

Auctions Meet Bandits: An Empirical Analysis

Mohammad Rashid*

University of Washington

Omid Rafieian*

Cornell University

Soheil Ghili*

Yale University

August 27, 2025

Abstract

Sponsored search positions are typically allocated through real-time auctions, where the outcomes depend on advertisers’ quality-adjusted bids—the product of their bids and quality scores. Although quality scoring helps promote ads with higher conversion outcomes, setting these scores for new advertisers in any given market is challenging, leading to the cold-start problem. To address this, platforms incorporate multi-armed bandit algorithms in auctions to balance exploration and exploitation. However, little is known about the optimal exploration strategies in such auction environments. We utilize data from a leading Asian mobile app store that places sponsored ads for keywords. The platform employs a Thompson Sampling algorithm within a second-price auction to learn quality scores and allocate a single sponsored position for each keyword. We empirically quantify the gains from optimizing exploration under this combined auction–bandit model and show that this problem differs substantially from the canonical bandit problem. Drawing on these empirical insights, we propose a customized exploration strategy in which the platform adjusts the exploration levels for each keyword according to its characteristics. We derive the Pareto frontier for revenue and efficiency and provide actionable policies, demonstrating substantial gains for the platform on both metrics when using a tailored exploration approach.

Keywords: auctions, bandits, Thompson Sampling, advertising, platforms, market thickness

*All authors contributed equally and are listed in reverse alphabetical order. We would like to thank the participants of the 2025 Marketing Science Conference. We also thank Garrett Johnson, Pallavi Pal, Valeria Stourm, and Yuyan Wang for their detailed comments, which have significantly improved the paper. Please address all correspondence to: soheil.ghili@yale.edu, or83@cornell.edu, and rashid98@uw.edu.

1 Introduction

The auction problem and the multi-armed bandit (MAB) problem are both fundamental problems in the fields of economics and computer science. In applied settings, firms often need to solve problems that combine elements from auctions and bandits. A prime example is sponsored search. In sponsored-search settings, for each search term, a platform auctions off advertisement slots to bidders whose ads have different probabilities of conversion (e.g., click/installment, etc.). These conversion probabilities are unknown to the platform a priori, rendering the problem one of multi-armed bandit, in which the platform needs to balance exploration against exploitation to achieve near optimal performance. At the same time, advertisers have heterogeneous valuations for each conversion, known to them but unknown to the platform. The bidding process based on which the object is awarded to the bidder based on their bid is an auction problem. The score for each bidder in such an auction is calculated by combining its bid amount with the platform’s most current belief about the bidder’s conversion probability. This renders the platform’s problem one that combines auction and bandit features.

Although the auction and bandit literatures are more or less separate, the overlap between the two is a small but growing body of research ([Devanur and Kakade 2009](#), [Babaioff et al. 2015](#), [Feng et al. 2023](#)). A common feature of all these studies is that they all investigate the theoretical properties of solutions to the auction-bandit problem. Our objective, however, is *empirical*. In this paper, we empirically examine how introducing the bandit component and fine-tuning the extent of exploration in these algorithms influence market outcomes such as revenue and efficiency. Specifically, we quantify the benefits a platform can achieve by optimizing its exploration strategy under a combined auction–bandit model and highlight how this setting differs from a bandit-only problem. To our knowledge, this is the first attempt to document the empirical dynamics of the auction–bandit problem and connect them to market outcomes in a real-world environment.

We theoretically characterize a general auction-bandit model, where a seller auctions uncertain events with probabilities that can be learned sequentially. This framework encompasses classical auction design and bandit problems as special cases. We then characterize how the exploration challenge differs from the standard bandit problem, as rewards (prices) depend not only on the winning buyer, but also on other players in a competitive market environment. As an empirical application of the general problem, we analyze large-scale data from sponsored search auctions run by a leading app store in a major Asian country, where we observe complete information about all advertisers competing for an impression. The app store employs a Thompson Sampling algorithm to update advertisers’ quality scores,

integrated with a second-price auction. The level of exploration in ad allocation is high enough to allow us to estimate advertisers’ conversion rates, while still preserving room for strategic bidding behavior. This enables us to infer advertisers’ valuations per conversion, assuming they are utility-maximizing. Together, our setting allows us to identify advertisers’ *private valuation for conversion* and *conversion rates* as the mechanism-invariant primitives, sufficient for evaluation of market outcomes under counterfactual mechanisms.

We focus on a class of Thompson Sampling Second-Price (TS-SP) auctions that only differ from the mechanism used in the data in the prior distribution of quality scores and evaluate market outcome under each. More specifically, we consider the optimal choice of a prior belief on the conversion rate (“CVR”) of any new bidder, and we study how this optimal choice is shaped by the auction-bandit problem. To this end, we compare two metrics. First, we calculate the maximum revenue that the platform can achieve by optimally choosing the prior within a given range, and the optimal prior that gives rise to this maximum revenue. This is the “full” problem in which both the auction and the bandit problems play roles in shaping the optimal choice. Second, we calculate the optimal prior, and the revenue it yields, under a conversion-rate prior that is optimized for efficiency, i.e., for allocating each object to the bidder who has valuation times conversion rate. We interpret this second optimization as one that treats the platform’s problem as essentially an bandit only. Comparisons between these two measures, hence will shed light on how the auction side of the problem impacts the bandit strategy.

We find that the interactions between the two sides of the platform’s problem are indeed of empirical relevance. Specifically, we find that to maximize efficiency, the platform optimally chooses a prior of 0.002 for entrants’(new ads) conversion rate. This level already captures the optimal degree of exploration that is necessary for solving a bandit problem. When we solve for the revenue-maximizing prior, however, the recommended prior is at the substantially larger level of 0.1. In addition, the ensuing expected revenue is about 32% higher than the revenue arising from a prior of 0.002. We interpret this empirical comparison as affirmative of the relevance of the combined auction-bandit problem.

We further study the mechanism by which the auction and bandit problems interact. In particular, we find that higher priors often help boost revenue by elevating the entrant’s score to the second place and close to that of the incumbent winner, thereby allowing the auctioneer to extract additional surplus from the winner. This revenue channel does not exist in pure bandit problems. We further confirm the empirical relevance of this revenue channel by decomposing the platform revenue into revenue accrued from auctions where (i) the entrant wins, (ii) the entrant finishes second, and (iii) the entrant finishes third or lower. Our results show that the revenue maximizing prior of 0.1 improves upon the profitability of

the efficiency maximizing prior of 0.002 by boosting the revenue in channel (ii) while revenue from channel (i) falls, though to a lesser degree.

We interpret our empirical finding as follows: a higher conversion rate prior in an auction-bandit problem helps make the market effectively “thicker” by pushing the second score closer to the first score. Consistent with this interpretation, we find that in auctions for keywords where the first and second scores are typically more distant (we term these keywords “thinner” markets), the revenue difference between the revenue-maximizing and efficiency-maximizing conversion-rate priors is the highest.

Further, we investigate the role of customizing the prior policy for each separate keyword on which auctions are run. We find that customization in the auction-bandit problem improves the performance not only by boosting revenue and efficiency individually (which happens by construction), but also by “easing” the trade-off between the two: targeting helps to boost revenue for auctions that benefit from higher priors without compromising on the efficiency in auctions where a high prior for entrant might adversely impact the allocation.

We extend our analysis by introducing Upper Confidence Bound Second-Price (UCB-SP) mechanisms as another common exploration method in auctions. Using the same setting and data, we find that a UCB-style exploration policy improves the revenue of the optimal TS-SP by about 26%, a gain we trace to UCB’s optimistic quality-score assignments to non-winner advertisers. Importantly, the revenue–efficiency tradeoff is similar to that of TS-SP: more exploratory UCB policies yield higher revenue at the cost of efficiency. As a robustness check to the main analysis, we relax our assumptions on advertisers’ budget and further limit bidder spending by imposing a daily cap equal to each bidder’s maximum observed spend. The results of this robustness show qualitative and quantitative patterns similar to our main analysis.

In summary, our paper makes important contributions to the literature while offering practical implications for managers and policymakers. Our key substantive contribution lies in documenting the nuances of the combined auction–bandit problem and how it differs from traditional bandit or auction settings. We empirically show that framing the problem of prior optimization as a combined auction–bandit problem leads to substantially higher priors and significantly greater revenue compared to treating it solely as a bandit problem. Further, we show that customizing priors for keywords is an actionable approach that balances the revenue–efficiency trade-off and generates significant improvements in market outcomes. While the primary managerial implications concern advertising platforms, our framework applies broadly to any setting where an auctioneer sells an uncertain event and must learn bidder relevance scores. Finally, our findings carry important policy implications for pricing transparency, showing that platforms can exploit exploration as a revenue-generating tool,

sometimes at the expense of overall market efficiency.

2 Related Literature

Our paper relates to the literature on online advertising platforms. Early theoretical studies in this area have examined the equilibrium properties of different auction formats in sponsored search (Edelman et al. 2007, Varian 2007, Lahaie et al. 2007, Lahaie and Pennock 2007). Empirical research has explored various aspects of advertising auctions, including measuring market outcomes under different auctions (Yao and Mela 2011, Athey and Nekipelov 2010, Kim and Pal 2025), incorporating consumer search behavior (Jeziorski and Segal 2015, Choi and Mela 2019), audience externalities and ad avoidance (Wilbur et al. 2013, Stourm and Bax 2017), reserve price optimization (Ostrovsky and Schwarz 2023), and dynamic allocation mechanisms (Rafieian 2020). Our study extends this body of work by empirically evaluating different bandit-style exploration strategies that platforms use to learn quality scores over time. In particular, we show that the choice of exploration strategy meaningfully affects market outcomes, highlighting it as an important managerial lever for advertising platforms.

Our work is also related to the market thickness and revenue-efficiency trade-off literature. Early studies have highlighted the critical importance of market thickness in determining market outcomes (Levin and Milgrom 2010). To increase market thickness, several methods have been proposed from a range of perspectives, such as utilizing randomized allocation (Celis et al. 2014), adjusting the exponent on quality scores in standard GSP auctions—a modification motivated by (Lahaie and Pennock 2007), reducing the level of targeting (Rafieian and Yoganarasimhan 2021), and giving extra exposure to new advertisers to reduce churn (Ye et al. 2022). We contribute to this literature by examining how exploration increases revenue by improving market thickness and reducing dominant players’ informational rents. We demonstrate this by empirically benchmarking the performance of two prominent bandit algorithms: Thompson Sampling and UCB.

Finally, our work is closely related to the literature on the interplay between auctions and bandits. Recent research has made significant theoretical advances in developing algorithms that improve regret bounds in this context (Babaioff et al. 2009, Devanur and Kakade 2009, Babaioff et al. 2015, Feng et al. 2023). Our work adds to this literature in two ways. First, rather than focusing on asymptotic regret bounds, we provide finite-sample empirical estimates of market outcomes using large-scale data and quantify the benefits of considering the full auction-bandit problem compared to the bandit problem alone. Second, we extend the analysis of auction-bandit problems to Thompson Sampling in second-price auctions, providing valuable insights for practitioners looking to leverage Thompson Sampling algorithms

due to their strong real-world performance (Chapelle and Li 2011).

3 Auction-Meet-Bandit Problem

We consider a general problem where a seller owns a sequence of items (e.g., advertising slots) and sells an uncertain event that can occur on these items (e.g., a click or conversion). There is a set of buyers who want to buy these events, and the seller wants to decide how to allocate items to these buyers. Let t denote the sequence of items and a denote each buyer. Each buyer a has a private valuation for the uncertain event in item t , denoted by $v_{a,t}$, and an event probability, denoted by μ_a , that is fixed for the sequence of items. The seller does not know $v_{a,t}$ and μ_a at the beginning but can learn μ_a over time as it observes whether the uncertain event occurs. The seller’s goal is to offer a mechanism $M(p, e)$ with an allocation rule p and a payment rule e that maximizes its revenue as follows:

$$\operatorname{argmax}_{p,e} \mathbb{E} \left[\sum_t \sum_a \mu_a p_a(v_t, \hat{\mu}_t) e_a(v_t, \hat{\mu}_t) \right], \quad (1)$$

where v_t is the vector of buyers’ private valuations, $\hat{\mu}_t$ is the vector of the seller’s beliefs about the event probabilities, $p_a(v_t, \hat{\mu}_t)$ is the allocation outcome for buyer a given the valuation vector v_t and belief vector $\hat{\mu}_t$, and $e_a(v_t, \hat{\mu}_t)$ is the payment for buyer a given the valuation and belief vectors.

