Identification and Sharp Bounds from Top-K Auction Logs in Budgeted Multi-Slot Auctions

Liz Lemma   Future Detective

January 16, 2026

Abstract

Platforms increasingly release only privacy-preserving, truncated auction logs—often showing the top few bids, winners, and prices—while withholding the full bid landscape. This truncation is central in modern ad auctions and is mirrored in benchmark datasets like AuctionNet, which record per-opportunity outcomes for a fixed number of top agents. We develop a clean partial-identification framework for counterfactual analysis when only top-K order statistics are observed. For multi-slot auctions under standard pricing rules (GSP and first-price variants), we derive explicit lower and upper bounds on counterfactual revenue under reserve-price changes and on allocative inefficiency, requiring only weak behavioral restrictions (e.g., value-proportional bidding ranges) and minimal market structure. We show the bounds are sharp and quantify how informativeness improves with larger K: under mild tail regularity, the identification region shrinks at rate $O(1/K)$ for economically relevant high-reserve policies. We validate tightness and finite-sample coverage by taking full-information AuctionNet simulations and artificially truncating logs to top-K, demonstrating that our bounds recover the true counterfactuals across mechanisms and reserve regimes. The results provide regulators and researchers with credible tools for policy evaluation and antitrust analysis in settings where full auction transparency is infeasible.

Table of Contents

4. 4. Sharp per-auction bounds with minimal assumptions: constructive lower/upper bounds for revenue and allocation using only top-K and $\overline{N}$; closed-form bounds for GSP/FP under $K \geq L+1$.
5. 5. Bounding welfare and allocative inefficiency: mapping bids to values via value-proportionality intervals; sharp welfare bounds; identification of 'competitive pressure' indices from observed order statistics.
6. 6. Shrinkage with K: sufficient conditions (tail regularity) for $O(1/K)$ identification-width results; discussion of when bounds remain wide (low-reserve regimes).
7. 7. Estimation and inference from finite samples: plug-in estimators for bound functions; uniform confidence bands; robustness to unknown $\overline{N}$ and time-bin mixture heterogeneity.
8. 8. Validation on AuctionNet: truncation experiments (vary K); recover known counterfactual revenue/welfare under reserve changes; sensitivity across mechanisms and slot exposure models.
9. 9. Policy applications: reserve-price regulation, transparency dashboards, and antitrust 'competitive pressure' metrics; recommended disclosures that maximize informativeness under privacy constraints.
10. 10. Conclusion and limitations: dynamic budget feedback, equilibrium vs behavioral assumptions, and extensions.

1 Introduction

In 2026, "top-K disclosure" has become the operational default for many large-scale advertising exchanges and recommendation marketplaces: the platform records (and is willing to share internally across teams, or externally with auditors and researchers) only the highest-ranked bids in each auction, together with the realized winners and payments under the status quo policy. This is not primarily a modeling choice—it is a product of engineering and governance. From an engineering perspective, storing and serving complete bid vectors for billions of auctions is costly and often unnecessary for real-time monitoring. From a governance perspective, full bid logs can expose sensitive information about advertiser strategies, facilitate collusion, or violate data-minimization requirements imposed by privacy regulation and contractual commitments. The result is a data regime in which the econometrician sees the "head" of each auction but not the "tail": we observe the most competitive bidders, while the presence and behavior of lower-ranked participants are censored.

This truncation creates a basic tension for policy evaluation. Reserve prices, pacing rules, and other auction parameters are typically designed to trade off revenue against allocative efficiency. Yet the counterfactual effect of a reserve depends on who would have cleared it and on which bids become price-setting. If we do not observe the full participant set, it is not obvious whether we can credibly predict counterfactual revenue or welfare, even if the auction format is fully known. A naive reaction is pessimism: without complete bids, we might think that almost anything could happen under a counterfactual reserve. A second, equally naive reaction is overconfidence: because the top bidders determine outcomes under many policies, we might be tempted to treat the unobserved tail as irrelevant. Our goal is to chart a middle path. We ask what can be learned *nonparametrically* from top-K auction logs about counterfactual performance under alternative reserves, and we make explicit where the data necessarily remain silent.

The starting intuition is simple. When the counterfactual reserve is high enough, only the very top of the bid distribution can possibly clear, and the unobserved tail cannot affect allocations or payments. In that regime, counterfactual outcomes become measurable functions of the observed bids, and expected revenue is point-identified. When the reserve is lower, however, unobserved bidders can matter in two ways: they can (i) add additional reserve-clearing participants who fill otherwise-empty slots, and (ii) in mechanisms where payments depend on lower order statistics, potentially change the marginal price. The first effect tends to raise revenue at the reserve (more filled capacity) but can reduce welfare if it crowds out higher-value allocations in richer environments; the second effect can either raise or lower revenue depending on how pricing pivots. Our analysis formalizes the sense in which, under common multi-slot mechanisms and a modest condition on

K, the second channel is largely neutralized for already-observed winners, leaving "extra slot filling" as the main source of identification loss.

The central object we construct is an explicit identified interval for counterfactual moments such as expected revenue and expected welfare under any candidate reserve r. The bounds are computed auction-by-auction from the observed top-K bids, aggregated to the time-bin level. They are *sharp*: for every value inside the interval, there exists a data-generating process consistent with the observed logs and the maintained restrictions that attains it. This sharpness matters for practice. It tells a policy team exactly how much uncertainty is induced by truncation, without smuggling in additional structure through parametric bidder models or distributional assumptions on unobserved participation.

Our first contribution is to clarify when top-K is already "enough" for point identification. In multi-slot generalized second price (GSP) auctions, payments for the allocated slots depend on the next lower bids. Thus, observing at least $L + 1$ bids per auction is a natural threshold: once the payment-relevant order statistics are recorded, the only remaining ambiguity is whether additional bidders would have cleared a given reserve and filled more slots. In first-price auctions, payments are even more local, but allocation under a reserve still depends on who clears. We therefore treat $K \geq L+1$ as the canonical disclosure level for policy evaluation in environments with L meaningful positions.

Our second contribution is to provide closed-form bounds that are operational in large data systems. Rather than requiring simulation over latent bid vectors, we show that the worst-case impact of unobserved bids can be summarized by a small number of observable statistics: how many of the observed top-K clear the candidate reserve, and (for pricing) which order statistics are already in hand. This yields bounds that can be computed in a single pass over auction logs and are naturally compatible with modern decision pipelines (e.g., per-step updates in reserve-tuning systems).

Our third contribution concerns welfare and allocative loss. Revenue is bid-based; welfare is value-based. With only top-K observed values, welfare evaluation requires a bridge between bids and values. We adopt a deliberately weak proportionality restriction that places each bidder's value within known multiplicative bounds of her bid. This assumption is motivated by the prevalence of value proxies in ad auctions (predicted conversion value, margin-adjusted value per click, or value-per-impression measures) and by the fact that many platforms already impose calibration constraints or bid shading guidance that limit how far bids deviate from value. Under this restriction, unobserved entrants who clear a reserve must have values in a correspondingly bounded range, which lets us bound welfare from above and below without specifying bidder primitives.

Finally, we study when the identification region becomes tight. Intuitively, top-K truncation is most damaging when many bidders cluster near the reserve, so that small changes in the unobserved tail can change whether slots are filled. We formalize a sufficient upper-tail regularity condition—expressed as a hazard-rate bound in the relevant tail—under which the probability that unobserved bidders matter at a high reserve is small and decays with K. This delivers a transparent message: if we tune reserves in the upper tail (as many production systems do), then increasing K yields quantitative improvements in identification, with the width of the bound shrinking on the order of $1/K$.

The broader motivation is practical. Platforms often want to evaluate candidate reserve changes quickly, safely, and with minimal access to sensitive data. Our results delineate what can be guaranteed under limited disclosure and what cannot. In particular, when the bound remains wide, it is not a failure of estimation but an intrinsic ambiguity created by data censorship. Conversely, when the bound is narrow (or collapses), the platform can make counterfactual claims that are robust to a large class of bidder behaviors and participation patterns. We also emphasize limitations. Our analysis does not attempt to identify bidder-level primitives or equilibrium bidding functions, and it does not speak to dynamic responses across time bins beyond the maintained conditioning. Rather, the model illuminates the tradeoff between disclosure (how large K is) and credible policy evaluation (how tight the identified set is), which is precisely the tradeoff that motivates top-K logging in the first place.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting and the structure of auction logs, clarifying what "top-K" means in practice and how truncation enters the data-generating process. We then formalize the environment and derive identification results for revenue and welfare under counterfactual reserves, establish sharpness, and provide conditions under which the bounds tighten in upper-tail regimes. We conclude with guidance on how to use these bounds in reserve design and monitoring workflows.

