prices and potentially distorted allocation.

It is important to emphasize that the collusion concern here is not the cartoon of explicit human communication. Rather, it is the possibility that repeated play plus shared observables allow *algorithmic* strategies to implement contingent responses that resemble classic collusive schemes. Public reporting can provide the state variables required for such contingencies. For example, if the platform publishes (even with delay) impression-share breakdowns, average winning prices by segment, or noisy signals of the top bid, then an algorithm can condition future aggressiveness on whether outcomes are "consistent" with a cooperative path. Conversely, if the platform only reveals each advertiser's own outcomes (bandit-style feedback), then deviations are harder to detect, punishments lose bite, and collusion is more difficult to sustain.

Privacy constraints make the design problem subtler, not simpler. Modern disclosure regimes frequently rely on aggregation, anonymization, thresholding, and noise (including variants of differential privacy). These tools are often motivated by user privacy, but they also directly shape strategic inference about the market. A reporting rule that adds Gaussian noise to a price statistic, for instance, is naturally interpreted as reducing the precision of a public signal. From a market-design perspective, the key question is not whether the signal is "private" in a legal sense, but how informative it is about underlying competitive conditions and about rivals' actions. In particular, privacy-preserving noise can have two opposing economic effects: it can slow down learning about demand and competition (reducing bids), while also blunting the ability to monitor and enforce coordination (reducing collusion).

We also stress that the platform's objective is not purely theoretical. Platforms routinely evaluate changes to reporting, auction rules, and bidding products through controlled experiments, and they increasingly deploy offline simulators and counterfactual estimators. This is where AuctionNet-style benchmarking becomes useful. AuctionNet introduced a disciplined way to compare bidder behavior to an internally computed auction-theoretic benchmark, adjusting for performance objectives such as target-CPA via multiplicative penalties. The broader methodological point is that we can operationalize the gap between observed bidding and a normative competitive benchmark by constructing a mapping from primitives (values, objectives, constraints) to implied bids, and then measuring systematic deviations. When a platform changes disclosure precision, we can observe (or estimate) how quickly bidders' multipliers converge toward the benchmark and whether cross-advertiser outcomes become more predictable in a way consistent with coordination. In this sense, the transparency question is empirically testable: disclosure policy shifts observable sufficient statistics (bid multipliers, spend shares, price distributions), and these shifts can be compared to the competitive predictions.

This perspective suggests a concrete policy relevance. Regulators often view transparency as a remedy for market power, while privacy regulators view opacity as a remedy for surveillance and discrimination. Our framework highlights a third dimension: transparency can itself *create* market power by enabling coordination among bidders. The same API that helps a small advertiser optimize can help a large advertiser stabilize a market-division scheme. This does not imply that transparency should be abandoned; rather, it implies that transparency should be *engineered*. The platform's choice is not binary (reveal everything versus reveal nothing). It includes the granularity of statistics, the amount of noise, the delay, the degree of aggregation across queries or audiences, and whether information is public or only privately visible to each bidder.

Finally, we acknowledge two limitations of any stylized approach. First, real ad auctions are multi-slot, involve quality scores and reserve prices, and interact with budget pacing and attribution systems. Second, the relevant "values" are derived from downstream conversion uncertainty and heterogeneous objectives, not literal private valuations drawn from a simple distribution. We do not claim to replicate these details. Instead, we aim to isolate a mechanism that is robust across institutional specifics: disclosure precision affects both (i) how quickly adaptive bidders learn the competitive environment and (ii) how easily bidders can monitor one another in repeated interaction. These two channels point in opposite directions for revenue and, in some cases, for welfare. The rest of the paper formalizes this tradeoff in a repeated auction model that is intentionally minimal but rich enough to accommodate both learning and collusion incentives.

2 Stylized model: repeated first-price auctions with learning and disclosure

We study a deliberately minimal repeated-auction environment that isolates the informational role of platform reporting. Time is discrete, with periods $t \in \{1, \dots, T\}$, and in each period a single indivisible impression (or "slot") is allocated via a first-price auction among $N \geq 2$ advertisers indexed by $i \in \{1, \dots, N\}$. The platform plays a dual role. First, it runs the auction and implements the allocation and payment rule. Second, it chooses what the market can publicly infer about competitive conditions and rivals' actions through a disclosure policy. Our goal is not to replicate the institutional details of modern ad exchanges, but to formalize a tradeoff that recurs across them: disclosures that help advertisers learn and optimize can simultaneously provide the public monitoring needed to sustain tacit coordination.

Values and market scale. In each period t, advertiser i privately observes a value $v_{i,t}$ for winning the slot. In the baseline specification, values are i.i.d. across advertisers and time with

$$v_{i,t} \sim \text{Unif}[0, \theta],$$

where the market scale (or competitiveness) parameter $\theta > 0$ is fixed over the horizon but unknown to bidders.[1] This "unknown θ" device lets disclosure affect behavior even when the per-period auction has a familiar benchmark under full information. In particular, advertisers face a state-learning problem: how aggressive should they be, given uncertainty about the distribution of opponents' values and hence about the bid they must beat?

Bidding, budgets, and the stage auction. Each period is a first-price auction. Advertiser i submits bid $b_{i,t} \geq 0$, the platform allocates the slot to the highest bidder, and the winner pays its bid. Formally, letting $x_{i,t} \in \{0, 1\}$ denote the allocation indicator,

$$x_{i,t} = \mathbf{1}\{b_{i,t} = \max_j b_{j,t}\},$$

with ties broken uniformly at random, and payments satisfy

$$p_{i,t} = x_{i,t} b_{i,t}.$$

Advertiser utility aggregates discounted per-period surplus,

$$U_i = \mathbb{E}\Big[\sum_{t=1}^{T} \delta^{t-1}\big(x_{i,t} v_{i,t} - p_{i,t}\big)\Big], \qquad \delta \in (0, 1).$$

We incorporate budgets to reflect a central operational constraint in ad markets: bids and spend are typically coupled through daily or campaign-level caps. Let $B_{i,t}$ denote advertiser i's remaining budget at the start of period t, with the feasibility constraint $b_{i,t} \leq B_{i,t}$ and the law of motion

$$B_{i,t+1} = B_{i,t} - p_{i,t}.$$

In the baseline analysis we treat budgets as nonbinding so as to focus on the information-design mechanism rather than on dynamic shadow values of

[1] In applications, θ can be interpreted as a reduced-form index of opportunity size and competitive intensity: larger θ implies both higher potential surplus and, in equilibrium, higher clearing prices. The assumption that θ is constant is a simplification that turns learning into a stationary inference problem; in practice, one could allow θ_t to drift and interpret our disclosure parameter as governing how quickly bidders can track that drift.

money. Nevertheless, retaining the budget state in the notation is useful because (i) it clarifies what information is potentially disclosable (e.g., whether rivals appear to be budget-constrained), and (ii) it anticipates extensions in which pacing and budget management interact with learning about θ.

Adaptive bidding via multipliers. Rather than allow arbitrary bid functions, we model algorithmic bidding as choosing a multiplier applied to the current private value. Specifically, each bidder selects $\alpha_{i,t} \in \mathcal{A} \subset (0, 1]$, where $\mathcal{A}$ is a finite action grid, and submits

$$b_{i,t} = \alpha_{i,t} v_{i,t}.$$

This captures a common feature of automated bidding products: they expose a small number of "aggressiveness" or "target" knobs, and the system maps those choices into bids that scale with predicted value. The finite grid allows us to treat bidders as running standard no-regret procedures (e.g., multiplicative weights or bandit variants) over a discrete action set. Importantly, the informational environment determines what feedback these learning algorithms receive, and hence how quickly multipliers adapt toward the best response implied by the underlying auction.

Platform disclosure as precision. At stage 0 the platform commits to a disclosure policy indexed by a scalar precision parameter $\kappa = 1/\sigma^2$. We treat disclosure as producing public observables that are informative about (i) the persistent state θ and (ii) rivals' realized behavior. The first component is a public signal

$$s_t = \theta + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma^2),$$

so higher κ corresponds to less noisy reporting about market scale. The second component is an additional public monitoring statistic m_t, released each period with informativeness governed by the same κ. We intentionally leave m_t abstract, because many real reporting products fit the same role: noisy top-bid estimates, noisy winning price, noisy rank/impression-share information, or anonymized summaries of competitors' bid multipliers. What matters for our purposes is that higher precision makes it easier to infer whether rivals behaved "as expected".

