Targeting Information in Ad Auction Mechanisms

Srinivas Tunuguntla* Carl F. Mela† Jason Pratt‡

January 9, 2026

Abstract

Digital advertising platforms and publishers sell ad inventory that conveys targeting information, such as demographic, contextual, or behavioral audience segments, to advertisers. While revealing this information improves ad relevance, it can reduce competition and lower auction revenues. To resolve this trade-off, this paper develops a general auction mechanism—the Information-Bundling Position Auction (IBPA) mechanism—that leverages the targeting information to maximize publisher revenue across both search and display advertising environments. The proposed mechanism treats the ad inventory type as the publisher's private information and allocates impressions by comparing advertisers' marginal revenues.

We show that IBPA resolves the trade-off between targeting precision and market thickness: publisher revenue is increasing in information granularity and decreasing in disclosure granularity. Moreover, IBPA dominates the generalized second-price (GSP) auction for any distribution of advertiser valuations and under any information or disclosure regime. We also characterize computationally efficient approximations that preserve these guarantees.

Using auction-level data from a large retail media platform, we estimate advertiser valuation distributions and simulate counterfactual outcomes. Relative to GSP, IBPA increases publisher revenue by 68%, allocation rate by 19pp, advertiser welfare by 29%, and total welfare by 54%.

Keywords: Display Advertising, Pricing, Auctions, Reserve Prices, Optimization, Real-Time Bidding (RTB), Programmatic Buying, Mechanism Design

*Assistant Professor of Marketing at the Fuqua School of Business, Duke University (email: srinivas.tunuguntla@duke.edu)

†The T. Austin Finch Foundation Professor of Business Administration at the Fuqua School of Business, Duke University (email: mela@duke.edu, phone: 919-660-7767).

‡General Manager and Senior Research Officer, Koddi (email:jason.pratt@koddi.com)

The authors thank seminar participants at the University of Miami, the University of Rochester, the 2025 Marketing Science Conference, Omid Rafieian, and Tulio Sousa for their constructive suggestions.

1 Introduction

Digital advertising is increasingly transacted through real-time programmatic auctions, with global spend on programmatic advertising expected to reach nearly \$1 trillion by 2032, with a 17% CAGR.[1] Such auctions underlie search advertising on retail media and search engines (typically sold using a generalized second price auction or GSP (Varian 2007)) and display ads (typically sold in open exchanges via a first price auction and walled gardens via a second price auction (Choi et al. 2020)).[2] Across these environments, impression opportunities differ in terms of the consumer viewing the ad—such as demographics, prior purchases, and browsing histories—and the context in which the ad is delivered, including the publisher's page (e.g., surrounding editorial content) and the ad's format (e.g., banner, tower). Advertisers' valuations depend on such targeting information as audience and context information directly affect the relevance and expected performance of an impression.

This paper considers how a publisher should optimally sell impressions when the publisher has private access to such targeting information. We model the publisher as selling impressions belonging to different inventory types, where an inventory type is defined by the underlying audience and contextual attributes of the impression. For each arriving impression opportunity, the publisher privately observes the realized inventory type, while advertisers do not. Our central contribution is to design an auction mechanism for such settings in which the value of an impression is type-dependent and the impression's inventory type is privately observed by the seller.

In current practice, publishers often disclose this targeting information so that advertisers can condition their bids on the realized type of the impression. This "full disclosure" regime improves match quality by allowing advertisers to avoid irrelevant impressions. Conventional wisdom holds that such disclosure should improve revenues by raising advertisers' effective bids. However, a growing body of work finds that disclosing granular targeting information can reduce publisher revenue by thinning competition within each auction (Ada, Abou Nabout, and Feit 2022; Chen, Narasimhan, and Zhang 2001; Iyer, Soberman, and Villas-Boas 2005; Levin and Milgrom 2010; Lu and Yang 2015; Yao and Mela 2011). Publishers therefore face a trade-off between targeting precision and market thickness—more granular targeting information improves relevance and advertiser willingness-to-pay,

[1] `https://www.themarketintelligence.com/market-reports/programmatic-advertising-market-4734`

[2] Display ads can be considered a special case of position auctions in which there is only one slot and the quality score is set to one.

but more granular disclosure fragments demand and reduces auction competitiveness.

Recent work (Rafieian and Yoganarasimhan 2021; Bergemann et al. 2022) studies this trade-off by holding the auction format fixed—typically GSP—and asking how much targeting information the publisher should reveal. In contrast, we take a mechanism-design perspective and redesign the auction itself, rather than holding the format fixed and optimizing disclosure within it.

We introduce a new auction mechanism—which we refer to as the Information-Bundling Position Auction (IBPA)—that integrates insights from mixed bundling in pricing with the marginal-revenue approach to mechanism design. Rather than disclosing the realized inventory type and soliciting a single bid, the publisher keeps this information private and solicits a multidimensional bid—one bid for each possible inventory type. In the single-advertiser case, the optimal mechanism reduces to a mixed-bundling menu of bundles and prices, a form of second-degree price discrimination. This constitutes the "information bundling" component of the mechanism. We extend this logic to multiple advertisers using the marginal-revenue framework (Bulow and Roberts 1989). For each advertiser, we construct a revenue curve, map the advertiser's multidimensional bid to a quantile on that curve, and compute the corresponding marginal revenue. The impression is then allocated to the advertiser with the highest marginal revenue, and payments follow a critical-quantile logic analogous to second price auctions. This ensures incentive compatibility: truthful multidimensional bidding is the advertiser's equilibrium strategy.

Our first theoretical result shows that the IBPA mechanism resolves the trade-off between targeting precision and market thickness. Revenue under IBPA is monotonically increasing in the publisher's information granularity and monotonically decreasing in disclosure granularity. Thus, the publisher benefits from maintaining granular targeting information and limiting disclosure. This decoupling of information collection and disclosure is feasible only because IBPA elicits multidimensional bids that capture valuation differences across inventory types. Traditional one-dimensional mechanisms such as GSP cannot achieve this decoupling, which explains their non-monotonic behavior in the existing literature.

Our second theoretical result establishes that IBPA dominates the GSP auction irrespective of the distribution of advertiser valuations or whether advertisers are symmetric. This dominance holds across all information and disclosure regimes, and is strongest under full information and null disclosure. We further show that computationally efficient approximations retain strong performance guarantees and continue to dominate GSP.

The IBPA mechanism is general, applicable to both display and search advertising, and requires only modest extensions to current platform infrastructure through the use of multidimensional bids. It is incentive compatible, ensures advertisers receive non-negative expected payoffs, improves publisher revenue relative to standard industry mechanisms, and admits computationally efficient approximations.

We empirically validate our mechanism using auction-level bidding data from a major online retailer and its ad-tech partner. Using the estimated distribution of advertiser valuations, inventory types, and click-through rate parameters, we simulate counterfactual outcomes under alternative mechanisms. We find that the proposed mechanism increases publisher revenue by 68%, allocation rate by 19pp, advertiser welfare by 29%, and total welfare by 54%, relative to GSP.

In the remainder of the paper, we first overview how our research builds on the existing literature. We then describe the bidding environment, formalize the publisher and advertiser problems, and introduce the proposed IBPA mechanism. Solving these problems, we establish theoretical guarantees for the proposed mechanism. We next turn to the empirical analysis, where we describe the data, estimate advertiser valuation distributions, and evaluate the counterfactual performance of alternative mechanisms in terms of publisher revenue, advertiser welfare, and total welfare. We conclude by summarizing the contributions of this research and outlining directions for future work.

2 Background and Related Literature

2.1 Digital Advertising

A large and growing literature examines real-time bidding in display advertising. Much of this work focuses on advertiser-side objectives, including the response or return to advertising (Gordon, Moakler, and Zettelmeyer 2023; Johnson 2023), the design of advertising content (Lee, Hosanagar, and Nair 2018), and optimal ad targeting policies (Rafieian and Yoganarasimhan 2021; Sayedi 2018; Tunuguntla and Hoban 2020). The publisher side has also received attention: Despotakis, Ravi, and Sayedi (2021) compare first- and second-price auctions; Choi and Mela (2023) study optimal reserve prices for publishers; and Balocco et al. (2025) consider optimal allocation of inventory across direct and exchange markets. This paper focuses on publisher selling strategies in digital advertising markets, including display and search advertising.

In display advertising, impressions on open exchanges are typically sold using either first-price or second-price auctions (Choi et al. 2020). In a second-price display auction, the highest bidder wins the impression and pays the second-highest bid. In contrast, search advertising takes place in position auctions, where multiple ordered slots are allocated simultaneously. The auction mechanism most commonly used in these settings is the generalized second-price (GSP) auction (Edelman, Ostrovsky, and Schwarz 2007; Varian 2007), which ranks advertisers by bid times quality score and charges per click. A standard display auction is a special case of a position auction: when there is a single slot, unit click-through rate, and quality scores equal to one, a position auction collapses to the display ad auction. We therefore treat position auctions as the unifying environment in which both display and search ads operate, while recognizing that GSP is only one specific mechanism that can be used in this environment.

Despite its widespread use to sell impressions, GSP has well-known limitations. It is not incentive compatible, so advertisers cannot bid their true valuations. Equilibrium analysis typically relies on envy-free equilibria (Edelman, Ostrovsky, and Schwarz 2007; Varian 2007), which are often non-unique and require strong informational assumptions (e.g., advertisers know all other advertisers' valuations). Under private information—arguably the empirically relevant case—equilibria are difficult to characterize, and existing analyses rely on restrictive symmetry assumptions (Gomes and Sweeney 2009). These issues complicate attempts to optimize publisher revenue under GSP. For instance, computing optimal reserve prices or quality-score adjustments becomes challenging due to equilibrium multiplicity and uncertain bidder strategies (Varian 2007). These limitations motivate the mechanism-design approach we pursue in this paper.

2.2 Information Asymmetries in Auctions

A broad literature exists on information asymmetries in auction settings and the related question of whether it is in the interest of the seller to disclose information about the quality of auctioned items. Broadly speaking, this literature can be categorized around vertical quality and horizontal match.

2.2.1 Vertical Quality Information Asymmetries. Sellers often list items of unknown vertical quality (Li, Srinivasan, and Sun 2009; Milgrom and Weber 1982). Bergemann and Pesendorfer (2007) consider whether to reveal information about the quality of the item being auctioned when bidders' decisions to acquire that information are endogenous. Shi (2012) build on this by analyzing how the choice of auction mechanisms—specifically, the informativeness of signals sent to bidders—affects their decisions to gather additional information and the resulting welfare implications for both buyers and sellers. Kim and Koh (2022) extend this work by not only considering whether bidders acquire information but also what type of information they choose to collect. Chen and Yang (2023) examine bidder outcomes when a regulator mandates seller disclosure of information to buyers who are uncertain about an item's value. Choi and Sayedi (2023) examine how intermediaries—display ad agencies bidding on behalf of advertisers—strategically coordinate bids to soften competition and extract higher returns from publishers. Their findings suggest that publishers have an incentive to withhold information from agencies to deter this behavior, highlighting the strategic value of information in advertising. In contrast to this stream of research, and because the sheer volume of purchased advertising inventory reduces advertiser uncertainty about its vertical quality, our emphasis lies on the match value of advertising.

2.2.2 Horizontal Quality Information Asymmetries. Horizontal information asymmetries arise when the value of an item depends on how well it matches the preferences of different buyers. In digital advertising, this corresponds to audience and contextual information that affects how well an impression matches an advertiser's targeting criteria. Disclosing such information can improve match quality and raise advertisers' willingness-to-pay by reducing uncertainty about the consumer viewing the ad. However, disclosing granular targeting information can also reduce competition among advertisers by narrowing the subset of advertisers interested in a particular impression. This reduction in bidder participation can offset the gains from improved targeting and make the publisher worse off (Levin and Milgrom 2010). Yao and Mela (2011), Ada, Abou Nabout, and Feit (2022), Lu and Yang (2015), Sousa (2025), Wu, Huang, and Li (2025), and Rafieian and Yoganarasimhan (2021) have explored this effect empirically in the context of digital advertising and found evidence supporting this trade-off. IBPA resolves this trade-off by allowing the publisher to retain granular inventory-type information while soliciting multidimensional valuations from advertisers, thereby preserving competition even when targeting information is highly granular.

2.2.3 Privacy in Advertising Targeting. Information asymmetries suggest sellers and sometimes advertisers can benefit from greater information granularity. However, there is an active discussion regarding the consumer welfare consequences of privacy (Goldfarb and Que 2023). As a result, publishers may be reticent to share granular information about consumers. The role of information granularity in advertising targeting in advertising auctions has seen limited attention. A notable exception is Rafieian and Yoganarasimhan (2021), who show that agglomerating information about customer types can improve revenue outcomes for ad-networks and advertisers. We build on this literature by distinguishing between seller side agglomeration (first party data) and buyer side agglomeration (third party information). We find that seller outcomes and total welfare improve when the seller uses its most granular data in its allocation and payment decisions but using its least granular data for advertiser bidding decisions. In other words, the approach we develop both preserves consumer data at the first-party level while enhancing auction efficiency.

2.3 Mixed-Bundle Pricing

A key insight underlying our mechanism is that the multi-advertiser auction can be decomposed into a collection of single-advertiser problems. Each such problem closely parallels a classic second-degree price discrimination problem faced by a multi-product monopolist. This connection provides both intuition and structure for our mechanism.

To build this intuition, imagine there is only one advertiser on the platform. The publisher must decide how to sell impressions belonging to different inventory types to this advertiser, knowing only the prior distribution over the advertiser's valuation vector. On the publisher side, this is equivalent to selling T distinct "goods" (inventory types), to a buyer whose valuations for these goods are random according to a known prior.

This problem mirrors the classical setting of a monopolist selling multiple goods to a market of consumers whose valuations for the goods are drawn from a joint distribution. Here, the goods correspond to inventory types, the consumer corresponds to a single advertiser, and the distribution of consumer valuations corresponds to the prior distribution over advertiser valuations.

In the multi-product monopolist setting, the optimal selling mechanism generally takes the form of second-degree price discrimination, implemented through a menu of bundles and prices (Adams and Yellen 1976; McAfee, McMillan, and Whinston 1989; Armstrong 2016). In the most general form—mixed bundling—the monopolist constructs a price for every possible subset of goods. With T goods, there are 2^T feasible bundles; each bundle can be represented by a binary vector of length T, with ones indicating the goods included. A consumer with a specific set of valuations chooses the bundle that yields the highest utility. The utility of a bundle is the sum of the consumer's valuations for the included goods minus the price of the bundle. The monopolist sets the menu of bundle prices to maximize expected revenue given the distribution of valuations.

Our single-advertiser mechanism generalizes this mixed-bundling logic. Instead of restricting bundles to be binary vectors, we allow randomized bundles, where the bundle is a vector in $[0,1]^T$, where the entry for type t is the probability of receiving that inventory type. These probability vectors correspond to lottery pricings, a well-established construct in mechanism design (Rochet 1985; Manelli and Vincent 2006; Daskalakis, Deckelbaum, and Tzamos 2015). Thus, a single-advertiser mechanism is a menu of lottery pricings: each menu item specifies a vector of allocation probabilities across inventory types and an associated expected payment. The advertiser, upon observing its valuation vector, chooses the lottery pricing that maximizes its utility.

Relative to mixed bundling—whose menu contains at most 2^T items—allowing lotteries yields a richer and potentially infinite menu, enabling finer second-degree price discrimination. This richer structure is what allows our mechanism to fully exploit heterogeneity in advertiser valuations.

Importantly, we show that this form of second-degree price discrimination does not suffer from the classic market-thickness problem (Levin and Milgrom 2010). Finer information about inventory type expands the set of feasible lotteries and enables more effective second-degree price discrimination, enhancing publisher profits.

2.4 The Marginal Revenue Framework

We extend the bundling-based pricing logic from the single-advertiser setting to the multi-advertiser environment through the marginal-revenue approach, a guiding principle in multi-agent auction mechanism design. Standard economic intuition suggests that when a monopolist sells a single good across multiple markets, each unit of supply should be allocated to the market with the highest marginal revenue, i.e., the derivative of the revenue curve mapping revenue to the quantity supplied. This allocation is optimal because total revenue can be expressed as the sum of marginal revenues across all units supplied.

Bulow and Roberts (1989) show that, in the single-item setting, the marginal revenue approach is exactly equivalent to Myerson's optimal auction with virtual valuations (Myerson 1981). In fact, Myerson's virtual values coincide with the marginal revenues derived from each agent's revenue curve. Building on this, Alaei et al. (2013) show that the marginal revenue approach is optimal more generally for any environment with single-dimensional agent types and quasi-linear utility in that type.

Multidimensional environments—where agent types are vectors of values rather than single parameters—pose additional challenges for implementing the marginal revenue approach (Alaei et al. 2013). First, when multiple resources must be allocated—for example, an auctioneer selling several items—the revenue curve itself may need to be multidimensional across the allocation of multiple items, and total revenue can no longer be expressed as the simple sum of marginal revenues for each item. Second, even if a one-dimensional revenue curve can be defined, the absence of a natural ordering among multidimensional types makes it difficult to construct a mapping from an agent type to a position on the revenue curve.

Our setting inherits these challenges of multidimensional environments. The publisher must allocate multiple ad slots across advertisers, and each slot may belong to one of many inventory types. Accordingly, each advertiser is characterized by a valuation vector with one component per inventory type. Hence, we advance the mechanism design literature by extending the marginal revenue approach to position auctions with multiple inventory types and, more broadly, to any environment where an item may assume multiple types, and the seller has private information about which type is realized.

The marginal-revenue framework implemented here can be viewed as a form of first-degree price discrimination. The publisher effectively tailors allocations and payments to each advertiser's full vector of valuations across inventory types by pricing impressions based on each advertiser's marginal revenue contribution, extracting surplus in a manner analogous to personalized pricing. In contrast to GSP, which compresses advertiser preferences into a single scalar bid, this approach exploits the full multidimensional structure of advertiser valuations.

3 Problem Setup and Proposed Mechanism

3.1 Model

Consider a publisher selling S ad slots on a page, indexed by $s \in [S]$.[3] Each page is characterized by an inventory type $t \in [T]$. The inventory type is fixed for the page and may reflect contextual

[3] We use the notation $[K] = \{1, 2, \dots, K\}$.

page attributes (e.g., content category) and/or consumer characteristics (e.g., demographics, past purchases, and browsing history). Let

$$\mathbf{p} = (p_1, p_2, \ldots, p_T), \quad \text{with} \quad \sum_{t=1}^{T} p_t = 1,$$

denote the probability distribution over inventory types.[4]

There are A advertisers, indexed by $a \in [A]$. Advertiser a has a per-click valuation $v_{at} \geq 0$ for an impression of type t; impressions that do not receive a click generate no value. Let $\mathbf{v}_a = (v_{a1}, \ldots, v_{aT})$ denote the advertiser's type (valuation vector).