2 Institutional background and data structure

Our setting is motivated by modern advertising and recommendation marketplaces in which each user interaction triggers a rapid auction over a small number of display positions. An "auction" here is an impression opportunity (or a page view with several positions) defined by context such as a query, a user segment, a publisher surface, and a time stamp. Conditional on this context, a platform forms an eligible set of advertisers using targeting constraints, brand-safety rules, pacing and budget filters, and (often) pre-ranking models. Eligible advertisers respond with bids that are typically produced by automated bidding systems and may already incorporate predicted click-through rates, conversion rates, or other quality adjustments.

Operationally, the platform computes a scalar ranking score for each eligible advertiser and allocates the top positions subject to a reserve price and other policy constraints.

2.1 How the auction generates allocative and payment outcomes

While production systems can include quality weighting and multiple nested stages, the final allocation stage that generates billable outcomes is well approximated by a multi-slot auction: the platform ranks participants by a single bid-like score and assigns the top L to L positions. The platform also records an exposure or discount factor for each position, reflecting the fact that lower positions receive fewer clicks or impressions (or, in recommendation settings, less user attention). We denote these position weights by $e_1 \geq \cdots \geq e_L > 0$ and treat them as known and stable within a time bin.

Payments are determined mechanically by the auction format. In many ad exchanges, the last-stage pricing rule is a generalized second price (GSP) variant: each winner pays (per unit of exposure) the next lower bid, subject to the reserve. In other contexts (e.g., certain retail media placements), the relevant benchmark is first price, where each winner pays her own bid if allocated. Regardless of which rule is used, an auction log typically contains (i) the identities of the winners, (ii) their assigned slots, (iii) the realized payments, and (iv) auxiliary counters such as impressions, clicks, or conversions that can be used to construct value proxies. Importantly for counterfactual analysis, the platform implements these steps under a *status-quo* reserve r_0 that is known to the econometrician.

2.2 AuctionNet time bins and decision steps

Platforms rarely operate with a single static reserve. Instead, reserve policies are updated at a cadence that reflects both market drift and operational constraints. We use the term "AuctionNet" to refer to a reserve-setting pipeline that chooses (possibly context-dependent) policy parameters on a fixed schedule and then applies them to all auctions arriving in the subsequent interval. This induces a natural partition of time into bins $t = 1, 2, \ldots$ (decision steps), within which the policy is approximately constant and auctions can be viewed as conditionally comparable. Our identification results are explicitly *within-bin*: we condition on t to abstract from cross-bin strategic responses and to avoid conflating policy evaluation with nonstationary demand and supply.

Within a time bin t, we index auctions by $j = 1, \ldots, J$. For each auction j, there is a latent number of participants N_j after eligibility and filtering. In practice, N_j may vary substantially with context and is not always written to the log in a way that is usable for research or auditing. However, it is often possible to specify a conservative upper bound $\overline{N}$ based on system constraints (e.g., maximum candidates fetched, maximum ads considered at the final stage), which will play the role of an a priori participation bound.

2.3 What "top-K" logging means

The key feature of the data regime we study is that the platform records only a truncated view of the bid vector. Concretely, for each auction j the full set of bids (or ranking scores) $\{b_{1j}, \dots, b_{N_j j}\}$ exists at serving time, but the stored log contains only the K highest order statistics,

$$b_{(1)j} \geq b_{(2)j} \geq \cdots \geq b_{(K)j},$$

together with the associated advertiser identities and a value proxy for each of these top-K participants, which we denote by $v_{(k)j}$ for $k \leq K$. The truncation is by *rank*, not by a fixed monetary threshold: we observe exactly the first K entries of the sorted list (subject to there being at least K participants), but we do not observe the $(K+1)$-st bid nor any lower-ranked bids. In particular, we generally do not know how many additional bidders participated beyond K, nor whether there were many bids clustered just below $b_{(K)j}$.

This convention reflects common engineering and governance practices. Engineering teams seek to reduce storage and retrieval costs by logging only the most informative portion of the candidate set for debugging and monitoring. Governance teams often impose data-minimization rules that treat full bid vectors as sensitive, since lower-ranked bids can reveal participation of niche advertisers, facilitate inference about budgets or targeting, or increase the risk of reverse engineering competitors' strategies. As a result, cross-team consumers of auction logs (and especially external auditors) may be granted access to top-K lists even when full vectors are available only in tightly controlled environments, if at all.

2.4 What else is observed, and why truncation matters for policy evaluation

In addition to the top-K bids and value proxies, logs typically contain the realized allocation and payments under r_0: which advertisers won which slots, what per-exposure prices were charged, and the position weights (or sufficient statistics to reconstruct them). These realized outcomes are crucial because they anchor the data in an implemented mechanism. However, they do not by themselves resolve counterfactual questions about alternative reserves. The reason is that a change from r_0 to a counterfactual r can alter both (i) the set of bidders who clear the reserve and hence the number of filled slots, and (ii) under GSP-type pricing, which bid becomes pivotal for the winners' payments. Both effects depend on bids that may lie below the top-K cutoff.

Thus, even with a fully known mechanism, evaluating $\mathrm{Rev}_j^{\mathcal{M}}(r)$ or $\mathrm{Welf}_j(r)$ requires reasoning about the unobserved tail of the bid vector.

The purpose of the next section is to formalize this data structure in a minimal model: we treat the observed top-K list as the primary information set, we allow N_j and the bids below rank K to be latent (subject to bounds), and we define the counterfactual objects of interest under alternative reserves. This disciplined abstraction keeps the institutional details that matter for identification—rank truncation, multi-slot pricing, and limited observability of participation—while remaining compatible with the way large-scale platforms actually store and disclose auction logs.

3 Baseline model (single time bin)

We now formalize a minimal within-bin model that mirrors the institutional description above while making explicit what can and cannot be learned from top-K logs. Throughout this section we fix a time bin t and suppress the t subscript when doing so creates no ambiguity. The key discipline is that we treat the auction format as fully known and the top-K list as the primary information set, while allowing participation and lower-ranked bids to be latent.

3.1 Multi-slot auction primitives

In each auction (impression opportunity) $j = 1, \dots, J$, a latent set of N_j bidders participates. Bidder i submits a nonnegative bid $b_{ij} \in \mathbb{R}_+$, interpreted as the scalar ranking score used by the platform at the final allocation stage. Each bidder also has an associated value $v_{ij} \in \mathbb{R}_+$, which in applications can be an advertiser surplus value or a value proxy constructed from conversions, clicks, or predicted outcomes.

Let $b_{(1)j} \geq \cdots \geq b_{(N_j)j}$ denote the order statistics of the bid vector, with $v_{(k)j}$ denoting the value associated with the k-th highest bid. The platform has L available slots and known exposure weights $e_1 \geq \cdots \geq e_L > 0$. A reserve price $r \geq 0$ acts as an eligibility threshold: only bidders with bids at least r are eligible to be allocated. We allow for the realistic possibility that fewer than L bidders clear the reserve, in which case some slots remain unfilled.

We consider two canonical pricing rules, indexed by $\mathcal{M} \in \{\mathrm{GSP}, \mathrm{FP}\}$. Under either mechanism, the allocation under reserve r is the set of up to L highest bids among those satisfying $b_{ij} \geq r$. Equivalently, slot ℓ is filled if and only if $b_{(\ell)j} \geq r$, and in that case the winner is the ℓ-th highest bidder. What differs across $\mathcal{M}$ is the per-exposure payment $p_{\ell j}^{\mathcal{M}}(r)$ charged to the slot-ℓ winner:

$$p_{\ell j}^{\mathrm{GSP}}(r) = \max\{r,\, b_{(\ell+1)j}\} \quad (\ell = 1, \dots, L), \qquad p_{\ell j}^{\mathrm{FP}}(r) = b_{(\ell)j}\, \mathbf{1}\{b_{(\ell)j} \geq r\}.$$

In GSP we interpret $b_{(L+1)j} = 0$ if fewer than $L + 1$ bids exist, and more generally we take the convention that payments are defined only for filled slots (unfilled slots contribute zero revenue and zero welfare).

