Adaptive Ad Sequencing

Omid Rafieian
University of Washington

This file contains a series of two job market papers on the topic of adaptive ad sequencing.

1. **Optimizing User Engagement through Adaptive Ad Sequencing** (Job Market Paper I, starting on page 2 of this file)
2. **Revenue-Optimal Dynamic Auctions for Adaptive Ad Sequencing** (Job Market Paper II, starting on page 49 of this file)

Optimizing User Engagement through Adaptive Ad Sequencing

Omid Rafieian*
University of Washington

For Most Updated Version: Click Here

June 30, 2019

*I am grateful to my advisor Hema Yoganarasimhan, and my committee members Arvind Krishnamurthy, Amin Sayedi, Simha Mummalaneni, and Jacques Lawarrée for their guidance and comments. I am also grateful to an anonymous firm for providing the data and to the UW-Foster High Performance Computing Lab for providing me with computing resources. Please address all correspondence to: rafieian@uw.edu.

Abstract

Mobile in-app advertising has grown exponentially in the last few years. In-app ads are often shown in a sequence of short-lived exposures for the duration of a user's stay in an app. The current state of both research and practice ignores the dynamics of ad sequencing and instead adopts a myopic framework to serve ads. In this paper, we propose a unified dynamic framework for adaptive ad sequencing that optimizes user engagement in the session, e.g., the number of clicks or length of stay. Our framework comprises of two components – (1) a Markov Decision Process that captures the domain structure and incorporates inter-temporal trade-offs in ad interventions, and (2) an empirical framework that combines machine learning methods such as Extreme Gradient Boosting (XGBoost) with ideas from the causal inference literature to obtain counterfactual estimates of user behavior. We apply our framework to large-scale data from the leading in-app ad-network of an Asian country. We document significant gains in user engagement from adopting a dynamic framework. We show that our forward-looking ad sequencing policy outperforms all the existing methods by comparing it to a series of benchmark policies often used in research and practice. Further, we demonstrate that these gains are heterogeneous across sessions: adaptive forward-looking ad sequencing is most effective when users are new to the platform. Finally, we use a descriptive approach to explain the gains from adopting the dynamic framework.

Keywords: advertising, personalization, adaptive interventions, dynamic policy, Markov Decision Process, machine learning, reinforcement learning

1 Introduction

1.1 Adaptive Interventions in Mobile Advertising

Consumers now spend a significant portion of their time on mobile devices. The average time spent on mobile devices by US adults has grown steadily over the last few years (eMarketer, 2019a). This demand expansion, in turn, has amplified marketing activities towards mobile users. In 2018, mobile advertising generated over $71 billion in the US, accounting for roughly double the share of its digital counterpart, desktop advertising (eMarketer, 2019b). Most of this growth in mobile advertising is attributed to in-app ads, i.e., ads shown inside mobile apps. Indeed, in-app advertising is now the dominant channel for mobile advertising, generating over 80% of ad spend in the mobile advertising category (eMarketer, 2018).

Two key features of mobile in-app ads have contributed to their growth. First, the mobile app ecosystem has excellent user tracking ability, thereby allowing "personalization" of ad interventions, i.e., targeting of users based on their prior behavioral history (Han et al., 2012). Second, in-app ads are usually short-lived and dynamic in nature: each ad intervention is shown for a fixed amount of time (e.g., 30 seconds or one minute) inside the app, and is then followed by another ad intervention. As such, a user can see multiple ad exposures within a session.[1] This is in contrast with the common practice in desktop advertising, where ads remain fixed throughout a session. Short-lived ads together with personalization potential make in-app advertising amenable to "adaptive interventions", i.e., targeting of ads based on time-varying behavioral information about users to maximize user engagement.

We illustrate the general schema for a mobile publisher's ad sequencing problem in Figure 1.[2] In this problem, the publisher has to decide which ad to serve in each period that the user stays in the app. We characterize three separate pieces of information that the publisher can incorporate when deciding which ad to show in any impression: (1) the pre-session information, which consists of user characteristics as well as the behavioral history of the user up until the current session, (2) the session-level information, which is the sequence of ads the user has seen so far within the session, as well as his response to each, and (3) the future information, which captures the publisher's expectations on how the sequence is likely to evolve in the next periods. The current research and practice mostly focus on the first two pieces and overlook the future information (Li et al., 2010; Lu et al., 2010; Rafieian and Yoganarasimhan, 2018b). One reason is that capturing future information often adds significantly to the complexity of the problem. In addition, the returns

[1] A session is an uninterrupted time that a user spends inside an app.

[2] In this paper, we use the publisher, ad-network, and platform interchangeably, when we refer to the agent who makes the ad placement decision.

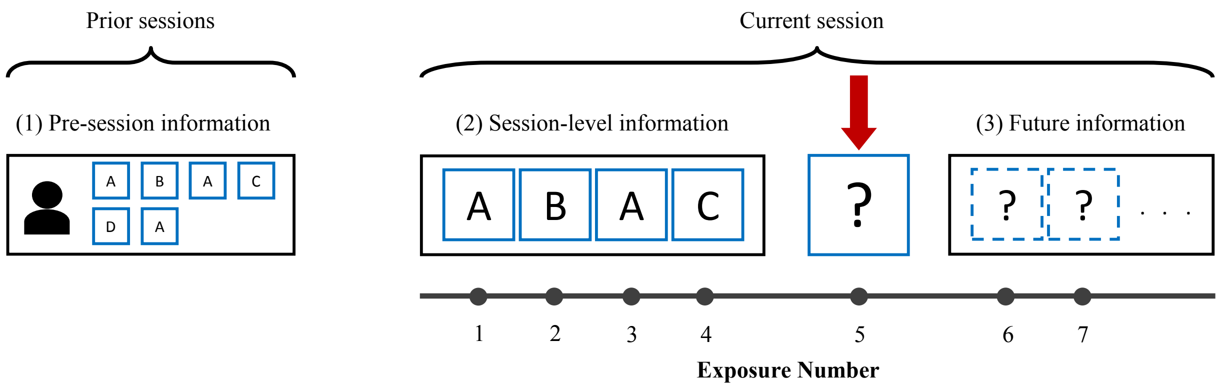


Figure 1: A visual schema the publisher's ad sequencing decision and different types of information available.

from adopting a forward-looking model are not clear. Thus, the publisher's decision on whether to use a dynamic framework boils down to whether incorporating future information helps her achieve a better outcome.

In principle, if there is no interaction between sequential ad interventions, incorporating future information will not improve the ad sequencing policy. However, at the advertiser level, the extant literature has documented various empirical effects that highlight such inter-temporal trade-offs. Prior findings on spillover and carryover effects of advertising rule out the independence of ads shown in a sequence (Rutz and Bucklin, 2011; Sahni, 2015a). In particular, in the context of mobile in-app advertising, Rafieian and Yoganarasimhan (2018a) find that a higher variety of previous ads shown within a sequence leads to a higher likelihood of click on the next ad. At the same time, they show that an ad will have a higher chance of being clicked if it has been shown more within the sequence of prior ads. These findings suggest a natural inter-temporal trade-off: higher variety by definition implies fewer repeat exposures of each ad. While these effects are well-established at the advertiser level, neither research nor practice has looked into how to collectively incorporate these findings to dynamically sequence ads to optimize publishers' outcomes. Thus, to the extent that these effects have significant impact on user's engagement with ads, it is important to develop a dynamic framework for ad sequencing and examine the gains from adopting such a framework.

1.2 Research Agenda and Challenges

In this paper, we propose a dynamic framework that incorporates the inter-temporal trade-offs in ad sequencing and develops an optimal policy that optimizes user engagement with ads in a session. User engagement with ads or the match between users and ads is particularly an important outcome for publishers, as it is the main channel through which the publisher can create value by ad sequencing. We can use different metrics to measure user engagement with ads depending on the context of advertising. In the context of mobile in-app advertising, click is a reasonably good metric for the user engagement with ads because of two reasons. First, all ads are mobile apps whose objective is to get more clicks and installs (e.g., performance ads). Second, clicks are directly linked to revenues as most in-app ad publishers employ cost-per-click or cost-per-install as their monetization strategy.