At the time of each impression, the publisher observes the realized inventory type t, whereas the advertisers only know the distribution $\mathbf{p}$. Each advertiser privately observes their own valuation vector $\mathbf{v}_a$, while the publisher knows only a prior distribution F_a over $\mathbb{R}_+^T$ for each advertiser a. We allow for heterogeneous priors $\{F_a\}_{a=1}^A$, and we assume that advertisers draw their types—valuation vectors—independently of each other.

This formulation follows the standard Bayesian mechanism design framework (Myerson 1981). The timing is as follows: the publisher first commits to a mechanism; then the publisher observes the realized inventory type while each advertiser privately observes its valuation vector; finally, the mechanism determines allocations and payments. Within this framework, ex-ante refers to the stage before the inventory type and advertiser valuations are realized. Interim refers to the perspective of a single player—advertiser or publisher—after observing its private information but before knowing the realizations of others, when expectations are taken over the types of the remaining players. Ex-post refers to the stage after both the inventory type and all advertisers' valuations are realized. We assume all agents are risk-neutral.

Without loss of generality, we focus on direct mechanisms (Myerson 1981), in which each advertiser a submits a bid vector $\mathbf{b}_a = (b_{a1}, \ldots, b_{aT})$, corresponding to bids for each inventory type. The mechanism then determines outcomes as a function of all submitted bids, denoted by $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_A)$.

A direct mechanism is characterized by allocation functions $x_{ast}(\mathbf{B})$ and an expected payment

[4] Although we present a single instance of selling slots on a page, one can equivalently view the publisher as selling a large number of i.i.d. impressions with a fraction p_t of them being type t.

$m_a(\mathbf{B})$ for each advertiser. An allocation $x_{ast} = 1$ indicates that advertiser a is assigned slot s for inventory type t, while $x_{ast} = 0$ indicates otherwise. More generally, $x_{ast} \in [0, 1]$ allows for randomized outcomes, in which case x_{ast} specifies the probability that advertiser a receives slot s for inventory type t. $m_a(\mathbf{B})$ denotes the expected payment of advertiser a, where the expectation is taken over the mechanism's randomization, inventory types, and click realizations.[5]

Feasibility requires that each advertiser is assigned at most one slot for each inventory type, i.e.,

$$\sum_s x_{ast} \leq 1 \quad \forall a, t$$

and that each slot is assigned to at most one advertiser, i.e.,

$$\sum_a x_{ast} \leq 1 \quad \forall s, t$$

3.2 Advertiser Problem

Given a mechanism—allocation and payment functions—advertisers submit bids that maximize their expected utility. Because valuations are specified per-click, advertiser utility depends on the click-through rate (CTR). It is standard practice in the literature to model the CTR of an advertiser a in slot s as a product of the advertiser's quality γ_a and the slot effect α_s (Athey and Nekipelov 2011). We further allow the CTR to vary with inventory type t using type effects β_t. Thus, the CTR of advertiser a in slot s for inventory type t is given by

$$\bar{c}_{ast} = \alpha_s \, \beta_t \, \gamma_a$$

We assume that slots are numbered so that $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_S$. Furthermore, to normalize the scale, we set the effect of the top slot to $\alpha_1 = 1$ and the effect of the first inventory type to $\beta_1 = 1$ without loss of generality.

[5] Only expected payments are needed to compute the publisher's revenue. The mechanism does not depend on how payments are disaggregated at the slot or type level. Thus, specific per-slot payments m_{ast} can be implemented in any convenient form as long as they aggregate to the expected payment m_a.

The expected utility of advertiser a, when bidding $\mathbf{b}_a$ with true valuation vector $\mathbf{v}_a$, is then

$$u_a(\mathbf{b}_a, \mathbf{v}_a) = \mathbb{E}\left[\sum_{t,s} p_t\, \alpha_s \beta_t \gamma_a v_{at}\, x_{ast}(\mathbf{b}_a, \mathbf{b}_{-a}) \;-\; m_a(\mathbf{b}_a, \mathbf{b}_{-a})\right],$$

where the expectation is taken over the bids of all other advertisers $\mathbf{b}_{-a}$. Each advertiser chooses bids to maximize this expected utility.

By the revelation principle (Myerson 1981), we can restrict our attention to mechanisms that are incentive compatible without loss of generality. A mechanism is Bayesian incentive compatible (BIC) if truthful bidding constitutes a Bayes-Nash equilibrium, that is,

$$u_a(\mathbf{v}_a, \mathbf{v}_a) \;\geq\; u_a(\mathbf{b}_a, \mathbf{v}_a) \quad \forall\, \mathbf{b}_a,\ \forall a.$$

We impose this constraint in our publisher optimization problem to ensure that bids equal true valuations, i.e., $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_A) = \mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_A)$.

In addition, mechanisms must satisfy a participation, or Bayesian individual rationality (BIR), constraint: each advertiser's expected utility—in equilibrium—must be non-negative,

$$u_a(\mathbf{v}_a, \mathbf{v}_a) \;\geq\; 0 \quad \forall a.$$

3.3 Publisher Problem

Given these incentive constraints, the publisher chooses a mechanism to maximize expected revenue, defined as the sum of expected payments across all advertisers[6]. The publisher's problem can thus be summarized as choosing allocation and payment rules to maximize expected revenue subject to

[6] More generally, the publisher's objective could involve a combination of revenue and advertiser utilities (e.g., to reflect competitive pressures across publishers), but the structure of the analysis remains unchanged.

feasibility, incentive compatibility, and individual rationality constraints:

$$
\begin{aligned}
\max_{\{x_{ast}(\cdot),\, m_a(\cdot)\}} \quad & \mathbb{E}_{\mathbf{V}}\left[\sum_a m_a(\mathbf{V})\right] \\
\text{s.t.} \quad & \sum_s x_{ast}(\mathbf{V}) \ \le\ 1, \quad \forall a, t,\ \forall \mathbf{V} \quad \text{(Feasibility: one slot per advertiser)} \\
& \sum_a x_{ast}(\mathbf{V}) \ \le\ 1, \quad \forall s, t,\ \forall \mathbf{V} \quad \text{(Feasibility: one advertiser per slot)} \\
& u_a(\mathbf{v}_a, \mathbf{v}_a) \ \ge\ u_a(\mathbf{b}_a, \mathbf{v}_a), \quad \forall \mathbf{b}_a,\ \forall a \quad \text{(BIC)} \\
& u_a(\mathbf{v}_a, \mathbf{v}_a) \ \ge\ 0, \quad \forall a \quad \text{(BIR)}.
\end{aligned} \tag{1}
$$

3.4 Information-Bundling Position Auction (IBPA)

We propose a solution to the publisher's problem using the marginal revenue approach that allocates each slot to the advertiser that maximizes the publisher's marginal revenue. The key challenge is that the publisher must allocate multiple ad slots, each subject to its own allocation/feasibility constraint and belonging to different inventory types, among multiple advertisers with incentive constraints. The marginal revenue approach simplifies the publisher problem considerably by reducing it from a joint slot-type-advertiser problem to a tractable set of single-advertiser, single ad-slot components.

The marginal-revenue approach we adopt generalizes the classic framework for allocating a single resource across many agents (Bulow and Roberts 1989; Alaei et al. 2013). For each advertiser, we construct a revenue curve, map the advertiser's valuation vector to a quantile position on that curve, and finally compute the derivative of the revenue curve at that quantile to obtain the advertiser's marginal revenue contribution to the publisher.

Implementing these steps in our setting, however, raises two challenges. First, unlike single-resource allocation, we must allocate multiple slots available at the same time, and slots may correspond to multiple possible inventory types (with the realized type observed only by the publisher). As a result, the single-advertiser problem that forms the building block of our approach appears multidimensional—both across inventory types and across slots. If the revenue curve were genuinely multidimensional, publisher revenue can no longer be expressed as a simple sum of marginal revenues, and the classical marginal-revenue approach breaks down (Alaei et al. 2013). We show, however, that the single-advertiser problem depends only on the highest slot available–the slot with the highest click-through rate. As a result, the problem can also be decoupled across slots by considering them sequentially in order of their effects on click-through rate, from highest to lowest. Moreover, because revenue curves are constructed ex ante, before the realization of the inventory type or advertiser valuations, we integrate over inventory types and obtain a one-dimensional revenue curve for each advertiser for each slot.

Second, in a single-resource allocation case (Alaei et al. 2013), each agent is described by a one-dimensional type, yielding a natural ordering of types and a natural quantile mapping: the quantile of a valuation is simply the mass of higher valuations under the prior. The advertiser types in our setting, however, are multidimensional valuation vectors. The absence of a natural ordering among multidimensional types makes it difficult to construct a quantile mapping analogous to the single-resource case. We resolve this by exploiting the publisher's private information. For any slot, which corresponds to a single inventory type observed only by the publisher, we can order advertiser types by their induced probability of receiving that slot under the prior distribution of advertiser valuations. Stronger types—those with higher allocation probabilities—receive lower quantiles. This yields a coherent inventory-type-specific quantile mapping from valuation vectors to positions on the revenue curve.

With revenue curves and quantile mappings in hand, and with advertisers truthfully reporting their valuation vectors (by incentive compatibility), the publisher computes each advertiser's marginal revenue contribution and allocates slots greedily in order of decreasing slot effect α_s, assigning each slot to the advertiser with the highest marginal revenue for that slot. Payments are then uniquely determined by the allocation rule via Myerson's payment identity (Myerson 1981), ensuring incentive compatibility across advertisers.

The remainder of this section formalizes these steps: Section 3.4.1 sets up the ex-ante constrained single-advertiser problems, Section 3.4.2 defines the revenue curve, Section 3.4.3 specifies the valuation-to-quantile mapping, Section 3.4.4 specifies the marginal-revenue allocation rule, and Section 3.4.5 derives the payment formula.

3.4.1 Single Advertiser Mechanisms. We begin with a single-advertiser version of the problem that forms the building block of our approach. With a single advertiser, the mechanism reduces to a menu of lotteries over inventory types, implementing second-degree price discrimination and illustrating how the mechanism exploits heterogeneity in valuations across types. The marginal-revenue approach then extends this logic to the multi-advertiser setting.

Consider a focal advertiser a_0 with quality effect γ_{a_0} on the click-through rate (CTR). Let $\tilde{S} \subset [S]$ denote the set of available slots. Among these, let s_0 be the slot with the highest slot effect α_{s_0}.

The publisher's problem is to select allocations $x_{a_0 st}\ \forall s \in \tilde{S},\ \forall t$ and an expected payment m_{a_0} in order to maximize revenue, subject to incentive compatibility (IC), individual rationality (IR), and the feasibility constraint that the advertiser cannot be assigned more than one slot.

It is well-known that any single-agent IC and IR mechanism can be represented as a *menu of lottery pricings* (Rochet 1985). Each menu item k—a lottery pricing—specifies probabilistic allocations across slots and inventory types, $\chi_{st}^{(k)}\ \forall s \in \tilde{S}, \forall t$, together with a payment $m^{(k)} = \alpha_{s_0}\gamma_{a_0}\mu^{(k)}$.[7]

Given its realized valuation vector $\mathbf{v}_{a_0}$, the advertiser chooses the lottery pricing that maximizes its utility. If the advertiser chooses item k, then $x_{a_0 st} = \chi_{st}^{(k)}\ \forall s, t$, and $m_{a_0} = m^{(k)}$. The advertiser's utility for lottery pricing k is thus given by

$$u_{a_0}(\mathbf{v}_{a_0}) = \mathbb{E}\left[\sum_{t,s} p_t\, \alpha_s \beta_t \gamma_{a_0} v_{a_0 t}\, \chi_{st}^{(k)} \ -\ \alpha_{s_0}\gamma_{a_0}\mu^{(k)}\right]. \tag{2}$$

Because the advertiser always selects the option that maximizes its utility, any menu of lottery pricings is incentive compatible. Moreover, because participation is voluntary, the menu always includes a "null lottery" that yields zero allocations and zero payment, ensuring individual rationality. The taxation principle (Rochet 1985) further establishes the converse: any single-agent mechanism that satisfies IC and IR can be equivalently represented as a menu of lottery pricings. Thus, without loss of generality, we can restrict attention to mechanisms of this form.

The single-advertiser mechanism can be simplified further. Intuitively, this is because, any lottery that assigns positive probability to lower slots can be replaced by an equivalent lottery that assigns probability only to the highest available slot s_0, with the probability rescaled by the ratio of slot effects so that the advertiser's CTR-weighted expected allocation is unchanged. Keeping the payment the same then leaves advertiser utility unchanged for every valuation vector, and therefore leaves

[7] We express the payment in this scaled form for convenience; multiplying by $\alpha_{s_0}\gamma_{a_0}$ is without loss of generality.

publisher revenue unchanged. The following proposition formalizes this idea.

Proposition 1. *For any mechanism M that includes a lottery pricing k with $\chi^{(k)}_{\tilde{s}\tilde{t}} > 0$ for some $\tilde{s} \neq s_0$, $\tilde{t}$, there exists another mechanism M' in which $\chi^{(k)}_{st} = 0 \; \forall s \neq s_0 \; \forall t$ such that the publisher's expected revenue under M' is equal to that under M.*

Proof. See Web Appendix A. □

Hence, without loss of generality, we can restrict attention to menus where each lottery pricing specifies only allocations for slot s_0 across inventory types $\chi_t \; \forall t$ and a corresponding payment $\alpha_{s_0}\gamma_{a_0}\mu^{(k)}$.

The single-advertiser mechanism is therefore fully characterized by allocations to a single slot—the highest available one. This simplification decouples the construction of revenue curves across slots, allowing us to construct the revenue curve for each advertiser one slot at a time, beginning with the highest slot and proceeding downward.

Formally, fix advertiser a_0 and let s_0 denote the highest available slot. The single-advertiser problem is to design a menu of lottery pricings—choosing both the menu length K and lotteries $\{\chi^{(k)}, \mu^{(k)}\}_{k\in[K]}$—to maximize expected publisher revenue. Given such a menu, the advertiser draws a valuation vector $\boldsymbol{v}_{a_0} \sim F_{a_0}$ and selects the lottery that maximizes its utility as defined in Equation 2. Let ξ_k denote the probability (under F_{a_0}) that the advertiser selects lottery pricing $k \in [K]$. These choice probabilities are induced by the menu and therefore are not decision variables.

If the advertiser selects lottery k, the publisher receives revenue $\alpha_{s_0}\gamma_{a_0}\mu^{(k)}$. The single-advertiser problem can therefore be written as

$$\max_{K,\{\chi^{(k)},\mu^{(k)}\}} \; \alpha_{s_0}\gamma_{a_0}\sum_{k=1}^{K}\xi_k\mu^{(k)}. \tag{3}$$

The ex-ante probability that slot s_0 is allocated under a menu of length K equals

$$\sum_{k=1}^{K}\xi_k\left(\sum_t p_t\chi_t^{(k)}\right).$$

The ex-ante constrained single-advertiser problem is identical to the unconstrained problem above, except that we impose an upper bound q on this ex-ante allocation probability. The constrained problem is therefore

$$\max_{K,\{\chi^{(k)},\mu^{(k)}\}} \alpha_{s_0}\gamma_{a_0}\sum_{k=1}^{K}\xi_k\mu^{(k)} \quad \text{subject to} \quad \sum_{k=1}^{K}\xi_k\left(\sum_t p_t\chi_t^{(k)}\right) \le q. \tag{4}$$

Solving this constrained problem for all $q \in [0,1]$ forms the basis of our marginal revenue approach: the resulting revenue as a function of q defines the advertiser's revenue curve.[8]

3.4.2 Revenue Curves. Solving the constrained single-advertiser problem in Equation 4 for all possible values of q maps how an advertiser's contribution to revenue evolves as its allocation probability increases. When performed for all advertisers, these revenue curves capture the trade-off on revenues inherent in raising one advertiser's allocation probability relative to others.

Formally, let $\mathcal{M}$ denote the class of mechanisms over which we optimize the single-advertiser problem in Equation 4. We require $\mathcal{M}$ to be closed under convex combinations[9] for the rest of the analysis. The revenue curve for advertiser a and slot s is defined as the optimal revenue $R_{as}(q)$ from this problem under allocation constraint $q \in [0,1]$.

The following proposition establishes that the revenue curve is non-decreasing and concave. This result is consistent with standard economic intuition: as the allocation probability increases, the mechanism exhibits diminishing returns, so marginal revenues decline with q.

Proposition 2. *When the class of mechanisms $\mathcal{M}$ is closed under convex combinations, the revenue curve $R_{as}(q)$ is non-decreasing and concave in q. Consequently, the marginal revenue $R'_{as}(q)$ is non-negative and weakly decreasing in q.*

Proof. See Web Appendix A. □

The following observation is an immediate consequence of Proposition 2. Since $R_{as}(q)$ is non-decreasing and concave on $[0,1]$, there exists a threshold $q^*_{as} \in [0,1]$ such that $R_{as}(q)$ is strictly increasing for $q < q^*_{as}$ and constant for $q \ge q^*_{as}$. Moreover, for all $q < q^*_{as}$, the allocation constraint is binding, i.e., the ex-ante probability of allocation of slot s is equal to q.

[8]Because revenue curves are constructed ex-ante, the allocation constraint averages not only over the advertiser's type distribution but also over inventory types. Once inventory type and advertiser valuations are realized, the marginal revenue mechanism computes marginal revenues by mapping each advertiser's realized valuation vector to a position on its revenue curve, conditional on the realized inventory type.

[9]A convex combination of mechanisms $M_1, \ldots, M_n$ is defined as selecting mechanism M_i with probability λ_i, where $\lambda_i \ge 0$ for all i and $\sum_{i=1}^{n}\lambda_i = 1$.

Concavity of the revenue curve is crucial because it provides a consistent way to map an advertiser's type (valuation vector) to a position on the revenue curve. Concavity implies that marginal revenue decreases with q. This property allows advertiser types to be ordered by their probability of receiving a slot and enables the definition of a quantile mapping. Stronger types—those more likely to be allocated a slot—are assigned to lower quantiles, which correspond to higher marginal revenues. This relationship is consistent with the logic of marginal-revenue maximization, where higher-valued types are allocated with higher probability. Thus, the monotonicity of marginal revenue implied by concavity ensures that the quantile mapping is both coherent and economically interpretable.