We characterize the problem above as the auction-bandit problem because the seller needs to deal with v_t through auction design and learn μ_a sequentially based on the observed events as in the bandits literature. To further illustrate this, we highlight two special cases: (1) if the μ_a ’s are known to the seller ex-ante, the problem simplifies to a standard auction design problem and can be solved using methods from the optimal auctions literature (Myerson 1981), and (2) if payments are fixed and equal for all buyers, the problem reduces to the classical bandits problem, extensively studied in prior work (Lattimore and Szepesvári 2020).¹

In the general case, uncertainty about μ_a necessitates the seller’s exploration of buyers through the allocation rule. Our goal is to determine the optimal level of exploration in this setting. In the canonical bandits problem, the optimal exploration strategy is shaped by two opposing forces: under-exploration may lead to missed opportunities in discovering high-quality buyers, while over-exploration can result in inefficient allocation to low-quality

¹It is worth noting that maximizing efficiency instead of revenue will be a weighted variant of the bandit problem, under a direct revelation mechanism where bidders truthfully report their valuations. Mathematically, in that case, we replace $e_a(v_t, \hat{\mu}_t)$ with v_t ’s that are reported by bidders.

buyers. The auction-bandit problem introduces an additional force—since prices are influenced not only by the winning bidder but also by market competition, exploration increases market thickness, which can translate into higher revenue for the seller through increased payments. In this paper, we empirically examine the interaction of these three forces in a sponsored search auction setting.

4 Setting and Data

4.1 Setting

Our data come from a leading Android app store in a large Asian country, which holds over 70% of the market share in the Android app store category. The app store has a total of 100 million users, with more than 40 million active users each month. On the supply side, the app store hosts nearly one million mobile apps developed by 80,000 developers. Like most app stores, its primary function is to serve as a matchmaker between users and apps. The main source of revenue for the app store comes from taking a share of transactions between users and apps facilitated through the platform. In this study, we focus on another key monetization strategy of the app store: sponsored search advertising.

4.1.1 Sponsored Search Auction

Users frequently search for apps by entering keywords, with nearly 5.5 million search sessions taking place on an average day. When a user submits a query, the app store’s algorithm provides a rank-ordered list of organic search results relevant to the query. In addition to the organic results, the app store offers a single sponsored slot, allowing apps to pay for greater visibility at the top of the search results. As is typical in search advertising, the app store uses an auction to allocate the sponsored slot to an app. As such, the app store is the auctioneer in our setting and mobile apps are advertisers. Throughout the paper, we use apps and ads interchangeably.

The app store in our study runs a pay-per-install second-price auction, which works as follows:

1. When a user searches for a keyword, the app store’s server receives a request to allocate a sponsored position to an app through an auction. The winning app in this auction will be displayed to the user as a sponsored app at the top of the organic search results.
2. Once the sponsored position for the keyword is identified, the app store sends a request to apps to submit their bids, indicating how much they are willing to pay for an app

install. We define app a 's bid as b_a , which represents the amount they are willing to pay for an install.

3. The app store assigns each participating app a a quality score q_a , which reflects the expected install probability or conversion rate for that app. The app store then calculates a quality-adjusted bid by multiplying the bid by the quality score: $b_a q_a$.
4. The app with the highest quality-adjusted bid is selected. Without loss of generality, assume $b_1 q_1 > b_2 q_2 > \dots > b_M q_M$. If the highest quality-adjusted bid, $b_1 q_1$, exceeds the reserve price, app 1 wins the auction and is shown to the user in the sponsored position, generating an impression.
5. App 1, the winning app, only pays if the user installs the app through the sponsored link. In our context, clicks that do not result in an install do not trigger payment. However, the winning app does not pay its own bid but instead pays the second price, which is the minimum amount required to win the auction. In this case, app 1 pays $b_2 q_2 / q_1$ per install.

Figure 1 provides a simple illustration of this process. In the example presented in this figure, a user searches for *games* and the app store displays both sponsored and organic app listings. The organic list is determined by the app store's algorithm, which ranks apps based on their relevance to the search query. There is a single sponsored slot, allocated through a pay-per-install second-price auction. Here, Apps B, D, and E compete for the sponsored slot by submitting bids of 2, 6, and 5, respectively. The app store then assigns quality scores to these apps, representing the probability that a user will install them. Using these scores, it calculates quality-adjusted bids for each app. The sponsored position is awarded to the highest quality-adjusted bid, provided it exceeds the reserve price. In this case, App E wins the auction but only pays if the user installs the app. The amount paid is not App E's original bid but the minimum bid required to outbid the second-highest quality-adjusted bid. In this example, App B has the second-highest quality-adjusted bid. If this bid is above the reserve price, App E's payment per install will be the smallest amount necessary to ensure its quality-adjusted bid remains higher than that of App B. If the second-highest bid is below the reserve price, the payment is calculated based on the amount required to exceed the reserve price.²

²If the reserve price is higher than the second-highest quality-adjusted bid, the bid is adjusted to ensure the winning quality-adjusted bid surpasses the reserve price.

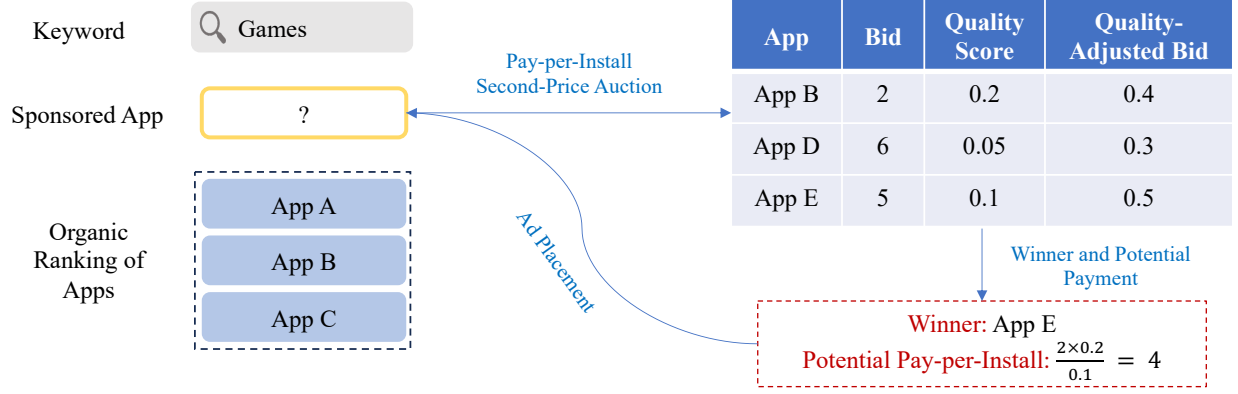


Figure 1. Sponsored position allocation through the auction used by the app store

4.1.2 Exploration through Thompson Sampling

Quality scores are essential in ensuring profitable allocation through auctions. In the example presented in Figure 1, the highest bid does not win the auction because its quality score is substantially lower than competitors. The advertising platforms often use massive amounts of data to fine-tune these quality scores. However, a major challenge is that the app store does not have data on the new apps who want to participate in the auction, leading to a cold-start problem.

To address the cold-start problem, the app store employs a Thompson Sampling approach that starts with a prior quality score distribution for the new app participating in the auction and continuously updates the quality score distribution based on the impressions allocated to this app in a Bayesian fashion. In particular, the app store in our study uses the Beta distribution $\text{Beta}(1, 9)$ to draw the initial quality score from. Suppose app a is new in an auction. We denote the quality score assigned to this app in time period t by $q_{a,t}$. In the first period, we have $q_{a,0} \sim \text{Beta}(\alpha_{a,0} = 1, \beta_{a,0} = 9)$, which has a mean $\alpha_{a,0}/(\alpha_{a,0} + \beta_{a,0}) = 0.1$. If the app wins an impression that does not lead to a conversion, the beta parameter in the next step will increase by one point, i.e., $\beta_{a,1} = \beta_{a,0} + 1$. However, only if this new impression leads to a conversion, the alpha parameter will be updated. That is, $\alpha_{a,1} = \alpha_{a,0} + 1$ if a conversion event happens and $\alpha_{a,1} = \alpha_{a,0}$. Repeating this process results in continuous updating of the quality score distribution using the past conversion history.

It is easy to verify that after a large number of impressions being allocated to an app, the quality score will be almost the same as its historical conversion rate. For example, suppose that an app has received 10,000 impressions and 200 conversions (historical conversion is 0.02), and the app store uses the prior policy where quality scores are drawn from $\text{Beta}(\alpha_{a,0} = 1, \beta_{a,0} = 9)$ in the first period. After 10,000 impressions, the posterior distribution of quality

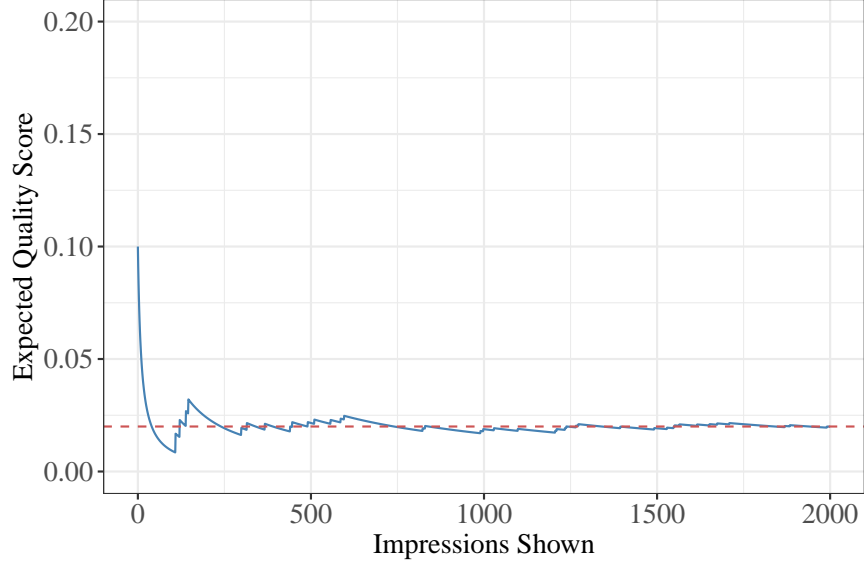


Figure 2. Posterior mean of quality score over time for one trajectory

score will be $\text{Beta}(\alpha_{a,10000} = 1 + 200 = 201, \beta_{a,10000} = 9 + 10000 - 200 = 9809)$, whose mean is $201/(201 + 9809) \approx 0.02$. Figure 2 shows the evolution of the posterior mean for one trajectory of events. As shown in this figure, the posterior mean starts from 0.1, fluctuates in early periods when one more conversion has a larger impact, and then converges to the true conversion rate 0.02.

4.2 Data

We have impression-level data from an app store over the span of six months. As discussed earlier, each impression is a sponsored position for a keyword searched by a user that is allocated to an app through the auction mechanism described above. A unique feature of our dataset is that, in addition to observing information about the winning bidder selected for the impression, we also observe details about the losing bidders. This is different from most high-frequency auction data sets where we only have information about the winning bidder, which allows us to conduct more robust counterfactual analysis. In §4.2.1, we describe the variables we observe for each impression. Next, in §4.2.2, we discuss our sampling strategy.

4.2.1 Variables

For each impression (auction) in our data, we observe the following sets of information:

- *Timestamp*: This presents the exact time of the impression.
- *Keyword-specific information*: This information set includes the information about the keyword searched by the user and whether it is a brand or non-brand keyword. Consistent

with the definition used in the literature, a brand keyword is a keyword or phrase that directly references a specific app in the app store.

- *Bidder-specific information*: For each bidder who participates in the auction, we observe their bid, the quality score drawn from their quality score distribution, and the α and β parameters of their quality score distribution.
- *Reserve price*: This is the price set by the app store as the minimum quality-adjusted bid needed for participating in the auction. This helps discard bidders whose quality-adjusted bid (product of bid and quality score) is lower than the reserve price.
- *Winning app*: This presents the bidder (app) who wins the auction. This is the bidder with the highest quality-adjusted bid (if any). There are auctions that are unallocated because no bidder’s quality-adjusted bid surpassed the reserve price.
- *Conversion*: This indicates whether the impression of the winning app resulted in an install or conversion event.
- *Potential revenue-per-conversion*: This is the amount that the winning app would be charged if a conversion event happens. In the second-price auction run by the app store, this is the lowest bid the winning app could have submitted to maintain the winning spot, as illustrated in Figure 1.
- *Revenue*: This information shows the amount of revenue earned by the app store for the impression. If the conversion event happens, this amount is equal to the potential revenue-per-conversion. Otherwise, it is equal to zero.

The detailed set of variables observed at the auction level allows us to evaluate counterfactual policies different from the one implemented in the data in terms of a variety of market outcomes.

4.2.2 Sample Construction

To reduce the computational burden of our analysis, we follow the convention in the literature and employ a sampling strategy (Athey and Nekipelov 2010, Rafeian and Yoganarasimhan 2021). We focus on the first week of our data which contains 38.6 million impressions and 446828 conversion events. We specifically focus on impressions generated by non-brand keywords.³ We focus on the top 1,800 non-brand keywords that account for 70% of the total non-brand keyword revenue and generate 5.18 million impressions during the first week we

³There is no specific restriction that prevents us from using the brand keywords. However, those keywords are naturally thin markets where there is a dominant player. As we will discuss later, market thickness is an important consideration in our analysis. Focusing on non-brand keywords allows us to share insights into how market thickness affects the app store’s decision on how much to explore. Non-brand keywords represent 52% of all impressions generated and 27% of the total revenue.

choose as our analysis sample.