Therefore, we use a click-maximizing objective and seek to answer the following three questions in this paper:

1. How can we develop a dynamic theoretical framework that incorporates the inter-temporal trade-offs in ad sequencing and designs a policy that maximizes the expected number of clicks per session?

2. How can we empirically evaluate the performance of this dynamic ad sequencing policy relative to other benchmark policies?

3. What are the gains from using a dynamic framework to allocate ads? What explains the differences in outcomes and interventions under a fully dynamic sequencing policy and other benchmark policies?

We need to overcome three major challenges to satisfactorily answer this set of questions. First, to capture the dynamics of this problem, we need a theoretical framework that incorporates inter-temporal trade-offs in the publisher's ad allocation decision. Second, to develop a click-maximizing dynamic sequencing policy, we need to obtain accurate personalized counterfactual estimates of user behavior that allows us to evaluate all feasible policies (and not only those implemented in the data). As such, we need a setting with enough randomization in the ad allocation process that enables us to exploit this randomization and obtain accurate counterfactual estimates of user behavior. Finally, to measure the gains from the dynamic framework and explore the differences in sequencing strategies across sessions, we need to have a solution concept that determines the optimal dynamic policy and an evaluation approach that accurately estimates the outcomes under the new optimal policy for each session.

1.3 Our Approach

In this paper, we present a unified three-pronged framework that addresses these challenges and develops a forward-looking adaptive sequencing policy to maximize user engagement with ads. We present an overview of our approach in Figure 2. As shown in the top row of this figure, we start with a theoretical framework that models the domain structure of our problem and allows us to

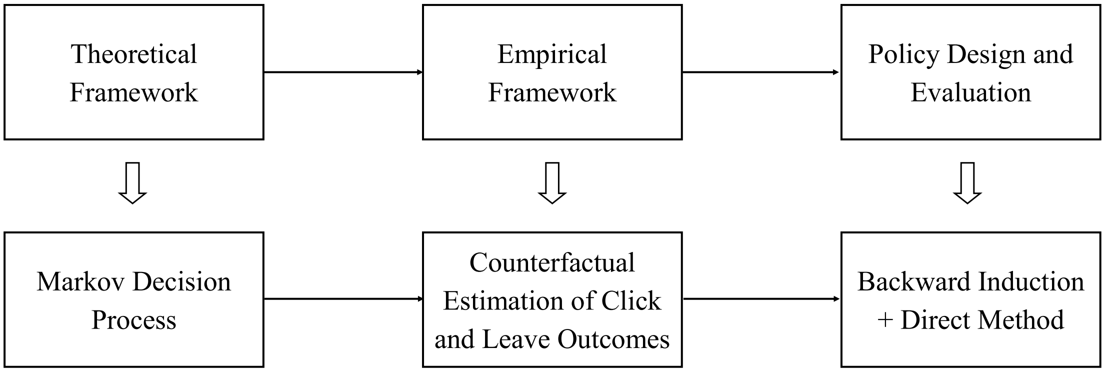


Figure 2: An overview of our approach. The top row presents our general framework and the bottom row shows the specific approach we take in this paper.

identify key empirical tasks required for the policy design and evaluation. The bottom row follows the same flow and illustrates the specifics of our approach to the ad sequencing problem in mobile in-app advertising:

Theoretical Framework: We start with a Markov Decision Process (MDP henceforth) that characterizes the structure of adaptive ad interventions. Starting from a theoretical framework allows us to address all the challenges discussed above. First, to capture the dynamics of the problem, we specify a domain-specific MDP with a rich set of state variables that incorporates the inter-temporal trade-offs identified in the literature. Our MDP characterizes the reward at any time period as well as how the state evolves in future periods, given any action taken by the publisher. Since our goal is to optimize the number of clicks per session, we define the reward as the expected probability of click on any ad selected. This probability is also informative of the available history in the next period, and thereby captures a probabilistic component of state transitions in our MDP. Another probabilistic factor that affects the future state is the expected probability of the user leaving the session after an intervention. It determines with what probability the user will be available to receive the next intervention. Together, these two probabilistic outcomes help us define the empirical tasks required in our problem: personalized counterfactual estimation of click and leave outcomes.

Empirical Framework: To obtain accurate counterfactual estimates for the click and leave outcomes, we use the filtering strategy proposed by Rafieian and Yoganarasimhan (2018b) that filters out the ads that *could have never been shown* in a session. This is because we cannot estimate the outcome reliably for these ads as they are not within the joint distribution of the training set that is used for model fitting. However, if an ad *could have been shown* within a session (i.e., has a non-zero propensity score), our outcome estimates will be accurate for this ad in this session as long as we control for all other covariates affecting the propensity of seeing this particular ad.

Further, to personalize these counterfactual estimates, we employ machine learning methods that can capture more complex relationships between the covariates and the outcome. In particular, we use the Extreme Gradient Boosting (XGBoost henceforth) method developed by Chen and Guestrin (2016), which is a fast and scalable version of Boosted Regression Trees (Friedman, 2001).

Policy Design and Evaluation: In order to develop the optimal dynamic policy, we can use our counterfactual estimates for click and leave probabilities and numerically solve for the value functions, given the state variables. In our case, we employ a non-stationary finite horizon MDP and determine the optimal dynamic policy using backward induction. Finally, we use a direct method evaluation method to evaluate the performance of any policy at the session level. This method uses the outcome estimates for both click and leave to simulate the stochasticity within the session and evaluate the outcome/outcomes of interest.

1.4 Findings and Contribution

We apply our framework on the data from a leading mobile in-app ad-network of a large Asian country. Our setting has some notable features that make it amenable to our research goals. First, the ad-network uses a short-lived ad format where ad interventions last for a short period of time and change within the session. Second, the extent of randomization in ad allocation is quite high because of two reasons: (1) the ad-network runs a quasi-proportional auction that employs a probabilistic allocation rule, and (2) the ad-network only allows limited targeting on broad categories. These two features help us estimate click and leave outcomes for a wide range of counterfactual ads. This task would not have been possible if the ad-network had run a deterministic mechanism such as the second-price auction or allowed micro-level targeting.

We first discuss the results from our first-stage machine learning models for both click and leave outcomes. We evaluate these predictive models on a hold-out test set, using various goodness-of-fit measures. We show that both the click and leave models achieve high out-of-sample predictive accuracy. We then show that dropping variables that are defined either at the ad- or session-level from the predictive model leads to a significant drop in the performance of both click and leave models. These results provide preliminary evidence for the gains from the dynamic framework since the publisher can actively influence these variables by adopting an adaptive forward-looking policy.

Next, we focus on the main goal in this paper – evaluating the gains from *adaptive forward-looking sequencing policy*. This policy incorporates all three pieces of information in Figure 1. To establish the performance of our adaptive forward-looking sequencing policy and identify where these gains come from, we define a set of benchmark policies that use a combination of different pieces of information:

- *Random sequencing policy*: This policy allocates exposures to ads randomly. It serves as a baseline for capturing how well we can do without any model.
- *Non-adaptive single-ad policy*: This policy uses the pre-session information to allocate ads within the session. Since this information does not change within the session, the optimal allocation based on only this information is to select a single ad that maximizes publisher's rewards. As such, this policy simulates the case where the ad slot is fixed throughout the session and helps us identify the opportunity costs of using a fixed ad slot.
- *Adaptive myopic sequencing policy*: This policy uses both pre-session and session-level information and allocates each impression to the ad with the highest probability of click in that impression, regardless of how it affects the expected future rewards.

We find that all model-based policies lead to substantial gains compared to the random sequencing policy, in terms of the expected number of clicks per session. In particular, the adaptive forward-looking sequencing policy increases the expected number of clicks by 80.36% relative to the random sequencing policy. We then show that the expected number of clicks is 7.87% higher under the adaptive forward-looking sequencing policy as compared to the non-adaptive single-ad policy. This finding demonstrates the opportunity cost of using a fixed ad slot throughout the session, which supports the current industry trend of using short-lived ad slots. Finally, we show that the adaptive forward-looking sequencing policy results in 1.50% increase in the expected number of clicks per session, compared to the adaptive myopic sequencing policy. This suggests that choosing the best match at any point will not necessarily create the best match outcome at the end of the session. Rather, the right action sometimes is to show the ad that is not necessarily the best match at the moment but transitions the session to a better state in the future. Together, these findings establish the benefits of adopting an adaptive and forward-looking approach to allocate ads. This has important implications for publishers and ad-networks, especially since the current practice in the industry overlooks the dynamics of ad sequencing.