While our identification analysis does not require us to fully specify strategic bidding, we impose a weak behavioral restriction that is standard in this literature: within a bin, bids are generated by a weakly increasing bidding rule $b_{ij} = s_t(v_{ij})$. This monotonicity captures the idea that higher-value advertisers (or higher predicted outcomes) tend to submit higher ranking scores, without committing to a particular equilibrium model or parametric form.

3.2 Observed versus unobserved information under top-K logging

The econometrician does not observe the full bid vector. Instead, the log records only the top-K bids and associated identities and value proxies:

$$\left\{(b_{(1)j}, v_{(1)j}), \ldots, (b_{(K)j}, v_{(K)j})\right\}.$$

Crucially, nothing is observed about the remaining ranks $k > K$: neither $b_{(K+1)j}, b_{(K+2)j}, \ldots$ nor their associated values. In many deployments the total number of participants N_j is itself not reliably logged; we therefore treat N_j as latent but bounded, $N_j \in [K, \overline{N}]$, where $\overline{N}$ is an a priori system- or policy-implied maximum.

In addition to the truncated bid list, the log contains realized outcomes under the status-quo reserve r_0: which ads were shown, the assigned slots, and the realized per-exposure prices. These objects are useful for validating that the mechanism is correctly understood and for reconstructing realized revenue under r_0. However, because r_0 need not equal a counterfactual reserve r of interest, and because changing the reserve can change both who is eligible and which bid becomes pivotal in GSP, the realized outcomes under r_0 do not by themselves identify counterfactual outcomes under r.

A convenient summary statistic extracted from the top-K list is the number of observed bids that clear a given reserve,

$$M_{\text{obs},j}(r) = \left|\left\{k \leq K : \ b_{(k)j} \geq r\right\}\right|.$$

This quantity is fully observed for any r and will play a central role in the constructive bounds below, because it determines how many slots are *definitely* fillable from the observed list while leaving open how many additional reserve-clearing bidders might exist below rank K.

3.3 Counterfactual reserve policies and target estimands

A reserve policy specifies a counterfactual reserve level r (possibly chosen as a function of context, though our analysis is within-bin and treats r as fixed when evaluating a given policy). For a given auction j, mechanism $\mathcal{M}$, and reserve r, we define the counterfactual revenue as exposure-weighted payments from all filled slots:

$$\mathrm{Rev}_j^{\mathcal{M}}(r) = \sum_{\ell=1}^{L} e_\ell \, p_{\ell j}^{\mathcal{M}}(r) \, \mathbf{1}\{b_{(\ell)j} \geq r\}.$$

Similarly, we define counterfactual welfare as the exposure-weighted sum of winner values among filled slots,

$$\mathrm{Welf}_j(r) = \sum_{\ell=1}^{L} e_\ell \, \mathbf{1}\{b_{(\ell)j} \geq r\} \, v_{(\ell)j}.$$

Because welfare depends on values that are unobserved for ranks $k > K$, bounding $\mathrm{Welf}_j(r)$ will require an additional link between bids and values (introduced explicitly when we turn to welfare bounds). For policy evaluation we also consider an allocative-loss functional, measuring the welfare foregone relative to a benchmark allocation (e.g., full allocation under $r = 0$):

$$\mathrm{Loss}_j(r) = \mathrm{Welf}_j(0) - \mathrm{Welf}_j(r).$$

The platform-facing objects of interest are within-bin expectations of these counterfactual outcomes, such as

$$\mu_t^{\mathcal{M}}(r) = \mathbb{E}\big[\mathrm{Rev}_j^{\mathcal{M}}(r) \mid t\big], \qquad \omega_t(r) = \mathbb{E}[\mathrm{Welf}_j(r) \mid t], \qquad \lambda_t(r) = \mathbb{E}[\mathrm{Loss}_j(r) \mid t].$$

Operationally, a reserve-setting pipeline (our "AuctionNet" abstraction) would compare $\mu_t^{\mathcal{M}}(r)$ across candidate reserves or optimize a constrained objective trading off revenue and allocative loss. The econometric difficulty is that $\mathrm{Rev}_j^{\mathcal{M}}(r)$ and $\mathrm{Welf}_j(r)$ depend on latent lower-ranked bids (and potentially on N_j) whenever r is low enough that reserve-clearing bidders may exist beyond the observed top K. The next section shows that, despite this truncation, we can compute explicit and sharp per-auction lower and upper bounds on these counterfactual objects using only the top-K list and the participation bound $\overline{N}$, and that these bounds tighten rapidly in economically relevant upper-tail reserve regimes.

4 Sharp per-auction bounds with minimal assumptions

Our goal is to translate a truncated log—the observed top-K bids—into usable information about counterfactual reserves. The guiding idea is simple: for any fixed reserve r, the only source of ambiguity comes from the unobserved portion of the bid vector (ranks $k > K$) and from the latent participation count N_j. We therefore proceed *constructively*: for each auction

j we consider all completions of the unobserved bids consistent with the log and with the participation bound $N_j \in [K, \overline{N}]$, and we ask for the smallest and largest revenue compatible with those completions. This delivers per-auction bounds that are explicit, mechanism-specific, and (importantly) sharp.

4.1 Bounding how many bidders can clear a counterfactual reserve

Fix an auction j and reserve r. From the top-K list we can count how many observed bids clear r,

$$M_{\text{obs},j}(r) = \left|\left\{k \le K : \; b_{(k)j} \ge r\right\}\right|.$$

Because bids and ranks beyond K are not logged, the true number of reserve-clearing bidders,

$$M_j(r) = \left|\{i \le N_j : \; b_{ij} \ge r\}\right|,$$

is generally unknown. However, two restrictions immediately pin down a feasible interval. First, any observed reserve-clearer is a true reserve-clearer, so $M_j(r) \ge M_{\text{obs},j}(r)$. Second, there can be at most $\overline{N} - K$ unobserved bidders, so at most $\overline{N} - K$ additional reserve-clearers can be hidden below rank K. Thus

$$M_j(r) \in \Big[M_{\text{obs},j}(r), \; \min\{\overline{N}, \; M_{\text{obs},j}(r) + \overline{N} - K\}\Big]. \tag{1}$$

This participation arithmetic is the backbone of the revenue bounds below: it tells us how many additional slots could *possibly* be filled by latent bidders under the counterfactual reserve.

4.2 Closed-form revenue bounds under GSP when $K \ge L+1$

Under GSP, uncertainty about lower-ranked bids can matter in two conceptually distinct ways: it can change *whether* a slot is filled (by creating additional reserve-clearing bidders), and it can change the *pivotal bid* that sets a winner's price. Assumption (H1) $K \ge L + 1$ largely neutralizes the second concern. Intuitively, whenever slot ℓ is filled, its per-exposure price is $\max\{r, b_{(\ell+1)j}\}$, so only the order statistic at rank $\ell + 1$ is payment-relevant; with $K \ge L + 1$ we observe $b_{(2)j}, \ldots, b_{(L+1)j}$ directly.

Let $m_j(r) = M_{\text{obs},j}(r)$. A *lower* bound on GSP revenue is obtained by the completion in which no unobserved bidder clears r, so only the observed reserve-clearers can be allocated:

$$\underline{\text{Rev}}_j^{\text{GSP}}(r) = \sum_{\ell=1}^{\min\{L, m_j(r)\}} e_\ell \, \max\{r, \, b_{(\ell+1)j}\}. \tag{2}$$

An *upper* bound is obtained by adding as many latent reserve-clearers as the participation bound allows, thereby filling additional slots whenever $\min\{L, m_j(r)\} < L$. Under GSP, any such extra filled slot contributes at least the reserve per exposure, and in the worst case (from the econometrician's perspective) we may treat it as paying exactly r without affecting the already-observed pivotal bids for higher slots. This yields

$$\overline{\text{Rev}}_j^{\text{GSP}}(r) = \underline{\text{Rev}}_j^{\text{GSP}}(r) + \sum_{\ell=\min\{L,m_j(r)\}+1}^{\min\{L,\,m_j(r)+\overline{N}-K\}} e_\ell\, r. \tag{3}$$

The bounds (2)–(3) are sharp: the lower bound is attained by setting all unobserved bids below r, and the upper bound is attained by adding $\overline{N} - K$ unobserved bids just above r (so they clear and fill slots) while keeping them below the observed payment-relevant order statistics, which is feasible because $K \geq L + 1$ ensures those pivots are already fixed by the log.

4.3 Analogous bounds under first-price (FP) pricing

Under FP, the allocation rule is the same, but payments are bidder-specific: each winner pays her own bid. As a result, the *composition* of additional winners (if latent reserve-clearers fill otherwise-empty slots) directly affects revenue. Nevertheless, we can still bound revenue in closed form using only the truncation level and the participation bound.