The revenue curve also exhibits additional structure that simplifies the marginal revenue approach.

Proposition 3. *For any mechanism class $\mathcal{M}$, the revenue curve factorizes as*

$$R_{as}(q) = \alpha_s \gamma_a \, \Phi_a(q),$$

where $\Phi_a(q)$ depends only on advertiser a's valuation distribution and not on slot s.

Proof. See Web Appendix A. $\square$

The term $\Phi_a(q)$ is obtained by solving the *slot-normalized* version of the single advertiser problem in Equation 4. It is given by

$$\Phi_a(q) \; = \max_{K, \{\chi^{(k)}, \mu^{(k)}\}} \sum_{k=1}^{K} \xi_k \, \mu^{(k)} \quad \text{s.t.} \quad \sum_{k=1}^{K} \xi_k \left(\sum_{t} p_t \, \chi_t^{(k)} \right) \leq q. \tag{5}$$

We therefore refer to $\Phi_a(q)$ as the advertiser's normalized revenue curve henceforth.

The factorization of the revenue curve into slot and advertiser components has two important implications. First, revenue curves for a given advertiser across different slots are simply scaled versions of one another, with slot effects entering only multiplicatively. As a consequence, when computing marginal revenues from advertiser types, the mapping to a quantile q on the revenue curve depends only on $\Phi_a(q)$, which is independent of the slot. Hence the quantile mapping is itself slot-independent.

Second, marginal revenues take the form $R'_{as}(q) = \alpha_s \gamma_a \Phi'_a(q)$. For a given advertiser, this expression implies that marginal revenue is higher for higher slots. Thus, to maximize marginal revenue, it is natural to allocate in order of slots from highest to lowest. At the same time, for a fixed slot, advertisers' marginal revenues are ranked by $\gamma_a \Phi'_a(q)$. This parallels the generalized second-price auction, where advertisers are ordered by the product of a quality score and a bid. In our formulation, γ_a plays the role of the quality score, while $\Phi'_a(q)$ can be interpreted as a virtual value—the derivative of the scaled revenue curve—similar to Myerson's optimal auction for a single item (Myerson 1981; Bulow and Roberts 1989). We refer to $\Phi'_a(q)$ as the advertiser's normalized marginal revenue or virtual value henceforth.

The revenue curves and their factorization therefore simplify the publisher's problem by decoupling advertisers and slots: a single normalized revenue curve characterizes each advertiser, while slot effects enter only as multiplicative scalars.

3.4.3 Quantile Mapping. The single-advertiser mechanisms that define revenue curves are constructed ex-ante, before advertiser types are realized. Once advertisers observe their realized valuation vectors, one can compute each advertiser's marginal revenue contribution. This is done by mapping each advertiser's valuation vector $\boldsymbol{v}_a$ to a quantile position $q_a \in [0, 1]$ on the advertiser's revenue curve.

Let $M_a(q)$ denote the mechanism (menu of lottery pricings) that solves the slot-normalized single-advertiser problem in Equation 5 under allocation constraint q, achieving the normalized revenue $\Phi_a(q)$.

To build intuition, in this section, we begin with the special case where the set of mechanisms $\{M_a(q)\}_{q\in[0,1]}$ exhibits additional structure. First, assume that allocations are deterministic—each advertiser type either receives an inventory type with probability 1 or with probability 0. Formally, for every menu option k, the allocation vector $\chi^{(k)} \in \{0, 1\}^T$ is binary across inventory types. Second, assume that these mechanisms satisfy nestedness, meaning that if $q_1 \leq q_2$, then the set of advertiser types that could be allocated under $M_a(q_1)$ is a subset of those that could be allocated under $M_a(q_2)$ for all inventory types t. Intuitively, higher value advertiser types will remain allocated when the allocation set is expanded to lower valuation advertiser types.

This assumption of nested deterministic allocations is strong: in general, the single-advertiser mechanisms $M_a(q)$ may involve randomization and need not generate nested sets of allocated types across values of q. Nevertheless, analyzing this special case is instructive, as it highlights the structure of the quantile mapping that extends to the general setting. The more general construction, without relying on nested deterministic allocations, is presented in Web Appendix B.

Assuming nested deterministic allocations for the moment, we now construct a mapping $q_{at} = Q_{at}(\boldsymbol{v}_a)$ from the advertiser's valuation vector to a position on its revenue curve. For an inventory type t, let $\Omega_{at}(q)$ denote the set of advertiser types allocated by the mechanism $M_{at}(q)$. Consider a small increase dq in the allocation constraint. The additional types allocated are those in $\Omega_{at}(q + dq)\backslash\Omega_{at}(q)$, and the corresponding increase in normalized revenue is $\Phi_a(q + dq) - \Phi_a(q) = \Phi'_a(q)dq$. Thus, all types in $\Omega_{at}(q+dq)\backslash\Omega_{at}(q)$ can be interpreted as having quantile q contributing a normalized marginal revenue of $\Phi'_a(q)$.

This yields a natural quantile assignment for any advertiser type $\boldsymbol{v}_a$: when the realized inventory type is t, the advertiser's quantile is defined as the smallest q for which it is allocated the inventory type,

$$q_{at} = Q_{at}(\boldsymbol{v}_a) = \inf\{q \in [0, 1] : \boldsymbol{v}_a \in \Omega_{at}(q)\}.$$

If an advertiser type $\boldsymbol{v}_a$ is never allocated for any $q < 1$, we set $q_{at} = 1$.

In theory, we could construct an advertiser valuation-to-quantile mapping separately for each slot s using the corresponding revenue curve $R_{as}(q)$. However, the slot-normalized and original single-advertiser problems in Equations 4 and 5 differ only by the multiplicative factor $\alpha_s\gamma_a$. Consequently, the mechanism $M_a(q)$ that achieves the normalized revenue $\Phi_a(q)$ also achieves $R_{as}(q) = \alpha_s\gamma_a\Phi_a(q)$ for any slot s. Because the underlying mechanism is identical across slots (up to this multiplicative scaling), the induced quantile mapping is likewise slot-independent and coincides with the mapping derived above.

The quantile mapping derived under nested deterministic allocations has the following property: for any convex combination of mechanisms $\{M_a(q)\}$, advertiser types that are more likely to be allocated (for a given inventory type) are assigned lower quantiles and therefore correspond to higher marginal revenues. In our full setting—where allocations in $M_a(q)$ may be randomized and need not be nested—we show that the advertiser's outcomes can still be represented as a well-defined convex combination of these mechanisms $\{M_a(q)\}$. This mixture is what we refer to as the *interim mechanism*; it captures the expected allocation probability that each advertiser type receives, where the expectation is taken over the distribution of competing advertisers' types. The general quantile mapping we construct (Web Appendix B) preserves the same ordering property for this interim mechanism: types with higher interim allocation probabilities are mapped to weakly lower quantiles. Alaei et al. (2013) show that satisfying exactly this property is sufficient for implementing marginal-revenue maximization in single-resource allocation problems. Our construction therefore extends the key structural feature of the nested deterministic case to the more general environment relevant for our setting.

3.4.4 Slot Allocations. With the advertiser revenue curves and valuation-to-quantile mappings in place, we now describe how the publisher allocates slots once all private information—inventory type and advertiser types—has been realized. Recall that revenue curves and quantile mappings are constructed ex-ante, before types are known. At the time of the auction, advertisers observe their valuation vectors, the publisher observes the realized inventory type, and—by incentive compatibility—advertisers truthfully report their valuations $\boldsymbol{v}_a$. For each advertiser a, the publisher computes the corresponding quantile $q_{at} = Q_{at}(\boldsymbol{v_a})$ using the inventory-type–specific mapping derived in Section 3.4.3. The marginal revenue of advertiser a for slot s is then $R'_{as}(q_{at}) = \alpha_s \gamma_a \Phi'_a(q_{at})$.

To maximize the total sum of marginal revenues across all slots, we use the observation established earlier: the marginal revenues across different slots differ only by the slot multiplier α_s. Consequently, a higher slot always yields a higher marginal revenue for any advertiser. This structure enables a greedy allocation rule: process slots in decreasing order of α_s, assigning each slot to the advertiser with the highest marginal revenue for that slot.

Next, for a given slot s, marginal revenues across advertisers are ordered by $\gamma_a \Phi'_a(q_{at})$. Therefore, to maximize the total sum of marginal revenues, we simply rank advertisers by this product and assign slots in that order until either all slots or all advertisers are exhausted, excluding any advertisers whose marginal revenues are non-positive.

This allocation rule closely parallels the generalized second-price (GSP) auction, in which advertisers are ranked by the product of their quality score and bid and assigned slots from highest to lowest. In our mechanism, the role of the bid is replaced by the advertiser's virtual value $\Phi'_a(q_{at})$, which is derived from the marginal revenue framework and depends on the advertiser's full valuation vector and the realized inventory type. The result is an allocation rule that shares the intuitive structure of GSP while improving on its revenue properties.

This allocation rule highlights a key distinction from standard mechanisms such as GSP. The quantile—and hence the marginal revenue—assigned to an advertiser depends on its entire valuation vector across inventory types, rather than on a single scalar bid. As a result, allocations are based on a richer set of information reflecting how the advertiser values different types of impressions. Moreover, because the quantile mapping is inventory-type-specific, the publisher's private information of the realized inventory type plays a central role in determining marginal revenues and assignments. This structure gives the publisher greater flexibility in allocating slots than mechanisms that rely on one-dimensional bids.

3.4.5 Payments. With the allocation rule in place, the final step is to determine payments for the advertisers who receive slots. Because the mechanism must satisfy Bayesian incentive compatibility, the payment of each winning advertiser is uniquely determined by the allocation rule (Myerson 1981).

Under the nested deterministic allocations case discussed in Section 3.4.3, the payment rule takes a particularly transparent form. Consider a winning advertiser a and the slot s it receives. Because the allocation rule assigns slots in decreasing order of marginal revenue, advertiser a receives slot s if and only if the advertiser's quantile q_{at} is below some threshold. Define the critical quantile for advertiser a as

$$\hat{q}_{at} = \sup\{q \in [0,1] : a \text{ would still receive slot s if its quantile is q}\}.$$

Let $\tilde{a}$ be the advertiser immediately below a in the marginal revenue ordering occupying slot $s+1$. The critical quantile for a is the largest quantile such that the marginal revenue of a for slot s is equal to the marginal revenue of $\tilde{a}$ for slot $s+1$. Formally, $\hat{q}_{at}$ satisfies

$$\gamma_a \Phi'_a(\hat{q}_{at}) = \gamma_{\tilde{a}} \Phi'_{\tilde{a}}(q_{\tilde{a}t}).$$

If a is the last winning advertiser and there is no advertiser below, then the critical quantile is instead the value at which the marginal revenue becomes zero, that is, when

$$\gamma_a \Phi'_a(\hat{q}_{at}) = 0.$$

In either case, the critical quantile identifies the boundary at which the advertiser transitions from receiving the slot to losing it.

The payment for advertiser a is then given by the single-advertiser mechanism with allocation constraint $\hat{q}_{at}$, i.e., the mechanism $M_a(\hat{q}_{at})$. Recall from the single-advertiser problem that any single advertiser mechanism—including $M_a(\hat{q}_{at})$—is implemented as a menu of lottery pricings, each specifying a vector of allocation probabilities across inventory types and an associated expected payment. For a realized type $\mathbf{v}_a$, the advertiser chooses the lottery that maximizes its utility. Let χ_{at} denote the allocation probability for inventory type t by the lottery pricing that advertiser a selects under $M_a(\hat{q}_{at})$, and let $\alpha_s\gamma_a\mu_a$ denote the associated expected payment when the realized type is $\mathbf{v}_a$. Because advertiser a is allocated slot s with certainty in our multi-advertiser mechanism, one natural implementation of her expected payment is to charge the amount

$$\frac{\alpha_s\gamma_a\,\mu_a}{\chi_{at}}$$

conditional on allocation. Scaling by $1/\chi_{at}$ ensures that the expected payment exactly matches the payment prescribed by $M_a(\hat{q}_{at})$. If payments are instead collected per click—the most common practice in sponsored search—we may divide by the realized click-through rate $\alpha_s\beta_t\gamma_a$. This yields a per-click payment of

$$\frac{\mu_a}{\chi_{at}\,\beta_t}$$

which again preserves the expected payment dictated by the single-agent mechanism while implementing it in the operational form used in practice.

This payment via critical quantiles is uniquely determined by the payment identity for incentive-compatible mechanisms (Myerson 1981) and follows the standard argument used throughout marginal-revenue mechanisms. It is directly analogous to the second-price auction where the winner pays the smallest value at which it would still win; here, the winner pays according to the largest (weakest) quantile at which it would still retain her allocated slot. The critical quantile therefore plays the same role as the critical value in single-dimensional settings.

In the general case, where the single-advertiser mechanisms $\{M_a(q)\}$ need not have nested or deterministic allocations, payments are determined by the advertiser's interim mechanism, as derived in Web Appendix B. Despite the additional technical steps required, the same principle applies: the payment is the unique transfer consistent with incentive compatibility and the marginal-revenue allocation rule.

3.4.6 Summary of IBPA. To summarize, the proposed mechanism proceeds in two stages. Ex-ante, for each advertiser a, solve the slot-normalized single-advertiser problem in Equation 5 for all $q \in [0, 1]$ to obtain the normalized revenue curve $\Phi_a(q)$. By Proposition 3, the revenue curve for slot s is $R_{as}(q) = \alpha_s \gamma_a \, \Phi_a(q)$.

Then, for each auction, after private information is realized, compute quantiles q_{at} using the inventory-type–specific mapping in Section 3.4.3. Compute marginal revenues $R'_{as}(q_{at}) = \alpha_s \gamma_a \Phi'_a(q_{at})$ and allocate slots greedily in decreasing order of α_s, ranking advertisers by $\gamma_a \Phi'_a(q_{at})$ and assigning each slot to the highest-ranked advertiser with positive marginal revenue (Section 3.4.4). Finally, for each winning advertiser, compute the critical quantile and charge the payment implied by the single-agent mechanism $M_a(\hat{q}_{at})$, as described in Section 3.4.5 (with the general case handled in Web Appendix B).

4 Theoretical Analysis

We now present our main theoretical results. A key strength of our mechanism arises from treating the inventory type as private information to the publisher while allowing advertisers to submit multidimensional bids—one valuation for each possible inventory type. This creates the opportunity for the publisher to implement second-degree price discrimination across inventory types, analogous to the mixed-bundling intuition for the single-advertiser mechanism.

This raises two natural questions about the robustness of the mechanism's revenue performance: (1) Information granularity: Should the publisher coarsen the set of inventory types it collects and internally distinguishes? This question is practically relevant because privacy regulations, data-collection restrictions, or modeling constraints may force the publisher to merge or suppress certain attributes and (2) Disclosure granularity: Should the publisher share any information with advertisers about the realized inventory type? This asks whether revealing partial targeting information (e.g., disclosing audience attributes) can increase competition sufficiently to boost revenue.

Our first main result shows that the publisher benefits from maximizing information granularity while minimizing disclosure granularity. Revenue increases when the publisher collects more refined inventory-type information and decreases when any part of that information is disclosed to advertisers. Intuitively, the publisher gains the most revenue when it has rich information privately and allows advertisers to bid on all possible inventory types without learning which specific type has realized—thus maximizing information asymmetry and enabling the richest form of second-degree price discrimination.

Our second main result shows that the proposed marginal-revenue mechanism dominates the widely used generalized second-price (GSP) auction in expected publisher revenue, under any combination of information and disclosure granularity.

Following these results, we analyze the computational aspects of implementing our mechanism, introduce tractable approximation classes, and show how the granularity and dominance results extend to these approximations.

4.1 Information and Disclosure Granularity

Following Rafieian and Yoganarasimhan (2021), we formalize granularity using the idea of a partition of the inventory types. This allows us to describe both information granularity, which captures how finely the publisher internally distinguishes inventory types, and disclosure granularity, which captures how much of this information is revealed to advertisers before bidding.

An information regime is defined by a partition $\mathcal{P}_{\text{info}}$ of the set of inventory types $[T]$ into I disjoint blocks. Let $g_{\text{info}} : [T] \to [I]$ denote the associated block-membership function. Under this regime, the publisher observes only the information block $g_{\text{info}}(t)$ to which the realized inventory type t belongs. Advertisers therefore cannot condition their bids on distinctions finer than these information blocks. An advertiser's valuation for a block i is the expected valuation over inventory types in that block. This formulation also captures scenarios in which privacy or data-collection constraints prevent the publisher from observing the more granular inventory types.

We compare the granularity of two partitions using set refinement (Rafieian and Yoganarasimhan 2021). Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be two partitions of $[T]$. We say that $\mathcal{P}_1$ is at least as granular as $\mathcal{P}_2$, i.e., $\mathcal{P}_1 \succeq \mathcal{P}_2$ if for every block $B \in \mathcal{P}_1$, there exists a block $C \in \mathcal{P}_2$ such that $B \subseteq C$. That is, $\mathcal{P}_1$ is more granular than $\mathcal{P}_2$ whenever it splits the inventory types more finely.

Similar to an information regime, a disclosure regime is a partition $\mathcal{P}_{\text{disc}}$ of $[T]$ into D disjoint blocks with membership function $g_{\text{disc}} : [T] \to [D]$. A pair $(\mathcal{P}_{\text{info}}, \mathcal{P}_{\text{disc}})$ is feasible if the publisher does not disclose more information than it observes internally, i.e., $\mathcal{P}_{\text{info}} \succeq \mathcal{P}_{\text{disc}}$. Under feasibility, each disclosure block is a union of information blocks.

When an impression of type t arrives, the publisher observes its information block $g_{\text{info}}(t)$ but discloses only the coarser disclosure category $g_{\text{disc}}(t)$ to advertisers. Advertisers may then bid only at the level of information blocks contained in that disclosure block. Let $\mathcal{I}(d) = \{i \in \mathcal{P}_{\text{info}} : i \subseteq d\}$ denote the set of information blocks contained within the disclosed block d. Advertiser a thus submits bids $\{b_a(i) : i \in \mathcal{I}(d)\}$. Under incentive compatibility, the bid for an information block is equal to the expected valuation over inventory blocks in that block.