Next, we explore the heterogeneity in the gains from the adaptive forward-looking sequencing policy compared to both non-adaptive single-ad and adaptive myopic sequencing policies, across the sessions. We examine how the relative gains from the adaptive forward-looking sequencing changes, as the number of prior sessions a user has participated in increases. On the one hand, we expect more data on the user to benefit the adaptive forward-looking sequencing policy more than other policies, as the publisher can take prior usage patterns of the user into account to improve the adaptive forward-looking sequencing policy. On the other hand, sequencing effects are shown to become smaller as the user becomes more experienced (Rafieian and Yoganarasimhan, 2018a). We find evidence for the latter: the relative gains from adopting adaptive forward-looking sequencing policy diminishes as the user participates in more sessions. We further provide some descriptive evidence that the source for these diminishing returns seems to be the number of distinct ads the user has seen: if a user has seen many distinct ads over time, it is harder to affect their decision through adaptive forward-looking sequencing of ads.

Finally, we present some descriptive analysis to better understand how and why the adaptive forward-looking and adaptive myopic sequencing policies differ. We find that the adaptive forward-looking sequencing policy tends to repeat the same ad in consecutive exposures more than the adaptive myopic sequencing policy. This is likely because repeating an ad consecutively is not the optimal decision if the publisher only takes the reward at that moment into account. However, it is the right decision if the publisher is forward-looking, as it increases the number of prior exposures for an ad and strengthens the session-level carryover effects.

In sum, our paper makes three contributions to the literature. First, from a methodological standpoint, we propose a unified dynamic framework that theoretically characterizes the domain structure of the mobile in-app advertising environment, and an empirical approach that allows us to break the problem into composite machine learning tasks. To our knowledge, this is the first paper to collectively incorporate temporal effects of advertising documented in the literature at the advertiser level, and propose a dynamic framework for the sequential allocation that characterizes optimal policy design for publishers. The generality of our framework makes it applicable to the contexts where advertisers have to make dynamic decisions such as the attribution problem. Second, from a substantive point-of-view, we establish the gains from an adaptive forward-looking sequencing policy as compared to other benchmarks that are often used in research and practice. This finding is of importance, as the current practice in this industry ignores the dynamics of ad allocation problem. Third, from a managerial perspective, we quantify the opportunity costs of using fixed ad slots. Our framework can help marketing practitioners and ad-networks to design the ad slot that is optimal in their context.

2 Related Literature

Our paper relates and contributes to several streams of literature.

First, our paper relates to the marketing literature on personalization and targeting in digital platforms. Early papers in this stream build Bayesian frameworks that exploit behavioral data and personalize marketing mix variables (Ansari and Mela, 2003; Manchanda et al., 2006; Arora and Henderson, 2007). Other recent papers in this area use a combination of feature generation and supervised machine learning frameworks to provide more scalable solutions for personalization and targeting policies for large-scale data (Yoganarasimhan, 2018; Rafieian and Yoganarasimhan, 2018b). While all these papers focus on prescriptive or substantive frameworks to study personalization, they all study this phenomenon from a static point-of-view. Our paper extends this literature by offering a scalable and dynamic framework to develop personalized targeting policies.

Second, our paper relates to the literature on the temporal effects of advertising, such as spillover effects in search advertising (Rutz and Bucklin, 2011; Sahni, 2016), carryover effects in display advertising (Johnson et al., 2016a), temporal interactions between multiple advertising channels (Li and Kannan, 2014), effects of temporal spacing in search advertising Sahni (2015b), and the effects of variety of previous ads in mobile in-app advertising context (Rafieian and Yoganarasimhan, 2018a). While these papers establish the presence of these temporal effects from advertisers' perspective, they do not address how a publisher can use this information to optimally show ads in sequences. In this paper, we view this problem from a publisher's perspective who wants to maximize users' engagement with ads and develop a dynamic framework that incorporates all possible temporal effects that have been documented in the literature.

Third, our paper relates to the literature on dynamic policy design in digital advertising. Given the complexity of solving a dynamic policy, prior works often simplify the problem to avoid the curse of dimensionality. Urban et al. (2013) focus on 16 cognitive-style segments and combine dynamic programming with a Bayesian framework to infer segment membership. In the context of linear video ads, Kar et al. (2015) incorporate ad-specific leave probability as the only source of inter-temporal trade-off and use a cascade model to obtain a dynamic policy. Using the context of mobile in-app advertising, Sun et al. (2017) theoretically examine the problem of short-lived ads in mobile in-app advertising and theoretically derive the dynamic policy by only focusing on certain aspects of the dynamics in this problem. Closely related to our problem, Theocharous et al. (2015) present a framework to develop and evaluate dynamic policies and study the gains at the advertiser level. Our paper differs from these papers since we are the first to present a dynamic framework that incorporates all established temporal effects of advertising and examines the gains from adopting an adaptive forward-looking sequencing policy, from the publisher's perspective.

Finally, our paper relates to the growing literature on machine learning applications in marketing. The vast majority of papers in this stream focus on prediction tasks and use various supervised or unsupervised learning algorithms to achieve a better predictive accuracy (Toubia et al., 2007; Hauser et al., 2010; Dzyabura and Yoganarasimhan, 2018). A narrower body of works in this area brings machine learning methods to policy design questions. Using a static approach, some recent papers develop machine learning methods to design optimal policies in various contexts such as pricing (Dubé and Misra, 2017), ad placement (Rafieian and Yoganarasimhan, 2018b), and CRM campaigns (Hitsch and Misra, 2018). Incorporating the dynamics of exploration-exploitation trade-off, Schwartz et al. (2017) offer a multi-armed bandit approach in a display advertising context.

Our paper adds to this literature by fully incorporating the dynamics of the ad sequencing problem through an MDP and linking it to smaller machine learning tasks.

3 Setting and Data

3.1 Setting

Our data come from a leading mobile in-app advertising network of a large Asian country that had over 85% of the market share around the time of this study. Figure 3 summarizes most key aspects of the setting. We number the arrows in Figure 3 and explain what each step of the ad allocation process in details below:

1. The ad-network designs an auction to sell ad slots. In our setting, the ad-network runs a quasi-proportional auction with a cost-per-click payment scheme. As such, for a given ad slot and a set of participating ads $\mathcal{A}$ with a bidding profile $(b_1, b_2, \ldots, b_{|\mathcal{A}|})$, the ad slot is allocated to ad a with the following probability:

$$q_a^p(b; z) = \frac{b_a z_a}{\sum_{j \in \mathcal{A}} b_j z_j}, \tag{1}$$

   where z_a is ad a's quality score, which is a measure reflecting the profitability of ad a. The ad-network does not customize quality scores across auctions.[3] The payment scheme is cost-per-click and is similar to Google's sponsored search auctions. That is, ads are first ranked based on their product of bid and quality score, and the winning ad pays the minimum amount that guarantees their rank if a click happens on their ad.

2. Advertisers participating in the auction choose: (a) design their banner, (b) specify the areas in which they want to show their ad, and (c) submit their bid. Figure 3 shows an example of auction with four different ads.

3. Whenever a user starts a new session in an app (in Figure 3, we use a messaging app as an example), a new impression is being recognized, and a request is sent to the publisher to run an auction.

4. The auction takes all the participating ads into account and selects the ad probabilistically based on the weights shown in Equation (1). It is worth noting that all the participating ads have the chance to win the ad slot. This is in contrast with more widely used deterministic

[3] In our data collection period, each ad just had one quality score.