Formally, bidder i's information at time t includes its private value and budget and the history of public disclosures and own outcomes:

$$\mathcal{I}_{i,t} = \big(v_{i,t}, B_{i,t}, s_{1:t}, m_{1:t}, x_{i,1:t-1}, p_{i,1:t-1}\big).$$

Low-precision regimes correspond to environments close to bandit feedback: bidders mostly observe their own wins and payments, with little ability to separate changes in θ from idiosyncratic auction noise. High-precision regimes move the game toward public monitoring, where realized outcomes become common knowledge and can support contingent strategies.

Two channels: learning efficiency and collusive monitoring. The model is designed to accommodate two opposing comparative statics in κ. First, disclosure can mitigate learning frictions. When bidders are uncertain about θ and receive only censored feedback, they may systematically underestimate the competitiveness of the auction and select multipliers that are too low. We summarize this channel by positing that the average multiplier induced by learning, $\bar{\alpha}(\kappa)$, increases with precision and approaches the full-information best-response level as κ grows. Second, disclosure can enable coordination. When m_t is sufficiently informative, bidders can condition future play on whether observed outcomes match a low-bid "understanding," making punishments credible in a repeated game. We represent this channel by a deviation-detection probability $q(\kappa)$, with $q'(\kappa) > 0$, which enters the incentive constraints for low-revenue perfect public equilibria.

We emphasize that these are not ad hoc effects but two sides of the same informational object: a richer public history helps bidders both *optimize* and *police* one another. The next sections use this structure to (i) establish a competitive benchmark under full information, against which learning-induced underbidding can be measured, and (ii) show how increases in public precision can shift the equilibrium set toward collusive outcomes, implying that platform revenue need not be monotone in transparency.

3 Competitive benchmark (full information): a closed-form multiplier and its revenue–welfare implications

We begin from the polar case in which bidders know the market scale parameter θ (equivalently, they know the value distribution) and there are no learning frictions. In this benchmark the platform's disclosure choice is payoff-irrelevant for the one-shot stage auction: conditional on θ, the period-t game is a standard independent private values first-price auction with risk-neutral bidders. The reason to establish this case is conceptual rather than empirical: it provides (i) a clean target toward which adaptive multiplier policies converge when disclosure improves inference, and (ii) a natural "high-transparency but non-collusive" upper envelope for revenue in the regions of κ where coordination is not sustainable.

Fix a period and suppress the time index. Let $v_i \sim \mathrm{Unif}[0,\theta]$ i.i.d. across i. Consider a symmetric equilibrium in strictly increasing bids $b(\cdot)$ with $b(0) = 0$. If bidder i with value v deviates to bidding as if it had value $\tilde{v}$, it wins whenever its bid exceeds opponents' bids, which under monotonicity is equivalent to $\tilde{v}$ exceeding opponents' values. Thus the winning probability is

$$\Pr(\text{win} \mid \tilde{v}) = F(\tilde{v})^{N-1} = \left(\frac{\tilde{v}}{\theta}\right)^{N-1},$$

and the expected payoff from bidding $b(\tilde{v})$ is

$$\pi(v,\tilde{v}) = \Big(\frac{\tilde{v}}{\theta}\Big)^{N-1}\big(v - b(\tilde{v})\big).$$

In a symmetric equilibrium, $\tilde{v} = v$ maximizes $\pi(v,\tilde{v})$ for each v. Differentiating with respect to $\tilde{v}$ and evaluating at $\tilde{v} = v$ yields the standard first-order condition

$$\frac{N-1}{v}\big(v - b(v)\big) = b'(v),$$

whose solution with boundary condition $b(0) = 0$ is

$$b(v) = \frac{N-1}{N}v.$$

Two features of this expression are central for what follows. First, it is linear, so the equilibrium can be summarized by a single *multiplier* $\alpha^* = (N-1)/N$. Second, α^* depends only on the number of competitors N, not on θ: knowing θ matters for welfare levels and revenue levels, but not for the equilibrium shading rate in this uniform specification.[2]

Because our bidders are constrained to proportional bidding $b_{i,t} = \alpha_{i,t}v_{i,t}$, this benchmark also clarifies what "competitive" means in the restricted action space. If the action grid $\mathcal{A}$ contains α^*, then α^* is the full-information symmetric equilibrium action; if $\alpha^* \notin \mathcal{A}$, then the relevant benchmark is the closest grid point(s) to α^*, and learning converges (in the no-regret sense) to that discretized best response. Either way, the comparative statics we emphasize later can be read as movements in the *average* chosen multiplier $\bar{\alpha}(\kappa)$ toward the competitive shading rate.

Revenue. In a first-price auction the platform's per-period revenue equals the winning bid. Let $v_{(1)} = \max_i v_i$ denote the top order statistic. Under the symmetric equilibrium above, the winning bid is $b(v_{(1)}) = \alpha^* v_{(1)}$, hence

$$R^{\text{comp}}(\theta) = \mathbb{E}\big[b(v_{(1)})\big] = \frac{N-1}{N}\,\mathbb{E}[v_{(1)}].$$

For $v_i \sim \text{Unif}[0,\theta]$, the expected maximum is well known:

$$\mathbb{E}[v_{(1)}] = \frac{N}{N+1}\theta,$$

so per-period expected revenue simplifies to

$$R^{\text{comp}}(\theta) = \frac{N-1}{N+1}\theta.$$

[2]This is a convenient knife-edge of the uniform family. With other distributions, θ-type parameters can enter bidding through hazard rates, but the logic remains: disclosure that helps bidders infer the relevant primitives pushes behavior toward the full-information equilibrium mapping.

Over a horizon of T stationary periods, the platform's expected total revenue is $T\,R^{\text{comp}}(\theta)$ if it does not discount, or $\sum_{t=1}^{T}\delta^{t-1}R^{\text{comp}}(\theta)=\frac{1-\delta^T}{1-\delta}R^{\text{comp}}(\theta)$ if we apply the same discount factor used in bidder utilities. These expressions will serve as a natural yardstick: in the learning-dominated region we can interpret increases in precision as raising revenue toward $R^{\text{comp}}(\theta)$, while in the collusion-dominated region equilibrium selection can drive revenue strictly below it.

Welfare and efficiency. Allocative efficiency is first-best in this benchmark. Since equilibrium bids are strictly increasing in values, the highest-value bidder wins each period. Thus expected per-period welfare (total surplus) equals

$$W^{\text{comp}}(\theta)=\mathbb{E}[v_{(1)}]=\frac{N}{N+1}\theta,$$

and bidder surplus is simply $W^{\text{comp}}(\theta)-R^{\text{comp}}(\theta)=\frac{1}{N+1}\theta$. The welfare comparison will matter later because collusive schemes in repeated games often involve bid suppression and/or market division (e.g., rotating winners), which can break the monotonicity between bids and values and thereby generate misallocation relative to the competitive benchmark.

Interpretation as a pacing/shadow-price analogue. Although we abstract from binding budgets in the baseline, the multiplier form admits a useful interpretation that connects the static benchmark to pacing behavior in practice. Suppose an advertiser faces an additional shadow cost of spend, captured by a Lagrange multiplier $\lambda_i \geq 0$ on expected payments (or, in a dynamic model, by the continuation value of remaining budget). In a reduced-form one-shot problem, the bidder behaves as if paying $(1+\lambda_i)$ per dollar of bid when it wins, so its objective resembles maximizing

$$\Pr(\text{win})\big(v-(1+\lambda_i)b\big).$$

Under the same uniform-IPV logic, the symmetric equilibrium effectively replaces v by $v/(1+\lambda_i)$ in the bid function, yielding

$$b(v;\lambda_i)=\frac{N-1}{N}\cdot\frac{v}{1+\lambda_i},\qquad\text{or}\qquad\alpha(\lambda_i)=\frac{\alpha^*}{1+\lambda_i}.$$

In this sense, the competitive multiplier α^* is the "no-pacing" baseline: any force that makes bidders behave as if their marginal dollar is more expensive—tight budgets, conservative pacing, or (crucially for our purposes) pessimistic beliefs about the competitiveness of the market—appears as a downward distortion in the chosen multiplier. This is precisely why a linear closed-form benchmark is useful here: it lets us translate informational frictions into a single statistic, $\bar{\alpha}(\kappa)$, that can be compared directly to α^* and interpreted through the same lens as a shadow-price wedge.