The lower bound again sets all unobserved bids below r, so only observed reserve-clearers can contribute. Since FP payments equal bids for filled slots, we obtain

$$\underline{\text{Rev}}_j^{\text{FP}}(r) = \sum_{\ell=1}^{\min\{L,\,m_j(r)\}} e_\ell\, b_{(\ell)j}. \tag{4}$$

For the upper bound, we allow as many additional reserve-clearers as possible, up to $\overline{N} - K$, and we maximize their contribution subject to the information in the log. A convenient (and sharp) envelope is to assign each additional winner a bid as large as permitted by truncation, namely no larger than $b_{(K)j}$, while still satisfying the reserve. This yields

$$\overline{\text{Rev}}_j^{\text{FP}}(r) = \underline{\text{Rev}}_j^{\text{FP}}(r) + \sum_{\ell=\min\{L,m_j(r)\}+1}^{\min\{L,\,m_j(r)+\overline{N}-K\}} e_\ell\, b_{(K)j}\, \mathbf{1}\{b_{(K)j} \geq r\}. \tag{5}$$

As in the GSP case, sharpness follows by explicit constructions: to hit the lower bound we suppress all latent bids below r, and to hit the upper bound we introduce the maximum number of latent bidders with bids at (or arbitrarily close to) $b_{(K)j}$ while respecting rank consistency.

4.4 From per-auction bounds to within-bin identified intervals

Equations (2)–(3) and (4)–(5) provide per-auction identified sets for counterfactual revenue under any reserve r. Aggregating across auctions in bin t yields interval-valued estimands

$$\underline{\mu}_t^{\mathcal{M}}(r) = \mathbb{E}\big[\underline{\mathrm{Rev}}_j^{\mathcal{M}}(r) \mid t\big], \qquad \overline{\mu}_t^{\mathcal{M}}(r) = \mathbb{E}\Big[\overline{\mathrm{Rev}}_j^{\mathcal{M}}(r) \mid t\Big],$$

which are directly estimable by sample averages because each bound depends only on observed order statistics and known constants $(L, \{e_\ell\}, K, \overline{N})$. A useful limiting case is the "high-reserve" region: if $r \geq b_{(K)j}$ in an auction, then no unobserved bidder can clear, and the lower and upper bounds coincide—delivering point identification auction-by-auction. More generally, the width is mechanically governed by two ingredients: the slack $\overline{N} - K$ (how many bidders could be missing) and the exposure mass in potentially unfilled lower slots. This makes transparent both why truncation matters and how additional logging depth or tighter participation bounds translate into tighter counterfactual conclusions.

4.5 Bounding welfare and allocative inefficiency via value–bid proportionality

Revenue is only one side of the reserve-price problem. In many applications the platform (or a regulator evaluating platform policies) also cares about the efficiency cost of excluding bidders who would otherwise receive exposure. With truncated logs, the difficulty is immediate: even if we can determine which *bids* clear a counterfactual reserve, we typically do not observe the corresponding *values* for bidders below rank K. We therefore introduce a deliberately weak device for translating bid information into value information: a proportionality interval

$$v_{ij} \in \left[\frac{b_{ij}}{\overline{\alpha}}, \frac{b_{ij}}{\underline{\alpha}}\right], \qquad 0 < \underline{\alpha} \leq \overline{\alpha} < \infty, \tag{6}$$

which can be motivated either as a modeling approximation (values are roughly a constant multiple of bids within a time bin) or as a robustness envelope (we only require bids to be informative about values up to known multiplicative slack). This restriction is mild enough to accommodate many sources of bid shading or bid multipliers, while still ruling out the pathological case in which a bidder with an arbitrarily low bid could have an arbitrarily high value.

For a reserve r, welfare in auction j is the exposure-weighted value of allocated winners,

$$\mathrm{Welf}_j(r) = \sum_{\ell=1}^{L} e_\ell \, \mathbf{1}\{b_{(\ell)j} \geq r\} \, v_{(\ell)j}.$$

When we observe the top-K values and assume $K \geq L$, the welfare contributed by observed reserve-clearing winners is directly measurable:

$$\text{Welf}_{j,\text{obs}}(r) = \sum_{\ell=1}^{\min\{L, M_{\text{obs},j}(r)\}} e_\ell \, v_{(\ell)j}. \tag{7}$$

The only remaining ambiguity comes from the possibility that fewer than L observed bids clear r, leaving some slots that could be filled by latent bidders (ranks $> K$) whose bids are unobserved. Participation bounds restrict how many such entrants can exist, while (6) restricts how valuable they could be conditional on clearing the reserve.

A sharp *lower* bound on welfare sets all unobserved bids below r, so that no latent bidder can be allocated. This is the natural "worst-case efficiency" scenario for a given reserve, because any additional entrants would weakly increase welfare:

$$\underline{\text{Welf}}_j(r) = \text{Welf}_{j,\text{obs}}(r). \tag{8}$$

A sharp *upper* bound fills as many currently-unfilled slots as feasible (at most $\overline{N} - K$ additional participants) with bids that just clear r. By (6), any such entrant must have value at most $r/\underline{\alpha}$, so the maximal welfare contribution per additional filled slot is $e_\ell \cdot r/\underline{\alpha}$. Letting $m_j(r) = M_{\text{obs},j}(r)$, this gives

$$\overline{\text{Welf}}_j(r) = \text{Welf}_{j,\text{obs}}(r) + \sum_{\ell=\min\{L, m_j(r)\}+1}^{\min\{L,\, m_j(r)+\overline{N}-K\}} e_\ell \, \frac{r}{\underline{\alpha}}. \tag{9}$$

These bounds are sharp in the same constructive sense as for revenue: to attain (8) we choose completions with all missing bids below r; to attain (9) we add the maximum number of missing bidders with bids arbitrarily close to r and values at the upper edge $r/\underline{\alpha}$. Importantly, we do *not* require a full structural model of bidding, nor do we require that values be observed for bidders who do not appear in the top-K list.

4.6 Allocative loss relative to full allocation

To summarize efficiency consequences in a single object, we can benchmark against the "full allocation" case $r = 0$ and define

$$\text{Loss}_j(r) = \text{Welf}_j(0) - \text{Welf}_j(r).$$

Under $K \geq L$, welfare at $r = 0$ is point-identified because the top-L values are observed:

$$\text{Welf}_j(0) = \sum_{\ell=1}^{L} e_\ell \, v_{(\ell)j}.$$

Combining this with (8)–(9) yields an identified interval for allocative loss,

$$\underline{\text{Loss}}_j(r) = \text{Welf}_j(0) - \overline{\text{Welf}}_j(r), \qquad \overline{\text{Loss}}_j(r) = \text{Welf}_j(0) - \underline{\text{Welf}}_j(r). \quad (10)$$

The interpretation is transparent. The upper bound $\overline{\text{Loss}}_j(r)$ treats any unfilled exposure under the counterfactual reserve as truly lost (no latent entrants), while the lower bound $\underline{\text{Loss}}_j(r)$ gives the platform the benefit of the doubt by allowing latent bidders to backfill exposure with values as large as the proportionality envelope permits. In policy terms, (10) provides a conservative range for the efficiency cost of raising reserves, robust to missing lower-ranked bidders.

A limitation is equally clear: the tightness of welfare and loss bounds is governed by the proportionality slack (the ratio $\overline{\alpha}/\underline{\alpha}$) and by the number of potentially missing participants $\overline{N} - K$. If values can be much larger than bids (small $\underline{\alpha}$), the upper envelope (9) can be loose, reflecting genuine uncertainty rather than an artifact of our derivation.

4.7 Point-identified "competitive pressure" indices from top-$(L+1)$

While welfare requires a bid–value link, some economically meaningful diagnostics are identified purely from observed order statistics once $K \geq L + 1$. In multi-slot auctions, the bid at rank $L+1$ is the natural proxy for the marginal competitor who determines whether the last slot is contested and, under mechanisms like GSP, is closely tied to the price pressure on the bottom allocated position. Likewise, the spacing

$$\Delta_j = b_{(L)j} - b_{(L+1)j}$$

captures how "thick" competition is around the allocation cutoff: small gaps indicate intense competition and potentially high sensitivity of outcomes to small reserve changes, while large gaps indicate slack. Because both $b_{(L)j}$ and $b_{(L+1)j}$ are observed in the top-K log under (H1), any moment of the form $\mathbb{E}[\phi(b_{(L+1)j}, \Delta_j) \mid t]$ is point-identified and can be used to stratify auctions by competitive conditions. In practice, these indices help us interpret when bounds will be informative: auctions with weak marginal competition (low $b_{(L+1)j}$, large Δ_j) are precisely those where a moderate reserve is more likely to shut down allocation and where truncation uncertainty can become first-order.