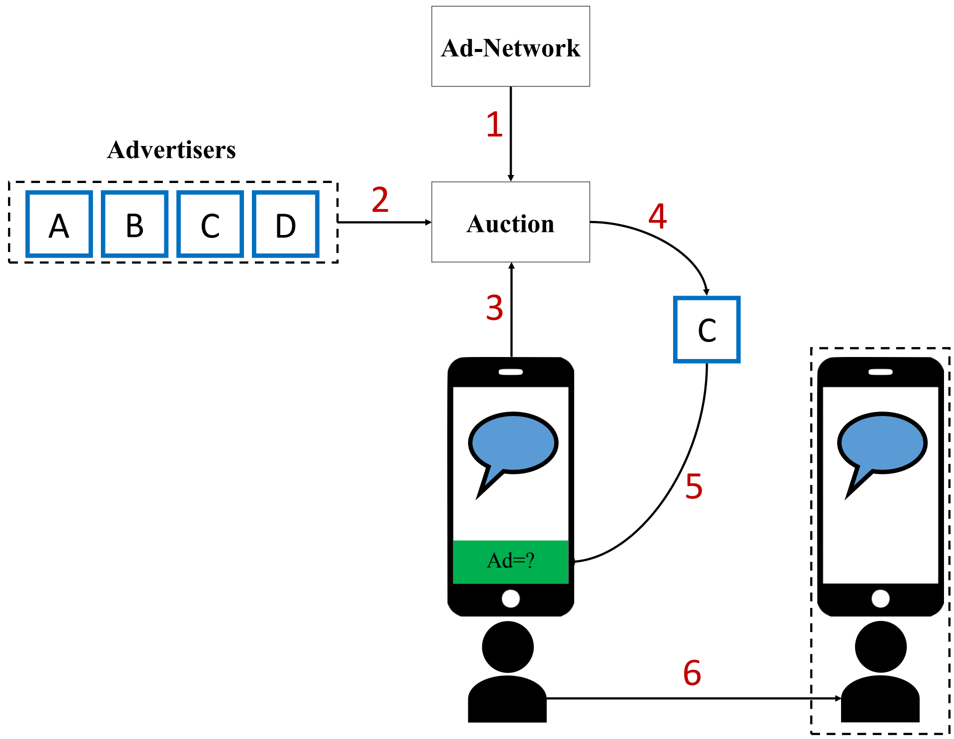


Figure 3: A visual schema of our setting

mechanisms like second-price auctions, where the ad with the highest product of bid and quality score always wins the ad slot.

5. The selected ad is placed at the bottom of the app, as shown in Figure 3.

6. Each ad exposure lasts one minute. During this time, the user makes two key decisions: (a) whether to click on the ad, and (b) whether to stay in the app or leave the app and end the session. If the user clicks on the ad, the corresponding advertiser has to pay the amount determined by auction. After one minute, if the user continues using the app, the ad-network treats the continued exposure as a new impression and repeat steps 3 to 6 are repeated until the user leaves the app. We assume that a user has left the app when the time gap until the next exposure exceeds 5 minutes. Consistent with this definition, we define a session as the time interval between the time a user comes to an app and the time she leaves the app.[4]

[4]There are obviously various ways to define a session based on the time gap between two consecutive exposures. We show that our results are robust to different definitions.

3.2 Data

We have data on all impressions and clicks for the one month period from 30 September 2015, to 30 October 2015. Overall, we observe 1,594,831,699 impressions along with 14,373,293 clicks in the data, implying a 0.90% CTR. We now describe our raw variables and sampling procedure.

3.2.1 Raw Variables

Each impression contains the following raw information:

- *Time and date:* The exact time-stamp of the impression.
- *App information:* The identifier for a mobile app that shows ads through the ad-network.
- *User information:* An identifier that is unique to each mobile device and serves as our user ID.
- *GPS information:* The exact latitude and longitude of the user at the time of the impression.
- *Targeting variables:* The set of variables that advertisers can target on. There are five main categories that advertisers can target: province, hour of the day, smartphone brand, connectivity type, and Mobile Service Provider (MSP). If an advertiser decides to exclude a certain sub-category within these variables (e.g., Samsung smartphones), his ad will not be shown in impressions in that sub-category.
- *Ad information:* The set of variables related to the ad shown in the impression. It consists of an ad identifier and the potential cost-per-click.[5][6]
- *Click outcome:* A binary variable indicating whether the user clicked on the ad. This is our primary outcome of interest to measure the match between users and ads. While click is generally an imperfect outcome for the effects of ads, it is reasonable to use it as our primary outcome for two reasons. First, all ads are mobile apps whose objective is to get more clicks or installs (e.g., performance ads). Second, the ad-network uses a cost-per-click auction, which directly links the click outcome to the ad-network's revenues.

3.2.2 Sampling Procedure

Our sampling procedure consists of two essential steps – (1) user sampling, and (2) app sampling. We describe each step in greater details below:

- *User sampling:* Since we want to optimally sequence ads within the session, our optimal intervention depends on users' past history. As such, we only focus on users for whom we can exploit their entire history. The challenge is that there is no variable in our data identifying new users. As illustrated in Figure 4, our approach is to split our data into two parts based on a date (October 22), and keep users who are active in the second part of the data (October 22 to October

[5]The potential cost-per-clik is the amount that the ad would have paid if the user had clicked on their ad.

[6]We do not have the data on the banner creatives and its format, i.e., whether it is a jpeg file or an animated gif.

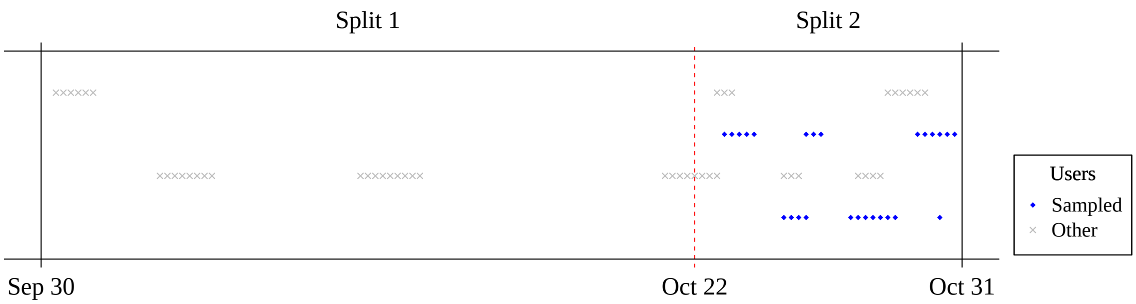


Figure 4: Schema for identification of new users.

30), but not in the first part (September 30 to October 22). This sampling scheme guarantees that the users who are identified as new users have not had any activity in the platform at least for the last three weeks. We drop all the other users from our data.

- *App sampling:* We only focus on the most popular mobile app in the platform, which is a messaging app that has over 30% share of total impressions. As such, we drop new users who do not use this app. There are a few reasons why we focus on this app. First, this is the only app whose identity is known to us. Second, we expect the sequencing effects to be context-dependent, and focusing on one app helps us perform a cleaner analysis. Finally, it takes users a relatively long time to learn how to use certain apps (e.g., games), and learning effects can interact with sequencing effects. However, this messaging app is widely popular in the country and easy to use, so we expect users to pay more attention to ads from the beginning.

Overall, our sampling procedure gives us a total of 8,323,778 impressions shown to a set of 94,884 unique new users. Over 40% of these users use other apps in addition to the messaging app. In our data, there are 6,955,995 impressions shown inside the focal messaging app that corresponds to 1,271,068 unique sessions. For our analysis, we only focus on the impressions shown in the messaging app. However, we use impressions shown in other apps for feature generation.[7]

3.3 Summary Statistics

We now present some summary statistics on the data. As mentioned earlier, these statistics correspond to the impressions shown in the messaging app.

[7]Our sampling procedure is almost identical to that of Rafieian and Yoganarasimhan (2018a). However, the number of impressions and sessions is slightly different, because we need to drop users with missing information on the latitude and longitude. Rafieian and Yoganarasimhan (2018a) use those impressions because latitude and longitude do not play a role in their analysis.

3.3.1 Shares of Categorical Variables

Since most raw variables presented in §3.2.1 are categorical, we cannot show the mean and standard deviation for them. Instead, we present the number of categories as well as the shares of the top three sub-categories within each variable in Table 1.