The mechanism described in Section 3 corresponds exactly to the full-information / null-disclosure pair $(\mathcal{P}^{\text{full}}_{\text{info}}, \mathcal{P}^{\text{null}}_{\text{disc}})$. Under full information regime, the publisher observes the exact inventory type of each impression $\mathcal{P}^{\text{full}}_{\text{info}} = \{\{t\} : t \in [T]\}$, while under null disclosure granularity, the publisher reveals nothing about this realized type to advertisers, $\mathcal{P}^{\text{null}}_{\text{disc}} = \{[I]\}$.

The following theorem shows that the publisher revenue under IBPA is monotonically increasing in the information granularity and monotonically decreasing in the disclosure granularity.

Theorem 4. *For any $\mathcal{P}^{(1)}_{\text{info}} \succeq \mathcal{P}^{(2)}_{\text{info}}$ and for all $\mathcal{P}_{\text{disc}}$ such that both pairs $(\mathcal{P}^{(1)}_{\text{info}}, \mathcal{P}_{\text{disc}})$ and $(\mathcal{P}^{(2)}_{\text{info}}, \mathcal{P}_{\text{disc}})$ are feasible,*

$$\text{Rev}_{\text{IBPA}}(\mathcal{P}^{(1)}_{\text{info}}, \mathcal{P}_{\text{disc}}) \;\geq\; \text{Rev}_{\text{IBPA}}(\mathcal{P}^{(2)}_{\text{info}}, \mathcal{P}_{\text{disc}}).$$

Similarly, for any $\mathcal{P}^{(1)}_{\text{disc}} \succeq \mathcal{P}^{(2)}_{\text{disc}}$ and for all $\mathcal{P}_{\text{info}}$ such that both pairs $(\mathcal{P}_{\text{info}}, \mathcal{P}^{(1)}_{\text{disc}})$ and $(\mathcal{P}_{\text{info}}, \mathcal{P}^{(2)}_{\text{disc}})$ are feasible,

$$\text{Rev}_{\text{IBPA}}(\mathcal{P}_{\text{info}}, \mathcal{P}^{(1)}_{\text{disc}}) \;\leq\; \text{Rev}_{\text{IBPA}}(\mathcal{P}_{\text{info}}, \mathcal{P}^{(2)}_{\text{disc}}).$$

Proof. See Web Appendix A. □

Taken together, the two monotonicity results imply that the publisher attains its maximum revenue under the marginal-revenue mechanism when operating under full information and null disclosure. The main insight underlying the proof is that reducing information granularity restricts the feasible set of single-advertiser mechanisms. Any single-advertiser mechanism designed under a coarser information partition can be replicated under a finer partition simply by imposing additional equality constraints across the types that have been merged. Thus, the feasible set under low information granularity is a subset of the feasible set under high granularity, implying revenue cannot increase when information granularity is reduced. An analogous argument establishes the disclosure result: increasing disclosure granularity imposes additional constraints on how allocations and payments can depend on the realized inventory type, thereby shrinking the feasible set and weakly reducing revenue.

4.1.1 Granularity in GSP vs. IBPA. Theorem 4 stands in contrast to prior work (Levin and Milgrom 2010; Yao and Mela 2011; Ada, Abou Nabout, and Feit 2022; Rafieian and Yoganarasimhan 2021) which shows that publisher revenue under (generalized) second-price auctions is not necessarily monotone in granularity. The reason is structural: under GSP, advertisers submit one-dimensional bids, which sharply limits the publisher's ability to perform effective second-degree price discrimination in the associated single-advertiser problems.

The distinction becomes clear when comparing the levers available in our mechanism versus GSP. In our mechanism, the publisher controls both information granularity and disclosure granularity, and these two levers can be adjusted independently. In particular, the publisher can simultaneously operate with full information granularity (distinguishing all inventory types internally) and null disclosure granularity (revealing none of this information), because advertisers submit multidimensional bids.

Under GSP, advertisers submit a single bid rather than a multidimensional vector. This bid reflects the advertiser's expected valuation for the disclosed block, which may include multiple information blocks. Consequently, if the publisher chooses null disclosure, the advertiser's bid must be based on the average valuation across all information blocks—making null disclosure effectively equivalent to null information. More generally, information and disclosure granularity are coupled under GSP: disclosing coarser blocks forces advertisers to collapse their multidimensional valuations into a single expectation, eliminating the publisher's ability to price discriminate across inventory types. Because these levers cannot be varied independently, the GSP revenue-maximizing regime need not occur at the extremes (full information or null disclosure), and may instead arise at intermediate combinations of the two, as documented by Levin and Milgrom (2010) and Rafieian and Yoganarasimhan (2021).

Levin and Milgrom (2010) interpret this non-monotonicity as arising from a trade-off between the thickness of competition and the publisher's ability to match advertisers with high-value inventory types. We offer a complementary lens: the non-monotonicity is fundamentally a consequence of the fact that GSP does not implement second-degree price discrimination over multidimensional valuations. Because GSP forces advertisers to compress their valuations into a single bid, richer information about inventory types cannot be fully exploited to extract revenue. In contrast, our mechanism solicits multidimensional bids and enables the publisher to leverage heterogeneity in valuations across inventory types through more flexible, type-contingent pricing. As a result, the granularity levers in our mechanism are uncoupled, and revenue is monotone in the direction predicted by second-degree price discrimination.

4.2 Revenue Comparison to GSP

Here, we show that the proposed IBPA mechanism dominates the widely used generalized second-price (GSP) auction in expected publisher revenue. While GSP is typically implemented under full information and full disclosure, our comparison is stronger: we allow GSP to operate under any combination of information and disclosure granularity. Even under this expanded set of regimes, our mechanism generates higher revenue. The following theorem formalizes this result.

Theorem 5. *For every feasible pair* $(\mathcal{P}_{\text{info}}, \mathcal{P}_{\text{disc}})$*,*

$$\text{Rev}_{\text{IBPA}}(\mathcal{P}_{\text{info}}^{\text{full}}, \mathcal{P}_{\text{disc}}^{\text{null}}) \geq \text{Rev}_{\text{GSP}}(\mathcal{P}_{\text{info}}, \mathcal{P}_{\text{disc}}).$$

Proof. The proof proceeds in two steps. First, we show that for GSP, it suffices to analyze regimes in which the publisher's information and disclosure partitions coincide, i.e., regimes of the form $(\mathcal{P}, \mathcal{P})$. Under a feasible GSP regime $(\mathcal{P}_{\text{info}}, \mathcal{P}_{\text{disc}})$, advertisers submit a single bid for each disclosed block. As a result, their bids depend only on the coarser disclosure partition $\mathcal{P}_{\text{disc}}$ and not on any finer distinctions contained in $\mathcal{P}_{\text{info}}$. Within any disclosed block d, GSP aggregates the advertiser's valuations over all inventory types in that block and treats them as a single unit. Therefore, any GSP regime $(\mathcal{P}_{\text{info}}, \mathcal{P}_{\text{disc}})$ is outcome-equivalent to the regime $(\mathcal{P}_{\text{disc}}, \mathcal{P}_{\text{disc}})$, obtained by collapsing the publisher's information partition down to the disclosure partition and disclosing exactly that partition. In the sense that both regimes induce the same allocation and payment rules for every profile of valuations, we can restrict attention, without loss of generality, to GSP operating under regimes of the form $(\mathcal{P}, \mathcal{P})$.

Second, fix an arbitrary partition $\mathcal{P}$, and consider IBPA under the regime $(\mathcal{P}, \mathcal{P})$. Because the publisher fully discloses its information in this regime, each advertiser is characterized by a single parameter—its valuation for the block of $\mathcal{P}$ containing the realized type. This is a standard single-parameter environment with quasi-linear preferences. As a result, the marginal-revenue approach results in the unique revenue-maximizing auction over all BIC and IR mechanisms (Myerson 1981; Bulow and Roberts 1989). In particular,

$$\mathrm{Rev}_{\mathrm{IBPA}}(\mathcal{P},\mathcal{P}) \geq \mathrm{Rev}_{\mathrm{GSP}}(\mathcal{P},\mathcal{P}).$$

Finally, by Theorem 4, publisher revenue under IBPA is weakly increasing in information granularity and weakly decreasing in disclosure granularity. Applying this result to the full-information / null-disclosure regime $(\mathcal{P}^{\mathrm{full}}_{\mathrm{info}}, \mathcal{P}^{\mathrm{null}}_{\mathrm{disc}})$ yields

$$\mathrm{Rev}_{\mathrm{IBPA}}(\mathcal{P}^{\mathrm{full}}_{\mathrm{info}}, \mathcal{P}^{\mathrm{null}}_{\mathrm{disc}}) \geq \mathrm{Rev}_{\mathrm{IBPA}}(\mathcal{P},\mathcal{P}),$$

which completes the argument. □

4.3 Computational Complexity and Approximation Classes

We now discuss the computational challenges of solving the ex-ante constrained single-advertiser problem. Solving the ex-ante constrained single-advertiser problem in Section 3 requires optimizing over the full class $\mathcal{M}$ of arbitrary lottery pricings, where each lottery specifies a probability vector $\chi \in [0,1]^T$ and an associated expected payment μ. While the problem is conceptually straightforward, the optimal menu can have arbitrary length; consequently, computation can become expensive when the number of inventory types T is large (Manelli and Vincent 2007; Hart and Nisan 2013).

To address this complexity, we consider two structured families of mechanisms that approximate $\mathcal{M}$ while retaining strong revenue guarantees. The first restricts allocation vectors to be binary, reducing the menu to a collection of bundles. The second further restricts payments to follow a simple additive form. Both classes retain strong performance guarantees by preserving the core price-discrimination structure of the full mechanism while also simplifying computation. We also discuss how the main theoretical results in Sections 4.1 and 4.2 extend to these approximations.

4.3.1 Binary-Allocation Menus ($\mathcal{M}^{\mathrm{bin}}$). Our first approximation solves the single-advertiser problem by restricting attention to menus of lottery pricings with binary allocation vectors. Formally, define

$$\mathcal{M}^{\text{bin}} = \{(\chi, \mu) \in \mathcal{M} : \chi_t \in \{0, 1\}\ \forall t\}.$$

A lottery in $\mathcal{M}^{\text{bin}}$ therefore corresponds to offering the advertiser a subset of inventory types together with an associated payment: the advertiser receives exactly those types for which $\chi_t = 1$. To ensure individual rationality, we normalize the payment for the zero-allocation vector to be $\mu(\mathbf{0}) = 0$.

Since there are at most 2^T binary allocation vectors and the zero vector is normalized, solving the optimal single-advertiser mechanism for any ex-ante allocation constraint q requires computing at most 2^{T-1} payments, one for each non-zero allocation vector. Although this growth is exponential in T, it is far more tractable than optimizing over the full mechanism class $\mathcal{M}$, whose optimal menus may require infinitely many (non-binary) lotteries.

Mechanisms in $\mathcal{M}^{\text{bin}}$ are closely related to the classic mixed bundling approach in multi-product monopoly pricing. Each binary allocation vector corresponds to a bundle of inventory types, and the mechanism offers a menu of such bundles with associated prices to implement second-degree price discrimination.

We denote the marginal-revenue extension of this restricted single-advertiser class by IBPA^{bin}.

Despite its simplicity, $\mathcal{M}^{\text{bin}}$ enjoys strong approximation guarantees for the single-advertiser problem: when valuations across inventory types are drawn independently (i.e., from a product distribution), the optimal mechanism in $\mathcal{M}^{\text{bin}}$ achieves a constant-factor approximation to the optimal revenue in $\mathcal{M}$, and under arbitrary correlations, it achieves an $O(\log T)$ approximation (see Proposition 8).

The following propositions show that the theoretical results from Sections 4.1 and 4.2 continue to hold for IBPA^{bin}.

Proposition 6. *For any $\mathcal{P}_{\text{info}}^{(1)} \succeq \mathcal{P}_{\text{info}}^{(2)}$ and for all $\mathcal{P}_{\text{disc}}$ such that both pairs $(\mathcal{P}_{\text{info}}^{(1)}, \mathcal{P}_{\text{disc}})$ and $(\mathcal{P}_{\text{info}}^{(2)}, \mathcal{P}_{\text{disc}})$ are feasible,*

$$\text{Rev}_{\text{IBPA}^{\text{bin}}}(\mathcal{P}_{\text{info}}^{(1)}, \mathcal{P}_{\text{disc}}) \ \geq\ \text{Rev}_{\text{IBPA}^{\text{bin}}}(\mathcal{P}_{\text{info}}^{(2)}, \mathcal{P}_{\text{disc}}).$$

Similarly, for any $\mathcal{P}_{\text{disc}}^{(1)} \succeq \mathcal{P}_{\text{disc}}^{(2)}$ and for all $\mathcal{P}_{\text{info}}$ such that both pairs $(\mathcal{P}_{\text{info}}, \mathcal{P}_{\text{disc}}^{(1)})$ and $(\mathcal{P}_{\text{info}}, \mathcal{P}_{\text{disc}}^{(2)})$

are feasible,

$$\mathrm{Rev}_{\mathrm{IBPA}^{\mathrm{bin}}}(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{(1)}) \;\leq\; \mathrm{Rev}_{\mathrm{IBPA}^{\mathrm{bin}}}(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{(2)}).$$

Coarsening information granularity imposes additional constraints across the types subsumed by the coarse blocks, shrinking the feasible set and thereby weakly reducing revenue. Similarly, increasing disclosure granularity imposes analogous constraints on how allocations may depend on the realized type, again shrinking the feasible set. Consequently, the result follows exactly as in the proof of Theorem 4.

Proposition 7. *For every feasible pair* $(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}})$*,*

$$\mathrm{Rev}_{\mathrm{IBPA}^{\mathrm{bin}}}(\mathcal{P}_{\mathrm{info}}^{\mathrm{full}}, \mathcal{P}_{\mathrm{disc}}^{\mathrm{null}}) \geq \mathrm{Rev}_{\mathrm{GSP}}(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}).$$

The argument mirrors the proof of Theorem 5. Under full disclosure, $\mathrm{IBPA}^{\mathrm{bin}}$ becomes a single-parameter marginal-revenue mechanism, since each advertiser has a single valuation for its information block. By Alaei et al. (2013), such a mechanism is optimal in any single-parameter environment. Its underlying single-advertiser mechanisms are simply convex combinations of deterministic menus, i.e., mechanisms within the deterministic class considered here. Hence,

$$\mathrm{Rev}_{\mathrm{IBPA}^{\mathrm{bin}}}(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{\mathrm{full}}) \geq \mathrm{Rev}_{\mathrm{GSP}}(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{\mathrm{full}}).$$

Finally, because

$$\mathrm{Rev}_{\mathrm{IBPA}^{\mathrm{bin}}}(\mathcal{P}_{\mathrm{info}}^{\mathrm{full}}, \mathcal{P}_{\mathrm{disc}}^{\mathrm{null}}) \geq \mathrm{Rev}_{\mathrm{IBPA}^{\mathrm{bin}}}(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{\mathrm{full}}),$$

the result follows.

4.3.2 Additive-Pricing Menus ($\mathcal{M}^{\mathrm{add}}$).

Our second approximation restricts payments to follow an additive structure:

$$\mu(\chi) = \rho_0 + \sum_{t\in[T]} p_t \chi_t \rho_t,$$

where ρ_t is a fixed price for targeting inventory type t and ρ_0 can be interpreted as an entry fee for participating in the mechanism. The resulting class, $\mathcal{M}^{\mathrm{add}} \subset \mathcal{M}^{\mathrm{bin}}$, retains the binary-allocation structure while imposing a simple, linear pricing rule. As before, to ensure individual rationality, we normalize the payment for the zero-allocation vector to be $\mu(\mathbf{0}) = 0$.

This structure is computationally convenient: menus need only specify a vector $(\rho_0, \rho_1, \ldots, \rho_T)$, and the resulting mechanism is implementable in time polynomial in T (Proposition 9).

We denote the marginal-revenue extension of this restricted single-advertiser class by IBPA^{add}.

Despite its simplicity, $\mathcal{M}^{\text{add}}$ achieves the same revenue-approximation guarantees as $\mathcal{M}^{\text{bin}}$ under both independent and correlated valuations (Proposition 8). In practice, $\mathcal{M}^{\text{add}}$ provides a highly scalable approximation that preserves the essential mixed-bundling logic of the full IBPA mechanism.

Proposition 8. *For the ex-ante constrained single-advertiser problem, the mechanism class $\mathcal{M}^{\text{bin}}$ of binary-allocation menus and its subclass $\mathcal{M}^{\text{add}}$ of additive-pricing menus each achieve a constant-factor approximation to the optimal revenue in $\mathcal{M}$ when valuations across inventory types are drawn independently (i.e., from a product distribution). More generally, when valuations may be arbitrarily correlated across types, both classes achieve an $O(\log T)$ approximation to the optimal revenue.*

Proof. See Web Appendix A. □

Proposition 9. *For the ex-ante constrained single-advertiser problem, optimization over the class $\mathcal{M}^{\text{add}}$ can be carried out, for any approximation tolerance $\epsilon > 0$, in time polynomial in both dimensions T and tolerance $1/\epsilon$. In particular, each step of an optimization routine can be implemented in $O(T)$ operations.*

Proof. See Web Appendix A. □

As shown in Section 4.2, when analyzing GSP it suffices to consider regimes in which the publisher's information and disclosure partitions coincide, i.e., regimes of the form $(\mathcal{P}, \mathcal{P})$. We now show that, for any information partition $\mathcal{P}$, the additive pricing approximation of IBPA with null disclosure weakly dominates GSP under the regime $(\mathcal{P}, \mathcal{P})$.

Proposition 10. *For every information partition $\mathcal{P}$,*

$$\text{Rev}_{\text{IBPA}^{\text{add}}}(\mathcal{P}, \mathcal{P}^{\text{null}}_{\text{disc}}) \geq \text{Rev}_{\text{GSP}}(\mathcal{P}, \mathcal{P}).$$

Proof. See Web Appendix A. □

5 Empirical Performance

This section evaluates the empirical performance of the proposed mechanism. While the theoretical guarantees establish the mechanism's relative performance, this section exemplifies the magnitude of the mechanism's potential gains when implemented. Using auction-level data from a large retail media platform, we estimate the key primitives required to implement the mechanism in practice, including the prior distribution of advertiser valuations, the distribution of inventory types, and slot effects on click-through rates. With these primitives in hand, we compute publisher revenue under the proposed IBPA mechanism and compare it against several benchmark mechanisms, including GSP under alternative information regimes.