Taken together, the full-information benchmark fixes a reference point for both revenue and welfare and provides an economically interpretable target for learning dynamics: higher-quality disclosure can push $\bar{\alpha}(\kappa)$ upward toward α^* (reducing "as-if shadow prices" induced by uncertainty), even as the same increase in public informativeness can enlarge the set of sustainable coordination outcomes in the repeated game.

4 Learning-efficiency channel (low transparency): limited feedback and downward-biased multipliers

We now turn to the opposite polar force from collusion: when disclosure precision is low, bidders face a *learning problem* about the effective competitiveness of the environment, and this can depress bids even absent any coordinated conduct. In our setting, low κ means that the public signal $s_t = \theta + \eta_t$ is very noisy and that any auxiliary monitoring statistic m_t (e.g., a noisy winning bid or noisy winning multiplier) is correspondingly uninformative. Each advertiser therefore relies heavily on its own realized experience—whether it won, what it paid, and the value it observed—to update its bidding policy over the finite multiplier grid $\mathcal{A}$.

Why limited feedback pushes multipliers down. A useful way to summarize the learning friction is that each advertiser is trying to infer an unobserved *price pressure* state from censored and noisy data. Even though the full-information competitive multiplier $\alpha^* = (N-1)/N$ is constant in our uniform benchmark, adaptive algorithms in practice do not directly solve for α^*. Instead they learn a mapping from state proxies (recent win rates, realized costs, noisy public signals) to a multiplier choice, and that mapping is typically *conservative* when the state is uncertain. We capture this conservatism as a downward bias in the average chosen multiplier relative to the competitive benchmark.

To make the mechanism concrete, suppose advertiser i maintains a posterior over θ using the signal history $s_{1:t}$. With Gaussian noise, the posterior mean $\hat{\theta}_t = \mathbb{E}[\theta \mid s_{1:t}]$ is a sufficient statistic for a large class of updates, and its posterior variance satisfies

$$\mathrm{Var}(\theta \mid s_{1:t}) \text{ decreases in } \kappa \qquad \text{(more precise disclosure shrinks uncertainty).}$$

Now suppose the bidder's learning algorithm implements a *regularized best response* or "soft" maximization over $\alpha \in \mathcal{A}$ based on estimated payoffs. A parsimonious reduced form is that the chosen multiplier can be written as a smooth function of the bidder's competitiveness estimate,

$$\alpha_{i,t} = g(\hat{\theta}_t),$$

where $g(\cdot)$ is increasing but concave.[3] Then uncertainty about θ lowers the average multiplier by Jensen's inequality:

$$\mathbb{E}[\alpha_{i,t}] = \mathbb{E}[g(\hat{\theta}_t)] \leq g(\mathbb{E}[\hat{\theta}_t]),$$

with strict inequality when $\hat{\theta}_t$ is nondegenerate. As κ rises, $\hat{\theta}_t$ becomes more precise, the dispersion of $g(\hat{\theta}_t)$ shrinks, and $\mathbb{E}[\alpha_{i,t}]$ increases. This provides a microfoundation for our reduced-form object

$$\bar{\alpha}(\kappa) = m(\kappa)\alpha^*, \qquad m(\kappa) \in (0, 1], \qquad m'(\kappa) > 0,$$

where $m(\kappa)$ summarizes how disclosure precision mitigates conservatism in adaptive multiplier choice.

A complementary (and empirically important) source of downward bias comes from *bandit feedback*. When m_t is coarse, bidders often observe essentially only their own realized payoff,

$$u_{i,t} = x_{i,t}\big(v_{i,t} - b_{i,t}\big) = x_{i,t}(1 - \alpha_{i,t})v_{i,t},$$

and do not observe counterfactual outcomes under alternative multipliers. In such settings, exploration of high multipliers is costly because it generates occasional large payments that dominate the realized utility signal; as a result, no-regret algorithms frequently behave as if they have an additional shadow cost of spending, selecting smaller multipliers until sufficient evidence accumulates. Higher κ improves the signal-to-noise ratio of public information and speeds up discrimination between multipliers, reducing the time spent in this low-bid "safe" region.

Revenue monotonicity in the learning-dominated region. Given proportional bids, the platform's per-period revenue is the winning bid. In the learning-dominated region we can treat the outcome as approximately competitive *conditional on* the average multiplier scale $\bar{\alpha}(\kappa)$, so that winning bids are scaled down relative to the full-information benchmark. In particular, under the uniform scaling property and our proportional-bidding restriction, the expected winning value remains $\mathbb{E}[v_{(1)}] = \frac{N}{N+1}\theta$, while the expected winning bid is approximately $\bar{\alpha}(\kappa)\mathbb{E}[v_{(1)}]$. Hence

$$R^{\text{learn}}(\kappa) \approx \bar{\alpha}(\kappa)\,\frac{N}{N+1}\theta = m(\kappa)\alpha^*\frac{N}{N+1}\theta = m(\kappa)\,R^{\text{comp}}(\theta).$$

[3] Concavity is a standard implication of regularization (e.g., entropy-regularized choice rules) and of cautious control rules that penalize overshooting. In practice, multiplicative-weights and bandit algorithms tend to damp extreme actions when payoff estimates are noisy, which is well approximated by concave response in the underlying competitiveness proxy.

Thus, on any range of κ where equilibrium selection is not distorted by collusion, revenue inherits the monotonicity of $m(\kappa)$:

$$\frac{d}{d\kappa}R^{\text{learn}}(\kappa) = m'(\kappa)R^{\text{comp}}(\theta) > 0.$$

Economically, precision raises revenue by correcting the "as-if shadow price" wedge induced by uncertainty and bandit learning: bidders stop behaving as though spending is unusually expensive and move multipliers upward.

Spend and payment volatility. Beyond mean revenue, a second operational implication of the learning-efficiency channel is the stabilization of spending over time. When disclosure is coarse, learning is slow and exploratory, so realized multipliers $\alpha_{i,t}$ exhibit substantial dispersion across advertisers and across time. Because payments equal winning bids, this translates into volatile spend at the platform level and volatile spend at the advertiser level (important in practice for pacing).

A simple decomposition illustrates the comparative statics. Let the winning payment in period t be $P_t = p_{\text{win},t} = \max_i \alpha_{i,t}v_{i,t}$. Write $\alpha_{i,t} = \bar{\alpha}(\kappa) + \varepsilon_{i,t}$, where $\mathbb{E}[\varepsilon_{i,t}] = 0$ and $\text{Var}(\varepsilon_{i,t}) = \sigma_\alpha^2(\kappa)$. Then a first-order approximation around $\bar{\alpha}(\kappa)$ yields

$$P_t \approx \bar{\alpha}(\kappa)\, v_{(1),t} + v_{(1),t}\, \varepsilon_{(1),t},$$

so that, suppressing higher-order order-statistic terms,

$$\text{Var}(P_t) \approx \bar{\alpha}(\kappa)^2\, \text{Var}(v_{(1),t}) \;+\; \mathbb{E}[v_{(1),t}^2]\, \text{Var}(\varepsilon_{(1),t}).$$

Higher κ affects this expression in two opposing but interpretable ways. First, $\bar{\alpha}(\kappa)$ rises, mechanically increasing the scale of payments (and therefore revenue). Second, and crucial for volatility, better disclosure reduces experimentation and cross-sectional disagreement about competitiveness, so $\sigma_\alpha^2(\kappa)$ falls. The second effect tends to dominate the variance term induced by multiplier dispersion, producing *more stable* payments and advertiser spending paths even as average spend rises. In this sense, moderate transparency can simultaneously raise revenue and improve pacing stability, aligning platform and advertiser operational objectives in the low-κ regime.

Summary and transition. In sum, when transparency is low, limited and noisy feedback leads adaptive bidders to choose multipliers below the competitive benchmark, which depresses revenue and typically increases spend volatility through prolonged exploration and heterogeneous beliefs. Increasing disclosure precision κ alleviates these learning frictions, pushing $\bar{\alpha}(\kappa)$ upward and stabilizing spending dynamics. The next section studies why this benign monotonicity need not persist: as κ becomes high enough that the public statistic m_t effectively monitors deviations, the same transparency that helps learning can also support perfect public equilibria with bid suppression and market division.