Variable	Number of categories	Share of top categories 1^{st}	2^{nd}	3^{rd}
Province	31	24.58%	9.54%	7.50%
Hour of the Days	24	8.48%	8.03%	7.25%
Smartphone Brand	7	44.72%	38.07%	10.11%
Connectivity Type	2	50.54%	49.46%	
MSP	3	50.15%	44.07%	5.77%
Ad	327	18.43%	8.03%	7.50%

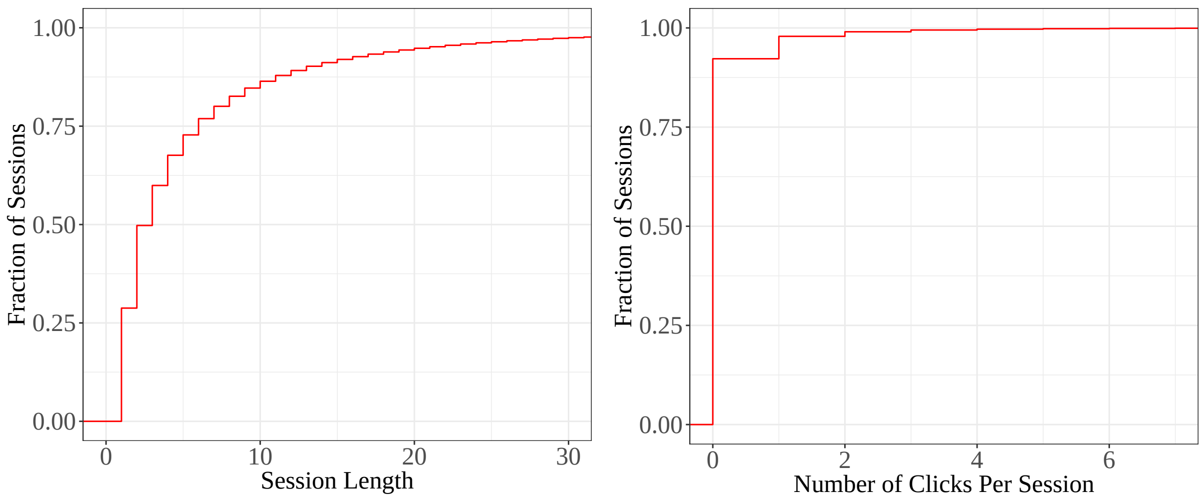
Table 1: Summary statistics of the categorical variables. This includes the number of categories and the percentage shares for the top sub-categories within each variable.

As shown in Table 1, one province accounts for a quarter of all impressions. The second row shows that there are certain hours of the day with more user activity. We find that these hours are late at the night when users are not at work. Table 1 shows that two major smartphone brands constitute over 80% of all impressions. Finally, we find that ads have different shares: the top three ads account for roughly 35% of all impressions. On the other hand, we observe that most ads have a very small share. This is mostly because these ads ran short campaigns. Later in §4.2.2, we show the full distribution of ad shares and provide a more detailed discussion.

3.3.2 Distribution of Session-Level Outcomes

Our goal in this paper is to examine how much we can improve session-level user engagement with ads through optimal sequencing of ads. As such, the key outcomes are defined at the session level. Figure 5 shows the empirical CDF of two main outcomes of interest in this study – session length, and the total number of clicks made in a session, which is our primary outcome of interest. We measure session length by the number of exposures shown within any session. Figure 5a shows how this outcome varies across sessions. As shown in this figure, around 50% of all sessions end in only two exposures. Further, the empirical CDF in Figure 5a shows that the vast majority of sessions do not last for more than 10 exposures and only a small fraction of them last for 30 or more exposures.

In Figure 5b, we show the empirical CDF for our primary outcome of interest – the total number of clicks per session. As expected, most sessions end with no clicks being made on ads shown within the session, and the percentage of sessions with at least one click amounts to 7.77%. This is a reasonably high percentage in this industry. Interestingly, there are sessions with more than one



(a) Session Length

(b) Number of Clicks Per Session

Figure 5: Empirical CDF of the session length and total number of clicks per session.

click. Further exploration suggests that these sessions are typically much longer than other sessions, with an average length of over 15 exposures.

3.3.3 Distribution of Micro-Interventions

A central piece of our study is the sequence of ads shown within the session. These sequences are determined by the publisher's ad placement decision at a given exposure. As such, each ad exposure can be treated as a micro-intervention that forms the whole sequence. We focus on three binary micro-interventions at any time period given the history of prior ads shown within the session – (1) *repeat*, (2) *breadth-increasing change*, and (3) *breadth-constant change*. Figure 6 illustrates these micro-interventions. This figure presents a case where the publisher wants to select the fourth ad. We define a micro-intervention as *repeat* if the publisher selects the last ad shown within the session. On the other hand, if the publisher shows any other ad, it is called a *change*, meaning that the current exposure shows a different ad from the last ad. In line with Rafieian and Yoganarasimhan (2018a), we then decompose *change* into two parts – (1) *breadth-constant change*, where the publisher changes the last ad but shows an ad that has been shown before in the session, and (2) *breadth-increasing change*, where the publisher changes the last ad and shows an ad that has not been shown before, thereby increasing the breadth of variety in the session. These three micro-interventions are mutually exclusive, i.e., only one of them takes the value one in any time period.

An important point to notice is that these binary micro-interventions together determine many characteristics of a sequence. For example, if prior interventions in a sequence are mostly breadth-

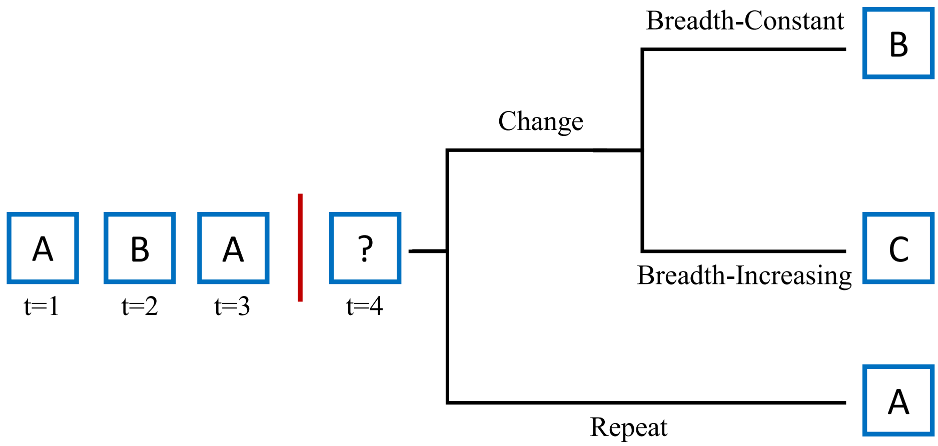


Figure 6: An example of three micro-interventions at exposure number 4

increasing changes, we will have a high breadth of variety in that sequence. We present the distribution of these micro-interventions in our data in Figure 7. Each line in this figure represents the percentage of a specific micro-intervention at different points within the session. A few patterns emerge from Figure 7. First, the lines for both *change* and *repeat* are flat over time, illustrating the independence and stability of the auction across different exposures. Further, repeat accounts for only 20% of micro-interventions, which is mostly due to the probabilistic nature of the auction: each ad has a chance of being shown proportional to the product of its bid and quality score. This creates a great degree of randomization in the ad allocation process. Finally, we observe a decreasing pattern in breath-increasing changes. By definition: a random change is less likely to show a completely new ad because we have a finite set of ads competing in a given auction. The fraction of breadth-increasing changes is over 30% across all exposures, indicating the extent of variation in ad allocation within the session.