An important feature of our data is the extent of exploration in ad allocation through the Thompson Sampling algorithm that allows us to observe the conversion rates for a wide range of apps. This feature enables us to estimate the true conversion rates and use them in our counterfactual analysis. Although we focus on the sample of auctions run in the first week of our data, we use the entire six months of data to estimate conversion rates to increase the reliability of our estimates and shrink the confidence bounds around them.

4.3 Summary Statistics

We now present some summary statistics on our data.

4.3.1 Keyword-level Characteristics

We begin by presenting the keyword-level characteristics. As outlined earlier, we focus on the top 1,800 non-brand keywords. For each keyword, we measure the following variables: (1) total number of impressions, (2) total number of conversions, (3) conversion rate, (4) total number of bidders, (5) number of bidders per impression, (6) total number of distinct ads shown, (7) total number of entrants, (8) total revenue, and (9) revenue per impression. Table 1 provides a summary of how these keyword-level characteristics are distributed across the top 1,800 keywords. It is important to note that the reported revenue figures have been scaled by a constant factor for confidentiality.

Variable	Mean	SD	Min	Median	Max
Total No. of Impressions	2880	5333	35	1434	99303
Total No. of Conversions	29.1	92.1	1	9	2592
Conversion Rate	0.022	0.048	0.0002	0.0058	0.49
Total No. of Bidders	38.2	7.7	18	37	74
Avg. No. of Bidders per Impression	11.6	7.7	1.1	10.2	40.6
Total No. of Distinct Ads Shown	31.4	8.2	2	32	69
Total No. of Entrant Bidders	21.6	11.7	3	20	64
Total Revenue	666	1468	8.3	233	29729
Avg. Revenue per Impression	0.93	2.65	0.0014	0.14	38.1

Table 1. Summary Statistics of the Keyword-level Characteristics.

As shown in Table 1, we observe substantial heterogeneity across keywords in their key characteristics. In particular, we note a few patterns in this table. First, conversion rates vary widely across keywords, ranging from as low as 0.0002 to as high as 0.49. Second, we observe that the number of distinct ads shown also reveals considerable variability, with auctions dominated by as few as two distinct ads shown and others that allocate to as many as

64 bidders. This pattern suggests a substantial variation in overall market competitiveness. Finally, a comparison of total bidders and entrant bidders per keyword shows that entrants constitute a significant share of participation, with the median number of entrants exceeding half of the median number of bidders. This pattern emphasizes the prominence of the cold-start problem in our empirical setting.

4.3.2 Exploration

We now turn to one of the key ingredients of our study: exploration. As shown in Table 1, many keywords have a large number of entrants who enter their market at some point. The app store’s strategy to overcome this cold-start problem is to assign a relatively high prior mean to the quality score distribution of entrants, so these entrants could win auctions and the app store could gradually converge to their true quality score. In this section, we show some evidence on the extent of exploration in our data and the motivation for exploration.

Figure 3 displays the empirical cumulative distribution function for the number of impressions received by entrants during the six-month period following their first bid. To capture the extent of exploration, we focus on all entrants who began bidding for the first time in the first week of data. They were then tracked over six months to record the total impressions allocated. As illustrated, 80% of these entrants amassed at least 100 impressions, and 20% exceeded 1,000 impressions, indicating a substantial level of exploration directed toward entrants. Later in our counterfactual analysis, we want to examine what would happen under different exploration policies. The extent of exploration in ad allocation allows us to use reliable estimates for entrants conversion rates.

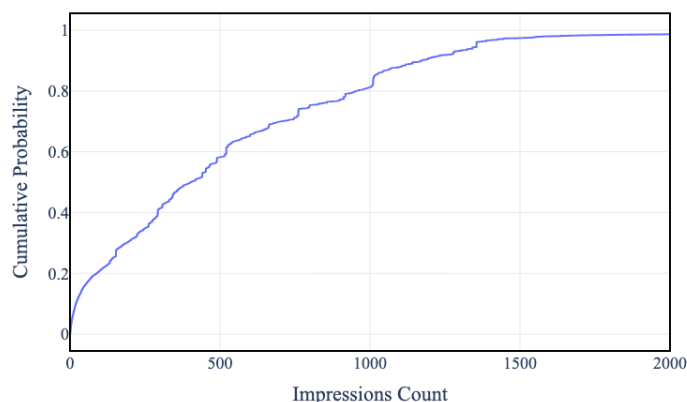


Figure 3. Empirical Cumulative Density Function (ECDF) of Impression Count for Entrants

Next, we turn to a more fundamental question: why does an app store engage in explo-

ration? As highlighted in our theoretical framework, the primary motivation for exploration is the potential to discover high-quality ads that create value for the app store. To empirically assess this argument, we compare the aggregated conversion rates of entrants and incumbents for each keyword. A higher conversion rate for entrants relative to incumbents indicates greater value derived from exploring new entrants. Figure 4 presents the empirical cumulative distribution function (CDF) of the difference in conversion rates between entrants and incumbents for each keyword. The differences in average conversion rates highlight the importance of exploration in the auction mechanism. Specifically, for 12% of keywords, entrants achieve a higher average conversion rate than incumbents. This finding suggests that while incumbents generally dominate due to their past performance and revenue generation, entrants can occasionally outperform them in certain contexts. These results underscore the necessity of continued exploration to identify high-performing entrants and prevent the auction mechanism from prematurely settling on suboptimal allocations driven solely by historical performance.

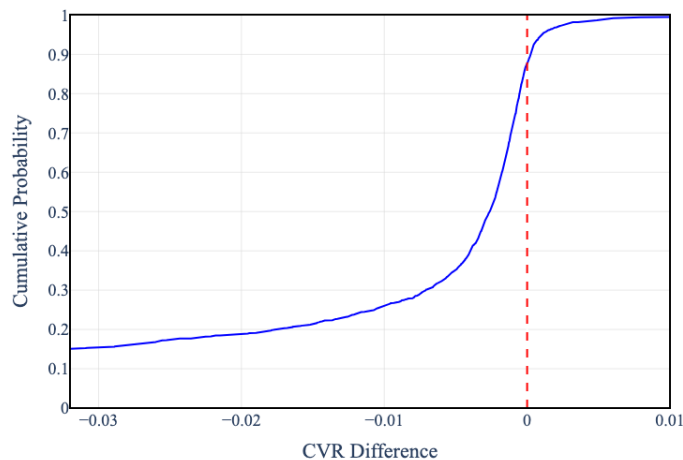


Figure 4. Empirical CDF of the Difference Between Average Conversion Rates of Entrants and Incumbents

4.3.3 Market Thickness

If a market is more competitive, entrants will face greater difficulty in securing impressions. As such, the choice of exploration policy becomes more sensitive to market thickness. In auction settings, market thickness can be quantified by the ratio of the second-highest bid to the highest bid, which reflects the fraction of the top bid that can be extracted as revenue in standard auction formats, such as second- or first-price auctions (Myerson 1981). In our

study, we define market thickness as the average ratio of the second-highest bidding score to the highest bidding score for a given keyword. By construction, thickness values lie between 0 and 1, with higher values indicating a more competitive auction environment.

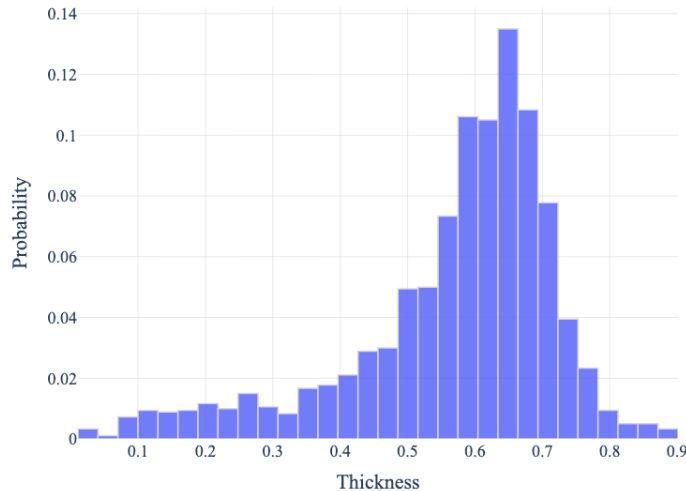


Figure 5. Histogram of the Market Thickness for Keywords

Figure 5 presents the histogram of market thickness, revealing substantial heterogeneity in its distribution. The largest concentration in the distribution lies between 0.6 and 0.7, indicating that, for a significant portion of keywords, the top bidder extracts an informational rent of approximately 30-40% of their valuation for an impression. We also observe instances of very thin markets, where the ratio of the second-highest bid to the highest bid is less than 0.1. Conversely, there are reasonably competitive markets with an average ratio close to 0.9. Using this measure of market thickness, we classify markets into three categories: the top 25% of values are labeled as “thick” markets, the bottom 25% as “thin” markets, and the interquartile range as “medium” markets—representing markets that are neither too competitive nor too thin.

4.3.4 Role of Entrants in Market Outcomes

We now present some statistics highlighting the role of entrants in determining key market outcomes, such as revenue. In principle, an entrant can impact market revenue in two ways. First, by taking the top place in the auction, the entrant wins the impression. Second, by finishing in the second place, the entrant determines the price the winning bidder must pay. Conversely, if the entrant ends up in the third place or lower, it does not influence the allocation or pricing of the auction.

To illustrate how entrants impact these allocation and pricing outcomes, Figure 6 shows the fraction of auctions in which entrants occupy each of the top five places. Although only the first place leads to an impression (allocation outcome), analyzing all top places reveals how frequently entrants appear in each rank, emphasizing their potential influence on pricing. The horizontal axis represents the rank (1 through 5), while the vertical axis shows the proportion of auctions in which an entrant holds that rank. Notably, entrants occupy approximately 18% of first places across all auctions, indicating they account for a substantial share of impressions. Additionally, entrants take around 20% of second places, which implies that their scores directly determine the auction’s revenue. These findings demonstrate the significant role entrants play in shaping key market outcomes and highlights the value of exploration policies for the app store.

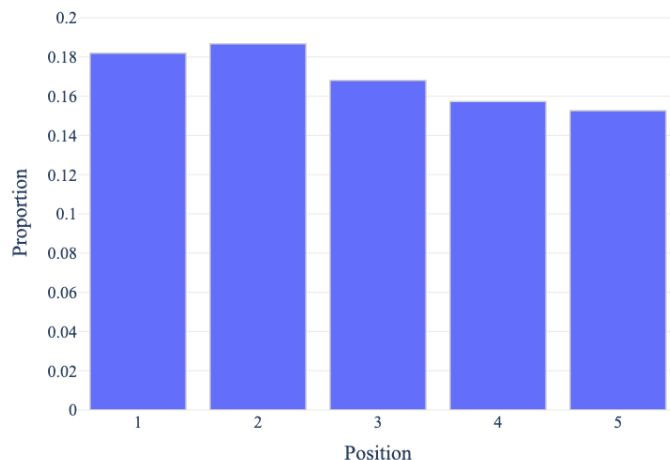


Figure 6. Proportion of Each Place Won by an Entrant

To illustrate how entrants influence keyword revenue, we categorize revenue into three groups: (1) revenue from impressions won by entrants (Entrant-Win), (2) revenue from impressions where an entrant finishes second and influences pricing (Entrant-Second), and (3) revenue from impressions where incumbents take the top two places, leaving entrants no role in allocation or pricing (Entrant-None). Figure 7 presents these results. Although the revenue share from entrants winning impressions is relatively small, entrants significantly affect pricing by finishing at the second place. Notably, the revenue share from entrants in second place is higher in thinner markets, where the gap between the highest and second-highest bidding scores is larger. The cross-keyword heterogeneity shown in Figure 7 underscores that the optimal exploration policy strongly depends on market thickness—a key concept we explore in our main analysis.

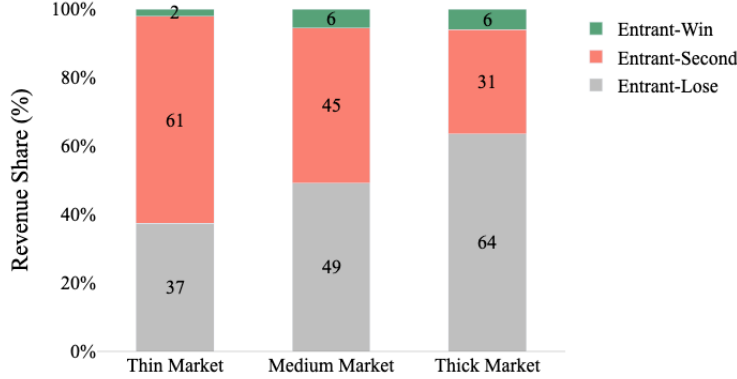


Figure 7. Revenue Share Based on Market Thickness and the Role of Entrants

5 Empirical Framework

5.1 Problem Definition

5.1.1 Data Structure

We start by defining some notation to characterize the market for sponsored keyword auctions in our data. In this market, there are advertisers, indexed by a , who compete for impressions associated with keywords, indexed by k . Let \mathcal{A} and \mathcal{K} denote the sets of all ads and keywords, respectively. We index impressions of a keyword $k \in \mathcal{K}$ with time period t and define $\mathcal{A}_{k,t}$ as the set of ads competing in the auction, such that $\mathcal{A}_{k,t} \subseteq \mathcal{A}$.