5 Monitoring-collusion channel (high transparency): perfect public equilibria and enforceability

We now analyze the second, opposing force: when disclosure precision is high, the public statistic m_t can become an effective *monitoring device* that allows bidders to coordinate on low-revenue outcomes and to discipline deviations. The economic point is not that advertisers explicitly "collude" in a legal sense, but that repeated interaction combined with sufficiently informative common signals can support *self-enforcing bid suppression* as a perfect public equilibrium (PPE). In modern ad auctions, where bidders are represented by adaptive bidding systems that observe similar dashboards and market reports, the relevant notion of coordination is often algorithmic and tacit: bidders can condition on shared observables and on predictable responses to those observables.

A canonical collusive scheme: bid rotation with suppressed multipliers. To make the mechanism concrete, consider a stationary, symmetric scheme in which bidders take turns winning the slot. Fix a low multiplier $\alpha^{\mathrm{L}} \in \mathcal{A}$ (with $\alpha^{\mathrm{L}} < \alpha^*$) and a public rotation rule (e.g., bidder i is the designated winner in periods $t \equiv i \pmod N$). On the collusive path, the designated winner bids $b_{i,t} = \alpha^{\mathrm{L}} v_{i,t}$ while all others bid an even lower multiplier (e.g., $\alpha^{\mathrm{L}} - \epsilon$ on the grid, or 0 if allowed), ensuring the designated bidder wins at a suppressed price. This scheme is attractive because it replaces competitive bidding with a predictable allocation of wins and a low payment level.

Such a path generally sacrifices allocative efficiency: the slot is not necessarily assigned to the highest-value bidder in each period, but rather to the designated winner. This is precisely why the scheme can reduce platform revenue and, depending on the environment, can also reduce welfare W.

Why monitoring matters: deviations are only deterred if they are detected. The obstacle to sustaining the scheme is the one-shot temptation to deviate. A non-designated bidder can raise its multiplier in a period when it is supposed to lose, capture the slot, and earn a positive surplus. The only force that can deter this is the prospect of future punishment *triggered by public evidence* of deviation. In a PPE, strategies must be measurable with respect to the public history, so punishments can only condition on what is publicly observed. If m_t is too noisy, a deviator can frequently "steal" wins without generating a detectable public trace.

We capture this by summarizing the monitoring environment with a deviation-detection probability $q(\kappa) \in [0, 1]$, increasing in precision κ. Intuitively, as the platform releases more granular or less noisy information about the winning bid, the winning multiplier, rank statistics, or spend shares, it becomes easier for rivals to infer whether the realized outcome is consistent with the prescribed low-bid rotation.

A simple Gaussian microfoundation for $q(\kappa)$. One convenient specification is that the platform releases a noisy public proxy for the winning multiplier,

$$m_t \;=\; \alpha_{\text{win},t} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_m^2), \qquad \kappa = \sigma_m^{-2}.$$

Under the collusive scheme, $\alpha_{\text{win},t} = \alpha^{\text{L}}$ deterministically (up to grid effects). A profitable deviation in a "losing" period entails bidding a higher multiplier, say $\alpha^{\text{D}} > \alpha^{\text{L}}$, and winning with positive probability; conditional on winning, the public signal shifts upward by $\Delta = \alpha^{\text{D}} - \alpha^{\text{L}}$. A natural public test is to flag a deviation when m_t exceeds a threshold τ chosen so that false positives are rare on the collusive path. Under this test, the detection probability conditional on a deviation-induced win is

$$q(\kappa) = \Pr\big(m_t > \tau \,\big|\, \alpha_{\text{win},t} = \alpha^{\text{D}}\big) = 1 - \Phi\left(\frac{\tau - \alpha^{\text{D}}}{\sigma_m}\right),$$

which is strictly increasing in κ because σ_m decreases. This illustrates the general point: even if deviations are not perfectly observed, increasing precision raises the likelihood that the public history contains enough evidence to coordinate punishments.

Public punishments and the PPE incentive constraint. Given any prescribed collusive path and any public punishment (e.g., grim-trigger reversion to competitive play, or a harsh phase where all bidders bid aggressively), sustainability is characterized by a familiar one-shot deviation constraint. Let u^{col} denote the on-path per-period expected utility for a representative bidder, u^{dev} the deviation payoff in the deviation period (optimizing over multipliers, taking others as collusive), and u^{pun} the expected per-period payoff during the punishment phase. Let V^{col} and V^{pun} be the corresponding continuation values. If a deviation is detected with probability $q(\kappa)$, the PPE constraint takes the form

$$(1-\delta)u^{\text{col}} + \delta V^{\text{col}} \;\geq\; (1-\delta)u^{\text{dev}} + \delta\Big((1-q(\kappa))V^{\text{col}} + q(\kappa)V^{\text{pun}}\Big),$$

equivalently,

$$\delta \;\geq\; \frac{u^{\text{dev}} - u^{\text{col}}}{q(\kappa)\,\big(V^{\text{col}} - V^{\text{pun}}\big)}.$$

This inequality makes transparent how monitoring and patience interact. Holding primitives fixed, higher precision raises $q(\kappa)$, which relaxes the incentive constraint and expands the set of (δ, κ) pairs for which collusion can be supported. Conversely, for any fixed δ, there is a threshold κ_c (possibly infinite) such that collusion is infeasible for $\kappa < \kappa_c$ and feasible for $\kappa \geq \kappa_c$. This is the monitoring-collusion channel: transparency does not merely inform learning; it also supplies the public correlation device needed to sustain low-revenue coordination.

Revenue implications at high κ. When a low-bid rotation (or a bid cap) is sustainable, platform revenue falls because winning bids are mechanically reduced relative to the competitive benchmark. Under proportional bidding, the winning payment is $P_t = \alpha_{\text{win},t} v_{\text{win},t}$, and collusion directly lowers the multiplier component $\alpha_{\text{win},t}$; under rotation it may also lower the value component by inducing misallocation. Thus, for sufficiently high κ, further increases in precision can *reduce* revenue by strengthening enforceability (raising $q(\kappa)$), even though those same increases continue to improve the learning environment.

Interpretation and design implications. This channel maps closely to practical disclosure choices: releasing bidder-level win reports, near-real-time clearing-price estimates, or fine-grained rank statistics can make it easier for bidders to infer who "stole" an impression and to condition future behavior accordingly. By contrast, aggregation, anonymization, or delay can reduce the effective $q(\kappa)$ even if bidders still learn θ reasonably well from coarser market signals. Our stylized PPE analysis therefore suggests a design tension: the platform may want to share enough information to mitigate conservative learning, but not so much that it creates a high-powered public monitoring technology.

Limitations. We emphasize that this is a reduced-form enforcement logic, not a claim that real markets literally implement grim-trigger strategies. In algorithmic environments, coordination can arise through shared features, common optimization objectives, and similar learning rules; the relevant question is whether public observables are sufficiently informative to make reactive strategies (or learned policies) stable. Moreover, equilibrium selection is crucial: high κ expands the set of feasible outcomes, but does not dictate which outcome is selected. In the next section we combine these two channels to show how an interior-optimal disclosure policy emerges once we allow transparency to simultaneously reduce learning frictions and increase the scope for coordinated bid suppression.

6 Main result: non-monotone disclosure incentives and an interior optimum

We can now combine the learning-efficiency channel (which pushes bids *up* as disclosure becomes more informative) with the monitoring-collusion channel (which can push bids *down* once public signals become sufficiently disciplining). The key implication is a *non-monotone* mapping from transparency to the platform objective: increasing precision is beneficial when it mainly corrects conservative learning, but harmful when it mainly supplies a high-powered public monitoring technology.

To state the result cleanly, it is useful to work with a reduced-form revenue decomposition that captures the equilibrium-selection switch discussed above. For κ below a critical precision level κ_c, collusive PPEs are not enforceable, and outcomes track the learning-dominated "competitive" branch. In this region, expected revenue inherits the monotonicity of the average multiplier,

$$R^{\text{learn}}(\kappa) \;=\; m(\kappa)\, R^{\text{comp}}(\theta), \qquad m'(\kappa) > 0,$$

so that $R^{\text{learn}\prime}(\kappa) > 0$. For κ above κ_c, low-revenue coordinated outcomes become feasible, and an equilibrium selection (or learning dynamics) may place positive weight on bid-suppression schemes with revenue $R^{\text{col}}(\kappa)$. When monitoring precision improves deviation detection, $q'(\kappa) > 0$, the enforceability of these schemes increases; in many natural constructions (e.g., rotation with threshold tests), this manifests as weakly lower selected revenue as κ rises further, i.e. $R^{\text{col}\prime}(\kappa) \le 0$.