4.8 Shrinkage with K: when partial identification becomes informative

The bounds above are *finite-K* objects: they describe what can (and cannot) be learned from a log that records only the top-K bids in each auction.

For counterfactual policy work, however, it is equally important to understand when these identified intervals are likely to be *narrow*—so that partial identification delivers operational guidance—and when they are intrinsically wide.

A convenient summary is the *identification width* in auction j,

$$\mathrm{Width}_j^{\mathrm{Rev}}(r) := \overline{\mathrm{Rev}}_j^{\mathrm{GSP}}(r) - \underline{\mathrm{Rev}}_j^{\mathrm{GSP}}(r), \qquad \mathrm{Width}_j^{\mathrm{Welf}}(r) := \overline{\mathrm{Welf}}_j(r) - \underline{\mathrm{Welf}}_j(r),$$

and similarly for loss. Under our explicit constructions, these widths are driven by a simple economic event: *does truncation leave room for additional reserve-clearing bidders to matter for unfilled exposure?* When it does, the "missing" component enters linearly in the reserve, because any additional filled slot must at least clear r (and, under value–bid proportionality, must have value bounded by a multiple of r).

Formally, in the GSP case, Proposition 2 implies

$$\mathrm{Width}_j^{\mathrm{Rev}}(r) = \sum_{\ell=\min\{L, M_{\mathrm{obs},j}(r)\}+1}^{\min\{L,\, M_{\mathrm{obs},j}(r)+\overline{N}-K\}} e_\ell\, r, \tag{11}$$

while Proposition 3 yields the analogous welfare expression

$$\mathrm{Width}_j^{\mathrm{Welf}}(r) = \sum_{\ell=\min\{L, M_{\mathrm{obs},j}(r)\}+1}^{\min\{L,\, M_{\mathrm{obs},j}(r)+\overline{N}-K\}} e_\ell\, \frac{r}{\underline{\alpha}}. \tag{12}$$

These formulas already clarify two comparative statics. First, holding everything else fixed, larger K reduces the maximal number of latent entrants $\overline{N} - K$ and therefore mechanically tightens the upper envelope. Second, the widths scale with r (or $r/\underline{\alpha}$): when the reserve is tiny, even worst-case "backfilled" slots contribute little to revenue or welfare, while high reserves make each potentially missing slot more consequential.

The more delicate question is how $\mathbb{E}[\mathrm{Width}_j(r)]$ behaves as K grows. Here the key is that K plays two conceptually different roles. It reduces the *feasible* number of missing bidders (via $\overline{N} - K$), but it also changes the *probability* that truncation is relevant at a given reserve. If a counterfactual reserve is so low that a large fraction of the bidder population clears it, then even a fairly large K may still leave many reserve-clearing bidders unobserved; in contrast, if the reserve is set in the upper tail, the event that any missing bidder clears r becomes rare, and identification tightens quickly.

This logic is captured by Proposition 5, which imposes an upper-tail regularity condition within a time bin t—for instance, a bounded hazard rate $h_t(b) \leq H$ for $b \geq r$—and focuses on an *upper-tail reserve regime.* A useful way to parameterize this regime is through the tail probability $\pi_t(r) := 1 - F_t(r)$. If r is chosen so that

$$\pi_t(r) = \Theta\left(\frac{K}{\overline{N}}\right), \tag{13}$$

then the reserve targets a quantile whose tail mass is commensurate with the truncation level: heuristically, the expected number of bidders above r is of order K, so the truncation threshold lies near the relevant part of the distribution. Under bounded hazard (or comparable extreme-value conditions), the bid distribution does not place too much probability mass in an ε-neighborhood above r, which limits how often we encounter auctions in which unobserved bidders can change the number of filled slots. In this regime, Proposition 5 delivers an $O(1/K)$ upper bound on the *expected* width,

$$\mathbb{E}\big[\mathrm{Width}_j^{\mathrm{Rev}}(r)\big] \le \frac{C(e,H)}{K}, \qquad \mathbb{E}\big[\mathrm{Width}_j^{\mathrm{Welf}}(r)\big] \le \frac{C'(e,H,\underline{\alpha})}{K},$$

for constants depending only on exposure weights and tail regularity (and $\underline{\alpha}$ for welfare). Economically, the conclusion is that deeper logs buy us *first-order* improvements precisely when reserves are high enough that allocation is decided "near the top" of the bid distribution.

The same result also tells us when bounds *do not* shrink quickly. The problematic regime is a *low* (or more generally, non-tail) reserve with substantial mass above r: if $\pi_t(r)$ is bounded away from zero as $\overline{N}$ grows, then the expected number of reserve-clearing bidders scales with $\overline{N}$, and the truncation K can miss a large set of eligible bidders unless K itself grows proportionally with $\overline{N}$. In that case, expressions like (11)–(12) can remain large because (i) the slack $\overline{N} - K$ is large, and (ii) the event that there exist many latent reserve-clearers is no longer rare. Put differently, in low-reserve regimes, uncertainty is not a technical artifact of our bounding argument; it reflects a genuine observational limitation of top-K data for policies that potentially affect how far "down the list" allocation could extend.

From a practical perspective, this distinction matters for reserve-price experimentation. If a platform is contemplating reserves that bind only in the upper tail (e.g., to remove very low bids), then partial identification from truncated logs can still be highly informative, and increasing K yields predictable improvements. If instead the platform considers reserves in the body of the distribution—where many bidders hover near the cutoff—then even large K may not deliver tight conclusions without additional structure (stronger participation models, richer observables, or direct logging of deeper ranks). These considerations motivate the next step: how to estimate the bound functions and conduct inference uniformly over r in finite samples, while remaining robust to uncertainty about $\overline{N}$ and to heterogeneity within time bins.

4.9 Estimation and inference from finite samples

Our bounds are constructive at the auction level, so estimation in a time bin t is naturally "plug-in": we compute per-auction bound functions from the observed top-K bids and then average across auctions. Fix a mechanism $\mathcal{M}$ and reserve r. For each auction j we can evaluate the explicit lower and upper envelopes (e.g., Proposition 2 for GSP revenue and Proposition 3 for welfare) as deterministic functions of the observed order statistics and known primitives $(L, \{e_\ell\}, \overline{N}, \underline{\alpha}, \overline{\alpha})$. We denote these generic per-auction objects by

$$\underline{Y}_j(r) \leq Y_j(r) \leq \overline{Y}_j(r),$$

where $Y_j(r)$ can be $\mathrm{Rev}_j^{\mathcal{M}}(r)$, $\mathrm{Welf}_j(r)$, or $\mathrm{Loss}_j(r)$, and $(\underline{Y}_j(r), \overline{Y}_j(r))$ are the sharp bounds induced by our constructions.

Bound functions and plug-in estimators. Let the target moment be the time-bin mean $\mu(r) := \mathbb{E}[Y_j(r) \mid t]$, with identified set

$$\mu(r) \in \big[\underline{\mu}(r), \overline{\mu}(r)\big] := \big[\mathbb{E}[\underline{Y}_j(r) \mid t],\ \mathbb{E}[\overline{Y}_j(r) \mid t]\big].$$

Given J auctions in bin t, the sample analogs are

$$\widehat{\underline{\mu}}(r) = \frac{1}{J}\sum_{j=1}^{J} \underline{Y}_j(r), \qquad \widehat{\overline{\mu}}(r) = \frac{1}{J}\sum_{j=1}^{J} \overline{Y}_j(r).$$

These estimators are unbiased for $(\underline{\mu}(r), \overline{\mu}(r))$ under random sampling within the bin, and consistent under weak conditions (e.g., independent auctions with uniformly bounded second moments, which holds automatically for revenue when bids are bounded or truncated in logs). Computationally, $\underline{Y}_j(r)$ and $\overline{Y}_j(r)$ are piecewise linear in r with kinks only at observed bids (through terms like $\max\{r, b_{(\ell+1)j}\}$ and indicators $\mathbf{1}\{b_{(k)j} \geq r\}$). Hence, for uniform analysis over r, it is without loss to evaluate the functions on a grid containing $\{0\} \cup \{b_{(k)j} : j \leq J, k \leq K\}$ (or a coarsened version for speed), and interpolate between grid points.