As discussed in §4.1, the allocation and revenue outcomes of the auction are determined by the product of each ad's bid and its quality score, where the quality scores are drawn probabilistically from the quality score distribution. For each ad a competing for impression t of keyword k , let $b_{a,k,t}$ denote the submitted bid by the advertiser, which is a function of their private valuation for a conversion $v_{a,k,t}$ as well as market characteristics. The quality score, denoted by $q_{a,k,t}$, is drawn from a quality score distribution that is continuously updated to reflect the advertiser's true conversion rate. This updating process is because the app store does not know the quality scores ex-ante and needs to explore and learn them over time. In our context, the app store uses a Beta distribution for the quality score distribution, given its nice conjugacy properties that simplify the updating process. Let $\alpha_{a,k,t}$ and $\beta_{a,k,t}$ denote the alpha and beta parameters of the quality score distribution for ad a in impression t of keyword k , that is, $q_{a,k,t} \sim \text{Beta}(\alpha_{a,k,t}, \beta_{a,k,t})$.

With bids and quality scores for all participating ads available, the app store runs a second-price auction to allocate the impression. The impression is then allocated to the highest product of the bid and quality score $(b_{a,k,t}q_{a,k,t})$ if it is higher than the reservation score set by the app store. We denote the winning ad that is shown in the impression for keyword k at time t by $A_{k,t}$, such that $A_{k,t} = \operatorname{argmax}_{a \in \mathcal{A}_{k,t}} b_{a,k,t}q_{a,k,t}$. Once the ad is shown to the user, the user decides whether to install the app, determining the conversion outcome $Y_{k,t}$. We follow the potential outcomes framework and define a set of potential outcomes for all possible ads for all impressions $\{Y_{k,t}(a)\}_{a \in \mathcal{A}}$, where $Y_{k,t} = Y_{k,t}(A_{k,t})$. Since the conversion outcome is binary, we associate it with a Bernoulli distribution with parameter $\mu_{a,k,t}$, such that $Y_{k,t}(a) \sim \operatorname{Bern}(\mu_{a,k,t})$.

Based on the ad allocation and conversion outcomes realized in each auction, the Bayes' updating rule for the parameters of the quality score distributions has a simple form as follows:

$$(\alpha_{a,k,t+1}, \beta_{a,k,t+1}) = (\alpha_{a,k,t}, \beta_{a,k,t}) + \mathbb{1}(A_{k,t} = a)(Y_{k,t}, 1 - Y_{k,t}) \quad (2)$$

According to this updating rule, if ad a does not win auction t in keyword k (that is, $a \neq A_{k,t}$), there will be no change in the parameters of the quality score distribution. However, if ad a wins, one unit will be added to either α or β depending on the conversion outcome $Y_{k,t}$.

As discussed in §4.1.2, the motivation for using posterior sampling and updating is to overcome the cold-start problem, wherein a new ad enters the market for a keyword. Let α_0 and β_0 denote the alpha and beta parameters of the prior distribution, respectively. In our setting, the app store sets the alpha and beta parameters of the prior distribution at 1 and 9, respectively. Using Equation (2), it is easy to map the parameters to the historical performance of each ad as follows:

$$(\alpha_{a,k,t}, \beta_{a,k,t}) = (\alpha_0 + C_{a,k,t}, \beta_0 + I_{a,k,t} - C_{a,k,t}), \quad (3)$$

where $C_{a,k,t}$ and $I_{a,k,t}$ are the total number of conversions and impressions ad a has received in keyword k up until impression t , respectively.

5.1.2 Exploration Policy

One of the main goals of our paper is to evaluate market outcomes under alternative exploration policies. To do so, we first need to formally define an exploration policy. As shown in Equation (3), for a given trajectory of allocation and conversion events, the only parameters that control the extent of exploration in the mechanism are the parameters of the prior quality score distribution. As such, we define the potential set of exploration policies as follows:

Definition 1. The *Thompson Sampling exploration policy* is characterized by two functions $\alpha^0 : \mathcal{A} \times \mathcal{K} \rightarrow \mathbb{R}$ and $\beta^0 : \mathcal{A} \times \mathcal{K} \rightarrow \mathbb{R}$ that set the initial parameters of the Beta distribution for an ad $a \in \mathcal{A}$ that enters the market for keyword $k \in \mathcal{K}$. That is, if ad a enters the market for keyword k in time period t , then we have $\alpha_{a,k,t} = \alpha^0(a, k)$ and $\beta_{a,k,t} = \beta^0(a, k)$.

In our data, the app store uses $\alpha^0(a, k) = 1$ and $\beta^0(a, k) = 9$ for any ad a and keyword k . For a fixed policy that does not vary with ads or keywords, the app store can change the parameters to increase or decrease the extent of exploration. Further, our definition is flexible and allows for customized exploration based on the ad and keyword.

Importantly, the exploration policy can be embedded within the allocation rule and payment rule in any auction mechanism, as it only controls the distribution from which quality scores are drawn. To be consistent with our empirical context, we focus on the second-price auction used in our setting and describe how the exploration policy is incorporated as follows:

Definition 2. The *Thompson Sampling Second-Price (TS-SP) auction mechanism* $M(p, e; \alpha^0, \beta^0)$ is characterized by allocation rule p , payment rule e , and the Thompson Sampling exploration policy with functions (α^0, β^0) . This mechanism first draws quality scores from the quality score distributions of all participating ads and then runs the same allocation and payment rule as in the second-price auction. We summarize these three steps as follows:

- *Quality Score Sampling:* The first step is quality score sampling. Let $C_{a,k,t}^M$ and $I_{a,k,t}^M$ denote the total number of conversions and impressions ad a receives in keyword k up until impression t under mechanism M . For each ad $a \in \mathcal{A}_{k,t}$, the quality score distribution is obtained by applying Equation (3), and the quality score $q_{a,k,t}^M$ is drawn as follows:

$$q_{a,k,t}^M \sim \text{Beta}(\alpha^0(a, k) + C_{a,k,t}^M, \beta^0(a, k) + I_{a,k,t}^M - C_{a,k,t}^M)$$

- *Allocation:* Let $b_{a,k,t}^M$ denote the bid submitted by ad a in auction t of keyword k under mechanism M . The allocation function takes all bids and quality score draws as inputs and determines the winner as the one with the highest product of bid and quality score draw as follows:

$$p(\{b_{a,k,t}, q_{a,k,t}\}_{a \in \mathcal{A}_{k,t}}) = \arg \max_{a \in \mathcal{A}_{k,t}} b_{a,k,t} q_{a,k,t}$$

- *Payment:* Let $A_{k,t}^M$ denote the ad selected in the allocation stage. The payment by this ad

would be calculated as follows:

$$e(\{b_{a,k,t}^M, q_{a,k,t}^M\}_{a \in \mathcal{A}_{k,t}}) = Y_{k,t}^M \left(\frac{\max_{a \in \mathcal{A}_{k,t} \setminus A_{k,t}^M} b_{a,k,t}^M q_{a,k,t}^M}{q_{A_{k,t}^M, k, t}^M} \right),$$

where $Y_{k,t}^M$ is the conversion outcome for the winning ad under mechanism M . In words, the payment is extracted from the winning ad only if a conversion happens and is equal to the minimum bid that guarantees winning the auction.

Unlike the standard second-price auction with fixed quality scores, the allocation outcome of the TS-SP auction is probabilistic. To characterize the allocation probability for each ad a , we define distribution π that depends on the vector of bids and past performance metrics. Let $B_{k,t}^M$, $I_{k,t}^M$, and $C_{k,t}^M$ denote the vector of bids, the total number of impressions, and the total number of conversions under mechanism M , respectively. We can define the conditional distribution π as follows:

$$\pi(a \mid B_{k,t}^M, I_{k,t}^M, C_{k,t}^M; \alpha^0, \beta^0) = \mathbb{E}_{Q_{k,t}^M \sim \text{Beta}(\alpha_{k,t}^M, \beta_{k,t}^M)} \left[\mathbb{1}(\arg \max_{a \in \mathcal{A}_{k,t}} b_{a,k,t}^M q_{a,k,t}^M = a) \right], \quad (4)$$

where $Q_{k,t}^M$, $\alpha_{k,t}^M$, and $\beta_{k,t}^M$ are all vectors of quality score draws and parameters of the Beta distributions for all ads.

5.1.3 Expected Market Outcomes

Our main goal is to evaluate the performance of any TS-SP auction $M \in \mathcal{M}$ in terms of the key market outcomes: revenue and efficiency. To do so, we first characterize these expected outcomes. For any TS-SP auction $M(p, e; \alpha^0, \beta^0)$, we denote the expected revenue by $\mathbb{E}[\text{Rev}(M)]$ and characterize it as follows:

$$\mathbb{E}[\text{Rev}(M)] = \mathbb{E} \left[\sum_{k \in \mathcal{K}} \sum_{t=1}^{T_k} Y_{k,t}^M \left(\frac{\max_{a \in \mathcal{A}_{k,t} \setminus A_{k,t}^M} b_{a,k,t}^M q_{a,k,t}^M}{q_{A_{k,t}^M, k, t}^M} \right) \mid M(p, e; \alpha^0, \beta^0) \right], \quad (5)$$

where the expectation is taken over the randomness induced by updating the quality score distribution and the conversion outcome $Y_{k,t}^M$ is drawn from a Bernoulli distribution with parameter $\mu_{A_{k,t}^M, k, t}$. It is worth emphasizing that the reason for the superscript M here is to emphasize that all those values—including bidding, the ad shown, and the conversion outcome—could change under the counterfactual TS-SP auction M .

We denote the expected efficiency or total surplus in the market for auction M by

$\mathbb{E}[\text{Eff}(M)]$ and define it as follows:

$$\mathbb{E}[\text{Eff}(M)] = \mathbb{E} \left[\sum_{k \in \mathcal{K}} \sum_{t=1}^{T_k} Y_{k,t}^M v_{A_{k,t}^M, k, t} \mid M(p, e; \alpha^0, \beta^0) \right], \quad (6)$$

where $v_{A_{k,t}^M, k, t}$ is the private valuation for a conversion for the ad winning impression t of keyword k under mechanism M , which is denoted by $A_{k,t}^M$.

In summary, to evaluate expected market outcomes under a counterfactual mechanism, we need to reliably simulate a trajectory of bids $b_{a,k,t}^M$, quality score draws $q_{a,k,t}^M$, ad allocation $A_{k,t}^M$ and conversion outcomes $Y_{k,t}^M$. In the next section, we discuss our empirical strategy to identify all the primitives needed to deliver this counterfactual evaluation.

5.2 Empirical Strategy

In this section, we outline our empirical strategy for estimating key market outcomes. To evaluate these outcomes, we need to estimate primitives from both the advertiser and consumer sides that enable us to simulate outcomes over a trajectory. We argue that by knowing each advertiser’s valuation for conversion, $v_{a,k,t}$, and their conversion rate, $\mu_{a,k,t}$, we can simulate all the elements necessary to estimate expected revenue and efficiency. This is because advertisers’ private valuations for conversion govern their bidding behavior, allowing us to simulate how they would bid in a counterfactual auction. Conversion rates, on the other hand, enable us to simulate $Y_{k,t}^M$, which appears in both expected revenue and efficiency calculations, and influences the state transitions of the quality score distributions. As such, our approach consists of two separate empirical tasks: (1) identifying $v_{a,k,t}$ and (2) identifying $\mu_{a,k,t}$. We provide details on our estimation of advertisers’ valuations in §5.2.1, and our strategy for estimating conversion rates in §5.2.2.

5.2.1 Estimation of Advertisers’ Valuation for Conversion

In this section, we estimate advertisers’ private valuations for conversions. A standard assumption in the auction estimation literature is that agents act to maximize their utility, enabling us to infer their private valuations directly from their bidding behavior (Guerre et al. 2000, Athey and Haile 2007). The specific version of this assumption that we adopt in our setting is as follows:

Assumption 1. *Bidders choose their bids to maximize their per-period expected utility.*

We now explain why this assumption is reasonable in our empirical context. An alternative assumption is that advertisers maximize utility over the long term rather than

on a per-period basis. The rationale for long-term utility maximization lies in the sequential updating of quality scores—an advertiser’s bid in the current period can influence how these scores evolve, thereby shaping future allocation and payment outcomes. To empirically test which assumption better aligns with the data, we consider different scenarios. If an advertiser’s bid does not affect quality score distributions, long-term utility maximization simplifies to per-period utility maximization. Thus, in settings where quality scores remain fixed over time (e.g., markets without new entrants), the distinction between these two utility-maximization assumptions becomes inconsequential.