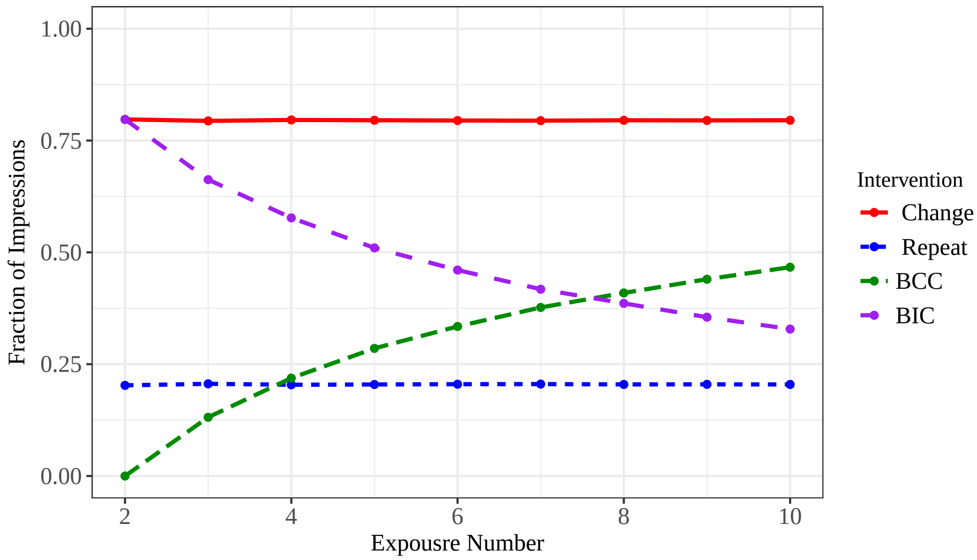


Figure 7: Distribution of micro-interventions at different exposure numbers within the session

4 Dynamic Framework for Sequencing of Ads

We now present our dynamic framework for sequencing of ads. This section proceeds as follows. We start with the motivation for the use of a dynamic framework in our context in §4.1. Next, in §4.2 we define the primitives of our framework and then specify an MDP that incorporates publishers' current and expected future rewards.

4.1 Motivation

The micro-interventions presented in §3.3.3, by and large, depend on the history of the sequence. Further, they determine how a sequence evolves, thereby shaping the history of the sequence for future exposures. Thus, in principle, if the within-sequence history affects which micro-intervention is most effective at any point, it is not clear whether that intervention is the optimal action from a dynamic point-of-view. In fact, the optimal action will be the one with the right balance between how effective it is now and where it transitions the entire sequence. For example, suppose that the publisher wants to fill two impressions with two ads A and B. Now, consider a case where ad A is generally a better ad, so the publisher's optimal decision in both impressions is to allocate them to ad A. However, showing ad A after ad B generates the best overall outcome, as it differentiates ad A and generates a significantly higher click probability. In that sense, the best possible myopic action at a point may not be the optimal decision from a forward-looking perspective.

The main question is whether the within-sequence history affects the outcomes that the publisher cares about. Prior literature has offered some evidence on such effects, including the effects of multiple exposures and temporal spacing of ads (Sahni, 2015a; Johnson et al., 2016b). In mobile in-app advertising setting, Rafieian and Yoganarasimhan (2018a) show that users' clicking behavior on the next ad depends on the variety of the prior sequence, as measure by the total number of changes and the breadth of variety (sum of all breadth-increasing changes). While they document the positive effects of variety, they also find that the number of prior exposures of an ad within the session positively affects users' clicking behavior. Given these two findings, it is not clear how the publisher should manage these micro-interventions at any given point. On the one hand, more change creates a higher variety in the sequence, thereby increasing the probability of click on the next ad. On the other hand, repeating an ad will increase the number of prior exposures of that ad which is shown to positively affect the probability of click. Thus, the publisher faces various inter-temporal trade-offs when allocating ads within the session.

We address this challenge by developing a dynamic framework that: (1) captures the inter-temporal trade-offs in publishers' ad placement decision in the session, and (2) uses both pre-session and adaptive session-level information to personalize the sequence of ads for the user in any given session. Our framework incorporates both how an action affects the outcome in the current period, as well as the externalities that influence future exposures. In the next sections, we present the details of our framework.

4.2 Model Setup

We specify an MDP that captures the inter-temporal trade-offs in publishers' decision problem by taking into account both current and expected future rewards. An MDP is characterized by a set of primitives that serve as inputs into the objective function that the decision-maker seeks to maximize. We present a generic definition of these primitives below and discuss the specifics of each in the next sections.

1. Time Period (t): The first component we need to specify is the time unit. Since exposures are shown sequentially in our case, we treat each ad exposure as a time period wherein the publisher needs to decide on her actions. $t = 1$ indicates the first exposure in a session.

2. State Space ($\mathcal{S}$): The state space consists of all the information the publisher has about an exposure, which affects her decision-making process.

3. Action Space ($\mathcal{A}$): The action space contains the set of actions the publisher can take. In our case, this action is to show one ad from the ad inventory every time an impression is recognized. As such, $\mathcal{A}$ is the full ad inventory in our problem.

4. Transition Function (P): This function determines how the current state transitions to the

future state given the action made at that point. As such, we can define $P : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$ as a stochastic function that calculates the probability $P(s' \mid s, a)$ where $s, s' \in \mathcal{S}$ and $a \in \mathcal{A}$. Note that this is a crucial component of an MDP since publishers cannot control the dynamics of the problem if the next state is not affected by the current decision. In §4.2.3, we discuss the components of the transition function in our problem in detail.

5. Reward Function (R): This function determines the reward for any action a at any state s. As such, we can define this function as $R : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$. This function can take different forms depending on the publisher's objective. In our case, since the publisher is interested in optimizing user engagement, she can use different metrics that reflect user engagement such as the probability that the user clicks on the ad. In §4.2.4, we discuss our choice of reward function in greater details.

6. Discount Factor (β): The rate at which the publisher discounts the expected future rewards. In other words, it is the weight that the publisher assigns to the future relative to the current period.

With all these primitives defined, we can now write the publisher's maximization problem as follows:

$$\underset{a}{\operatorname{argmax}} \left[R(s, a) + \beta \mathbb{E}_{s'|s,a} V(s') \right], \tag{2}$$

where $V(s')$ is the value function incorporating expected future rewards at state s' if the publisher selects ads optimally. Following Bellman (1966), we can write this value function for any state $s \in \mathcal{S}$ as follows:

$$V(s) = \max_{a} R(s, a) + \beta \mathbb{E}_{s'|s,a} V(s') \tag{3}$$

As shown in Equation (2), the optimization problem consists of two key elements – the current period reward and the expected future rewards. The publisher chooses the ad that maximizes the sum of these two elements.

4.2.1 State Variables

The state variables contain all the information that the publisher can use for any given exposure in a session. As discussed earlier, the publisher can take two pieces of information into account: (1) pre-session information, and (2) session-level information. Pre-session information contains any data on the user up until the current session, including his demographic variables and behavioral history. For any session i, we denote the pre-session state variables by X_i. It is important to notice that the pre-session variables are not adaptive, i.e., it does not change within the session and hence not have t subscript.

On the other hand, session-level variables are adaptive and change within the session. At any given point in a session, this information captures the information about the prior sequence of ads shown to the user as well as user's actions after seeing each ad. Let $G_{i,t}$ denote the session-level state variables. We can write:

$$G_{i,t} = \langle A_{i,1}, Y_{i,1}, A_{i,2}, Y_{i,2}, \ldots, A_{i,t-1}, Y_{i,t-1} \rangle, \tag{4}$$

where $A_{i,s}$ denotes the ad shown in exposure number s and $Y_{i,s}$ denotes whether the user clicked on this ad. As a result, $G_{i,t}$ is the sequence of all ads and actions within the session up to the current time period. Overall, we define the state variables as $S_{i,t} = \langle X_i, G_{i,t} \rangle$, i.e., a combination of both pre-session and session-level variables.

4.2.2 Action Space: Ad Inventory

As mentioned in §4.2, the publisher's action space in our case is the ad inventory. At each point of time, the publisher chooses one ad from the inventory to show to the user. We only focus on the top 15 ads in the ad inventory due to four key reasons. First, as shown in Figure 8, the top 15 ads generate over 70% of all impressions in the focal messenger app. So they collectively account for a significant portion of our observed data. Second, these top ads are more stable in terms of their budget. Since we want to run counterfactual policies, it is important to make sure that advertisers' budgets do not cause any problem for the reliability of our results. Third, we have more data for these ads as these are shown in a wide variety of state variables. This makes our estimates for their outcomes more accurate and less noisy. Finally, limiting the action space makes the problem computationally more tractable when we want to numerically solve for the optimal policy.