A compact way to represent this logic is the piecewise objective

$$R(\kappa) = \begin{cases} R^{\text{learn}}(\kappa) & \kappa < \kappa_c, \\ \min\bigl\{R^{\text{learn}}(\kappa),\, R^{\text{col}}(\kappa)\bigr\} & \kappa \ge \kappa_c, \end{cases}$$

together with a strict "gap" at the onset of enforceability, $R^{\text{col}}(\kappa_c) < R^{\text{learn}}(\kappa_c)$. This gap is economically intuitive: the very first moment at which bidders can credibly punish deviations, the platform becomes exposed to equilibria that mechanically depress the winning multiplier (and, under rotation/market division, may also depress the realized winning value).

Proposition (interior-optimal disclosure). Suppose (i) $R^{\text{learn}\prime}(\kappa) > 0$ on $[0, \kappa_c)$; (ii) $R^{\text{col}}(\kappa_c) < R^{\text{learn}}(\kappa_c)$ and $R^{\text{col}\prime}(\kappa) \le 0$ on $[\kappa_c, \infty)$; and (iii) R^{learn} is continuous. Then platform revenue $R(\kappa)$ is strictly non-monotone and admits an interior maximizer $\kappa^* \in (0, \kappa_c)$.

Proof sketch. On $[0, \kappa_c)$, revenue is strictly increasing, so any maximizer must lie weakly above κ_c *if* the branch does not switch. But at κ_c the selected outcome weakly drops to $R^{\text{col}}(\kappa_c)$, which is strictly below the limit from the left. Hence $\kappa = \kappa_c$ cannot be optimal, and no $\kappa > \kappa_c$ can dominate the best point just below κ_c when R^{col} is nonincreasing. Therefore the maximizer lies strictly below κ_c.

Welfare and mixed objectives. The same logic extends to a revenue–welfare objective $J(\kappa) = \lambda R(\kappa) + (1-\lambda)W(\kappa)$. Under competitive play, welfare is first-best because the highest value wins each period; under rotation or market division, welfare typically falls due to misallocation. Thus for sufficiently high precision, transparency can simultaneously reduce revenue and welfare by supporting coordination on allocations that ignore realized values. When $\lambda < 1$, the welfare loss makes the "high-κ" region even less attractive, strengthening the case for an interior optimum. Importantly, our conclusion does not require that *all* high-precision equilibria are collusive—only that the set expands and that plausible equilibrium selection (including algorithmic learning dynamics) can place weight on bid-suppression outcomes once they become stable.

Comparative statics: concentration, patience, and heterogeneity. The location of the interior optimum is governed by where the enforceability threshold κ_c sits relative to the learning gains from raising κ.

First, greater market concentration (smaller N) tends to *lower* κ_c and thereby *lower* the optimal precision κ^*. With fewer bidders, coordinating on a rotation is simpler, and the temptation to deviate can be punished more effectively because the continuation value from future assigned wins is large relative to the one-shot gain. Formally, in the PPE constraint

$$\delta \;\geq\; \frac{u^{\text{dev}} - u^{\text{col}}}{q(\kappa)\left(V^{\text{col}} - V^{\text{pun}}\right)},$$

a smaller N often increases $(V^{\text{col}} - V^{\text{pun}})$ (collusion promises larger future rents per bidder), so a lower detection probability $q(\kappa)$ suffices, implying a smaller κ_c.

Second, greater bidder patience (higher δ) also lowers κ_c and thus lowers κ^*. Patience magnifies the disciplining effect of any given monitoring signal, because future punishments loom larger. In environments where the platform cannot perfectly decouple learning-relevant disclosure from monitoring-relevant disclosure, this comparative static implies that highly repeated markets (stable demand, long-lived advertisers, or high-frequency auctions) warrant *less* granular public reporting.

Third, greater heterogeneity in values, budgets, or objectives tends to *raise* κ_c and thereby *raise* κ^*. Asymmetries increase the incentive to deviate from a symmetric market-division plan (high types want more share; constrained types cannot sustain prescribed actions), and they make punishments less credible or less coordinated. In the PPE constraint, heterogeneity can raise $u^{\text{dev}} - u^{\text{col}}$ and/or reduce the effective continuation gap $V^{\text{col}} - V^{\text{pun}}$, so higher $q(\kappa)$ (hence higher κ) is required to support coordination. From a design perspective, markets with diverse advertisers are therefore naturally "self-protecting" against collusion, permitting the platform to share more information to alleviate learning frictions.

Extensions: binding budgets and target-CPA objectives. While budgets are nonbinding in our baseline, binding budgets introduce a second intertemporal linkage that interacts with disclosure. On the pro-transparency side, better information can improve pacing and reduce conservative under-delivery, raising effective multipliers and hence revenue in the low-κ region. On the anti-transparency side, disclosure of spend shares, win rates, or (especially) bidder-level budget exhaustion can act as an additional monitoring device: it helps bidders infer whether a rival "stole" share or violated a market-division plan, effectively increasing $q(\kappa)$ even if bid data remain noisy. Moreover, budgets can make punishments harsher (a punished bidder may be forced to spend inefficiently), which can further expand enforceable collusion. These forces reinforce the central non-monotonicity: transparency that is useful for pacing can simultaneously strengthen coordination.

A related extension is target-CPA bidding, where realized CPA enters payoffs through a penalty (as in U_i^{CPA}). Such objectives can change both channels. They can amplify learning benefits (because bidders must infer competitive pressure to hit CPA targets without overbidding), but they can also alter deviation incentives: aggressive deviations that win more often may worsen realized CPA and reduce the effective gain from deviating. In other words, CPA constraints can *stabilize* low-bid coordination by shrinking $u^{\text{dev}} - u^{\text{col}}$, again pushing κ_c downward. This suggests that policy or product changes that standardize bidder objectives (e.g., widespread target-CPA automation) may increase the importance of limiting high-frequency, bidder-identifying monitoring disclosures.

These results motivate an empirical design that varies disclosure regimes in a controlled way and measures both learning-driven bid adjustment and the emergence of stable bid-suppression patterns, which we implement next via explicit logging and observation-policy changes in AuctionNet.

7 AuctionNet empirical design: implementing disclosure regimes and measuring coordination

Our empirical strategy uses AuctionNet as a controlled "information design" testbed: we hold fixed the allocation and payment rule (a repeated first-price single-slot auction unless otherwise stated), and vary only what the platform reveals through the observation and logging interfaces. This mirrors the theoretical object κ in a way that is both implementable and auditable. Concretely, we implement a family of disclosure regimes as *observation policies* that map the platform's internal state at time t into (i) bidder-facing observations and (ii) researcher-facing logs. Bidder observations are the sole inputs to bidding algorithms; researcher logs are never observed by bidders and serve only for ex post measurement. This separation is important: it lets us study how transparency affects strategic and learning dynamics without inadvertently changing bidders' action space or the auction outcome function.

Disclosure regimes via observation/log APIs. We parameterize disclosure along two dimensions. First, *precision* is controlled by adding Gaussian noise with variance $\sigma^2 = 1/\kappa$ to scalar statistics, matching the signal structure $s_t = \theta + \eta_t$. Second, *granularity* controls which statistic is released (and at what temporal resolution), ranging from minimal bandit-like feedback to near-public monitoring. In our baseline family, bidders always observe their own realized value $v_{i,t}$, budget state $B_{i,t}$ (if budgets bind), and own outcome $(x_{i,t}, p_{i,t})$. We then layer on public components (s_t, m_t) that vary by regime: (i) *Low disclosure*: bidders observe only $(x_{i,t}, p_{i,t})$ and a coarse aggregate s_t about θ with low κ (high noise), yielding highly censored feedback about rivals. (ii) *Aggregate price statistics*: release a noisy winning bid (or noisy top bid) as m_t, optionally delayed by d periods to reduce high-frequency monitoring. (iii) *Rank/quantile statistics*: release noisy rank information (e.g., "you were top-k") or anonymized quantiles of bids, which can improve learning about competitiveness while limiting bidder-identifying monitoring. (iv) *High disclosure*: release a near-public monitoring signal such as noisy realized winning multiplier, noisy spend shares, or bidder-level win indicators (an anonymized identity code), which sharply increases the scope for contingent punishments. In all cases, the auction rule is unchanged; only the observation map changes. To ensure comparisons are not driven by different randomness, we use common random numbers: for each experimental replicate, the full sequence of value draws $\{v_{i,t}\}$ and signal noise draws $\{\eta_t\}$ is held fixed across disclosure regimes.