However, the analysis becomes more nuanced when entrants enter the market and quality scores evolve dynamically. In these settings, current bidding behavior can influence the rate at which quality scores adjust, potentially affecting future auction outcomes. As such, long-term utility-maximization can lead to *incumbents’ bid-shading incentive* and *entrants’ overbidding incentives*. In the main text, we focus on the former and discuss the latter in Appendix A.

Overestimating entrants’ quality scores can result in situations where the winning incumbent ends up paying more than anticipated. In response, incumbents may reduce their bids temporarily, allowing entrants to win additional impressions. This, in turn, provides the auctioneer with more data to refine the entrants’ quality scores. Over time, this improved accuracy in quality assessment is expected to lead to lower future prices for incumbents. Fortunately, we can empirically test these theoretical claims and determine whether bidders maximize their utility statically or dynamically. We begin with the incumbents’ bid-shading incentive. According to this hypothesis, winning incumbents should reduce their bids when their costs suddenly increase, as a result of an entrant entering the market with a higher prior mean. To test this, we use data from such events and estimate the following model:

$$\bar{b}_{a,k,h} = \beta_0 + \beta_1 \bar{b}_{a,k,h-1} + \beta_2 \text{CPC}_{a,k,h-1} + \epsilon_{a,k,h}, \quad (7)$$

where $\bar{b}_{a,k,h}$ is incumbent a ’s average bid in keyword k at hour h , and $\text{CPC}_{a,k,h-1}$ is his average cost-per-conversion (CPC) in keyword k in the previous time period $h - 1$. We estimate this model and present the results in Column (1) of Table 2. As shown in this column, the CPC from the previous round does not appear to influence the incumbent’s bid in the current period. Since the CPC in the prior period may be endogenous, we further investigate this relationship using an Instrumental Variable (IV) approach. Specifically, we use the entrant’s quality score mean as the instrument to isolate the effect of the change in CPC due to the entrant’s prior quality score. The results of the 2SLS model are presented in Column (2) of Table 2, and we find the same conclusion: the increase in CPC resulting

Table 2. Regression Specifications to Test Static vs. Dynamic Bidding

	<i>Dependent variable: Bid ($\bar{b}_{a,k,h}$)</i>	
	(1) OLS	(2) 2SLS
Lagged Bid ($\bar{b}_{a,k,h-1}$)	0.9952*** (0.000)	0.9950*** (0.000)
Lagged Cost-per-Conversion ($\text{CPC}_{a,k,h-1}$)	0.0012 (0.001)	0.0019 (0.013)
R^2	0.991	0.991
No. of Obs.	92,750	92,750
<i>Note:</i>	*p<0.05; **p<0.01; ***p<0.001	

from the entrant’s higher-quality score does not lead to bid-shading behavior by the winning incumbent. This finding provides empirical evidence that supports our assumption of per-period utility maximization.

The regression analysis in Table 2 together with the analysis in Appendix A provide empirical support for Assumption 1. Based on this, we present the following proposition, which underpins our identification of advertisers’ private valuations:

Proposition 1. *Under Assumption 1, advertisers bid truthfully, that is $b_{a,k,t} = v_{a,k,t}$.*

The proof of this proposition relies on two key arguments. First, the bid placed by an advertiser at time $b_{a,k,t}$ does not influence the quality score distributions at that period, meaning the analysis is equivalent to assuming fixed quality scores. Second, for any given set of fixed quality scores, it is straightforward to demonstrate that no bidder has an incentive to deviate from truthful bidding, using the same reasoning as in Vickrey (1961).

For our estimation, we make another simplifying assumption as follows:

Assumption 2. *Advertisers’ private valuations for conversion do not change over time, that is, $v_{a,k,t} = v_{a,k}$ for any advertiser a in any keyword k .*

Therefore we adopt a simple estimation strategy that estimates advertisers’ private valuation as the average bid they submitted in that keyword as follows:

$$\hat{v}_{a,k} = \frac{\sum_{t=1}^{T_k} \mathbb{1}(a \in \mathcal{A}_{k,t}) b_{a,k,t}}{\sum_{t=1}^{T_k} \mathbb{1}(a \in \mathcal{A}_{k,t})} \quad (8)$$

Our identification strategy follows the approach used in prior literature that leverages data from second-price auctions (Celis et al. 2014). For robustness, one could modify the averaging criteria to focus only on impressions where quality score distributions are stable (i.e., low posterior variance), ensuring that dynamic utility maximization aligns with per-period utility

maximization. Our analysis indicates minimal differences between these two approaches in estimating advertisers’ private valuations for conversions.

5.2.2 Estimation of Conversion Rates

We now turn to the second empirical task: identifying $\mu_{a,k,t}$. The presence of exploration in our data allows us to reliably estimate conversion rates across a broad range of ads. However, we first introduce some simplifying assumptions:

Assumption 3. *Advertisers’ conversion rate remains fixed over time, that is, $\mu_{a,k,t} = \mu_{a,k}$.*

This assumption is widely adopted in the multi-armed bandit (MAB) literature and provides the foundation for incorporating bandit frameworks into auctions to address the cold-start problem (Feng et al. 2023). We estimate conversion rates using the following sample analogue estimator:

$$\hat{\mu}_{a,k} = \frac{\sum_{t=1}^{T_k} \mathbb{1}(a = A_{k,t}) Y_{k,t}}{\sum_{t=1}^{T_k} \mathbb{1}(a = A_{k,t})} \quad (9)$$

The estimator remains unbiased if impressions allocated to an ad are randomly drawn from the overall impression distribution. This assumption is reasonable given the posterior sampling-induced randomization in our setting. As an alternative approach, one could relax this assumption and employ Inverse Propensity Score (IPS) estimator that is shown to be unbiased (Horvitz and Thompson 1952) and other variants of IPS designed to handle adaptively collected data (Hadad et al. 2021, Zhan et al. 2021). We present the details of the IPS approach in Appendix B.

5.3 Counterfactual Policy Evaluation

As discussed earlier, our primary objective is to assess market outcomes under a counterfactual TS-SP auction that utilizes a different prior distribution for entrants. Since the estimated primitives—advertisers’ valuation per conversion ($v_{a,k}$) and conversion rate ($\mu_{a,k}$)—are invariant to the auction mechanism, they remain unchanged under a mechanism different from the one observed in the data.

Algorithm 1 outlines how we use these estimated primitives and the mechanism M as inputs to evaluate market outcomes, such as expected revenue and efficiency. As detailed in the algorithm, we exogenously fix the number of impressions per keyword (T_k) and maintain the set of ads competing in each impression ($\mathcal{A}_{k,t}$) as observed in the data.⁴ The algorithm

⁴For smaller advertisers, we can impose a cap preventing them from receiving more impressions than they

describes the process of quality score sampling, ad allocation based on sampled quality scores and advertisers' truthful bidding, conversion simulation, and the calculation of efficiency and revenue. Truthful bidding follows from the assumption that advertisers maximize their per-period expected utility. To ensure this assumption holds in our counterfactual analysis, we restrict our policies to those that are no more exploratory than the one observed in the data.

Let a^{rp} denote the index for the reserve price among ads such that $a^{\text{rp}} \in \mathcal{A}_{k,t}$. The app store defines the reserve price as a combined score RP, so we set $\hat{v}_{a^{\text{rp}},k} q_{a^{\text{rp}},k}^M = \text{RP}$. We further set $\hat{\mu}_{a^{\text{rp}},k} = 0$ so it reflects the fact that the app store does not receive any efficiency or revenue from cases where the impression is not allocated.

Algorithm 1 Counterfactual Auction Simulation Algorithm of keyword k

Input: Mechanism $M(p, e; \alpha^0, \beta^0)$, data $\{\{\mathcal{A}_{k,t}\}_t, \{I_{a,k,1}, C_{a,k,1}\}_a, T_k\}$ and estimates: $\{\hat{v}_{a,k}, \hat{\mu}_{a,k}\}_{a \in \mathcal{A}}$

Output: Estimated market outcomes: $\{\widehat{\text{Eff}}_{k,t}, \widehat{\text{Rev}}_{k,t}\}_t$

$I_{a,k,1}^M \leftarrow I_{a,k,1}$

$C_{a,k,1}^M \leftarrow C_{a,k,1}$

for $t = 1$ **to** T_k **do**

for $a \in \mathcal{A}_{k,t}$ **do**

$q_{a,k,t}^M \sim \text{Beta}(\alpha^0(a, k) + C_{a,k,t}^M, \beta^0(a, k) + I_{a,k,t}^M - C_{a,k,t}^M)$

end

$A_{k,t}^M \leftarrow \text{argmax}_{a \in \mathcal{A}_{k,t}} \hat{v}_{a,k} q_{a,k,t}^M$

$Y_{k,t}^M \sim \text{Bern}(\hat{\mu}_{A_{k,t}^M, k})$

$\widehat{\text{Eff}}_{k,t} \leftarrow \hat{v}_{A_{k,t}^M, k} Y_{k,t}^M$

$\widehat{\text{Rev}}_{k,t} \leftarrow \frac{\max_{a \in \mathcal{A}_{k,t} \setminus A_{k,t}^M} \hat{v}_{a,k} q_{a,k,t}^M}{q_{A_{k,t}^M, k, t}^M} Y_{k,t}^M$

for $a \in \mathcal{A}_{k,t}$ **do**

$I_{a,k,t+1}^M \leftarrow I_{a,k,t}^M + \mathbb{1}(a = A_{k,t}^M)$

$C_{a,k,t+1}^M \leftarrow C_{a,k,t}^M + \mathbb{1}(a = A_{k,t}^M) Y_{k,t}^M$

end

end

$\widehat{\text{Eff}}_k \leftarrow \frac{1}{T_k} \sum_{t=1}^{T_k} \widehat{\text{Eff}}_{k,t}$

$\widehat{\text{Rev}}_k \leftarrow \frac{1}{T_k} \sum_{t=1}^{T_k} \widehat{\text{Rev}}_{k,t}$

did in the data to account for potential budget constraints and/or inaccuracies in conversion rate estimates. We verify that this constraint has minimal impact on the estimated market outcomes.

6 Results

The objective of this section is to examine, qualitatively and quantitatively, how the optimal exploration policy is impacted by the fact that the problem is not a standard bandit problem, but a combined auction-bandit problem.

Our strategy to formally capture the role of auctions involves comparing two metrics: efficiency and revenue. Our measure of efficiency in a given auction is defined as the total value generated in the auction that is shared between the platform and the advertisers. One can show that maximizing efficiency is equivalent to solving a weighted bandit problem. This is because the efficiency metric is a function of bidders’ valuation only and invariant to the price of the item. As such, under a direct revelation mechanism where bidders truthfully report their valuations, efficiency maximization reduces to a weighted bandit problem: the auctioneer can treat the reported valuations as weights for arms with unknown rewards (CVR in our setting). On the other hand, maximizing the revenue depends not only on which bidder the item is allocated to, but also how much the bidder eventually pays. This “full” version of the problem, hence, involves both the auction and the bandit aspects. As a result, comparing the revenue-optimal exploration policy (i.e., optimal prior on entrants) against the efficient policy would allow us to isolate the role of the auctions in our auction-bandit problem.

We carry out the comparative analysis in two stages. First, we assume the firm has a uniform prior across each group of thin, thick, and mid-sized auctions. We will then turn to studying how the comparison between revenue- and efficiency-maximizing policies is impacted by allowing the firm to customize the prior in a more granular fashion.

6.1 Market Outcomes Under Uniform Priors

6.1.1 Revenue-Efficiency Tradeoff

Figure 8a shows how efficiency varies as we vary the prior mean $1/(1 + \beta)$ on the CVR, by varying β . As can be seen from the figure, the efficiency-maximizing prior mean is 0.002. To see whether this policy involves a degree of exploration, observe that the average CVR among entrants in our dataset is 0.0035 and the median is 0.0013. Our results show that, as expected in multi-arm bandit problems, limited exploration yields some gains. But over-exploration of entrants sharply reduces the efficiency by mis-allocating the position.

Figure 8b shows how revenue varies as we vary the prior mean. As can be seen from the figure, the revenue-maximizing prior mean is 0.1, at the end of the range examined.⁵ This

⁵As the figure suggests, the revenue is expected to further increase if we set a prior mean above 0.1. That said, we exclude prior means above 0.1 from our revenue analysis. The reason is that sufficiently high prior

stands in contrast with the efficiency-maximizing prior.