4.2.3 Transition Function

We now characterize the law-of-motion, i.e., how state variables transition given the publisher's action at any point. As mentioned earlier, we are interested in the probability of the next state being s', given that action a is taken in state s in the current period, i.e., $P(s' \mid a, s)$. Suppose that the user is in state $S_{i,t} = \langle X_i, G_{i,t} \rangle$ at exposure t in session i. The only time-varying factor in $S_{i,t}$ that can transition is $G_{i,t}$, which is the history of the sequence. Given the definition of $G_{i,t}$ in Equation (4), we can determine the next state if we know user's decision to click on the current ad and/or continue staying in the session. There are three mutually exclusive possibilities for state transitions:

1. *Click and stay:* If the user clicks on ad $A_{i,t}$ and stays in the session, we can define the next state as follows:

$$S_{i,t+1} = \langle X_i, G_{i,t}, A_{i,t}, Y_{i,t} = 1 \rangle, \tag{5}$$

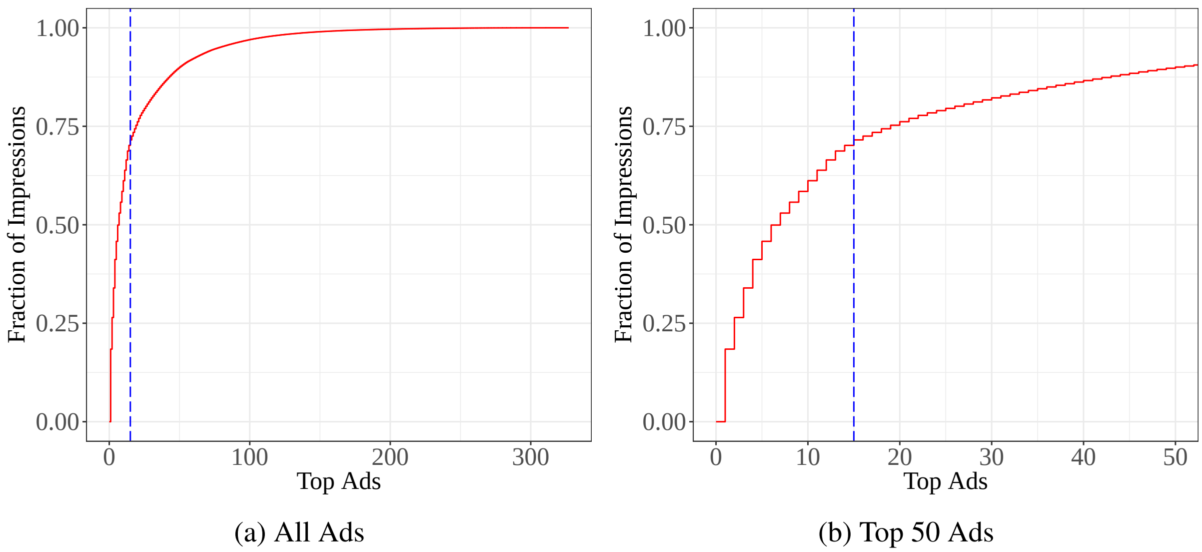


Figure 8: Cumulative fraction of impressions associated with the top ads. The figure in the left shows the distribution for all 327 ads. The figure in the right zooms in the top 50 ads.

where $Y_{i,t} = 1$ indicates that the user has clicked on the ad shown in exposure number t.

2. *No click and stay:* If the user does not click on ad $A_{i,t}$ and stays in the session, we can similarly define the next state as follows:

$$S_{i,t+1} = \langle X_i, G_{i,t}, A_{i,t}, Y_{i,t} = 0 \rangle, \tag{6}$$

where $Y_{i,t} = 0$ indicates that the user has not clicked on the ad shown in exposure number t.

3. *Leave:* Regardless of user's clicking outcome, if the user decides to leave, the entire session is terminated and there is no more decision to be made. Thus, we can write:

$$S_{i,t+1} = \varnothing \tag{7}$$

Figure 9 visually presents the three possibilities presented above. This figure illustrates an example where the publisher shows an ad in the fourth exposure in a session. It shows three possibilities and how each forms the next state. Based on this characterization, we can now define

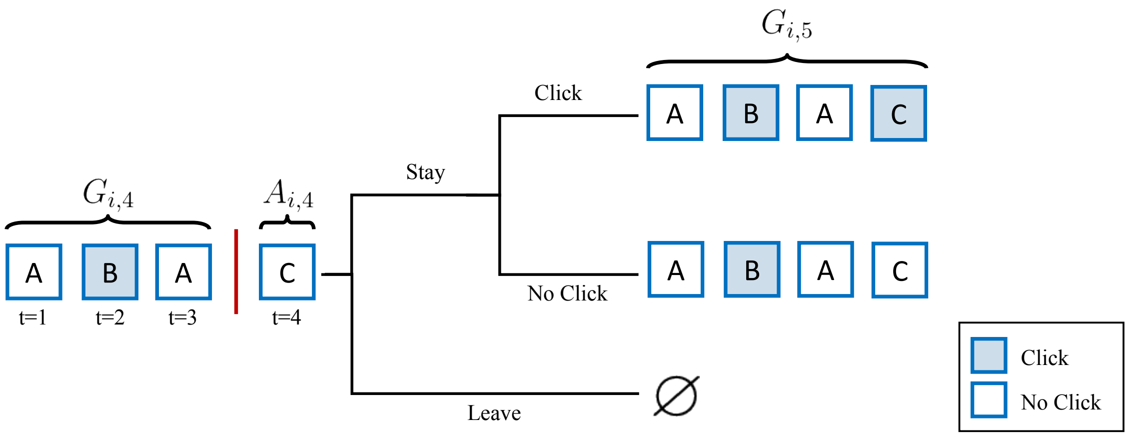


Figure 9: An example illustrating the state transitions.

the transition function for any pair of action and state as follows:

$$P(S_{i,t+1} \mid a, S_{i,t}) = \begin{cases} \big(1 - P(L_{i,t} \mid a, S_{i,t})\big) P(Y_{i,t} \mid a, S_{i,t}) & \text{Equation (5)} \\ \big(1 - P(L_{i,t} \mid a, S_{i,t})\big)\big(1 - P(Y_{i,t} \mid a, S_{i,t})\big) & \text{Equation (6)} \\ P(L_{i,t} \mid a, S_{i,t}) & \text{Equation (7)} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Equation (8) illustrates that the publisher needs to accurately estimate two user-level outcomes given any ad shown – click and leave probabilities. In §5, we discuss our approach to obtain these estimates.

4.2.4 Reward Function

Another piece of an MDP that needs to be defined is the reward function. The reward function can take different forms that vary with the publisher's objective and main outcome of interest. We primarily focus on the number of clicks per session as our main objective. In our case, clicks are particularly good measures of the user engagement with ads because of two reasons. First, all ads in our study are mobile apps that want more clicks and installs. In the literature, this type of ads is called performance ads and their match value is generally assumed to be the probability of click (Arnosti et al., 2016). Second, click is the main source of revenue for the publisher, since the advertiser only pays when a click happens.

There are other advantages in using the click as the main outcome of interest. First, it is a well-recorded outcome that is realized immediately in the data. Second, click is a function of users' behavior, whereas other outcomes usually involve other players such as advertisers (e.g., publisher's revenues). Thus, the publisher's optimization problem only depends on inferring users' behavior, which is a feasible task in a data-rich environment.

Given that publishers want to maximize the number of clicks made per session, we can define the reward function as the probability of click for a pair of state and action. For exposure number t in session i, we can write:

$$R_t(a; S_{i,t}) = P(Y_{i,t} \mid a, S_{i,t}) \tag{9}$$

This is the probability of click on ad a if shown in the current state.

4.2.5 Discount Factor

The discount factor in Equation (2) reflects the relative importance of the expected future rewards compared to the current period reward from the publisher's perspective. As such, it generally takes a positive value close to one, if the publisher is forward-looking, i.e., they incorporate the expected future rewards in their decision problem. If the discount factor is zero, it means that the publisher is myopic and only cares about the reward in the current period.

In our dynamic framework, the publisher's decision is to select an ad for each impression in a session. The entire session happens in just a few minutes, depending on the number of exposures the user chooses to stay for. Given the short time horizon of the optimization problem, a risk-neutral publisher must value the current and expected future rewards equally, indicating that β is very close to 1.