Standardizing bidders: common action grids and learning budgets. To isolate the effect of disclosure, we standardize bidder algorithms along the dimensions that most directly interact with the theory: the action space $\mathcal{A}$ of multipliers and the feedback structure. All bidders choose $\alpha_{i,t} \in \mathcal{A} \subset (0, 1]$ and bid $b_{i,t} = \alpha_{i,t} v_{i,t}$, with budget feasibility enforced by truncation when budgets bind. The default grid $\mathcal{A}$ is chosen to contain the competitive benchmark $\alpha^* = (N-1)/N$ and a range of lower multipliers that can support bid suppression. We then evaluate several algorithm families, each trained under identical compute and data budgets: (i) a *pacing+bandit* baseline that maintains a pacing multiplier (for budgets) and a no-regret bandit update over $\mathcal{A}$ using own payoff feedback; (ii) *independent Q-learning (IQL)* over discretized states that include recent public signals and own outcomes; (iii) *behavior cloning (BC)* and *decision transformer (DT)* variants trained on offline trajectories generated under a mixture of disclosure regimes, then deployed online under a fixed regime. Across all families, we enforce symmetry ex ante (same architecture class, same hyperparameter search budget, same random seed schedule) to avoid mechanically creating "leader–follower" asymmetries. When we introduce heterogeneity (e.g., different budgets or value scalings), we do so explicitly and record it as a treatment dimension.

Operationalizing "collusion" in a learning environment. Because algorithms need not converge to a stationary equilibrium and may coordinate implicitly, we define collusion through a set of outcome-based and behavior-based metrics that are meaningful in repeated auctions. The central outcome metric is a *revenue suppression index* relative to the competitive benchmark:

$$\mathrm{RSI}(\kappa) \;=\; 1 - \frac{\widehat{R}(\kappa)}{\widehat{R}^{\mathrm{comp}}},$$

where $\widehat{R}^{\mathrm{comp}}$ is computed either from known θ in simulation or from a high-disclosure calibration run where learning frictions are minimal. To distinguish "benign" low revenue driven by underbidding during learning from coordinated suppression, we also measure whether allocation departs from value-ranking. Let $v_{(1),t}$ denote the highest value at t and $v_{\mathrm{win},t}$ the winning value; then a welfare loss proxy is

$$\mathrm{Misalloc}(\kappa) \;=\; \frac{1}{T}\sum_{t=1}^{T} \big(v_{(1),t} - v_{\mathrm{win},t}\big),$$

which is near zero under competitive play but positive under rotation or market division.

Behaviorally, we track *rotation* and *market-division* signatures in the sequences $\{x_{i,t}\}$ and $\{\alpha_{i,t}\}$. Rotation implies that wins cycle across bidders with low within-bidder variance in multipliers; we quantify it by (i) the entropy of the win-share vector, (ii) autocorrelation and spectral mass of each bidder's win indicator, and (iii) a "turn-taking" statistic measuring how often the identity of the winner changes conditional on similar realized values. Market division implies persistent identity-based specialization: bidder i wins disproportionately in a subset of contexts even after controlling for values. We measure this by fitting a predictive model for $x_{i,t}$ using realized values and public signals, and testing whether adding bidder identity materially increases predictive power; large incremental fit indicates identity-based allocation beyond value-based competition.

Stability, punishments, and empirical detection probability. A key theoretical object is deviation detection. Empirically, we capture its analogue in two ways. First, we compute *stability* of the action profile as the within-window variance of multipliers and the rate of switching across $\mathcal{A}$; collusive coordination typically exhibits low variance punctuated by abrupt "punishment" episodes. Second, we run controlled deviation probes: in a small fraction of periods (unknown to the deviator's policy but known to the experimenter), we override one bidder's action with an aggressive multiplier (e.g., the upper tail of $\mathcal{A}$) and record whether subsequent public histories trigger statistically detectable punishment responses by rivals (e.g., sustained increases in their multipliers or targeted win-denial behavior). The frequency with which these probes lead to a punishment classification provides an empirical counterpart to $q(\kappa)$, and allows us to connect disclosure regimes to enforceability without assuming equilibrium play.

Implementation details and limitations. All metrics are computed on held-out evaluation horizons after a training burn-in, and we report uncertainty using across-seed variation and block bootstrap over time to account for serial correlation. We emphasize that our empirical definitions are intentionally conservative: we label a run as "collusive" only when revenue suppression coincides with either misallocation or strong rotation/identity signatures and high stability. This reduces false positives in regimes where learning is slow. At the same time, we acknowledge a limitation: in algorithmic environments, coordination can be subtle and may not resemble textbook grim-trigger behavior. Our multi-metric approach is designed to detect a broad class of stable bid-suppression patterns while remaining interpretable for platform policy, which we then evaluate in the results section by tracing revenue–welfare curves across disclosure levels.

8 Experimental results: revenue–welfare curves by disclosure level

We now turn to the central empirical object predicted by the model: how platform outcomes vary with disclosure precision κ once bidders are adaptive and can condition on public histories. Our main outputs are revenue $\widehat{R}(\kappa)$, welfare $\widehat{W}(\kappa)$, and the derived diagnostics $\mathrm{RSI}(\kappa)$ and $\mathrm{Misalloc}(\kappa)$ defined above. We report results after a burn-in phase, and we average across seeds with common random numbers so that differences across disclosure regimes are not mechanically driven by different value realizations.

Revenue–welfare curves are non-monotone in precision. Across algorithm families (bandit pacing, IQL, and offline-trained DT/BC policies), we observe a robust *inverted-U* pattern in revenue as a function of disclosure precision. At low κ (high noise / weak monitoring), revenue is depressed relative to the competitive benchmark; the revenue suppression index $\mathrm{RSI}(\kappa)$ is high even though our behavioral collusion indicators are largely absent. As κ rises from this region, $\widehat{R}(\kappa)$ increases sharply while $\mathrm{Misalloc}(\kappa)$ remains near zero. This is the learning-efficiency channel: with more informative (s_t, m_t), bidders more quickly infer competitiveness and select larger multipliers, pushing outcomes toward the competitive shading benchmark $\alpha^* = (N-1)/N$.

Beyond an intermediate precision range, however, revenue peaks and then declines as κ continues to increase. Importantly, the revenue decline is accompanied by *increasing* misallocation and increasingly strong rotation/market-division signatures. In this high-κ region, the joint outcome pattern is inconsistent with "benign" learning dynamics (which would predict low revenue *and* competitive allocation failures to vanish over time as learners improve). Instead, we see persistent revenue suppression that co-moves with stable, history-dependent behavior. In the language of the theory, the data trace the switch from the learning-dominated branch $R^{\text{learn}}(\kappa)$ to a collusion-selected branch $R^{\text{col}}(\kappa)$ once monitoring becomes sufficiently informative.

For welfare, we see a corresponding tradeoff. In intermediate regimes, welfare is essentially first-best because the allocation remains approximately value-ranked. At high κ, the welfare curve bends downward: average welfare falls and $\text{Misalloc}(\kappa)$ becomes economically meaningful, consistent with rotation or market division that occasionally allocates the slot to a bidder who does not have the highest realized value.

Separating learning from coordination using dynamics. To distinguish underbidding due to slow learning from coordination sustained by monitoring, we examine the time profile of outcomes. Under low disclosure, revenue is lowest early in the horizon and rises gradually; multipliers exhibit high switching rates across $\mathcal{A}$, and stability metrics indicate ongoing exploration. This is the signature of informational frictions. Under high disclosure, by contrast, revenue suppression emerges *after* a short transient, following which multipliers become extremely stable within long windows. When deviations occur (either naturally through exploration noise or through our controlled probes), the response is episodic: rivals temporarily increase aggressiveness or adopt exclusionary patterns, and then return to the low-multiplier regime. This "punishment-and-return" behavior is difficult to explain through independent myopic learning but is natural under repeated-game logic with public histories.