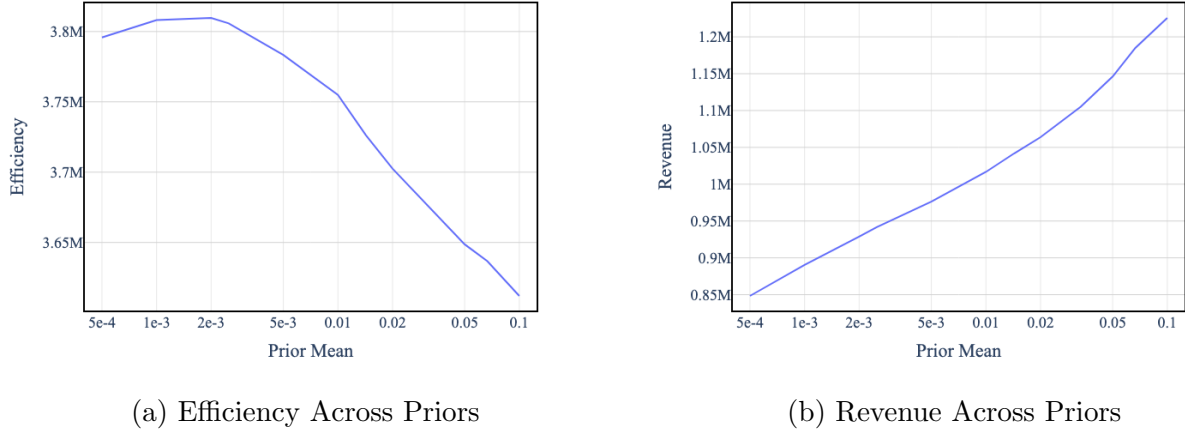


Figure 8. (a) Efficiency and (b) Revenue Across Priors (Log Scale).

The contrast between the prior means suggests auction component of the auction-bandit problem plays an empirically non-trivial role. To the extent that the result of our analysis may generalize to other online advertising settings, this result is suggestive that the auction-bandit problem is of general empirical relevance, and that further theoretical and empirical studies of this problem would be of applied value.

This contrast between the revenue- and efficiency-maximizing priors is indicative that there is a trade-off between maximizing the two objects. Quantifying this trade-off is a task that we turn to next.

Figure 9 examines possible trade-offs between revenue and efficiency. This figure plots efficiency against revenue as we vary the prior. As shown by the figure, and as expected from the previous analysis, there is no trade-off for lower priors. If a prior is below the efficiency-maximizing level, increasing it improves both the efficiency and the revenue. The trade-off appears once we reach an efficiency-maximizing level. Past this threshold, an increase in the prior improves the revenue at the expense of efficiency.

It is worth examining the empirical magnitude of the revenue-efficiency tradeoff. As Figure 9 shows, the range of the revenue as we trace a range of priors is wider than that

means impacts incentives by incumbent bidders who pay prices dictated by high scores given to entrants. Such incumbents may find it optimal to bid lower and lose a few rounds so that the entrant wins and, subsequently, gets its score negatively updated. The incumbent will then have the incentive to return to bidding truthfully, after having lowered the second-highest bid score. Such dynamic incentives impact the accuracy of our predictions, in which we assume all bidders bid truthfully in every auction. Given that we did not find evidence for such strategic and forward-looking bidding behavior under the observed prior mean of 0.1, we find the results in figure 8b convincing. Predicting revenue under higher ranges for the prior mean, however, would require a more complex model of bidder behavior.

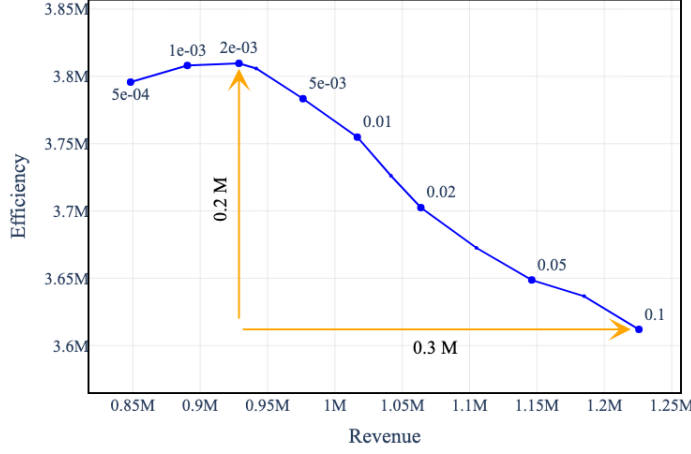


Figure 9. Points Represent Efficiency-Revenue Achieved if Using Different Priors

of the efficiency. Efficiency ranges between 3.6M and 3.8M whereas revenue ranges between 0.8M and 1.2M. The latter range is wider than the former both in absolute terms (0.4M v.s. 0.2M) and in relative terms (50% v.s. 6%). This, once again, is suggestive that the auction component of the auction-bandit problem is of substantial empirical relevance. This observation motivates a deeper analysis of the mechanisms behind the tradeoff, a task which we turn to next.

6.1.2 Discussion of the Underlying Mechanisms

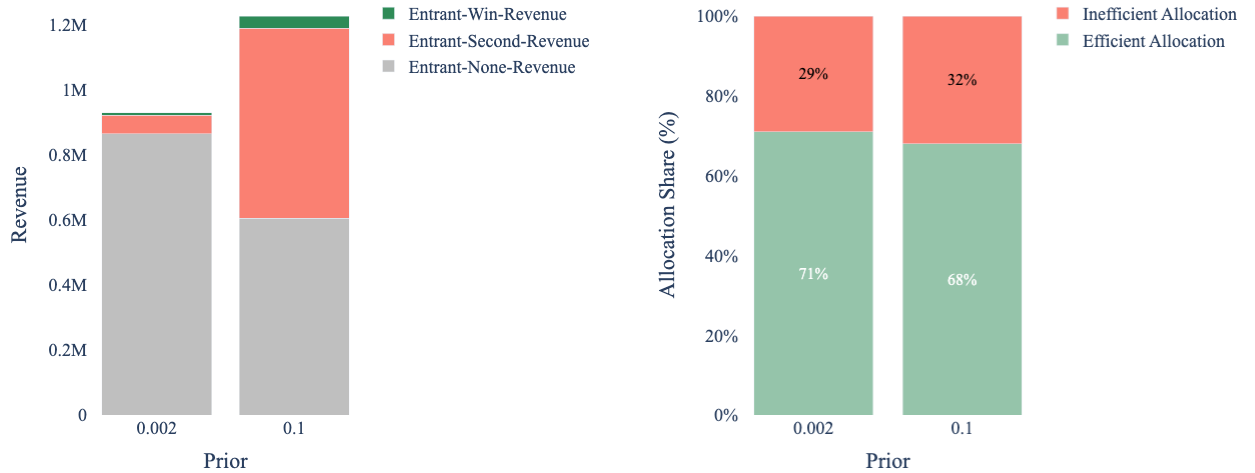
Why does a higher prior mean on entrants' CVR achieve a higher revenue? And why does it come at the expense of efficiency?

Figure 10a illustrates part of the answer. The figure presents stacked bar plots of total revenue under revenue- and efficiency-maximizing priors. Each total-revenue amount is decomposed into three components: (i) revenue from auctions in which the entrant was allocated the object, (ii) revenue from auctions in which the entrant was not allocated the object but finished with the second highest score and, hence, influenced the revenue extracted from the winner, and (iii) auctions in which the entrant ranked 3rd or lower, thereby impacting neither the allocation of the ad nor the revenue extracted.

A comparison between the efficiency- and revenue-maximizing policies shows that the latter generates significant revenue by placing the entrant in the second place, thereby impacting the revenue without changing the allocation. This is consistent with the theoretical observation by Akbarpour and Li (2020) in a different context, that the auctioneer has the

incentive to induce payment rules akin to a first-price auction in a setting where the auction rule is second-price. The difference relative to the analysis by [Akbarpour and Li \(2020\)](#) is that they study lack of truthfulness by the auctioneer, whereas in our case the seller truthfully runs a second-price auction, tuning the entrant prior optimally.

Relatedly, for the auctioneer, it is a consequence to truthfully amplify entrant bids as a means to extract more revenue from the winner. The consequence entails the possibility that the entrant may indeed win, leading to an increased likelihood for inefficient allocation. This can be seen in Figure 10b. Of course such a compromise in efficiency indirectly hurts the revenue. That said, as long as the revenue effect of the efficiency compromise does not outweigh the direct revenue boost from the cases where the entrant is placed second, a policy of increased prior for the entrants may indeed be revenue-optimal, as we find in our empirical application.



(a) Decomposition of Total Revenue in Revenue-Maximizing and Efficiency-Maximizing TS-SP Auctions.

(b) Proportion of Auctions with Ex-Post Efficient Allocation in Revenue-Maximizing and Efficiency-Maximizing TS-SP Auctions.

Figure 10. Revenue and allocation efficiency under revenue-maximizing and efficiency-maximizing TS-SP auctions.

Implications for market thickness. We finish this discussion by noting that the above mechanism points to a possible role to be played by market thickness. In thinner markets, there is a higher likelihood that the score for the winner and the rest of the bidders are far apart. This, in turn, makes it more likely that a boost to the prior on entrants can place them second and extract a higher revenue from the winner, without posing as pronounced a risk of damaging the efficiency of the allocation. Figure 11 confirms this expectation. We partition

the dataset into thin, mid-size, and thick markets; and we replicate the revenue-components analysis in Figure 11 for each of these partitions. As the figure shows, the revenue boost from changing the prior is indeed the highest in thinner markets, confirming our expectation.



Figure 11. Decomposition of Total Revenue in Revenue-Maximizing and Efficiency-Maximizing TS-SP Auctions Across Thin, Medium, and Thick Markets

6.2 Customized Exploration

Our analysis of market thickness in the previous section suggests that the distinction between the auction-bandit problem and the standard bandit problem may be more pronounced depending on the features of the auction (such as thickness). Thus, it would be worth studying how further customization (beyond thickness) may impact the performance of the solution on the fronts of efficiency and revenue as well as the tradeoff between them.

To this end, we build on our analysis of Pareto frontiers in the revenue-efficiency space. More specifically, for each of our three market-thickness categories, we plot the frontier emerging from uniform policies alongside the frontier arising from full customization. The former frontier can be traced by varying β in the prior and obtaining a revenue-efficiency pair for each level of β examined. The latter frontier is more nuanced. For each keyword, we form a convex combination $V(\tau, \beta_k^0)$ of revenue and efficiency:

$$V(\tau, \beta_k^0) := \tau \times \mathbb{E}[\widehat{\text{Rev}}_k(M(p, e; \alpha^0, \beta_k^0))] + (1 - \tau) \times \mathbb{E}[\widehat{\text{Eff}}_k(M(p, e; \alpha^0, \beta_k^0))]$$

where τ is the weight on revenue in the convex combination and β_k^0 is the keyword specific

prior policy that we vary to optimize $V(\tau, \beta_k^0)$.⁶

For each value of τ , we choose β_k^0 such that the convex combination $V(\tau, \beta_k^0)$ is maximized. We record the resultant $\widehat{\text{Rev}}_k$ and $\widehat{\text{Eff}}_k$. We do this by tracing values of τ from 0 to 1, which yields a Pareto frontier curve over attainable revenue and efficiency levels.

The three panels of Figure 12 present the results. There are three lessons worth describing from these graphs. First, customization can deliver Pareto improvement in the efficiency-revenue at any thickness level and upon any uniform prior. This is perhaps not surprising; but it is interesting to observe that non-trivial gains can be made even for thick markets. As the right panel of the figure shows, customization can deliver, roughly, a 5% revenue increase upon the revenue-maximizing uniform prior. Note that this improvement in revenue cannot be entirely attributed to the advantage of the customized solution in matching keyword-specific priors on the CVR (i.e., the advantage of the customized solving the bandit problem). To see why, note that the most efficient point on the customized Pareto curve delivers an about 5% *lower* revenue relative to the revenue-maximizing uniform prior. As a result, the distinction between the auction-bandit problem and the pure bandit problem is relevant even in the case of thick markets.



Figure 12. Pareto Curve for Customized and Uniform Priors Across Thickness Groups

The second lesson is that the customized Pareto curves in all three thickness groups are visibly more concave relative to the uniform Pareto curves. We interpret this to mean that customization “eases” the tension between the auction component and the bandit component of the auction-bandit problem. In other words, customization allows to increase the prior only for auctions where such increase helps with revenue, and to keep the prior selection

⁶Similar to the main analysis, we maintain all $\alpha_{a,k}^0$ at $\alpha_0 = 1$. Also note that we are varying $\beta_{a,k}^0$ only by keyword k and keeping it fixed within each keyword across ads a .

focused on efficiency in other cases. To visually see this ease of tension and its relation to concavity, observe that on the customized Pareto curve, it is easier to pick a point (i.e., a “prior policy”) that achieves a revenue approximating that curve’s maximum revenue and, at the same time, an efficiency level approximating the curve’s maximum efficiency.