5.1 Data

Our empirical analysis draws on auction-level data from a large retail media platform collected during the period May 17, 2025 to July 14, 2025. The dataset contains all position auctions conducted on the platform during this interval for one product category, totaling 9.6 million auctions. For each auction, we observe the realized inventory type, the number of ad slots available, and information on participating advertisers, including their submitted bid and quality score (γ).

The platform classifies impressions into inventory types that combine audience and device information. Specifically, inventory types are defined by (i) consumer behavioral segment (three categories), (ii) demographic segment (eight categories), and (iii) device type (three categories), yielding $3 \times 8 \times 3 = 72$ distinct inventory types.[10]

The platform employs the generalized second-price (GSP) auction format under the full-targeting regime. Advertisers therefore observe the realized inventory type before bidding and submit a bid specific to that type. Auction outcomes follow the standard GSP rule: advertisers are ranked by the product of their bid and quality score, slots are assigned in that order, and each advertiser pays the minimum bid needed to retain its assigned slot.

To estimate click-through rate (CTR) parameters, we use a separate dataset covering the same time period as the auction data. This dataset contains daily-level observations for each advertiser, reporting the number of impressions and clicks received at each slot position. These daily aggregates

[10] Although additional audience attributes are available, we focus on these dimensions because the platform considers them the most relevant for advertiser targeting.

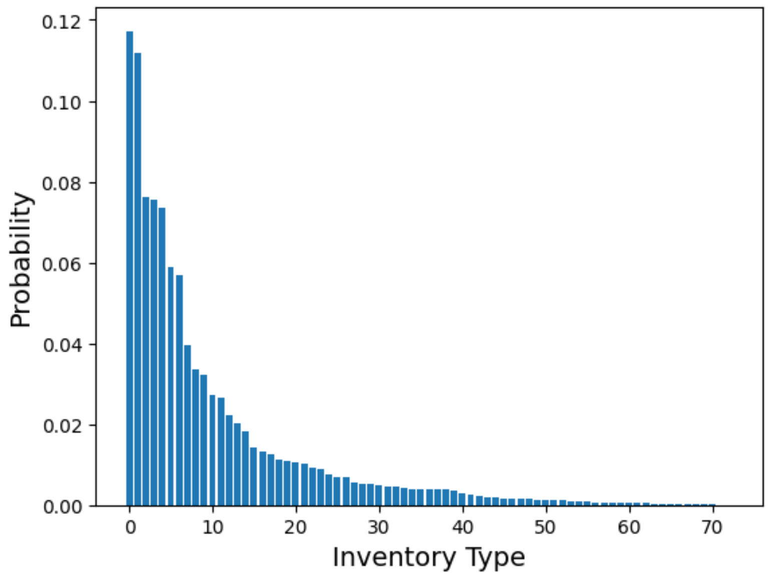


Figure 1: Distribution of inventory types

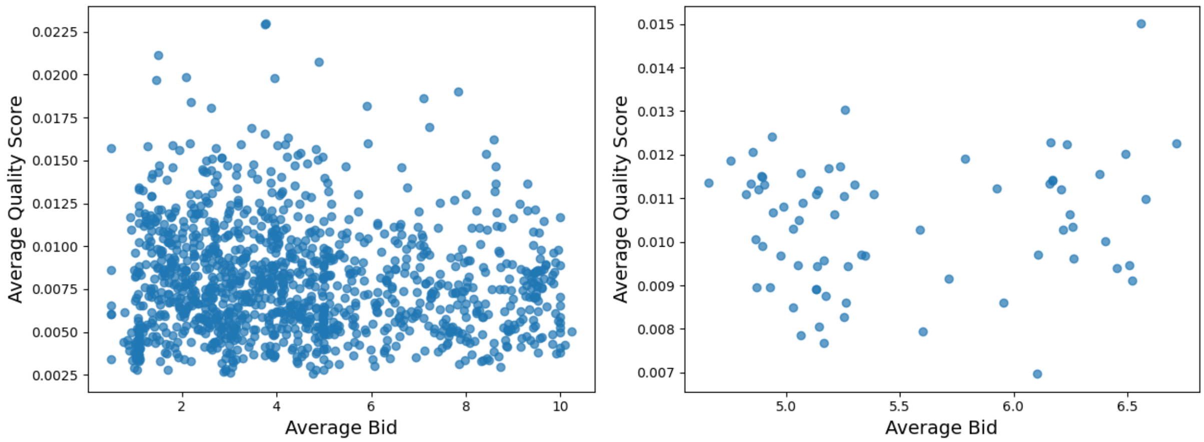
allow us to identify slot effects on CTR following the standard approach in Athey and Nekipelov (2011), which exploits variation in realized slot positions across advertisers over time. The resulting estimates provide the slot multipliers α_s.

5.1.1 Data Summary. Across the 9.6M auctions spanning 72 inventory types, we observe participation from 1,207 distinct advertisers. The majority of auctions contain either 8 slots (4.0M auctions) or 20 slots (2.38M), while 1-slot and 3-slot auctions appear with 1.39M and 1.04M occurrences, respectively; the remaining slot configurations account for a small fraction of the data.

Figure 1 plots the distribution of the 72 inventory types, ordered by their empirical frequency. The most common type occurs in roughly 12% of auctions, and these empirical frequencies correspond to the inventory-type probabilities p_t used in the mechanism.

Advertiser bids and quality scores exhibit substantial heterogeneity. The average bid is \$5.31 per-click, and the average quality score (γ) is 0.0096. Figure 2 shows scatter plots of average bids and average quality scores by advertiser and by inventory type. The figure illustrates the degree of variation in valuations and quality scores across advertisers and inventory types—the two key inputs to the mechanism.

From the CTR dataset, we observe 11.4M impressions and 198,609 clicks during the sample period, yielding an average click-through rate of 0.0174. Figure 3 plots the empirical CTR by slot posi-



(a) Advertiser-level **(b)** Inventory-type-level

Figure 2: Distribution of average bid vs. quality score

tion. While slot effects (α_s) are expected to decline monotonically with slot index, CTRs themselves need not exhibit such monotonicity because the CTR is the product of slot effects and advertiser quality. If advertisers with higher quality scores tend to occupy lower slots, this composition effect can mask the monotonic structure of the underlying slot effects, which we estimate separately.

5.2 Estimating Primitives

To evaluate the performance of the marginal-revenue mechanism and the benchmark mechanisms, we require several primitives: (i) the distribution of advertiser valuations $\{F_a\}_{a\in[A]}$, (ii) the distribution of advertiser quality scores $\{\gamma_a\}$, (iii) slot effects on click-through rates $\{\alpha_s\}$, and (iv) the distribution of inventory types $\{p_t\}$. The distribution of inventory types follows directly from the empirical frequencies computed in Section 5.1. Similarly, the empirical distribution of advertiser quality scores is used directly. The remaining primitives—slot effects and valuation distributions—are identified from the data using the framework of Athey and Nekipelov (2011) and the equilibrium structure of generalized second-price auctions. We describe each in turn.

5.2.1 Slot Effects. Click-through rate (CTR) in the model is multiplicatively separable into slot, inventory type, and advertiser effects:

$$\mathrm{CTR}_{ast} = \alpha_s \beta_t \gamma_a.$$

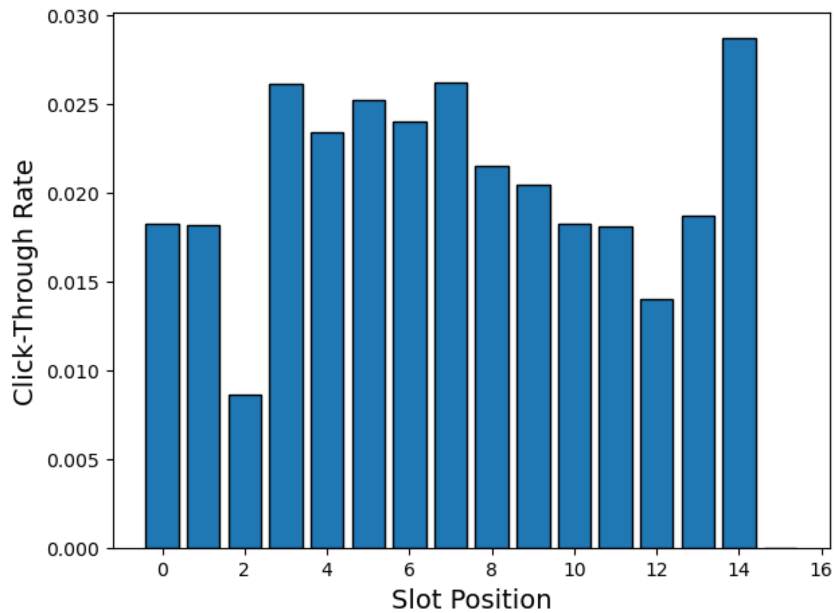


Figure 3: Variation of click-through rate by slot

Following Athey and Nekipelov (2011), the slot effects $\{\alpha_s\}$ and advertiser effects $\{\gamma_a\}$ are identified from variation in realized positions across advertisers over time. Because our CTR data do not break out clicks by inventory type, we normalize $\beta_t = 1$ for all t. This normalization applies uniformly across all counterfactual mechanisms and therefore preserves the fairness of comparisons.

Let y_{asd} denote the observed CTR for advertiser a in slot s on day d. We assume a multiplicative error term ϵ_{asd} and estimate:

$$y_{asd} = \alpha_s \gamma_a \, \epsilon_{asd}.$$

Taking logs yields a two-way additive fixed-effects representation:

$$\log y_{asd} = \log \alpha_s + \log \gamma_a + \log \epsilon_{asd}.$$

We estimate this model by least squares, treating $\{\log \alpha_s\}$ as slot fixed effects and $\{\log \gamma_a\}$ as advertiser fixed effects. Consistent with the structure of the mechanism, we impose the monotonicity constraint $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_S$ during estimation.

The regression achieves $R^2 = 0.908$, indicating that the assumptions of multiplicative separability and monotonicity of slot effects provide a good approximation to the CTR data. The resulting slot effects $\{\alpha_s\}$ are used both in the estimation of distribution of advertiser valuations below and in counterfactual mechanism comparisons.

5.2.2 Distribution of Advertiser Valuations. Because GSP is not truth-telling, advertiser bids do not equal valuations. Nevertheless, Edelman, Ostrovsky, and Schwarz (2007) and Varian (2007) show that equilibrium bidding behavior in GSP, which they refer to as envy-free equilibria, imposes sharp inequality constraints that bound advertiser valuations. Let advertisers be ordered by bid-rank ($b_a \times \gamma_a$) so that advertiser a_s occupies slot s. The envy-free conditions require:

$$\gamma_{a_s} v_{a_s} \quad \geq \quad \frac{\gamma_{a_{s+1}} b_{a_{s+1}} \alpha_s - \gamma_{a_{s+2}} b_{a_{s+2}} \alpha_{s+1}}{\alpha_s - \alpha_{s+1}} \quad \geq \quad \gamma_{a_{s+1}} v_{a_{s+1}}.$$

The middle term defines the *incremental cost per click* (ICC) of moving from slot $s+1$ to slot s:

$$ICC_{s,s+1} = \frac{\gamma_{a_{s+1}} b_{a_{s+1}} \alpha_s - \gamma_{a_{s+2}} b_{a_{s+2}} \alpha_{s+1}}{\alpha_s - \alpha_{s+1}}.$$

In equilibrium these ICCs must satisfy the monotonicity condition

$$ICC_{s,s+1} \geq ICC_{s+1,s+2}.$$

Because observed bids may violate monotonicity, we follow Varian (2007) and compute *weighted* ICCs $\{ICC^{d^*}_{s,s+1}\}$ that minimize the weighted deviation from monotonicity:

$$ICC^{d}_{s,s+1} = \frac{\gamma_{a_{s+1}} b_{a_{s+1}} \alpha_s d_s - \gamma_{a_{s+2}} b_{a_{s+2}} \alpha_{s+1} d_{s+1}}{\alpha_s d_s - \alpha_{s+1} d_{s+1}},$$

with weights d chosen to minimize $\sum_s (1 - d_s^2)$ subject to the ICCs being monotone. The optimal weights d^* yield a monotone sequence of ICCs.

These monotone ICCs deliver valuation bounds:

$$\gamma_{a_s} v_{a_s} \in \left[ICC^{d^*}_{s,s+1},\ ICC^{d^*}_{s-1,s}\right], \qquad s = 2, \ldots, S-1.$$

For slot 1, the lower bound is identified; for the upper bound we set $\overline{v} = 2b_{\max}$, where $b_{\max}$ is the largest bid observed. For slot S, the upper bound is identified; the lower bound requires the highest losing bid, which we do not observe, so we conservatively set it to zero.

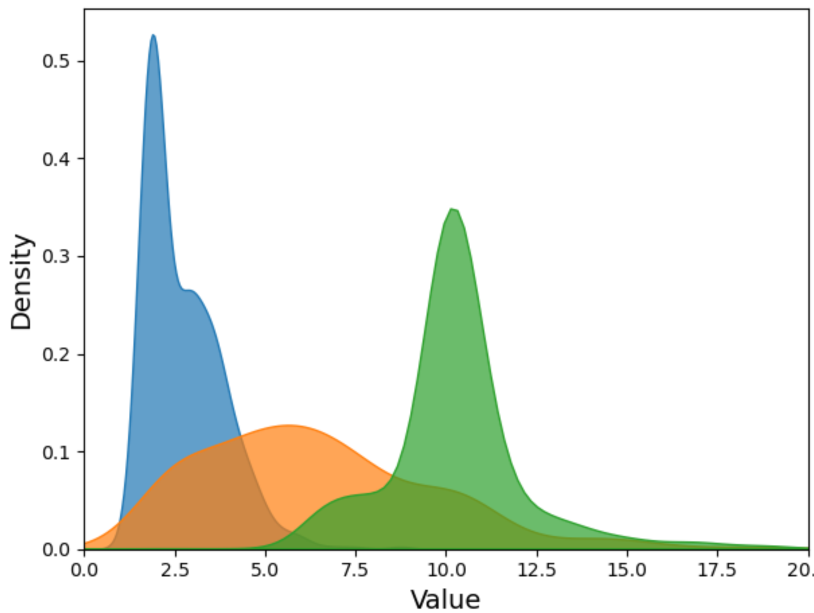


Figure 4: Distribution of valuations for randomly sampled inventory-types

Thus each advertiser in each auction contributes an interval $(L, U]$ containing the true valuation v_{at}. To estimate the valuation distribution F_{at}, we pool the intervals across auctions and compute the non-parametric maximum likelihood estimator using the Turnbull estimator (Turnbull 1976). Because all auctions in our sample belong to a single product category, we assume that advertisers are ex ante symmetric within an inventory type, and that valuations are independent and identically distributed draws from a common distribution F_t. This symmetry assumption permits pooling interval observations across advertisers and auctions to recover a single valuation distribution for each inventory type. We implement the estimator using the EM algorithm, which yields a piecewise-uniform CDF with data-driven support.

Figure 4 plots estimated valuation densities for three randomly sampled inventory-types, illustrating substantial heterogeneity in valuation levels and dispersion. These estimated valuation distributions, together with the CTR primitives above, are used in the counterfactual mechanism comparisons in Section 5.3.

5.3 Comparison of Mechanisms

We now empirically evaluate the performance of the proposed mechanism. Using the primitives estimated in Section 5.2, we compute publisher revenue under the marginal-revenue mechanism and compare it against several benchmark mechanisms, including GSP under alternative information and disclosure regimes. We first describe the mechanisms considered and then outline our counterfactual simulation procedure.

5.3.1 IBPA. Our primary mechanism is under $\left(\mathcal{P}_{\text{info}}^{\text{full}}, \mathcal{P}_{\text{disc}}^{\text{null}}\right)$, which is computationally efficient and serves as a conservative approximation to IBPA under the same information and disclosure regimes. We refer to this mechanism by IBPA-FI-ND (full-information/null-disclosure). The single-advertiser revenue problems are solved by optimizing the constrained problem in Equation 5. Although several optimization routines are available, we employ a genetic algorithm to account for the possibility of multiple local optima. Constraints are incorporated directly into the objective with a large penalty for violations. Results are robust to algorithmic tuning parameters; we use: population size 100, 50 parents per generation (tournament selection), 5 elites, crossover probability 0.8 (blending), and mutation probability 0.2 (random mutation).

For each advertiser, we solve the constrained problem on a grid of allocation probabilities $q \in [0, 1]$. Results are robust for grids of 10–100 points. The normalized revenue curve $\Phi_a(q)$ is then approximated by a piecewise-linear interpolation across this grid. This yields a conservative approximation, as the true revenue curve is concave and therefore weakly above the piecewise-linear fit. Quantile mappings follow Section 3.4.3, and auction outcomes (allocations and payments) are computed as described in Sections 3.4.4 and 3.4.5.

5.3.2 GSP under alternative information regimes. Section 4.2 establishes that, for GSP, it is without loss of generality to analyze regimes of the form $\left(\mathcal{P}, \mathcal{P}\right)$, i.e., regimes in which the publisher's information and disclosure partitions coincide. Thus, to characterize counterfactual GSP performance, it suffices to vary the information partition $\mathcal{P}$ over inventory types $[T]$.

Ideally, one might evaluate GSP under all possible partitions of the 72 inventory types, but this is computationally infeasible. Instead, we exploit the structure of the inventory attributes: consumer behavior (1 attribute), demographics (3 attributes), and device type (1 attribute). By toggling each attribute on or off, we generate $2^5 = 32$ meaningful information regimes.

In our empirical context, the null-disclosure partition $\mathcal{P} = \mathcal{P}_{\text{disc}}^{\text{null}}$ generates the highest GSP revenue among these regimes. We therefore report results for both this regime and the commonly deployed full-disclosure version of GSP with $\mathcal{P} = \mathcal{P}_{\text{disc}}^{\text{full}}$. We refer to these specifications as GSP-NI-

ND (null information and disclosure) and GSP-FI-FD (full information and disclosure), respectively.

As GSP is not incentive compatible, advertiser bids follow envy-free equilibria (Edelman, Ostrovsky, and Schwarz 2007; Varian 2007). Because such equilibria are not unique, we compute the envy-free upper bound—the equilibrium that yields the highest publisher revenue—ensuring a conservative benchmark for comparison with the IBPA mechanism.