These dynamics map cleanly into the detection-probability object $q(\kappa)$. Empirically, when we inject a deviation probe (forcing one bidder to play a high multiplier for a short burst), we can classify subsequent histories as "punishment" episodes with high probability under high κ and low probability under low κ. Thus, the *empirical* detection frequency is increasing in precision, providing direct support for the monitoring-collusion channel.

Evidence for rotation and market division. Our behavioral metrics identify two dominant coordination modes under high disclosure. First, in many runs we see *rotation*: win shares are close to equal, the identity of the winner changes frequently even conditional on similar realized values, and each bidder's $\{x_{i,t}\}$ exhibits strong periodic structure. Rotation is accompanied by low within-bidder variance in $\alpha_{i,t}$ and a small set of frequently revisited multipliers, consistent with a tacitly coordinated "low bid, take turns" pattern.

Second, in other runs we see *market division*: win probabilities become strongly bidder-identity dependent even after controlling for $v_{i,t}$ and public signals. In these cases, the incremental predictive power of bidder identity for $x_{i,t}$ is large, and the allocation deviates from value ranking in a systematic way. Market division becomes especially salient when we introduce mild heterogeneity (e.g., slightly different budgets or value scalings): rather than eliminating coordination, heterogeneity often shifts the form of coordination from symmetric rotation to asymmetric specialization.

Comparative statics in the laboratory: N and patience. While our simulations are not a structural estimation exercise, the comparative statics align with the theory. Holding algorithms fixed, the onset of coordination (as measured by the joint event of high RSI(κ), low switching, and high punishment classification) occurs at *lower* precision when N is smaller. Likewise, when we increase effective patience in the training objective (either via a higher discount factor in RL-style updates or via longer-horizon training with less restart noise), coordination becomes more prevalent at a given κ. Both patterns are consistent with the prediction that the "collusion-feasible" region expands with concentration and patience, shifting the revenue-maximizing precision downward.

Robustness across auction formats: first-price versus GSP-like rules. We replicate the disclosure treatments under two alternative auction environments. First, we implement a GSP-style payment rule (winner pays the second-highest bid, with suitable tie-breaking), keeping the same observation policies. The qualitative transparency tradeoff persists: increasing κ initially improves revenue through better learning of competition, but high κ again enables coordination that suppresses effective prices. The behavioral manifestation differs slightly: instead of suppressing the winning bid directly (as in first price), bidders often coordinate on keeping the *runner-up* bid low, which is the pivotal price in GSP. Accordingly, under high disclosure the gap between the top two bids widens, and the second-order statistic (the "price-setting" bid) becomes unusually stable.

Second, we consider a hybrid "value-based" bidding interface where bidders choose multipliers but the platform computes scores that mimic quality adjustment. The same patterns emerge provided that the disclosed monitoring statistic is sufficiently aligned with the payoff-relevant object (e.g., revealing a noisy score or rank). This reinforces a practical lesson: it is the informativeness of the monitoring signal about rivals' strategic state, not the specific payment formula, that governs the feasibility of contingent punishments.

Robustness to multi-slot structures. Finally, we extend to multiple slots with position effects and either (i) independent first-price per slot or (ii) a standard separable click model with rank-by-bid allocation. Multi-slot environments amplify the scope for market division: bidders can implicitly "claim" particular positions (or ranges of positions), yielding misallocation both across bidders and across slots. Revenue non-monotonicity remains, but the welfare losses at high κ are larger because coordination can distort the entire rank ordering rather than only the top allocation. These results motivate our governance discussion: fine-grained, high-frequency disclosure that reveals rank, price, and identity information can unintentionally provide exactly the public monitoring needed for sustained coordination in rich allocation environments.

8.1 Implications for platform governance and antitrust: disclosure design and auditing

Our results place platform "transparency" in a different category from the usual market-design levers (reserve prices, scoring rules, or auction format). In the environment we study, disclosure is itself a strategic instrument: increasing precision κ simultaneously (i) alleviates learning frictions, raising bids through $m(\kappa) \uparrow 1$, and (ii) strengthens public monitoring, raising deviation detection $q(\kappa)$ and thereby enlarging the set of sustainable low-revenue perfect public equilibria. The governance implication is that a platform cannot treat richer reporting as unambiguously pro-competitive. What matters is not only how disclosure helps an individual bidder optimize, but also how it creates *common knowledge* of rivals' strategic states, which is precisely the input needed for contingent punishments.

Design principle 1: decouple "market-scale learning" from "rival-action monitoring." A practical reading of the model is that the platform often wants to provide information about fundamentals (e.g., the market scale θ, seasonality, demand shifts) without providing a high-frequency public record of rivals' bid multipliers. In our notation, this corresponds to choosing high effective precision for s_t (to push $m(\kappa)$ upward) while keeping the monitoring statistic m_t deliberately coarse so that $q(\kappa)$ remains low. Concretely, this suggests that "transparency" should be implemented as *fundamental guidance* (forecast ranges, aggregate demand indices, or lagged clearing-price bands) rather than as *action disclosure* (who bid what, who won, at what multiplier, and with what remaining budget).

Design principle 2: treat identity and timing as first-order policy levers. Repeated-game coordination typically needs a mapping from deviations to punishments, which in turn requires stable identifiers and timely, payoff-relevant public signals. Two low-cost levers therefore matter disproportionately: (i) *anonymization* (removing bidder identities from public reports) and (ii) *delay* (publishing statistics with lags long enough to break period-by-period contingencies). In the language of our detection object, both reduce the effectiveness of punishments by lowering the probability that a deviation is detected *and* attributed in time to respond, thereby reducing $q(\kappa)$ even when the underlying data are eventually released.

Recommended disclosure menu. A governance takeaway is not "be opaque" but "choose the coarsest disclosure that solves the intended learning or budgeting problem." In practice we can implement this through a menu of design choices:

- **Noise / coarsening (controlled κ):** release price or competition indicators only in bins (e.g., deciles) or with calibrated noise. From our perspective, the goal is to preserve enough signal to improve estimates of θ while preventing precise inference of rivals' $\alpha_{j,t}$.
- **Delay and batching:** publish aggregates weekly (or over random-length windows) rather than per auction. Batching preserves long-run guidance while removing the tight linkage between a deviation today and a punishment tomorrow.
- **Aggregation across segments:** aggregate over sufficiently broad sets of queries or audiences. Fine segment-level transparency can inadvertently create many small "repeated games" where rotation is easier and more stable.
- **Anonymization and suppression of rank identities:** avoid public reports that let bidders reconstruct who is typically the runner-up, who is pacing out, or whose budget is binding. Even when payments are not disclosed, rank and identity information can make punishments targeted and therefore credible.
- **Private rather than public feedback:** provide each bidder rich *private* diagnostics (own spend, win rate, marginal returns), while keeping cross-bidder comparables public only in coarse form. This shifts learning from public monitoring to idiosyncratic optimization.
- **Randomized rounding / throttling of public statistics:** when publishing a statistic like "top bid" or "clearing multiplier," apply stochastic rounding so that small deviations do not deterministically change

the public signal. This directly disrupts trigger strategies that hinge on thresholds.

These interventions are complements: for example, anonymization without delay may still permit inference via stable behavioral fingerprints, while delay without coarsening may still permit coordination at the batch frequency.

Antitrust framing: disclosure as a potential facilitating practice. Antitrust doctrine often focuses on explicit communication among rivals, but information design by a central intermediary can create a functionally equivalent coordination channel. In our setting, high-precision m_t resembles an "information exchange" mechanism: it supplies a shared scoreboard that makes deviations observable and punishable. This is especially salient in concentrated markets (small N) or where bidders are effectively patient (high δ), precisely the environments where our comparative statics predict a lower optimal κ^*. A conservative governance stance is therefore to treat fine-grained, high-frequency, identity-linked reporting as presumptively risky in concentrated verticals, even when the platform's intent is benign (e.g., "help advertisers optimize").

Auditing protocol: measuring $q(\kappa)$ and separating collusion from slow learning. Because the same low-revenue outcome can arise from either channel, governance requires audits that distinguish "underbidding due to uncertainty" from "suppression sustained by monitoring." We recommend that platforms (and, where relevant, regulators) institutionalize three layers of diagnostics.

(i) Outcome diagnostics. Track revenue and welfare jointly, not revenue alone. Persistent revenue suppression coupled with measurable misallocation is a red flag because learning improvements should tend to restore value-ranking, whereas rotation/market division can depress both revenue and allocation quality. In our empirical vocabulary, sustained elevation in $\mathrm{RSI}(\kappa)$ together with increases in $\mathrm{Misalloc}(\kappa)$ is more consistent with coordination than with benign learning.