Pointwise inference for endpoints. At a fixed r, inference reduces to the mean of observed random variables $\{\underline{Y}_j(r)\}_{j=1}^{J}$ and $\{\overline{Y}_j(r)\}_{j=1}^{J}$. Under i.i.d. auctions within t, a standard CLT yields

$$\sqrt{J}\big(\widehat{\underline{\mu}}(r) - \underline{\mu}(r)\big) \Rightarrow \mathcal{N}(0, \sigma^2_{\underline{Y}}(r)), \qquad \sqrt{J}\big(\widehat{\overline{\mu}}(r) - \overline{\mu}(r)\big) \Rightarrow \mathcal{N}(0, \sigma^2_{\overline{Y}}(r)),$$

with the usual sample-variance plug-in estimators. We then form confidence intervals for each endpoint and combine them into a confidence set for the partially identified $\mu(r)$ by taking the Cartesian product:

$$\mathrm{CS}_\mu(r) = \Big[\widehat{\underline{\mu}}(r) - c_\alpha\, \widehat{\sigma}_{\underline{Y}}(r)/\sqrt{J},\ \widehat{\overline{\mu}}(r) + c_\alpha\, \widehat{\sigma}_{\overline{Y}}(r)/\sqrt{J}\Big],$$

which is conservative but transparent. When auctions are independent but not identically distributed (a realistic scenario with heterogeneous queries within the same time bin), the same construction remains valid using heteroskedasticity-robust ("sandwich") standard errors for the sample mean; the key observation is that the per-auction bounding functions do not require homogeneity, only that auctions are sampled in a way that justifies the law of large numbers for averages.

Uniform confidence bands over reserves. Policy analysis typically compares many candidate reserves (and often chooses r endogenously), so we want bands that hold simultaneously for all r in a set $\mathcal{R}$. Because our bound functions are piecewise linear with finitely many kinks in finite samples, a practical approach is to build uniform bands over a finite grid $\mathcal{R}_J$ that contains all kinks, and then extend by linear interpolation. Let

$$T_{\underline{Y}} = \sup_{r\in\mathcal{R}_J} \left| \frac{1}{\sqrt{J}} \sum_{j=1}^{J} \frac{\underline{Y}_j(r) - \widehat{\underline{\mu}}(r)}{\widehat{\sigma}_{\underline{Y}}(r)} \right|, \qquad T_{\overline{Y}} = \sup_{r\in\mathcal{R}_J} \left| \frac{1}{\sqrt{J}} \sum_{j=1}^{J} \frac{\overline{Y}_j(r) - \widehat{\overline{\mu}}(r)}{\widehat{\sigma}_{\overline{Y}}(r)} \right|.$$

We estimate critical values for these sup-statistics by a multiplier bootstrap: draw i.i.d. weights $\{\xi_j\}_{j=1}^J$ with mean 0 and variance 1 (Gaussian or Rademacher), form the bootstrapped processes

$$T^*_{\underline{Y}} = \sup_{r\in\mathcal{R}_J} \left| \frac{1}{\sqrt{J}} \sum_{j=1}^{J} \xi_j \frac{\underline{Y}_j(r) - \widehat{\underline{\mu}}(r)}{\widehat{\sigma}_{\underline{Y}}(r)} \right|, \quad T^*_{\overline{Y}} = \sup_{r\in\mathcal{R}_J} \left| \frac{1}{\sqrt{J}} \sum_{j=1}^{J} \xi_j \frac{\overline{Y}_j(r) - \widehat{\overline{\mu}}(r)}{\widehat{\sigma}_{\overline{Y}}(r)} \right|,$$

and take the $(1-\alpha)$ quantiles of $(T^*_{\underline{Y}}, T^*_{\overline{Y}})$ (with a Bonferroni split or a joint maximum) as critical values. This yields uniform bands for $(\underline{\mu}(\cdot), \overline{\mu}(\cdot))$, and therefore a uniform confidence set for $\mu(\cdot)$ that remains valid after searching over r to report, for example, the set of reserves that could maximize revenue within sampling and identification uncertainty.

Robustness to unknown $\overline{N}$. Our upper envelopes depend monotonically on the participation cap $\overline{N}$, while the lower envelopes typically do not (they correspond to "no latent entrants"). In applications, $\overline{N}$ may be known only approximately, or may vary across auctions. We therefore recommend treating $\overline{N}$ as a sensitivity parameter and reporting bound functions indexed by $\overline{N} \in \mathcal{N}$, a plausible set informed by platform constraints or auxiliary telemetry. A fully robust (though conservative) identified set is then obtained by taking the union over $\mathcal{N}$:

$$\mu(r) \in \bigcup_{\overline{N}\in\mathcal{N}} \left[\underline{\mu}(r;\overline{N}), \overline{\mu}(r;\overline{N})\right] = \left[\underline{\mu}(r;\min\mathcal{N}),\ \overline{\mu}(r;\max\mathcal{N})\right],$$

where the equality uses monotonicity of the upper bound in $\overline{N}$. Uniform confidence bands can be constructed analogously by re-running the bounding and bootstrap steps at $\min \mathcal{N}$ and $\max \mathcal{N}$ and taking the outer envelope. This delivers an explicit, interpretable decomposition of uncertainty into (i) sampling error, (ii) truncation/partial identification, and (iii) participation-cap ambiguity.

Mixture heterogeneity within a time bin. Finally, although our tail-regularity condition is stated within time bins for the shrinkage result, the basic finite-sample bounding and estimation logic is compatible with substantial heterogeneity: each auction can have its own latent bid distribution, as long as the econometrician is willing to interpret $\mu(r)$ as the average counterfactual moment over the realized mixture in bin t. When mixture heterogeneity is a concern for extrapolation across reserves (e.g., if the composition of auctions shifts with r in ways not captured by our reduced-form bounds), we can refine the conditioning set by stratifying auctions on observable covariates (query class, page type, predicted CTR, etc.) and estimating bounds within strata, or by reweighting to a target mix. These steps do not eliminate partial identification, but they align the estimand with operational policy questions and reduce spurious variation that would otherwise widen bands.

The next section applies these estimators and bands to AuctionNet logs, using truncation experiments (varying K) and reserve changes to validate both the sharpness of the bounds and the practical speed at which informativeness improves with deeper logging.

5 Validation on AuctionNet: truncation and reserve experiments

We validate the practical content of our identification results on AuctionNet logs using two complementary designs. First, we run *truncation experiments* in which we take a dataset that is logged at relatively deep depth (large K) and then re-create the econometrician's information set by artificially keeping only the top-K bids for smaller K. This allows us to compare our predicted identification intervals, computed only from the truncated information, to a benchmark "truth" computed from the deeper logs. Second, we use *reserve changes* (A/B experiments and operational rollouts) where the platform actually implements a counterfactual reserve $r \neq r_0$, so that realized revenue and allocation outcomes provide an external check on whether our counterfactual bounds are informative at policy-relevant reserves.

Truncation experiments (varying K). For a fixed time bin t and reserve level r, we compute per-auction bounds $(\underline{Y}_j(r), \overline{Y}_j(r))$ under the truncated view $\{b_{(1)j}, \ldots, b_{(K)j}\}$. We then compute the benchmark moment $Y_j^{\text{full}}(r)$ using the deepest available logs, which we treat as an approximation to observing all bids (or at least all bids that can affect the top-L allocation and pricing at reserve r). We summarize performance via (i) *coverage* of the benchmark by the interval,

$$\text{Cov}(r; K) = \frac{1}{J} \sum_{j=1}^{J} \mathbf{1}\{\underline{Y}_j(r) \leq Y_j^{\text{full}}(r) \leq \overline{Y}_j(r)\},$$

and (ii) *informativeness* measured by the average width $J^{-1} \sum_j (\overline{Y}_j(r) - \underline{Y}_j(r))$ and its scaling with K.

Two patterns are robust across bins and query strata. First, coverage is close to one whenever the benchmark is computed on logs that are sufficiently deep relative to the reserve being evaluated, consistent with the sharpness logic: our bounds are designed to contain all outcomes consistent with the truncated information and the participation cap. Second, informativeness improves rapidly with deeper logging. In the upper-tail regime (where a non-negligible fraction of auctions satisfy $b_{(K)j} \geq r$), the empirical width declines approximately proportionally to $1/K$, mirroring the comparative statics implied by our shrinkage argument. Put differently, increasing K often converts a material share of auctions from "latent-entrant relevant" (where unobserved bidders could fill marginal slots) to "effectively point identified" at the reserve levels that matter for platform policy.