5.3.3 Marginal-revenue mechanisms under full disclosure. The proof of Theorem 5 establishes that, for any information partition $\mathcal{P}$,

$$\mathrm{Rev}_{\mathrm{IBPA}}\left(\mathcal{P}, \mathcal{P}\right) \;\geq\; \mathrm{Rev}_{\mathrm{GSP}}\left(\mathcal{P}, \mathcal{P}\right).$$

Since we report GSP under full and null disclosure, we also report the corresponding $\mathrm{IBPA}\big(\mathcal{P}^{\mathrm{full}}_{\mathrm{info}}, \mathcal{P}^{\mathrm{full}}_{\mathrm{info}}\big)$ and $\mathrm{IBPA}\big(\mathcal{P}^{\mathrm{null}}_{\mathrm{info}}, \mathcal{P}^{\mathrm{null}}_{\mathrm{info}}\big)$. We refer to these mechanisms as IBPA-FI-FD and IBPA-NI-ND, respectively.

5.3.4 Counterfactual Performance Comparison. To compare mechanisms in a common environment, we simulate 10 million auctions using the primitives from Section 5.2. For each simulated auction, we (1) set the number of slots to 8 and use the estimated slot effects $\{\alpha_s\}$, (2) draw the number of advertisers and their quality scores γ_a from the empirical distributions, (3) draw a valuation vector v_a for each advertiser from the estimated Turnbull distributions $\{F_{at}\}$, and (4) draw the realized inventory type t from the empirical distribution $\{p_t\}$.

Figure 5 reports publisher revenue across all mechanisms. The ordering of revenues matches our theoretical predictions:

$$\mathrm{IBPA-FI-ND} \;>\; \mathrm{IBPA-FI-FD} \;>\; \mathrm{GSP-FI-FD}$$

and

$$\mathrm{IBPA-FI-ND} \;>\; \mathrm{IBPA-NI-ND} \;>\; \mathrm{GSP-NI-ND}.$$

The overall revenue gain from GSP–FI–FD to IBPA–FI–ND is 68%. The improvement from GSP-FI-FD to IBPA-FI-FD (22%) reflects the gain from optimal marginal-revenue allocation under full disclosure. The further improvement from IBPA–FI–FD to IBPA–FI–ND (38%) reflects the additional value created when the publisher withholds inventory-type information and exploits second-degree price discrimination across types.

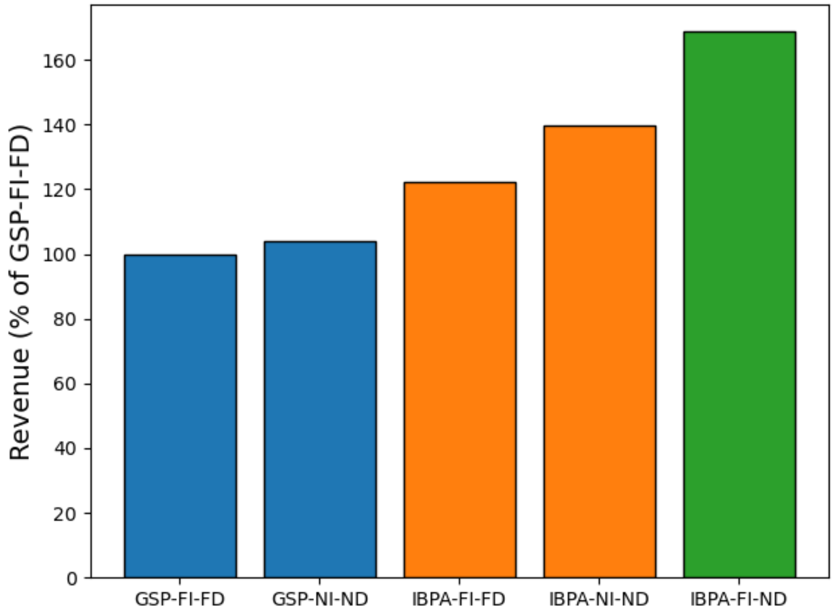


Figure 5: Revenue Comparison

Similarly, the revenue gain from GSP–NI–ND to IBPA–FI–ND is 62%. The improvement from GSP-NI-ND to IBPA-NI-ND (34%) reflects the gain from optimal marginal-revenue allocation under null disclosure. The additional improvement from IBPA–NI–ND to IBPA–FI–ND (20%) reflects the revenue benefit of finer internal information granularity.

Although our theoretical guarantees concern revenue, we also measure advertiser welfare (expected sum of utilities) and total welfare (publisher revenue plus advertiser welfare). Table 1 reports percentage changes in revenue, advertiser welfare, and total welfare, along with percentage-point changes in allocation rate, relative to the GSP–FI–FD benchmark across alternative mechanisms. While welfare improvements are not theoretically guaranteed for arbitrary primitives, in our empirical setting, both metrics improve under the proposed mechanism relative to the benchmark—advertiser welfare by 29% and total welfare by 54%. The improvements appear to be driven by allocation rate—in our simulations, the probability of an impression being sold under the proposed mechanism is 47% compared to 28% under the benchmark. Higher allocation rate indicates that impressions are more frequently matched to advertisers who value them, improving outcomes for all participants.

Metric	GSP-FI-FD	GSP-NI-ND	IBPA-FI-FD	IBPA-NI-ND	IBPA-FI-ND
Revenue	+0%	+4%	+22%	+40%	+68%
Advertiser Welfare	+0%	-36%	+17%	+29%	+29%
Total Welfare	+0%	-11%	+20%	+36%	+54%
Allocation Rate	+0pp	+10pp	+4pp	+18pp	+19pp

Table 1: Performance of mechanisms relative to GSP-FI-FD

6 Discussion & Conclusion

This paper considers whether and how publishers should share private information about impression types when auctioning digital ads. Current auction mechanisms such as the GSP induce a potential trade-off that arises when targeting options become more granular. On the one hand, granularity enables advertisers to match with higher value audiences raising advertiser bids. On the other hand, finer targeting reduces the potential competition set, lowering revenue. The trade–off is an artifact of the requirement inherent in most internet ad sales that audience and contextual information about the impression are revealed to the advertisers ex-ante. There is no operational reason for this constraint to exist.

We develop a general auction mechanism, applicable to both search (position) and display environments, that treats inventory type as the publisher's private information and does not reveal this type before bidding. For any given advertiser, the publisher offers a menu of prices that corresponds to mixed bundles of inventory types—effectively implementing second-degree price discrimination within each advertiser. Across advertisers, the mechanism applies the marginal-revenue framework to allocate each impression to the advertiser with the highest marginal revenue, thereby layering first-degree price discrimination across advertisers on top of second-degree price discrimination within advertiser types, with finer second-degree discrimination directly improving the effectiveness of first-degree discrimination.

The resulting Information-Bundling Position Auction (IBPA) possesses several desirable properties. First, it is incentive compatible—truth-telling is the advertiser's optimal strategy—and guarantees non-negative expected advertiser utility. Second, it improves publisher revenue relative to standard industry mechanisms such as GSP. Third, it is straightforward to implement requiring only modest extensions to existing bidding interfaces and is computationally scalable.

Using data from a retail media platform, we estimate advertiser valuations and forecast outcomes under IBPA. Relative to GSP, IBPA raises publisher revenue by approximately 68% and increases advertiser welfare by 29%. While the revenue gains are driven by improved price discrimination, the welfare gains arise primarily from improved allocation rate—impressions are more often matched to advertisers who value them.

Several considerations merit discussion. A first pertains to advertiser participation. In practice, advertisers typically multi-home across many publishers. One potential concern is that if a single publisher adopts a more revenue-optimal mechanism such as IBPA—thereby increasing advertiser payments for impressions—advertisers may reallocate spend toward competing publishers. In the long run, however, such displacement is unlikely: every publisher has an incentive to adopt IBPA, since it weakly increases revenues regardless of the distribution of advertiser valuations. Thus, widespread adoption is an equilibrium outcome, and relative competitive positioning across publishers is preserved.

In the short run, even if one publisher adopts IBPA while others do not, advertisers continue to obtain weakly non-negative utility from participating because the mechanism satisfies individual rationality. Multi-homing concerns arise primarily for budget-constrained advertisers, who may reduce expenditure if their marginal return on spend declines relative to other platforms. Our framework allows publishers to accommodate such advertisers. In particular, the publisher can either modify the individual-rationality constraint to require that each advertiser obtain at least some minimum utility level, or equivalently incorporate advertiser welfare directly in the objective function—for example, by maximizing a linear combination of revenue and advertiser welfare. These two formulations are mathematically equivalent by standard Lagrangian arguments, and both can be implemented straightforwardly within IBPA, since the mechanism is expressed entirely in terms of single-agent revenue curves to which such constraints or modifications can be directly applied.

A second consideration pertains to general-equilibrium effects. While a representative publisher adopting IBPA may substantially increase advertiser payments and improve allocation rate, these gains do not necessarily scale proportionally when all platforms adopt the mechanism. In equilibrium, industry-wide advertiser budgets are fixed or slow-moving, so aggregate spending cannot expand at the same rate as a single publisher's revenue forecast. Even at the level of a single platform, advertisers may lack the budget to absorb large increases in spend, irrespective of the mechanism's allocation rate. Relatedly, our empirical findings show that IBPA increases the likelihood that impressions are allocated—raising the platform's ad load. A higher ad load can affect consumer welfare and engagement with the platform, as users may experience more ad exposures. Platforms routinely face this trade-off between short-run monetization and long-run user experience. A publisher can manage both budget constraints on the advertiser side and attention constraints on the consumer side by reducing the supply of impressions, thereby preserving consumer welfare while still benefiting from improved pricing efficiency.

More broadly, publishers solve a multi-objective dynamic optimization problem, balancing short-run revenue from auctions and ad load, long-run consumer welfare and engagement, advertiser welfare and retention, and advertiser budget constraints. Auction design is only one component of this broader optimization problem. Crucially, however, nothing in this larger objective weakens the rationale for adopting IBPA over GSP. Whatever the platform's optimal choice of ad supply, objective function, or pacing rules, IBPA generates higher revenue conditional on those choices. The mechanism is therefore a dominant design choice within the class of feasible auction formats. These considerations—including multi-homing, budget constraints, and consumer welfare effects—are not unique to IBPA. They arise equally when platforms adjust reserve prices, modify pacing rules, or optimize quality scores under GSP. A complete general-equilibrium analysis of publisher competition, advertiser budget allocation, and consumer engagement is an important but distinct research direction and lies outside the scope of this paper.

A third consideration concerns implementation of multidimensional bids. Although IBPA requires advertisers to submit valuations over all inventory types, this requirement does not pose a practical burden. In most advertising platforms—including retail media—advertisers already submit structured bids consisting of a base bid and a set of bid modifiers tied to audience or contextual attributes (e.g., device type, demographic segment, behavioral signals). These modifiers naturally define a multidimensional bid space, allowing the advertiser's per-type valuation vector to be constructed from a small number of attribute-level parameters. Thus, the solicitation of multidimensional bids under IBPA is fully compatible with existing bidding tools and does not require advertisers to manage an unwieldy number of independent type-specific bids.

A fourth consideration pertains to transparency and trust. Although IBPA is incentive compatible—truth-telling is the advertiser's optimal strategy—advertisers increasingly seek clarity regarding how auctions are executed, how prices are determined, how quality is assessed, and how competitor behavior affects outcomes.[11] At the heart of these concerns is uncertainty about the publisher's commitment to adhering to the mechanism's rules (Akbarpour and Li 2020). This issue is not unique to IBPA: even GSP requires trust in the publisher's computation of quality scores and ranking rules, and platforms may optimally adjust these internal parameters in ways that are opaque to advertisers. An important direction for future work therefore concerns the role of commitment, transparency, and verifiability in ad auction design. Publishers could, for example, disclose ex-post allocations and payments, or allow advertisers to audit rule compliance. Advertisers can also verify that truthful bidding is optimal by appealing to the convexity of expected utility in private valuations—a general property of quasi-linear mechanism design environments (Daskalakis, Deckelbaum, and Tzamos 2015). Enhancing trust through greater transparency is thus a complementary pathway alongside adopting more efficient mechanisms such as IBPA.

Overall, we hope this research deepens understanding of the sell-side economics of the rapidly growing digital advertising market. IBPA offers a theoretically grounded, practically implementable, and welfare-improving alternative to current auction formats. We anticipate that the approach will be widely adopted and will stimulate further research on mechanism design, transparency, platform competition, and long-run market dynamics in the trillion-dollar online advertising ecosystem.

[11] `https://tinyurl.com/ad-sales-transparency`

References

Ada, Sila, Nadia Abou Nabout, and Elea McDonnell Feit (2022), "Context Information Can Increase Revenue in Online Display Advertising Auctions: Evidence from a Policy Change," *Journal of Marketing Research*, 59 (5), 1040–1058.

Adams, William James and Janet L. Yellen (1976), "Commodity Bundling and the Burden of Monopoly," *The Quarterly Journal of Economics*, 90 (3), 475–498.

Akbarpour, Mohammad and Shengwu Li (2020), "Credible Auctions: A Trilemma," *Econometrica*, 88 (2), 425–467.

Alaei, Saeed, Hu Fu, Nima Haghpanah, and Jason Hartline "The Simple Economics of Approximately Optimal Auctions," "2013 IEEE 54th Annual Symposium on Foundations of Computer Science," pages 628–637 (2013).

Armstrong, Mark (2016), "Nonlinear Pricing," *Annual Review of Economics*, 8 (Volume 8, 2016), 583–614.

Athey, Susan and Denis Nekipelov (2011), "A Structural Model of Sponsored Search Advertising Auctions," *Working Paper, University of California, Berkeley.*

Babaioff, Moshe, Nicole Immorlica, Brendan Lucier, and S Matthew Weinberg (2020), "A Simple and Approximately Optimal Mechanism for an Additive Buyer," *Journal of the ACM (JACM)*, 67 (4), 1–40.

Balocco, Francesco, Yixin Lu, Ting Li, and Alok Gupta (2025), "Lemon Ads: Adverse Selection in Multichannel Display Advertising Markets," *Management Science.*

Bergemann, Dirk, Tibor Heumann, Stephen Morris, Constantine Sorokin, and Eyal Winter (2022), "Optimal information disclosure in classic auctions," *American Economic Review: Insights*, 4 (3), 371–388.

Bergemann, Dirk and Martin Pesendorfer (2007), "Information Structures in Optimal Auctions," *Journal of Economic Theory*, 137 (1), 580–609.

Bulow, Jeremy and John Roberts (1989), "The Simple Economics of Optimal Auctions," *Journal of political economy*, 97 (5), 1060–1090.

Chawla, Shuchi, Yifeng Teng, and Christos Tzamos (2019), "Buy-many Mechanisms are Not Much Better than Item Pricing," *Proceedings of the 2019 ACM Conference on Economics and Computation.*

Chen, Yi-Chun and Xiangqian Yang (2023), "Information Design in Optimal Auctions," *Journal of Economic Theory*, 212, 105710.

Chen, Yuxin, Chakravarthi Narasimhan, and Z John Zhang (2001), "Individual Marketing with Imperfect Targetability," *Marketing Science*, 20 (1), 23–41.

Choi, Hana and Carl F. Mela (2023), "Optimizing Reserve Prices in Display Advertising Auctions," *SSRN Electronic Journal.*

Choi, Hana, Carl F. Mela, Santiago R. Balseiro, and Adam Leary (2020), "Online Display Advertising Markets: A Literature Review and Future Directions," *Information Systems Research*, 31 (2), 556–575.

Choi, W. Jason and Amin Sayedi (2023), "Agency Bidding in Online Advertising," *SSRN Electronic Journal.*

Daskalakis, Constantinos, Alan Deckelbaum, and Christos Tzamos (2015), "Strong Duality for a Multiple-good Monopolist," *Proceedings of the Sixteenth ACM Conference on Economics and Computation.*

Despotakis, Stylianos, R. Ravi, and Amin Sayedi (2021), "First-Price Auctions in Online Display Advertising," *Journal of Marketing Research*, 58 (5), 888–907.

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz (2007), "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars," *American Economic Review*, 97 (1), 242–259.

Goldfarb, Avi and Verina F. Que (2023), "The Economics of Digital Privacy," *Annual Review of Economics*, 15 (Volume 15, 2023), 267–286.

Gomes, Renato D and Kane S Sweeney (2009), "Bayes-Nash Equilibria of the Generalized Second Price Auction," *Proceedings of the 10th ACM conference on Electronic Commerce.*

Gordon, Brett R., Robert Moakler, and Florian Zettelmeyer (2023), "Close Enough? A Large-Scale Exploration of Non-Experimental Approaches to Advertising Measurement," *Marketing Science*, 42 (4), 768–793.

Hart, Sergiu and Noam Nisan (2013), "The Menu-size Complexity of Auctions," *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*.

Iyer, Ganesh, David Soberman, and J Miguel Villas-Boas (2005), "The Targeting of Advertising," *Marketing Science*, 24 (3), 461–476.

Johnson, Garrett A. (2023), "Inferno: A Guide to Field Experiments in Online Display Advertising," *Journal of Economics and Management Strategy*, 32 (3), 469–490.

Kim, Kyungmin and Youngwoo Koh (2022), "Auctions with Flexible Information Acquisition," *Games and Economic Behavior*, 133, 256–281.

Lee, Dokyun, Kartik Hosanagar, and Harikesh S. Nair (2018), "Advertising Content and Consumer Engagement on Social Media: Evidence from Facebook," *Management Science*, 64 (11), 5105–5131.

Lee, Jasper CH and Paul Valiant (2016), "Optimizing Star-Convex Functions," *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*.

Levin, Jonathan and Paul Milgrom (2010), "Online Advertising: Heterogeneity and Conflation in Market Design," *American Economic Review*, 100 (2), 603–607.

Li, Shibo, Kannan Srinivasan, and Baohong Sun (2009), "Internet Auction Features as Quality Signals," *Journal of Marketing*, 73 (1), 75–92.

Lu, Shijie and Sha Yang (2015), "A Two-Sided Market Analysis of Behaviorally Targeted Display Advertising," *SSRN Electronic Journal*.

Manelli, Alejandro M and Daniel R Vincent (2006), "Bundling as an Optimal Selling Mechanism for a Multiple-good Monopolist," *Journal of Economic Theory*, 127 (1), 1–35.

Manelli, Alejandro M and Daniel R Vincent (2007), "Multidimensional Mechanism Design: Revenue Maximization and the Multiple-good Monopoly," *Journal of Economic Theory*, 137 (1), 153–185.

McAfee, R. Preston, John McMillan, and Michael D. Whinston (1989), "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values," *The Quarterly Journal of Economics*, 104 (2), 371–383.

Milgrom, Paul R. and Robert J. Weber (1982), "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50 (5), 1089–1122.