(ii) Behavior diagnostics. Monitor switching rates over $\mathcal{A}$, periodicity in win shares, and identity-predictability of allocations. Collusive regimes typically exhibit low within-bidder variance in $\alpha_{i,t}$, long stretches of stability, and sharp, episodic responses that resemble punishments rather than exploration.

(iii) Probe-based detection tests. To directly estimate an operational analogue of $q(\kappa)$, platforms can run controlled "deviation probes" (e.g., temporarily perturb the bidding interface for a small, consented subset, or inject small randomized shocks into rankings) and measure whether the public-history response resembles a punishment episode. Importantly, such probes should be designed to avoid harming auction outcomes materially and should

be auditable ex post (logged, pre-registered thresholds, and clear governance over when probes are permitted).

Governance process: dynamic disclosure with guardrails. Finally, because bidder algorithms evolve, the safe level of disclosure today need not be safe tomorrow. We therefore view disclosure as a policy that should be revisited under a formal change-management process: before increasing precision or adding new statistics, the platform should run an "information impact assessment" that stress-tests candidate disclosures in simulation and, when feasible, in small-scale randomized experiments with predefined stopping rules. A practical guardrail is to require that any change that materially increases the informativeness of m_t (hence plausibly increases $q(\kappa)$) be paired with compensating reductions in identifiability or timeliness, keeping the system on the learning-improving side of the tradeoff rather than crossing into the collusion-feasible region.

8.2 Limitations and extensions: heterogeneity, entry, multi-slot environments, and privacy constraints

Our framework is deliberately stylized in order to isolate a single economic mechanism: disclosure precision κ simultaneously improves *individual learning* (pushing behavior toward the competitive benchmark) and improves *public monitoring* (making coordinated punishments feasible). The simplicity that yields clear comparative statics also creates limitations that matter for interpretation and for policy translation. We view these limitations not as defects, but as a map of where the tradeoff is likely to become sharper, and where it may be attenuated by additional forces.

Richer bidder heterogeneity. The baseline assumes i.i.d. values $v_{i,t} \sim \mathrm{Unif}[0, \theta]$ and (in the benchmark) nonbinding budgets, which together deliver a convenient scalar sufficient statistic for competitive shading, $\alpha^* = (N - 1)/N$. Real ad markets exhibit heterogeneity in at least four dimensions: (i) value distributions (different conversion rates and margins), (ii) budget processes (daily pacing constraints and seasonality), (iii) objectives (e.g., value maximization versus target CPA), and (iv) algorithmic sophistication (different learning rates and exploration policies). Each dimension changes the relative strength of our two channels.

On the learning side, heterogeneity can *increase* the value of informative signals by reducing the portability of one's own experience: with asymmetric types, a bidder's own win/loss feedback may be a poor guide to the competitive pressure it faces in states where its value distribution has little overlap with rivals'. On the collusion side, asymmetry tends to *destabilize* simple rotation or bid-suppression schemes because the temptation to deviate is larger for high-value or deep-pocket bidders. A promising extension is therefore to endogenize $u^{\mathrm{dev}} - u^{\mathrm{col}}$ and $V^{\mathrm{col}} - V^{\mathrm{pun}}$ as functions of type dispersion, rather than treating them as reduced-form constants in the incentive constraint. Likewise, when budgets bind, the monitoring statistic m_t may reveal budget exhaustion and thereby enable more sophisticated intertemporal coordination (e.g., "let rival j spend out today, punish tomorrow"); conversely, rapidly fluctuating budgets can break the predictability needed for targeted punishment. Understanding when binding budgets amplify versus dampen the collusion channel is, in our view, a first-order open question.

Endogenous entry and participation dynamics. We take N as fixed, but in many settings participation is itself a dynamic decision: advertisers enter and exit based on realized performance, and platforms may throttle participation through quality scores, targeting constraints, or minimum spend requirements. Endogenous entry can interact with disclosure in two opposing ways. More informative signals about θ (or about competition) can attract entry by reducing uncertainty, potentially raising revenue even if incumbent bidders coordinate. At the same time, transparency that improves monitoring among incumbents can deter entry if entrants anticipate that deviations will be punished or that market division will prevent them from gaining scale.

A natural extension is a two-stage dynamic model: each period begins with a participation decision (or a stochastic arrival process) followed by the auction. In such an environment, κ affects not only per-period bidding but also the evolution of the effective N_t and, therefore, the long-run competitiveness of the marketplace. This extension would also allow one to study whether the platform can use disclosure as a *screening device*, encouraging entry by small bidders while limiting the common knowledge needed for incumbent coordination.

Multi-slot auctions and externalities across positions. We study a single-slot first-price auction, whereas display and search settings are typically multi-slot with position effects, and many platforms run variants of generalized second price (GSP) or hybrid pay-as-bid mechanisms. Multi-slot environments introduce two additional features that can strengthen the monitoring channel. First, there are more margins on which to coordinate: bidders may rotate not only *winning* but *rank positions*, using the vector of allocations as a richer public signal. Second, multi-slot auctions generate externalities in learning: a bidder's observed return from a given multiplier depends on the entire distribution of rival multipliers and on position-dependent click-through rates, making the mapping from outcomes to best responses more complex and potentially increasing the value of public information.

At the same time, multi-slot settings can destabilize coordination because the deviation set is larger and misreports are harder to detect when many reallocations are consistent with noise. Extending our detection object $q(\kappa)$ to multi-dimensional monitoring (e.g., noisy rank vectors, noisy per-position prices) would clarify when additional disclosed statistics act as "coordination focal points" versus "confounding noise." More broadly, multi-slot models raise welfare questions that are muted in the single-slot case: collusion can distort not only who wins, but also which advertisers occupy marginal positions where incremental welfare may be low or negative once user experience is considered.

Privacy, confidentiality, and regulatory constraints on disclosure. We treat the platform as choosing a precision parameter κ and a statistic m_t, but in practice disclosure is constrained by privacy law, contractual confidentiality, and the platform's own commitments to protect proprietary information. These constraints can be incorporated formally as a feasible set of information structures, for example by requiring that the release mechanism satisfy a differential privacy constraint or a mutual-information cap. Doing so changes the design problem from choosing "how transparent" to choosing *which features* of the state and actions may be safely revealed.

A key conceptual point is that privacy protections can cut both ways with respect to competition. Strong privacy can reduce $q(\kappa)$ by limiting identifiability and thereby deter collusion, but it can also reduce $m(\kappa)$ by limiting bidders' ability to infer fundamentals, potentially depressing bids through persistent uncertainty. This suggests an open design problem: identify disclosure mechanisms that are *privacy-preserving yet competition-promoting*, such as statistics that reveal market-scale information about θ while provably obfuscating any single bidder's action in a way that prevents stable attribution.

Algorithmic learning dynamics and equilibrium selection. We model learning through an average multiplier $\bar{\alpha}(\kappa)$ and capture coordination through the existence of low-revenue perfect public equilibria. A deeper treatment would explicitly model the learning algorithm (e.g., multiplicative weights, UCB, policy-gradient updates) and study which equilibria are selected under realistic adaptation. This matters because algorithmic bidders may coordinate without explicit "strategies" in the repeated-game sense: public feedback can synchronize updates, creating emergent collusion even when bidders are individually optimizing against stationary assumptions. Characterizing the mapping from disclosure to long-run outcomes in such adaptive systems remains largely open, especially when the learning process itself depends on the platform's reporting API, logging granularity, and delays.

Empirical identification and open problems. Finally, while our theory motivates measuring objects like $m(\kappa)$ and $q(\kappa)$, identifying them in field data is nontrivial because both channels affect revenue in the same direction over some ranges of κ. A central open problem is to design empirical tests that can separately recover learning improvements versus coordination facilitation, ideally using quasi-experimental variation in disclosure regimes. Another open question is normative: when the platform's objective is $J(\kappa) = \lambda R(\kappa) + (1 - \lambda) W(\kappa)$, what disclosure is optimal once we account for downstream product-market effects, user experience externalities, and the possibility that transparency changes entry and innovation among bidders? Addressing these questions requires combining information design, dynamic mechanism design, and the economics of algorithmic interaction—precisely the intersection where we expect future work to be most valuable.