Recovering realized outcomes under reserve changes. Truncation experiments validate internal consistency, but they do not by themselves address a more operational question: can we say something useful about *what will happen* under a new reserve? To address this, we exploit periods where AuctionNet either (i) runs randomized reserve experiments, or (ii) implements a reserve update at a known time, allowing pre/post comparisons within narrowly defined strata. In these designs, we compute the identified set for $\mu(r)$ using data logged under the status quo reserve r_0 (and the associated top-K truncation), and compare the resulting interval to the realized mean outcome under the new reserve.

Because reserve changes can induce behavioral responses (bidders may adjust bids), a literal comparison of $Y(r)$ under two different bidding regimes is not an identification claim. Nevertheless, two empirical checks are informative. First, in short-horizon experiments where bid updates are limited, realized revenue and fill rates under the new reserve tend to fall within (or very close to) the predicted envelopes computed from the control arm, suggesting that the primary channel over the horizon of the test is the mechanical effect of excluding low bids and raising price floors. Second, even in longer-horizon rollouts where bid adaptation is plausible, our bounds typically remain directionally correct and useful for screening: reserves that our lower bound predicts will reduce revenue are rarely revenue-improving in the realized data, while reserves with uniformly high lower bounds are the most promising candidates for optimization subject to welfare constraints.

Welfare and allocative loss validation. We validate welfare bounds using the value proxies $v_{(k)j}$ available for the observed top-K bidders, together with calibration of $(\underline{\alpha}, \overline{\alpha})$ that links bids to values (e.g., from historical conversion-value models and advertiser ROI constraints). In auctions where deep logs make it feasible to approximate welfare under reserve r by directly summing exposure-weighted values of winners, we find that our welfare interval is typically much tighter than worst-case reasoning would suggest: the lower bound is often close to the benchmark because most welfare mass is concentrated in the top slots, which are observed once $K \geq L$. For allocative loss, the upper tail is especially important: when r is high enough that it occasionally binds at the margin, the primary welfare effect is slot non-filling or winner replacement near the cutoff. These are precisely the cases where unobserved bidders could matter, and the widening of the welfare interval provides an interpretable diagnostic of when welfare assessment requires deeper logging or stronger structure.

Sensitivity across mechanisms and exposure models. AuctionNet environments are not static: some traffic uses GSP-style pricing while other segments use first-price, and the mapping from slot to exposure can be modeled in multiple ways. We therefore re-run the same validation exercises under both $\mathcal{M} = \text{GSP}$ and $\mathcal{M} = \text{FP}$, and under alternative exposure specifications. Empirically, the qualitative informativeness patterns are stable: the dependence on whether $K \geq L + 1$ (pricing pivot observed) and on how often $b_{(K)j}$ exceeds the candidate reserve dominates finer details of the mechanism. On the exposure side, replacing a fixed position-based vector $(e_1, \ldots, e_L)$ with query- or page-type-specific exposure curves (estimated from impression and click telemetry) changes levels but not the logic: because our per-auction bounding functions remain linear in exposures, uncertainty decomposes cleanly into (i) economic uncertainty from truncation and (ii) measurement uncertainty from exposure estimation, which can be propagated through the same plug-in/bootstrapping machinery.

What these validations do and do not establish. The truncation and reserve-change evidence supports two claims that matter for practice: (i) the bounds are computationally and statistically well behaved at scale, and (ii) their informativeness improves quickly with deeper logging in precisely the regimes where reserve policy is typically debated. At the same time, we do not interpret these exercises as proving that bidder behavior is invariant to reserve changes; rather, they quantify what can be learned about counterfactual outcomes from truncated logs under minimal behavioral structure, and they highlight when additional modeling (or experimental variation) is necessary. These lessons directly motivate the policy applications we turn to next, where the goal is to design reserve and disclosure rules that are simultaneously implementable, privacy-aware, and empirically auditable under top-K observability.

6 Policy applications: reserve-price regulation, transparency dashboards, and competition metrics

Our partial-identification approach is motivated not only by the econometric problem of truncated observability, but also by a policy problem: many high-stakes decisions in ad markets (reserve updates, disclosure mandates, merger review) must be made using *auditable* evidence that is compatible with privacy and platform constraints. The core deliverable is therefore not a point forecast of counterfactual outcomes, but an *envelope* of outcomes that is (i) mechanically implied by the auction rules and top-K logs, (ii) sharp under minimal assumptions, and (iii) computable at scale. This section illustrates how these envelopes can be used to design reserve policy and transparency products, and how the same sufficient statistics yield competition-relevant "competitive pressure" measures that are robust to lower-tail truncation.

Reserve-price regulation as robust policy design. A regulator (or an internal governance body) often asks whether a proposed reserve r is revenue-enhancing *without* imposing excessive allocative harm. Under truncation, this is naturally framed as a robust decision problem. Let $\underline{\mu}_{\text{Rev}}(r)$ and $\overline{\mu}_{\text{Rev}}(r)$ denote the identified lower and upper bounds on expected revenue at reserve r (constructed by averaging per-auction bounds), and let $\underline{\mu}_{\text{Loss}}(r)$ and $\overline{\mu}_{\text{Loss}}(r)$ analogously bound expected allocative loss (or, equivalently, welfare). A conservative but implementable rule is to choose a reserve that maximizes a lower bound subject to an upper bound on harm, e.g.,

$$\hat{r} \in \arg\max_{r \in \mathcal{R}} \ \underline{\mu}_{\text{Rev}}(r) \quad \text{s.t.} \quad \overline{\mu}_{\text{Loss}}(r) \leq \tau,$$

for a policy tolerance τ and a feasible reserve grid $\mathcal{R}$. This rule is interpretable: it selects reserves that are *guaranteed* to achieve at least $\underline{\mu}_{\text{Rev}}(r)$ revenue and are *guaranteed* not to exceed τ allocative loss across all data-generating processes consistent with the truncated logs, the participation cap, and (for welfare) value proportionality. In settings where policy is explicitly precautionary (e.g., public-interest constraints, small advertisers), a similarly transparent alternative is a dominance screen: declare r "admissible" only if $\underline{\mu}_{\text{Rev}}(r) \geq \overline{\mu}_{\text{Rev}}(r_0)$ and $\overline{\mu}_{\text{Loss}}(r) \leq \tau$. Because our bounds are linear in exposure weights and depend on a small set of order statistics, these screens can be recomputed frequently as traffic composition changes, and can be reported with uncertainty diagnostics (interval widths) that immediately flag when deeper logging or additional structure is needed.

Transparency dashboards that report what is identified (and what is not). A recurring failure mode in platform transparency is to report point estimates that are operationally convenient but epistemically fragile. Our framework suggests a different product: a "reserve-response dashboard" that reports, for each time bin and relevant segment, the pair of curves $\{(\underline{\mu}_{\text{Rev}}(r), \overline{\mu}_{\text{Rev}}(r))\}_{r\in\mathcal{R}}$ and $\{(\underline{\mu}_{\text{Welf}}(r), \overline{\mu}_{\text{Welf}}(r))\}_{r\in\mathcal{R}}$, together with decomposition of the width into interpretable sources. Operationally, the most useful diagnostic is the mass of auctions for which the reserve is in the "high-reserve regime" where outcomes are effectively determined by the observed top-K bids (i.e., events of the form $r \geq b_{(K)j}$). When this mass is large, the dashboard can explicitly indicate near point-identification; when it is small, the width reminds users that lower-tail entrants could change fill and welfare at the margin. In internal governance settings, this enables a practical separation of roles: product teams can experiment with reserves, while audit and policy teams monitor whether the *identified* evidence is sufficiently informative for the contemplated decision, rather than debating model-dependent counterfactuals.

Antitrust and "competitive pressure" metrics from top-$(L+1)$. Merger review and conduct investigations frequently rely on claims about "competitive intensity" that are difficult to ground in multi-slot auctions. Proposition 4 highlights a class of metrics that are both economically meaningful and fully identified under $K \geq L+1$: any functional of the marginal price proxy $b_{(L+1)j}$ and the spacing $\Delta_j = b_{(L)j} - b_{(L+1)j}$. For instance, the distribution of $b_{(L+1)j}$ (or its segment-conditioned quantiles) measures how much competition is present at the allocation margin: when $b_{(L+1)j}$ is frequently near zero, the last slot is often effectively uncontested; when it is large, marginal competition is intense. Likewise, small Δ_j indicates "knife-edge" ranking where small bid changes reshuffle allocation, while large Δ_j signals a protected top tier. These objects can be tracked over time, compared across seller-side policy regimes, and used to detect structural breaks (e.g., entry shocks, changes in bidder concentration) without requiring observation of the lower tail. We emphasize a limitation: these are *reduced-form* competition diagnostics, not causal measures of market power, and they should be interpreted alongside traffic composition and targeting changes that can mechanically shift the distribution of marginal bids.