Myerson, Roger B (1981), "Optimal Auction Design," *Mathematics of Operations Research*, 6 (1), 58–73.

Rafieian, Omid and Hema Yoganarasimhan (2021), "Targeting and Privacy in Mobile Advertising," *Marketing Science*, 40 (2), 193–218.

Rochet, Jean-Charles (1985), "The Taxation Principle and Multi-time Hamilton-Jacobi Equations," *Journal of Mathematical Economics*, 14 (2), 113–128.

Sayedi, Amin (2018), "Real-Time Bidding in Online Display Advertising," *Marketing Science*, 37 (4), 553–568.

Shi, Xianwen (2012), "Optimal Auctions with Information Acquisition," *Games and Economic Behavior*, 74 (2), 666–686.

Sousa, Tulio (2025), "Competition and Targeting in Auctions for Online Ads," *Working Paper, Duke University*.

Tunuguntla, Srinivas and Paul R. Hoban (2020), "A Near-Optimal Bidding Strategy for Real-Time Display Advertising Auctions," *Journal of Marketing Research*, 58 (1), 1–21.

Turnbull, Bruce W (1976), "The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data," *Journal of the Royal Statistical Society: Series B (Methodological)*, 38 (3), 290–295.

Varian, Hal R. (2007), "Position Auctions," *International Journal of Industrial Organization*, 25 (6), 1163–1178.

Wu, Ruiqi, Yufeng Huang, and Nan Li (2025), "Platform Information Design and Competitive Price Targeting," *Available at SSRN 4619584*.

Yao, Song and Carl F. Mela (2011), "A Dynamic Model of Sponsored Search Advertising," *Marketing Science*, 30 (3), 447–468.

Web Appendix

Web Appendix A: Omitted Proofs

Proof of Proposition 1. Consider any single-advertiser IC/IR mechanism M, represented (by the taxation principle) as a menu of lottery pricings $\{(\chi^{(k)}, \mu^{(k)})\}_k$, where each lottery k specifies allocation probabilities $\chi_{st}^{(k)} \in [0,1]$ for slots $s \in \tilde{S}$ and inventory types $t \in [T]$, and an expected payment $m^{(k)} = \alpha_{s_0}\gamma_{a_0}\mu^{(k)}$. Fix a lottery pricing k.

Define, for each inventory type t, the CTR-weighted (slot-collapsed) allocation probability

$$\bar{\chi}_t^{(k)} \equiv \sum_{s \in \tilde{S}} \frac{\alpha_s}{\alpha_{s_0}} \chi_{st}^{(k)}.$$

Now construct a modified lottery pricing $(\chi^{(k)\prime}, \mu^{(k)\prime})$ that allocates only the highest available slot s_0 as follows:

$$\chi_{s_0 t}^{(k)\prime} = \bar{\chi}_t^{(k)} \quad \forall t, \qquad \chi_{st}^{(k)\prime} = 0 \quad \forall s \neq s_0, \forall t, \qquad \mu^{(k)\prime} = \mu^{(k)}.$$

For each t, since $\alpha_s \leq \alpha_{s_0}$ for all $s \in \tilde{S}$,

$$\bar{\chi}_t^{(k)} = \sum_{s \in \tilde{S}} \frac{\alpha_s}{\alpha_{s_0}} \chi_{st}^{(k)} \;\leq\; \sum_{s \in \tilde{S}} \chi_{st}^{(k)} \;\leq\; 1,$$

where the last inequality follows from the feasibility constraint that the advertiser receives at most one slot for any given inventory type. Hence the modified lottery pricing is also feasible.

For any valuation vector $v_{a_0} = (v_{a_0 1}, \ldots, v_{a_0 T})$, the expected utility from lottery k under M is

$$U^{(k)}(v_{a_0}) = \sum_t \sum_s p_t\, \alpha_s \beta_t \gamma_{a_0} v_{a_0 t} \chi_{st}^{(k)} - \alpha_{s_0}\gamma_{a_0}\mu^{(k)}.$$

Under the modified lottery pricing,

$$U^{(k)\prime}(v_{a_0}) = \sum_t p_t\, \alpha_{s_0} \beta_t \gamma_{a_0} v_{a_0 t} \chi_{s_0 t}^{(k)\prime} - \alpha_{s_0}\gamma_{a_0}\mu^{(k)\prime}.$$

By construction, $\alpha_{s_0}\chi_{s_0 t}^{(k)\prime} = \alpha_{s_0}\bar{\chi}_t^{(k)} = \sum_s \alpha_s \chi_{st}^{(k)}$ for every t, and $\mu^{(k)\prime} = \mu^{(k)}$. Therefore $U^{(k)\prime}(v_{a_0}) = U^{(k)}(v_{a_0})$ for all valuation vectors v_{a_0}.

Thus replacing lottery pricing k by $(\chi^{(k)\prime}, \mu^{(k)\prime})$ leaves the utility of that menu item unchanged pointwise in v_{a_0}, so the advertiser's menu choice behavior is unchanged. Because the payment parameter is unchanged, publisher revenue from this modified lottery pricing is also unchanged. Repeating this argument for every lottery pricing in the mechanism establishes the result. $\square$

Proof of Proposition 2. For notational convenience, we drop the advertiser and slot subscripts and write $R(q)$ for the optimal expected revenue of the ex-ante single-advertiser problem with allocation constraint q. Let $M(q)$ denote an optimal mechanism that achieves $R(q)$, and let $\Lambda(q)$ be its ex-ante allocation probability. By feasibility, $\Lambda(q) \leq q$.

To prove monotonicity, take any $q_1, q_2 \in [0, 1]$. Any mechanism feasible under constraint q_1—in particular $M(q_1)$—satisfies $\Lambda(q_1) \leq q_1 \leq q_2$ and is therefore feasible at constraint q_2. Since $R(q_2)$ is the maximum revenue over all mechanisms feasible at constraint q_2, we obtain $R(q_2) \geq R(q_1)$. Thus, $R(q)$ is non-decreasing in q, and consequently, $R'(q) \geq 0 \, \forall q$.

Take any $q_1, q_2 \in [0, 1]$ with $q_1 < q_2$ and any $\nu \in [0, 1]$. Let $q = \nu q_1 + (1 - \nu) q_2$. Consider the convex combination of mechanisms

$$\widetilde{M} \;\equiv\; \nu\, M(q_1) + (1 - \nu)\, M(q_2).$$

The ex-ante allocation probability under $\widetilde{M}$ is

$$\widetilde{\Lambda} \;=\; \nu\, \Lambda(q_1) + (1 - \nu)\, \Lambda(q_2) \;\leq\; \nu q_1 + (1 - \nu) q_2 \;\equiv\; q,$$

so $\widetilde{M}$ is feasible for the problem with allocation constraint q.

Moreover, the expected revenue of $\widetilde{M}$ is

$$\text{Revenue}(\widetilde{M}) \;=\; \nu\, R(q_1) + (1 - \nu)\, R(q_2).$$

Since $R(q)$ is defined as the maximum expected revenue among all mechanisms feasible at constraint q, it follows that

$$R(q) \;\geq\; \text{Revenue}(\widetilde{M}) \;=\; \nu\, R(q_1) + (1-\nu)\, R(q_2).$$

This inequality holds for every $q_1, q_2 \in [0,1]$ and every $\nu \in [0,1]$ with $q = \nu q_1 + (1-\nu) q_2$. Hence $R(\cdot)$ is concave on $[0,1]$. Concavity immediately implies that the (right) derivative $R'(q)$ exists almost everywhere and is weakly decreasing wherever it exists. □

Proof of Proposition 3. Fix an advertiser a and a slot s. Recall that the ex-ante single-advertiser problem with allocation constraint q chooses a menu of lotteries $\{(\chi^{(k)}, \mu^{(k)})\}_k$ to maximize

$$R_{as}(q) \;=\; \max_{K,\{\chi^{(k)},\mu^{(k)}\}} \; \alpha_s \gamma_a \sum_{k=1}^{K} \xi_k\, \mu^{(k)} \quad \text{s.t.} \quad \sum_{k=1}^{K} \xi_k \left(\sum_t p_t\, \chi_t^{(k)} \right) \leq q,$$

where ξ_k is the probability that lottery k is chosen ex-ante under the prior.

Define $\Phi_a(q)$ to be the optimal value of the *slot-normalized* problem that keeps the same feasible menus and allocation constraint but omits the multiplicative factor $\alpha_s \gamma_a$ from the objective:

$$\Phi_a(q) \;=\; \max_{K,\{\chi^{(k)},\mu^{(k)}\}} \; \sum_{k=1}^{K} \xi_k\, \mu^{(k)} \quad \text{s.t.} \quad \sum_{k=1}^{K} \xi_k \left(\sum_t p_t\, \chi_t^{(k)} \right) \leq q,$$

To prove the claim, it is sufficient to show that $\Phi_a(q)$ is independent of the slot s.

Consider any fixed menu $\{(\chi^{(k)}, \mu^{(k)})\}_k$. Take two options k_1, k_2 from this menu. The advertiser's expected utility difference between k_1 and k_2 when the highest remaining slot is s is

$$\sum_t p_t\, \alpha_s \beta_t \gamma_a\, v_{at} \left(\chi_t^{(k_1)} - \chi_t^{(k_2)} \right) \;-\; \alpha_s \gamma_a \Big(\mu^{(k_1)} - \mu^{(k_2)} \Big).$$

Factoring out $\alpha_s \gamma_a > 0$, the sign of this difference—and thus the choice between k_1 and k_2—is determined by

$$\sum_t p_t\, \beta_t\, v_{at} \left(\chi_t^{(k_1)} - \chi_t^{(k_2)} \right) \;-\; \Big(\mu^{(k_1)} - \mu^{(k_2)} \Big),$$

which is independent of s. Therefore, for any menu, the induced choice probabilities do not depend on s. The allocation constraint $\sum_k \xi_k \big(\sum_t p_t \chi_t^{(k)} \big) \leq q$ is also independent of s.

Consequently, the feasible set of menus for the normalized problem is the same for every s, and the only appearance of (α_s, γ_a) in the original objective is as a multiplicative factor. It follows that the optimal value satisfies

$$R_{as}(q) \;=\; \alpha_s \gamma_a \, \Phi_a(q),$$

where $\Phi_a(q)$ is independent of the slot s. This proves the claim. $\square$

Proof of Theorem 4. Fix a disclosure partition $\mathcal{P}_{\text{disc}}$. Let $\mathcal{P}^{(1)}_{\text{info}}$ and $\mathcal{P}^{(2)}_{\text{info}}$ be two information partitions such that

$$\mathcal{P}^{(1)}_{\text{info}} \succeq \mathcal{P}^{(2)}_{\text{info}} \succeq \mathcal{P}_{\text{disc}},$$

so that both pairs $(\mathcal{P}^{(1)}_{\text{info}}, \mathcal{P}_{\text{disc}})$ and $(\mathcal{P}^{(2)}_{\text{info}}, \mathcal{P}_{\text{disc}})$ are feasible.

Fix an advertiser a, a slot s, and the disclosure regime $\mathcal{P}_{\text{disc}}$. Under a given information partition $\mathcal{P}_{\text{info}}$, the ex-ante constrained single-advertiser problem (Equation (5)) chooses a menu of lottery pricings $\{(\chi^{(k)}, \mu^{(k)})\}_{k\in[K]}$ to maximize expected revenue subject to incentive compatibility, individual rationality, and an ex-ante allocation constraint.

Because the publisher cannot distinguish between inventory types within an information block, each lottery specifies the same allocation probability for all types in that block. If $\mathcal{P}^{(1)}_{\text{info}} \succeq \mathcal{P}^{(2)}_{\text{info}}$, then each block of $\mathcal{P}^{(2)}_{\text{info}}$ is a union of blocks of $\mathcal{P}^{(1)}_{\text{info}}$. Thus, any lottery feasible under $\mathcal{P}^{(2)}_{\text{info}}$ can be implemented under $\mathcal{P}^{(1)}_{\text{info}}$ simply by imposing equality of allocation probabilities across the constituent finer blocks. This preserves expected revenue, expected utility, and the ex-ante allocation constraint.

Hence, the feasible set of single-advertiser mechanisms under $\mathcal{P}^{(2)}_{\text{info}}$ is a subset of that under $\mathcal{P}^{(1)}_{\text{info}}$, and therefore for every advertiser a, slot s, and $q \in [0,1]$,

$$R^{(1)}_{as}(q) \;\geq\; R^{(2)}_{as}(q),$$

where $R^{(\ell)}_{as}(q)$ denotes the optimal single-advertiser revenue curve under $\mathcal{P}^{(\ell)}_{\text{info}}$ for $\ell \in \{1,2\}$.

Thus, the optimal single-advertiser revenue curves under $\mathcal{P}^{(1)}_{\text{info}}$ weakly dominate those under $\mathcal{P}^{(2)}_{\text{info}}$. Alaei et al. (2013) show that the marginal-revenue mechanism preserves such orderings: if one collection of single-advertiser problems yields weakly higher revenue curves than another, then the corresponding multi-advertiser marginal-revenue mechanisms also yield weakly higher expected revenue. Applying this result to the two information partitions under consideration, we obtain

$$\mathrm{Rev}_{\mathrm{IBPA}}\left(\mathcal{P}_{\mathrm{info}}^{(1)}, \mathcal{P}_{\mathrm{disc}}\right) \;\geq\; \mathrm{Rev}_{\mathrm{IBPA}}\left(\mathcal{P}_{\mathrm{info}}^{(2)}, \mathcal{P}_{\mathrm{disc}}\right).$$

This completes the proof of information monotonicity.

Now, fix an information partition $\mathcal{P}_{\mathrm{info}}$. Let $\mathcal{P}_{\mathrm{disc}}^{(1)}$ and $\mathcal{P}_{\mathrm{disc}}^{(2)}$ be two disclosure partitions such that

$$\mathcal{P}_{\mathrm{info}} \succeq \mathcal{P}_{\mathrm{disc}}^{(1)} \succeq \mathcal{P}_{\mathrm{disc}}^{(2)},$$

so that both pairs $(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{(1)})$ and $(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{(2)})$ are feasible.

To illustrate the construction, begin with the full-information regime $\mathcal{P}_{\mathrm{info}}^{\mathrm{full}} = \{\{1\}, \ldots, \{T\}\}$ and consider two disclosure partitions: a finer partition with two blocks $D_1 = \{1, \ldots, T'\}$ and $D_2 = \{T'+1, \ldots, T\}$, and the coarser partition $\mathcal{P}_{\mathrm{disc}}^{(2)} = \{\{1, \ldots, T\}\}$. Let δ_1 and δ_2 denote the probabilities that the realized type falls in D_1 and D_2 under the prior.

Under the finer disclosure regime $\mathcal{P}_{\mathrm{disc}}^{(1)}$, a single-advertiser mechanism specifies one menu of lottery pricings for impressions in D_1 and another menu for impressions in D_2. For any pair of lotteries (ℓ_1, ℓ_2) from these two menus, we may define a single lottery ℓ under the coarser disclosure regime $\mathcal{P}_{\mathrm{disc}}^{(2)}$ whose allocation probabilities over types in D_1 and D_2 are those of ℓ_1 and ℓ_2, respectively, and whose payment equals the probability-weighted average $\delta_1 \cdot \mu(\ell_1) + \delta_2 \cdot \mu(\ell_2)$. Because both the advertiser's utility and the publisher's expected revenue are linear in allocations and payments, ℓ replicates exactly the ex-ante expected allocation and payment of choosing ℓ_1 when D_1 is realized and ℓ_2 when D_2 is realized. Hence, any single-advertiser mechanism feasible under $(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{(1)})$ can be simulated under $(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{(2)})$ by offering the combined menu of all such lotteries. In particular, the feasible set of single-advertiser mechanisms under the finer disclosure partition is a subset of the feasible set under the coarser partition.

Therefore, letting $R_{as}^{(j)}(q)$ denote the optimal single-advertiser revenue curve for advertiser a, slot s, and allocation constraint q under disclosure regime $j \in \{1, 2\}$, we obtain

$$R_{as}^{(1)}(q) \;\leq\; R_{as}^{(2)}(q) \quad \text{for all } a, s, q.$$

Alaei et al. (2013) show that the marginal-revenue mechanism preserves such orderings: if one environment generates single-advertiser revenue curves that are everywhere weakly lower than those of another environment, then the corresponding multi-advertiser marginal-revenue mechanism yields weakly lower expected revenue. Applying this result to the two disclosure regimes and using the inequality $R_{as}^{(1)}(q) \leq R_{as}^{(2)}(q)$ yields

$$\mathrm{Rev}_{\mathrm{IBPA}}\Big(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{(1)}\Big) \;\leq\; \mathrm{Rev}_{\mathrm{IBPA}}\Big(\mathcal{P}_{\mathrm{info}}, \mathcal{P}_{\mathrm{disc}}^{(2)}\Big).$$

The argument above was presented for expositional simplicity with two disclosure blocks and the full-information partition. The construction extends directly to any finite pair of disclosure partitions $\mathcal{P}_{\mathrm{disc}}^{(1)} \succeq \mathcal{P}_{\mathrm{disc}}^{(2)}$ and any information partition $\mathcal{P}_{\mathrm{info}}$ such that $\mathcal{P}_{\mathrm{info}} \succeq \mathcal{P}_{\mathrm{disc}}^{(1)}$. This completes the proof of disclosure monotonicity. $\square$

Proof of Proposition 8. Fix a single advertiser a_0 and the highest available slot s_0. Because the multiplicative factor $\alpha_{s_0}\gamma_{a_0}$ is irrelevant for approximation ratios, we work with the slot-normalized problem in Equation (5). A single-advertiser mechanism is a menu of options $\{(\chi^{(k)}, \mu^{(k)})\}_{k=1}^{K}$, where $\chi^{(k)} \in [0,1]^T$ specifies, for each inventory type t, the probability $\chi_t^{(k)}$ that the advertiser is served when type t realizes, and $\mu^{(k)}$ is the corresponding normalized expected payment. Let ξ_k denote the ex-ante probability (under the advertiser's prior) that option k is chosen. For any such menu M, define its normalized expected revenue and its ex-ante allocation probability:

$$\mathrm{Rev}(M) \;\equiv\; \sum_{k=1}^{K} \xi_k\, \mu^{(k)}, \qquad \mathrm{Alloc}(M) \;\equiv\; \sum_{k=1}^{K} \xi_k \Big(\sum_{t=1}^{T} p_t\, \chi_t^{(k)} \Big).$$

For a mechanism class $\mathcal{C} \subseteq \mathcal{M}$, let

$$\mathrm{OPT}_{\mathcal{C}}(q) \;\equiv\; \sup\big\{ \mathrm{Rev}(M) : M \in \mathcal{C},\; \mathrm{Alloc}(M) \leq q \big\}$$

denote the optimal value at ex-ante cap q. Our goal is to show that for every $q \in [0,1]$, $\mathrm{OPT}_{\mathcal{M}^{\mathrm{bin}}}(q)$ and $\mathrm{OPT}_{\mathcal{M}^{\mathrm{add}}}(q)$ approximate $\mathrm{OPT}_{\mathcal{M}}(q)$ within a constant factor under independent values and within $O(\log T)$ under arbitrary correlations.