Relatedly, our third lesson is that, in the case of all thickness groups, the customized Pareto curve seems to have an almost vertical part. That is, there is a non-trivial degree of revenue improvement that can be achieved at a negligible cost to efficiency. We expect that this happens because there are auctions with strong incumbents with substantially larger “true” scores than all other participants, including entrants. In such auctions, revenue can be improved with little efficiency cost. We suspect that as one goes from the bottom to the top of the almost-vertical part of each Pareto curve, the optimal policy is changing course only for these specific types of auctions. This hypothesis is consistent with the observation that the Pareto frontier for thin markets (which by definition have strong incumbents) seems to be more pronounced than that in the other two frontiers: for those auctions, revenue can be almost tripled at an efficiency cost of less than 1%.

In sum, our analysis recommends customization. It does so not only because customization (by construction) helps improve on revenue and efficiency individually, but also because it helps ease the trade-off between the two. Customization especially helps mitigate the trade-off under thinner markets.

7 Extensions and Robustness Checks

7.1 Alternative Bandit Algorithm: Upper Confidence Bound

So far, our analysis has focused on Thompson Sampling (TS) as the primary bandit algorithm for learning quality scores in the auction–bandit problem. In particular, we have shown that TS’s exploration policy can have a non-trivial impact on market outcomes. However, TS is only one approach to balancing the exploration-exploitation trade-off. In this section, we extend our analysis to another prominent bandit algorithm with strong regret guarantees: the Upper Confidence Bound (UCB) algorithm. Our goal is to compare market outcomes under a UCB-based auction with our main findings for TS-SP. We begin by formally defining the UCB exploration policy and describing its integration into an auction. We then compare the performance of auctions using UCB and TS and examine the mechanisms underlying their differences. Similar to TS, we first define the exploration policy under UCB:

Definition 3. *The **UCB Exploration Policy** with exploration parameter $\rho \geq 0$ is defined*

by the quality score estimate for each ad $a \in \mathcal{A}$ on keyword $k \in \mathcal{K}$ at time t :

$$q_{a,k,t} = \underbrace{\frac{C_{a,k,t}}{I_{a,k,t} + 1}}_{\text{empirical conversion rate}} + \underbrace{\rho \sqrt{\frac{\log(t+1)}{I_{a,k,t} + 1}}}_{\text{exploration bonus}}$$

where $I_{a,k,t}$ and $C_{a,k,t}$ are the total number of impressions and conversions the ad a has received on keyword k up to time $t - 1$, respectively, and ρ controls the degree of exploration (with $\rho = 0$ reducing to pure exploitation).

In the generic bandit formulation, the Upper Confidence Bound (UCB) algorithm follows the principle of *optimism in the face of uncertainty*, selecting the action with the highest upper bound on its estimated reward. This upper bound combines the mean estimate with an exploration bonus, encouraging the algorithm to try actions with high potential or little data. As data accumulate, the bounds narrow and UCB stops exploring an arm once its upper bound is lower than the lower bound of another arm. We now describe how we can integrate it in a second-price auction as follows:

Definition 4. Under the **UCB Second-Price (UCB-SP)** auction mechanism, quality score estimates are according to the UCB Exploration Policy and the allocation and payment rule is the same as defined under TS-SP in Definition 2.

While TS is probabilistic, UCB is deterministic and explores in face of optimism. As such, we expect TS to tend to explore for a longer period of time, which can have important consequences in terms of market outcomes. Under Thompson Sampling, the sampled quality scores asymptotically converge to the conversion rates for every ad as posterior distributions concentrate; formally $\mathbb{E}[q_{a,k,t}^{\text{TS}}] \rightarrow \mu_{a,k}$. Under UCB, the winner receives almost all impressions asymptotically, so its bonus vanishes and the quality score q^{UCB} converges to its μ ; however, non-winners are pulled only until their upper confidence bounds fall just below the winner's mean with high probability. Their bonuses therefore remain strictly positive on average over finite horizons, so their quality scores tend to be upward biased relative to μ . In a second-price mechanism, this compresses the gap between the top two scores and raises prices; hence, we expect higher revenue under UCB.

Figure 13 plots the revenue-efficiency frontier for UCB (varying ρ) against TS (varying the prior). To maintain the validity of our counterfactual estimates, we only display the UCB policies that explore the entrants no more than what we observed in the data (i.e., 18% entrant impression share; see Figure 6). Across the range we consider, UCB Pareto-dominates TS. While efficiency is slightly improved under UCB, the revenue gain is large in

both percentage and magnitude. UCB with parameter $\rho = 0.04$ achieves about 25% more revenue compared to TS with prior 0.1, while both allocate 18% of impressions to entrants. To understand the dominance, we compare TS with the mean prior 0.1 (same prior as in the

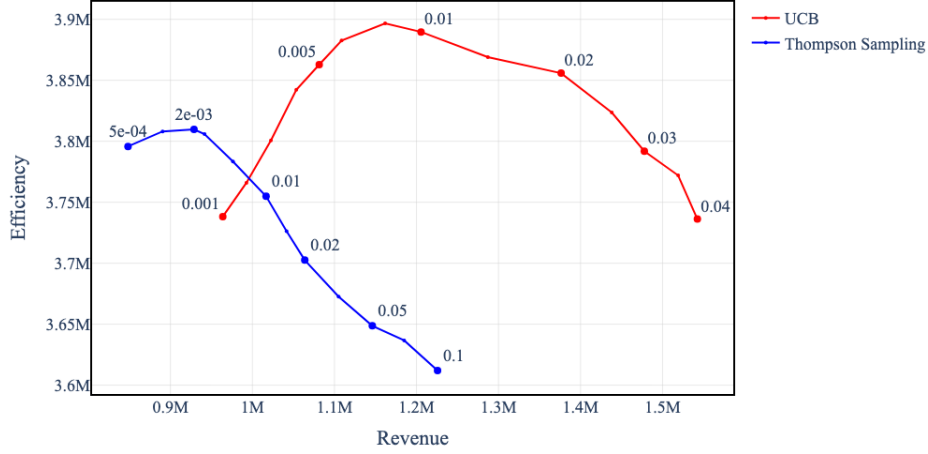


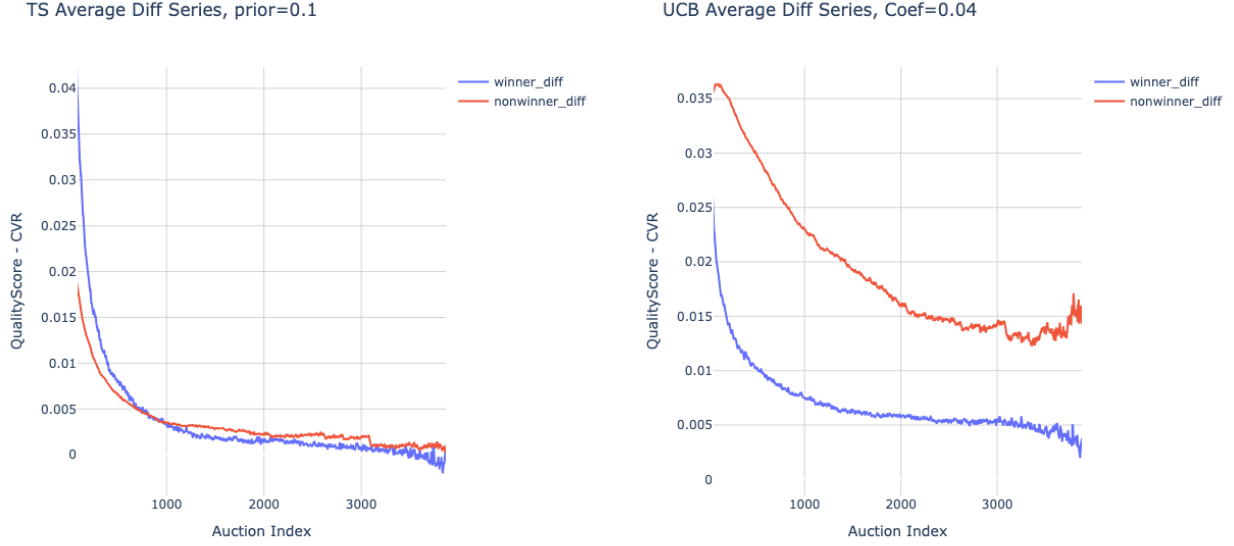
Figure 13. Revenue-efficiency tradeoff for UCB (red) and TS(blue). Labels represent prior mean under TS and ρ parameter under UCB.

data) and UCB with $\rho = 0.04$. In particular, we examine how biased the quality scores are relative to the CVR values under the two mechanisms. We define the winner bias and the nonwinner bias as follows:

Quality Score Bias: Let $\mu_{a,k}$ be the true CVR and $q_{a,k,t}^M$ the quality score under the mechanism M . With $A_{k,t}^M$ the winner on keyword k at t , define average winner and nonwinner biases

$$\text{w-bias}_t^M = \frac{1}{|\mathcal{K}|} \sum_k (q_{A_{k,t}^M, k, t}^M - \mu_{A_{k,t}^M, k}), \quad \text{nw-bias}_t^M = \frac{1}{|\mathcal{K}|} \sum_k \frac{1}{|A_{k,t}| - 1} \sum_{a \neq A_{k,t}^M} (q_{a,k,t}^M - \mu_{a,k})$$

Figure 14 plots these time series: Thompson Sampling in the left panel (prior mean 0.1) and UCB in the right panel ($\rho = 0.04$). The y -axis reports the average bias, the difference between quality scores and CVR averaged across keywords. Under TS, both biases decay toward zero, consistent with the posterior concentration. Under UCB, the winner bias also decays, but the *nonwinner bias remains positive* throughout the horizon—precisely the optimism that elevates second-highest scores and prices. This result explains why UCB achieves higher revenue than TS in our setting.



(a) Thompson Sampling with Prior Mean = 0.1

(b) UCB with $\rho = 0.04$

Figure 14. Winner and nonwinner quality score bias

7.2 Robustness Check: Budget Constraint

In the baseline analysis, we evaluated Thompson Sampling Second-Price (TS-SP) policies under truthful bidding and without explicit budget limits. This section examines the robustness of our findings when advertisers face daily spending caps.⁷ We do not observe advertisers' true budget capacities. To impose a conservative and nonparametric constraint, we set, for each ad-keyword pair (a, k) , a daily spending cap equal to the *maximum* daily expenditure observed for (a, k) over the sample window. Formally, letting $S_d(a, k)$ denote the total payment by advertiser a on keyword k and day d , the cap for (a, k) is:

$$\bar{B}_{a,k} = \max_d S_d(a, k)$$

We assume that (i) any daily budget, if present, is weakly above the largest observed daily spent for that pair within our sample, and (ii) daily budgets, if they exist, are constant across the 7 days window. We rerun the same counterfactuals as in section 6, varying the entrants CVR prior, while adding one extra rule: within each day, we track cumulative spend for every (a, k) . When ad a reaches $\bar{B}_{a,k}$ on keyword k , we remove a from subsequent auctions for keyword k for the remainder of that day and re-admit it at the start of the next day with its posterior carried forward. All other components of the algorithm are unchanged.

⁷Throughout, we maintain the per-period truthful-bidding assumption

Figure 15 displays two panels showing revenue (left) and efficiency (right) across the prior mean for: (i) the baseline unconstrained Thompson Sampling projections (blue; identical to the main results in Section 6), and (ii) the Robust Thompson Sampling results under the empirically observed daily caps (green). Qualitatively, our main conclusion is unchanged: higher priors raise revenue at the cost of efficiency.

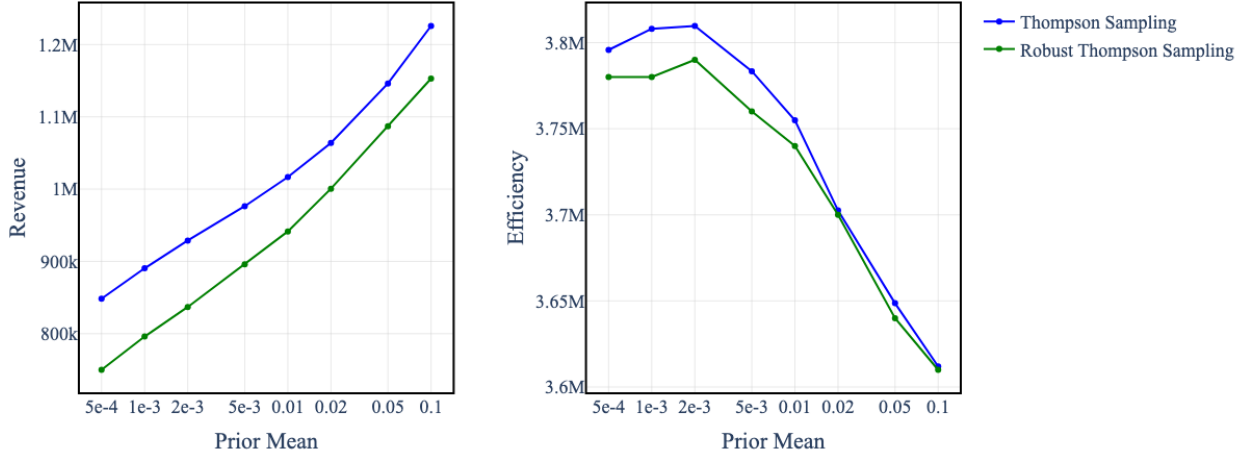


Figure 15. Revenue (left) and efficiency (right) vs. prior mean under the baseline, unconstrained Thompson Sampling policy (blue) and the Robust Thompson Sampling policy with empirically observed daily spending caps (green).