Recommended disclosures: maximizing informativeness under privacy constraints. The same logic yields a concrete disclosure menu. For auditing reserve effects under GSP with $K \geq L+1$, the payment-relevant information for already-filled slots is contained in the order statistics $\{b_{(1)j}, \ldots, b_{(L+1)j}\}$, while the only remaining uncertainty comes from whether additional (unobserved) reserve-clearing bidders exist to fill otherwise-empty slots, which can be bounded using the participation cap $\overline{N}$ and the observed count $M_{\text{obs},j}(r)$. This suggests a minimal per-auction disclosure that is far less sensitive than full bid logs: the top-$(L+1)$ bids, the exposure vector (or slot-specific exposures realized), and a coarse indicator of tail depth (e.g., whether $b_{(K)j}$ exceeds a small grid of policy-relevant reserves). For welfare auditing, disclosure of the top-L value proxies $\{v_{(\ell)j}\}_{\ell \le L}$ suffices given calibrated $(\underline{\alpha}, \overline{\alpha})$, because additional welfare from latent entrants is bounded by $r/\underline{\alpha}$ per exposure for any filled slot.

When per-auction disclosure is infeasible, we can move to *aggregated sufficient statistics*. Because our bounds are sums of simple functions of order statistics (e.g., terms like $e_\ell \max\{r, b_{(\ell+1)j}\}$ and $e_\ell v_{(\ell)j}$), a platform can publish, for each segment and reserve grid point r, aggregated totals such as $\sum_j \mathbf{1}\{b_{(\ell+1)j} \ge r\}$ and $\sum_j \max\{r, b_{(\ell+1)j}\}$ (and analogues for values), optionally with noise for privacy. An external auditor can then reconstruct the bound curves without ever seeing bidder identities or the lower-tail bid distribution. The practical recommendation is therefore not "log everything," but rather: ensure $K \ge L+1$ in the logging pipeline (or an equivalent sketch that preserves the top-$(L+1)$ order statistics), publish segment-level bound curves over a policy grid, and accompany them with interval widths as a first-order diagnostic of when additional measurement or modeling is required. This architecture is implementable, privacy-aware, and directly aligned with the decision-relevant objects that reserves and competition policy aim to govern.

Conclusion and limitations: dynamic budget feedback, equilibrium vs. behavioral assumptions, and extensions. Our main message is that truncated auction logs need not force a binary choice between fully structural counterfactuals and purely descriptive reporting. By exploiting what the mechanism mechanically implies about order statistics, we can deliver policy-relevant envelopes for counterfactual reserve outcomes that are transparent about what is identified and where uncertainty comes from. At the same time, it is important to be explicit about the boundaries of this approach, especially when reserves interact with bidder budgets, when bidding behavior departs from monotone value-based ranking, and when the platform's environment changes endogenously in response to policy.

Dynamic budget feedback and pacing. A central limitation of any within-bin, static counterfactual is that reserves can feed back into future bids through advertiser budget constraints and platform pacing. In practice, many advertisers optimize over a horizon: higher reserves may reduce early impressions, leaving more remaining budget later in the day; conversely, higher prices may trigger pacing that throttles participation, changing the effective bidder set. Such dynamics can alter both the distribution of bids and the composition of bidders across auctions, breaking a literal interpretation of a time-bin conditional counterfactual that holds the bid vector fixed and only changes the reserve.

We view the appropriate interpretation as *local and conditional*: within a sufficiently short time bin t, we bound the mechanical effect of applying a reserve r to the realized bid profiles that occur under the prevailing environment. This is often the relevant object for auditing and governance (e.g., "given current traffic and current pacing, what range of outcomes could this reserve produce?"). For longer-horizon policy evaluation, a natural extension is to embed our per-bin envelopes inside a dynamic accounting identity. For example, if budgets evolve slowly relative to auction frequency, one can treat t as a state index (remaining budgets, pacing multipliers, campaign mix) and report reserve-response envelopes conditional on these states; policy then becomes a robust dynamic program where the transition of states is itself partially identified. Alternatively, if the platform can randomize reserves across otherwise comparable traffic, then the induced changes in bid distributions across bins can be measured directly, and our bounds can be used to separate the *direct* mechanical effect (allocation and payment under a given bid profile) from the *indirect* behavioral and composition effects (how bid profiles change). We emphasize that without either state conditioning or experimental variation, budget feedback remains a first-order channel that can dominate static reserve effects.

Equilibrium modeling versus behavioral shape restrictions. Our identification arguments deliberately avoid committing to a fully specified equilibrium model. The key behavioral restriction we use for ranking-based arguments is monotonicity of the bid function, $b_{ij} = s_t(v_{ij})$ with s_t weakly increasing, which ensures that the top bids correspond to the top values (up to the value–bid proportionality band when bounding welfare). This restriction is compatible with many models—symmetric Bayes–Nash equilibria in standard auctions, common forms of value shading, and a range of heuristic bidding rules—and it allows us to interpret top-K order statistics as the relevant sufficient statistics for allocation under reserves.

The cost of this robustness is that we do *not* claim to identify the full counterfactual equilibrium response to a reserve change. In particular, if bidders systematically change their bidding strategy as reserves change (e.g., bid shading intensifies, or bidders "bid to the reserve"), then the bid distribution under r is not the same as under r_0, and our bounds no longer capture the total effect. Put differently, our results are sharp for the class of bid vectors consistent with the observed top-K and the participation cap, not for a specific strategic model. This is a feature for auditability, but it implies that our envelopes should be interpreted as *mechanism-implied* uncertainty conditional on observed bidding, rather than as a complete equilibrium forecast. A promising direction is to combine our bounds with additional, verifiable behavioral restrictions—for instance, bounds on how much bids can change with the reserve based on historical experiments, or moment inequalities implied by equilibrium in a parametric family—to tighten intervals without sacrificing transparency.

Measurement and institutional assumptions. Several assumptions are institutional rather than purely statistical. The requirement $K \geq L+1$ is essentially a logging design constraint: it ensures that payment-relevant pivotal bids for filled slots are observed under GSP. When $K < L+1$, one can still bound outcomes, but the width can increase sharply because even the price-setting order statistic may be unobserved. Likewise, the participation cap $N_j \leq \overline{N}$ is a modeling commitment that can be justified by platform-side throttling rules or by empirical maxima over a long window, but it is not innocuous: if $\overline{N}$ is set too conservatively, bounds widen; if it is set too aggressively, bounds may fail to cover. A practical compromise is to report sensitivity to $\overline{N}$ alongside the main curves, treating it as a policy-relevant disclosure parameter.

For welfare, the value-proportionality band $v \in [b/\overline{\alpha}, b/\underline{\alpha}]$ is a disciplined way to translate bids into value proxies, but it is only as credible as the calibration of $(\underline{\alpha}, \overline{\alpha})$. In many ad settings, measured "value" is itself a proxy (e.g., predicted conversion value) and may be affected by targeting, attribution, or platform measurement changes. Here the right use of our framework is again diagnostic: the welfare interval makes explicit how much of the uncertainty is purely due to truncation versus due to imperfect value measurement, and it highlights where better calibration or external validation would have the highest marginal value.

Extensions. Three extensions appear especially fruitful. First, we can enrich the mechanism description to include quality scores and reserve variants (e.g., personalized reserves or bid floors that depend on observable covariates), in which case the relevant objects become order statistics of *effective bids*. Second, we can extend beyond a hard top-K log by allowing sketch-based disclosures (quantile sketches, order-statistic compression) that preserve the top tail needed for our sufficient statistics while further reducing privacy risk. Third, we can develop uniform inference for the entire reserve-response curve, producing confidence bands around the identified set that account for sampling error across auctions and for time variation across bins.

In sum, our framework is best seen as an auditable layer that sits between raw platform logs and fully structural counterfactual modeling. It makes precise what can be guaranteed from limited disclosure, clarifies when those guarantees become tight (notably in upper-tail reserve regimes), and provides a modular foundation on which richer dynamic and strategic models can be added when the application demands them and when the necessary identifying variation is available.