For any mechanism class $\mathcal{C} \subseteq \mathcal{M}$ and any multiplier $\lambda \geq 0$, define the Lagrangian value

$$\Pi_{\mathcal{C}}(\lambda) \;\equiv\; \sup_{M \in \mathcal{C}} \big\{ \mathrm{Rev}(M) - \lambda \, \mathrm{Alloc}(M) \big\}.$$

For any $\lambda \geq 0$ and any mechanism M feasible at cap q (i.e., $\mathrm{Alloc}(M) \leq q$),

$$\mathrm{Rev}(M) \;\leq\; \mathrm{Rev}(M) - \lambda \mathrm{Alloc}(M) + \lambda q \;\leq\; \Pi_{\mathcal{M}}(\lambda) + \lambda q.$$

Taking the supremum over all $M \in \mathcal{M}$ feasible at q yields the weak-duality bound

$$\mathrm{OPT}_{\mathcal{M}}(q) \;\leq\; \Pi_{\mathcal{M}}(\lambda) + \lambda q \qquad \forall \lambda \geq 0. \tag{6}$$

We now show that the single-advertiser problem is isomorphic to a seller offering multiple items (one per type t) to a single buyer with additive per-item values. For a realized valuation vector $v = (v_1, \ldots, v_T)$, the (normalized) expected utility of option k is $U^{(k)}(v) = \sum_{t=1}^{T} p_t \, \beta_t \, v_t \, \chi_t^{(k)} \;-\; \mu^{(k)}$. Define the transformed "item values" $v'_t \equiv p_t \beta_t v_t$ for $t \in [T]$. Then $U^{(k)}(v) = \sum_t v'_t \chi_t^{(k)} - \mu^{(k)}$, which is exactly the utility of an *additive* buyer in a multi-item problem who receives (possibly randomized) quantities $\chi_t^{(k)}$ of each item t and pays $\mu^{(k)}$. In particular: allowing $\chi^{(k)} \in [0,1]^T$ corresponds to allowing lotteries/randomized bundles (the full class $\mathcal{M}$), restricting $\chi^{(k)} \in \{0,1\}^T$ corresponds to deterministic bundles (the class $\mathcal{M}_{\mathrm{bin}}$), and imposing additive pricing $\mu(\chi) = \rho_0 + \sum_t p_t \chi_t \rho_t$ corresponds to item pricing (with an entry fee) for an additive buyer (the class $\mathcal{M}_{\mathrm{add}}$).

Fix $\lambda \geq 0$. For any menu option (χ, μ), the Lagrangian objective contributes

$$\mu \;-\; \lambda \sum_{t=1}^{T} p_t \chi_t \;=\; \mu \;-\; \sum_{t=1}^{T} c_t \, \chi_t, \qquad \text{where } c_t \equiv \lambda p_t.$$

Thus $\Pi_{\mathcal{M}}(\lambda)$ is exactly the optimal expected profit in the additive multi-item problem in which serving "item" t incurs (expected) cost c_t per unit of χ_t. Equivalently, define a shifted payment $\tilde{\mu} \equiv \mu - \sum_t c_t \chi_t$ and shifted item values $\tilde{v}_t \equiv v'_t - c_t$. Then buyer utility is unchanged:

$$\sum_t v'_t \chi_t - \mu \;=\; \sum_t \tilde{v}_t \chi_t - \tilde{\mu},$$

and seller profit becomes exactly the shifted revenue $\tilde{\mu}$. Hence maximizing profit with costs $\{c_t\}$ is isomorphic to revenue maximization for an additive buyer with (shifted) values $\{\tilde{v}_t\}$; shifting by constants preserves independence/correlation structure across items.

Known approximation results for additive single-buyer multi-item mechanism design therefore apply: Babaioff et al. (2020) show that when valuations for items are drawn independently (from a product distribution), the better of selling items separately (per-item pricing) and selling the grand bundle achieves a universal constant-factor approximation to optimal revenue. In particular, the optimal revenue is at most six times the revenue from the better of these two mechanisms. When valuations are arbitrarily correlated, Chawla, Teng, and Tzamos (2019) show that selling items separately guarantees an $O(\log T)$ approximation.

In our setting, both item pricing and grand-bundle pricing are contained in $\mathcal{M}_{\text{add}}$ (as special cases of additive pricing menus). Separate good pricing corresponds to setting $\rho_0 = 0$ in the class of mechanisms with binary allocations and additive pricings, and grand-bundle pricing corresponds to setting $\rho_t = 0\ \forall t$. Thus, the revenue achievable under binary allocations and additive pricings inherits the same constant-factor and $O(\log T)$ approximation guarantees. In particular, there exists a factor c such that, for every $\lambda \geq 0$,

$$\Pi_{\mathcal{M}}(\lambda) \ \leq\ c\, \Pi_{\mathcal{M}_{\text{add}}}(\lambda). \tag{7}$$

Because $\mathcal{M}^{\text{add}}$ (and $\mathcal{M}^{\text{bin}}$) is closed under convex combinations, the curve $q \mapsto \text{OPT}_{\mathcal{M}^{\text{add}}}(q)$ is concave by the same argument as Proposition (2). Fix any $q \in (0,1)$ and choose any subgradient $\lambda_q \in \partial \text{OPT}_{\mathcal{M}^{\text{add}}}(q)$. Concavity implies that q maximizes the function $q' \mapsto \text{OPT}_{\mathcal{M}^{\text{add}}}(q') - \lambda_q q'$, and consequently

$$\Pi_{\mathcal{M}^{\text{add}}}(\lambda_q) \ =\ \text{OPT}_{\mathcal{M}^{\text{add}}}(q) - \lambda_q q. \tag{8}$$

Combining (6) (with $\lambda = \lambda_q$), (7), and (8) gives

$$\text{OPT}_{\mathcal{M}}(q) \ \leq\ \Pi_{\mathcal{M}}(\lambda_q) + \lambda_q q \ \leq\ c\, \Pi_{\mathcal{M}^{\text{add}}}(\lambda_q) + \lambda_q q \ \leq\ c\big(\Pi_{\mathcal{M}^{\text{add}}}(\lambda_q) + \lambda_q q\big) \ =\ c\, \text{OPT}_{\mathcal{M}^{\text{add}}}(q),$$

where the penultimate inequality uses $c \geq 1$. Thus, for every $q \in [0,1]$, $\mathcal{M}^{\text{add}}$ achieves a c-approximation to the *ex-ante constrained* optimum in $\mathcal{M}$ (constant c under independent values and $c = O(\log T)$ under arbitrary correlations). Since $\mathcal{M}^{\text{add}} \subseteq \mathcal{M}^{\text{bin}}$, the same guarantee holds for $\mathcal{M}^{\text{bin}}$ as well. $\square$

Proof of Proposition 9. Results such as Lee and Valiant (2016) show that polynomial-time global optimization is possible for broad classes of non-convex objectives in both the number of dimensions T and the tolerance parameter $1/\epsilon$. To establish the stated result, it therefore suffices to show that a single step of such an optimization routine—namely, the evaluation of the objective and the constraint—can be carried out in $O(T)$ operations.

Because the objective and the constraint are expectations over the advertiser's valuation distribution, we approximate them using sample averages. Replacing exact expectations with sample averages increases the number of iterations required to reach a given tolerance level, but does not affect the complexity of each step with respect to T. Thus, it remains to show that, for any given sample of the advertiser's valuation vector, the optimal bundle can be identified in $O(T)$ time.

Let σ be a binary vector (excluding the all-zero vector) representing a bundle of inventory types, where $\sigma_t = 1$ indicates inclusion of type t. The advertiser's utility from bundle σ is

$$U(\sigma) = -\rho_0 + \sum_t p_t \sigma_t (v_t - \rho_t).$$

We determine the optimal bundle by examining each inventory type t as follows:

Types with large surplus. If $v_t > \rho_t + \frac{\rho_0}{p_t}$ for some type t, i.e., $-\rho_0 + p_t(v_t - \rho_t) > 0$, then the bundle containing only type t yields positive utility. In addition, adding t to any other bundle strictly increases utility, so the optimal bundle must include t.

Types with negative surplus. If $v_t < \rho_t$, then adding t decreases utility in any bundle. Therefore, the optimal bundle does not contain such types.

Intermediate types. If $\rho_t < v_t < \rho_t + \frac{\rho_0}{p_t}$, then type t is not profitable on its own, but if another type with large surplus is already included, adding t increases utility. In this case, the optimal bundle must include both the large-surplus type and t.

Residual case. If no type satisfies the large-surplus condition, the only candidate optimal bundle is the set of all types with $v_t > \rho_t$. Denote this bundle by σ. If

$$-\rho_0 + \sum_t p_t \sigma_t (v_t - \rho_t) > 0,$$

then σ is optimal; otherwise, the null bundle is optimal.

This procedure requires only a single pass through the T inventory types, and thus runs in $O(T)$ time. Combined with the guarantees of derivative-free optimization, this establishes that the single-advertiser problem with binary allocations and additive pricings can be solved in time polynomial in T and $1/\epsilon$. $\square$

Proof of Proposition 10. Fix an arbitrary partition $\mathcal{P}$ and consider the regime $(\mathcal{P}, \mathcal{P})$, in which the publisher both observes and discloses the block of $\mathcal{P}$ containing the realized inventory type. Under this regime, each advertiser is characterized by a single parameter—its valuation for the block of $\mathcal{P}$ containing the realized type. This is a standard single-parameter environment with quasi-linear preferences. As a result, the marginal-revenue approach results in the unique revenue-maximizing auction over all BIC and IR mechanisms (Myerson 1981; Bulow and Roberts 1989). In particular, by the argument in Theorem 5,

$$\mathrm{Rev}_{\mathrm{IBPA}}(\mathcal{P}, \mathcal{P}) \geq \mathrm{Rev}_{\mathrm{GSP}}(\mathcal{P}, \mathcal{P}).$$

Next we show that the additive class $\mathcal{M}_{\mathrm{add}}$ is rich enough to implement the single-advertiser mechanisms of the Myerson-optimal mechanism in the regime $(\mathcal{P}, \mathcal{P})$. Under full disclosure, the single-agent problem for each type t reduces to offering a posted price for an impression of type t: the optimal mechanism allocates the impression to advertiser a if and only if v_{at} exceeds a type-specific reserve price r_t, and charges r_t per click upon allocation. Such a mechanism is contained in $\mathcal{M}_{\mathrm{add}}$ by setting the entry fee $\rho_0 = 0$ and choosing the per-type coefficient ρ_t equal to the reserve price r_t. Consequently, the single-advertiser mechanisms underlying the Myerson-optimal mechanism—and hence IBPA$(\mathcal{P}, \mathcal{P})$—belongs to $\mathcal{M}_{\mathrm{add}}$. Because IBPA$(\mathcal{P}, \mathcal{P})$ is already optimal among all BIC and IR mechanisms in this single-parameter environment, restricting attention to $\mathcal{M}_{\mathrm{add}}$ cannot raise its revenue, and the supremum over $\mathcal{M}_{\text{add}}$ coincides with the supremum over $\mathcal{M}$. Therefore,

$$\text{Rev}_{\text{IBPA}^{\text{add}}}(\mathcal{P}, \mathcal{P}) = \text{Rev}_{\text{IBPA}}(\mathcal{P}, \mathcal{P}).$$

Finally, we compare the regimes $(\mathcal{P}, \mathcal{P})$ and $(\mathcal{P}, \mathcal{P}^{\text{null}}_{\text{disc}})$ under IBPA^{add}. Under $(\mathcal{P}, \mathcal{P}^{\text{null}}_{\text{disc}})$, the publisher still observes the realized inventory type t and the information block in $\mathcal{P}$, but discloses only the null disclosure block to advertisers. Any additive mechanism that is feasible under $(\mathcal{P}, \mathcal{P})$ can be implemented under $(\mathcal{P}, \mathcal{P}^{\text{null}}_{\text{disc}})$ by offering the same menu of additive prices and allocations, independent of t, and setting $\rho_0 = 0$. Because advertisers observe less information about t, the incentive constraints are weakly relaxed, so the feasible set of additive mechanisms under $(\mathcal{P}, \mathcal{P}^{\text{null}}_{\text{disc}})$ weakly contains that under $(\mathcal{P}, \mathcal{P})$. In addition, under null disclosure there is no need to impose the restriction $\rho_0 = 0$; allowing ρ_0 to vary can only enlarge the feasible set further. Hence,

$$\text{Rev}_{\text{IBPA}^{\text{add}}}(\mathcal{P}, \mathcal{P}^{\text{null}}_{\text{disc}}) \geq \text{Rev}_{\text{IBPA}^{\text{add}}}(\mathcal{P}, \mathcal{P}).$$

Combining the results,

$$\text{Rev}_{\text{IBPA}^{\text{add}}}(\mathcal{P}, \mathcal{P}^{\text{null}}_{\text{disc}}) \geq \text{Rev}_{\text{IBPA}^{\text{add}}}(\mathcal{P}, \mathcal{P}) = \text{Rev}_{\text{IBPA}}(\mathcal{P}, \mathcal{P}) \geq \text{Rev}_{\text{GSP}}(\mathcal{P}, \mathcal{P}),$$

which establishes the proposition. $\square$

Web Appendix B: Quantile Mapping and Payments

In Section 3.4.3, we derived a valuation-to-quantile mapping under the simplifying assumption that the single-advertiser mechanisms $\{M_a(q)\}_{q\in[0,1]}$ exhibit nested deterministic allocations. Here, we develop a general mapping without relying on that assumption. Our construction mirrors that of Alaei et al. (2013) for single-resource allocation mechanisms.

B.1 Preliminaries

Alaei et al. (2013) show that, for the marginal-revenue mechanism to be well defined and incentive compatible, it is necessary and sufficient that the mapping from advertiser type $\boldsymbol{v}_a$ to quantile q_a satisfy the following property: when $\boldsymbol{v}_a$ is drawn from its prior distribution F_a, the resulting quantile q_a is distributed uniformly on $[0, 1]$. We therefore begin by assuming the existence of such a mapping, and we show how to construct one.

Fix a slot s. When the quantile mapping is applied across advertisers, each advertiser draws a quantile $q_a \sim \mathrm{Unif}[0, 1]$. Given these quantiles, the marginal-revenue allocation rule described in Section 3.4.4 assigns the slot to the advertiser with the largest value of $\gamma_a \Phi'_a(q_a)$, where Φ'_a is the derivative of the normalized revenue curve.

This marginal revenue allocation rule induces, for each advertiser a, an allocation function

$$x_a^{MR}(q) \in [0, 1],$$

defined as the probability that advertiser a receives the slot when its quantile is q and all other advertisers draw quantiles independently from the uniform distribution.

B.2 Interim Mechanism as a Convex Combination of $\{M_a(q)\}$

Alaei et al. (2013) show that $x_a^{MR}(q)$ completely determines the interim mechanism faced by advertiser a. In particular, the interim mechanism is a convex combination of the optimal single-agent mechanisms $\{M_a(q)\}_{q\in[0,1]}$, with weights determined by the derivative $\frac{-dx_a^{MR}(q)}{dq}$.

Two observations support this construction. First, $x_a^{MR}(q)$ is weakly decreasing in q. This follows from Proposition 2. Because $\Phi'_a(q)$ is decreasing in q, an advertiser with a higher quantile has a weakly lower marginal revenue and therefore a weakly lower probability of receiving the slot. Second, we can normalize the allocation rule so that $x_a^{MR}(0) = 1$ (the strongest type always wins) and $x_a^{MR}(1) = 0$ (the weakest type never wins).

Thus, $-dx_a^{MR}(q)/dq$ is a valid probability density on $[0, 1]$. Drawing a quantile $\tilde{q}$ from this density and offering the advertiser the single-agent mechanism $M_a(\tilde{q})$ produces an interim mechanism whose allocation and payment rules match those induced by the marginal-revenue allocation rule.

B.3 Constructing the Quantile Mapping

The interim mechanism yields, for each type v_a, a vector of allocation probabilities $x_{at}(v_a)$ across inventory types t. For a fixed inventory type t, these allocation probabilities induce an ordering over advertiser types. This ordering allows us to define the inventory-type-specific quantile mapping:

$$Q_{at}(v_a) = \Pr_{v'_a \sim F_a}\left[x_{at}(v'_a) \geq x_{at}(v_a)\right].$$

This definition assigns lower quantiles to types with higher interim allocation probabilities and ensures:

$$Q_{at}(v_a) \sim \text{Unif}[0, 1] \quad \text{when } v_a \sim F_a.$$

One remaining subtlety is that revenue curves may contain regions where $\Phi'_a(q)$ is constant (i.e., intervals $[q_1, q_2]$ over which marginal revenue does not change). To ensure coherence of the quantile mapping, Alaei et al. (2013) show that all types whose quantiles fall in such an interval should be treated symmetrically. This is accomplished by resampling uniformly within the interval:

$$q \in [q_1, q_2] \quad \Longrightarrow \quad q \leftarrow q_1 + U \cdot (q_2 - q_1),\ U \sim \text{Unif}[0, 1].$$

This resampling preserves uniformity of the final quantiles and ensures that types receiving the same marginal revenue also receive the same quantile ex post.

B.4 Payments

Finally, the payment rule follows directly from the interim mechanism. For a fixed inventory type t and advertiser type v_a, the interim mechanism provides an allocation probability $x_{at}(v_a)$ and an expected payment $m_a(v_a)$. These can be implemented by charging

$$\frac{m_a(v_a)}{x_{at}(v_a)}$$

whenever advertiser a receives a slot of type t. As in Section 3.4.5, this produces the correct expected payment. If payments are collected on a per-click basis, we divide by the realized click-through rate, exactly as described in Section 3.4.5.

This completes the general valuation-to-quantile mapping and payment construction without relying on nested deterministic allocations.