8 Managerial and Policy Implications

Our empirical analysis has key implications for managers and regulators. On the managerial front, the main lesson for platforms is that the assignment of a prior to an entrant’s CVR in ad auctions should be done with an eye on the auction aspect of the problem. In more specific terms, as we show empirically, it might be profitable for a platform to set the CVR prior at a level at least an order of magnitude larger than the CVR range that the majority of bidders are empirically observed to have. The profitability of having a CVR prior so disparate from the true CVR range is difficult to explain with learning-based (i.e., bandit theory) mechanisms only. But as shown in our analysis, auction-based reasons for high CVR priors are theoretically expected and empirically relevant. In this light, our empirical findings may be interpreted to suggest that “prior setting” in the cold-start problem is an *implicit pricing tool* at the disposal of the platforms. Such a view toward prior-setting is largely absent from the literature, which tends to focus on (explicit) reserve prices.

Our empirical findings also carry policy implications. From a regulatory perspective, a

CVR prior set substantially above the true empirical CVR range could be viewed as an “abuse of the learning process.” As discussed earlier, the reason higher priors can be optimal is that they enable platforms to charge the winner an amount close to its bid, thereby effectively approximating a first-price auction within a second-price auction environment where bidders submit their true valuations without shading. A direct version of this adverse incentive by auctioneers (i.e., running a first-price auction in stealth while announcing to bidders that the auction is second-price) has been mentioned in the literature (Akbarpour and Li 2020). Our empirical results suggest that a regulatory concern about such direct credibility issues should also extend to concerns about the role of priors in the generalized auction-bandit problem that the platform solves.

If “excessively high” levels of CVR priors are to be considered concerning, what regulatory actions constitute a solution? A formal answer to this question is beyond our scope. We speculate, however, that one possible direction is transparency. A possible regulation is to require platforms to disclose a minimal information content that would shed light on whether it is setting priors to boost revenue. The determination of what information is minimal but useful to disclose is an avenue for future research. One should also note that informational regulations may be met with resistance. Platforms often avoid sharing detailed information about auctions citing concerns that bidders might use the information to “game the system”.

We leave a formal and comprehensive treatment of such regulatory issues to future research. We do emphasize, however, that our empirical results serve to motivate such future studies and a broader policy discussion around whether/how to regulate exploration strategies by platforms. We expect the policy debate around the pricing transparency of advertising auctions to only grow in importance with the rise of Generative AI (due to an expected increase both in the number and types of advertisers and in the types of ads displayed on platforms), which amplifies the need for effective cold-start strategies. This should further increase the importance of the regulatory oversight of online ads, which includes (but is not limited to) the cold start problem.

9 Conclusion

This paper examined the interactions between the bandit problem and the auction problem. Applied to a context of online advertisements on a platform based in Asia, our framework provides the first empirical analysis of the combined auction-bandit problem. We find that the platform’s revenue-optimal level of exploration in this problem is substantially larger than the exploration level if only the bandit aspect of the problem were considered. We further investigate the mechanism behind this additional exploration. We find that the

additional exploration (i.e., assigning a higher prior to the “conversion rate” of an entrant bidder) boosts revenue by increasing the “effective market thickness,” through elevating the overall score of the entrant bidder and placing it below the winning incumbent by a small margin.

Our analysis suggests that the auction-bandit problem is empirically relevant, thereby providing further motivation for future research in the area, both theoretical and empirical. Nevertheless, our analysis has some limitations that point to fruitful directions for future research. First, we abstract away from budget pacing and autobidding mechanisms. In our empirical context, advertisers managed budgets manually, and we found no evidence of automated pacing behavior. However, pacing is now common in many platforms, and its interaction with exploration could meaningfully alter both efficiency and revenue outcomes. Understanding how auction-bandit mechanisms perform in the presence of pacing algorithms remains an important open question.

Second, our study has focused primarily on exploration policies that are no more exploratory than those observed in our data. This choice ensures the credibility of our counterfactuals but leaves unexplored the performance of more aggressive exploration strategies. In reality, if exploration were pushed to significantly higher levels, incumbents could have an incentive to bid low in early rounds, deliberately letting entrants win so that their quality scores are revised downward. Once weakened, such entrants would pose less of a competitive threat, enabling incumbents to restore higher bids in later periods. We do not find evidence of such behavior in our data, but it remains a potential concern under more exploratory policies. Extending our framework to a dynamic model that captures these forward-looking incentives would allow for assessing a wider range of exploration regimes in the auction-bandit problem.

Competing Interests Declaration

Author(s) have no competing interests to declare.

References

- Akbarpour, M. and S. Li (2020). Credible auctions: A trilemma. *Econometrica* 88(2), 425–467.
- Athey, S. and P. A. Haile (2007). Nonparametric approaches to auctions. *Handbook of econometrics* 6, 3847–3965.
- Athey, S. and D. Nekipelov (2010). A structural model of sponsored search advertising auctions. In *Sixth ad auctions workshop*, Volume 15, pp. 5.

- Babaioff, M., R. D. Kleinberg, and A. Slivkins (2015). Truthful mechanisms with implicit payment computation. *Journal of the ACM (JACM)* 62(2), 1–37.
- Babaioff, M., Y. Sharma, and A. Slivkins (2009). Characterizing truthful multi-armed bandit mechanisms. In *Proceedings of the 10th ACM conference on Electronic commerce*, pp. 79–88.
- Celis, L. E., G. Lewis, M. Mobius, and H. Nazerzadeh (2014). Buy-it-now or take-a-chance: Price discrimination through randomized auctions. *Management Science* 60(12), 2927–2948.
- Chapelle, O. and L. Li (2011). An empirical evaluation of thompson sampling. *Advances in neural information processing systems* 24.
- Choi, H. and C. F. Mela (2019). Monetizing online marketplaces. *Marketing Science* 38(6), 948–972.
- Devanur, N. R. and S. M. Kakade (2009). The price of truthfulness for pay-per-click auctions. In *Proceedings of the 10th ACM conference on Electronic commerce*, pp. 99–106.
- Edelman, B., M. Ostrovsky, and M. Schwarz (2007). Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American economic review* 97(1), 242–259.
- Feng, Z., C. Liaw, and Z. Zhou (2023). Improved online learning algorithms for ctr prediction in ad auctions. In *International Conference on Machine Learning*, pp. 9921–9937. PMLR.
- Guerre, E., I. Perrigne, and Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica* 68(3), 525–574.
- Hadad, V., D. A. Hirshberg, R. Zhan, S. Wager, and S. Athey (2021). Confidence intervals for policy evaluation in adaptive experiments. *Proceedings of the national academy of sciences* 118(15), e2014602118.
- Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association* 47(260), 663–685.
- Jeziorski, P. and I. Segal (2015). What makes them click: Empirical analysis of consumer demand for search advertising. *American Economic Journal: Microeconomics* 7(3), 24–53.
- Kim, D. and P. Pal (2025). Nonparametric estimation of sponsored search auctions and impact of ad quality on search revenue. *Management Science*.
- Lahaie, S. and D. M. Pennock (2007). Revenue analysis of a family of ranking rules for keyword auctions. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, pp. 50–56.
- Lahaie, S., D. M. Pennock, A. Saberi, and R. V. Vohra (2007). Sponsored search auctions. *Algorithmic game theory* 1, 699–716.
- Lattimore, T. and C. Szepesvári (2020). *Bandit algorithms*. Cambridge University Press.
- Levin, J. and P. Milgrom (2010). Online advertising: Heterogeneity and conflation in market design. *American Economic Review* 100(2), 603–607.
- Myerson, R. B. (1981). Optimal auction design. *Mathematics of operations research* 6(1), 58–73.

- Ostrovsky, M. and M. Schwarz (2023). Reserve prices in internet advertising auctions: A field experiment. *Journal of Political Economy* 131(12), 3352–3376.
- Rafieian, O. (2020). Revenue-optimal dynamic auctions for adaptive ad sequencing. Technical report, Working paper: Cornell University.
- Rafieian, O. and H. Yoganarasimhan (2021). Targeting and privacy in mobile advertising. *Marketing Science* 40(2), 193–218.
- Stourm, V. and E. Bax (2017). Incorporating hidden costs of annoying ads in display auctions. *International Journal of Research in Marketing* 34(3), 622–640.
- Varian, H. R. (2007). Position auctions. *international Journal of industrial Organization* 25(6), 1163–1178.
- Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance* 16(1), 8–37.
- Wilbur, K. C., L. Xu, and D. Kempe (2013). Correcting audience externalities in television advertising. *Marketing Science* 32(6), 892–912.
- Yao, S. and C. F. Mela (2011). A dynamic model of sponsored search advertising. *Marketing Science* 30(3), 447–468.
- Ye, Z., D. J. Zhang, H. Zhang, R. Zhang, X. Chen, and Z. Xu (2022). Cold start to improve market thickness on online advertising platforms: Data-driven algorithms and field experiments. *Management Science*.
- Zhan, R., V. Hadad, D. A. Hirshberg, and S. Athey (2021). Off-policy evaluation via adaptive weighting with data from contextual bandits. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp. 2125–2135.

Web Appendix

A Examining Entrants' Overbidding Incentives

In this section, we discuss *entrants' overbidding incentives*, which can be justified under long-term utility maximization by bidders. High-quality entrants may bid above their true valuation in the early stages, especially if their prior quality score underestimates their actual conversion rate, in order to accelerate the revelation of their true quality. This strategy helps them win impressions, which leads to the auctioneer learning their true quality scores. As a result, they can lower their bid in later periods and extract higher total utility compared to a scenario in which they only maximize per-period utility.

We empirically examine the overbidding incentives of entrants, which suggest that an entrant with a quality score higher than the prior mean may have an incentive to overbid in order to allow the auctioneer to learn its true quality score. To test this empirically, we first identify the entrants in our data whose quality scores exceed the prior mean. We then estimate the following regressions:

$$\bar{b}_{a,k,h} = \alpha + \delta \cdot \text{Hour}_h + \tau_{a,k} + \epsilon'_{a,k,h} \quad (10)$$

$$\bar{b}_{a,k,h} = \alpha + \gamma \cdot \bar{q}_{a,k,h} + \tau_{a,k} + \epsilon''_{a,k,h} \quad (11)$$

where $\text{Hour}_{k,h}$ is the numerical variable for the hour that measures the time trend, and $\bar{q}_{a,k,h}$ is the average quality score draw for ad a in keyword k at hour h . If high-quality entrants' overbidding is indeed a deliberate strategy, we would expect these entrants to place higher bids in the early periods and subsequently lower their bids over time (indicated by a negative δ) or when assigned lower prior quality scores (indicated by a negative γ). We present the estimates for these regressions in Columns (1) and (2) of Table 3. As shown in these columns, the coefficients are not statistically significant, which provides support for the per-period utility maximization assumption over the long-term utility maximization hypothesis.

Table 3. Regression Specifications to Test Static vs. Dynamic Bidding Based on Entrants' Overbidding Incentives

	<i>Dependent variable: Bid ($\bar{b}_{a,k,h}$)</i>	
	(1) OLS	(2) OLS
Hour	-2.2934 (2.340)	
Mean Quality Score		-670.7890 (1921.237)
Keyword-Ad FE	✓	✓
R^2	0.905	0.905
No. of Obs.	12,641	12,641
<i>Note:</i>	*p<0.05; **p<0.01; ***p<0.001	

B Inverse Propensity Score Estimator for Conversion Rates

In §5.2.2, we presented a sample analogue estimator for conversion rates and discussed that the estimator would be unbiased if impressions allocated to an ad are randomly drawn from the overall impression distribution. In this section, we relax this assumption and employ an Inverse Propensity Score (IPS) estimator:

$$\hat{\mu}_{a,k}^{\text{IPS}} = \frac{1}{T_k} \sum_{t=1}^{T_k} \frac{\mathbb{1}(a = A_{k,t})Y_{k,t}}{\hat{\pi}(a \mid B_{k,t}, I_{k,t}, C_{k,t}; \alpha^0, \beta^0)}, \quad (12)$$

where $\hat{\pi}(a \mid B_{k,t}, I_{k,t}, C_{k,t}; \alpha^0, \beta^0)$ represents the estimated propensity score for ad a to be shown in impression t of keyword k . We confirm that both approaches yield highly similar estimates, reinforcing the validity of our assumption that impressions allocated to each ad approximate a random selection of impressions.