4.2.6 Policy Definition

We characterize a dynamic policy as a mapping $\pi : \mathcal{S} \times \mathcal{A} \rightarrow [0, 1]$, that assigns a probability to any action $a \in \mathcal{A}$ taken in any given state $s \in \mathcal{S}$. For a deterministic policy, $\pi(a \mid s)$ will take value one only for one ad for any given state. However, we define the policy as a probability function to allow for non-deterministic policies as well.

In sum, the main goal of our framework is to develop a policy π^* that maximizes the expected reward given the initial set of state variables and the transition function. We later discuss how we develop and evaluate this policy in §6.2 and §6.3, respectively.

5 Empirical Strategy

In this section, we present our empirical strategy to estimate the primitives of the dynamic framework defined in §4. This involves the estimation of probabilistic components of the transition function as characterized in §4.2.3, as well as estimation of unknown components in the reward function as presented in §4.2.4. Together, this gives us two estimands – (1) leave outcome, and (2) click outcome. The task in both cases is to accurately predict the outcome for a pair of state and action. However, given the broader task of taking the best action at any state that we are interested in, we need to obtain these outcome estimates not only for the ad that is shown in the data but also for the counterfactual ads that are not shown in the data. Below is a formal definition of these two tasks:

Task 1: For any set of state variables observed in the data, we want to accurately estimate the click probability for all ads if shown in that impression. That is:

$$\hat{y}_{i,t}(a; S_{i,t}) = P(Y_{i,t} \mid a, S_{i,t}), \forall a \in \mathcal{A}_i \tag{10}$$

Task 2: For any set of state variables observed in the data, we want to accurately estimate the leave probability for all ads if shown in that impression. That is:

$$\hat{l}_{i,t}(a; S_{i,t}) = P(L_{i,t} \mid a, S_{i,t}), \forall a \in \mathcal{A}_i \tag{11}$$

We discuss our empirical strategy for Tasks 1 and 2 and helps us set the scope of our predictive model in §5.1. We then describe the learning algorithm that we use and its advantages in §5.3. Finally, in §5.4, we present the estimation results on our click and leave probability model.

5.1 Filtering Strategy for Counterfactual Estimation

For the task of outcome prediction like ours, we can use machine learning methods that capture more complex relationships between the covariates and outcomes (Mullainathan and Spiess, 2017). However, these methods are only able to predict the outcome for observations within the joint distribution of the observed data. This implies that if an ad could have never been shown in an observation in the actual data, our outcome estimate for that ad may not be reliable. Thus, if there is no randomization in the ad allocation process, we cannot reliably estimate the outcome for counterfactual ads.

This informs our empirical strategy for counterfactual estimation of click and leave outcomes. While ads are selected through a deterministic allocation rule in most commonly used auctions such as second-price, the quasi-proportional auction in our setting has the advantage of using a probabilistic rule that induces randomization in the ad allocation process. The key issue with a deterministic allocation rule is that the ad that is actually shown in the data is the only ad that *could have been shown*, and other counterfactual ads *could have never been shown*. We argue that if an ad could have been shown in a session (i.e., has non-zero propensity score), this observation could have been generated within the joint distribution of the training data. Thus, we are able to accurately estimate the outcome for these ads in both factual or counterfactual situations.

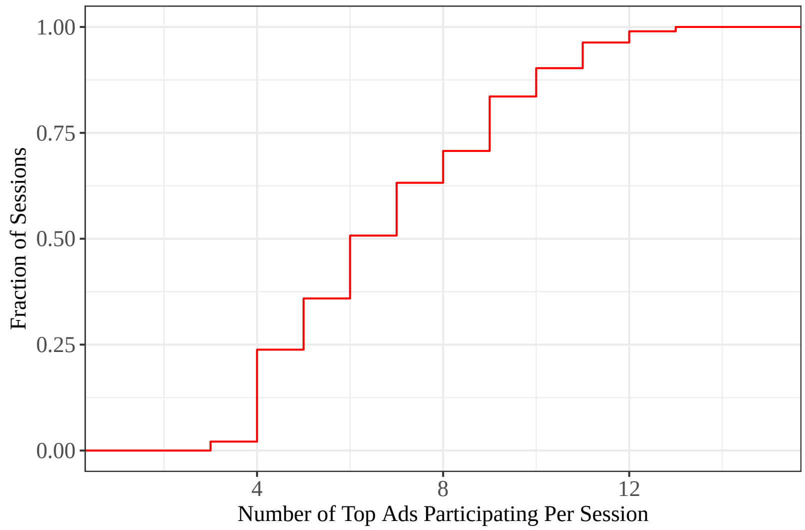


Figure 10: Empirical CDF of the number of ads competing per session.

To reflect this idea, we employ a filtering strategy similar to that in Rafieian and Yoganarasimhan (2018b). Here, for each session i, we identify the set of ads that could have been shown in that session and call it $\mathcal{A}_i$. We filter out the set of ads that are not in $\mathcal{A}_i$. This step sets the scope of our ability to generate counterfactual estimates. Figure 10 shows the empirical CDF of the number of ads participating in the auction for a session. As shown in this figure, the number of ads competing for each impression is quite variable across sessions.

Further, the extent of randomization in our problem allows us to make the unconfoundedness assumption: for any exposure number t in session i, the set of potential outcomes for all ads in $\mathcal{A}_i$ is independent of the actual ad that is shown in that exposure, conditional on the state variables. We can write:

$$\{Y_{i,t}(a)\}_{a \in \mathcal{A}_i} \perp\!\!\!\perp A_{i,t} \mid S_{i,t},$$

where $Y_{i,t}(a)$ is the potential outcome at the exposure t in session i when ad a is being shown. We can even weaken this assumption as only conditioning on demographic variables D_i yields the conditional independence presented above. This is because the ad allocation is random, controlling for the propensities determined by auction. Since only targeting variables can affect propensities in an auction, we only need to control for them. Thus, if we properly control for D_i in our predictive model, our outcome estimates preserve consistent treatment effects.

5.2 Feature Generation

As discussed earlier, our goal is to estimate click and leave outcomes for any combination of ad and state variables, as shown in Equations (10) and (11). A major challenge in estimating these

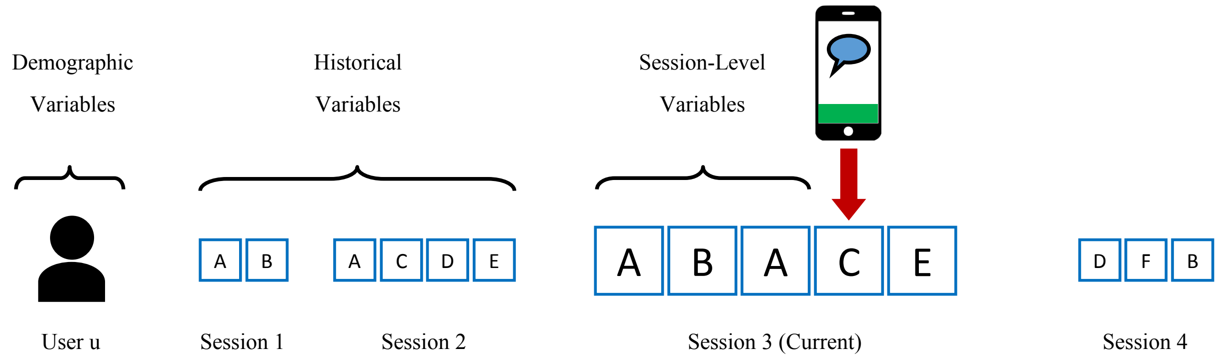


Figure 11: A visual schema for our feature generation and categorization.

equations is that the set of inputs is quite large, containing the entire sequence of prior ads shown to the user. In this section, we present a feature generation framework that maps a combination of state variables and ads ($\langle S_{i,t}, a\rangle$) to a set of meaningful features that we can give as inputs to our learning algorithm. Ideally, we need our final set of features to fully represent $\langle S_{i,t}, a\rangle$ in a lower dimension without any information loss. Thus, we generate a set of features that help us predict users' clicking behavior and app usage based on the prior literature on advertising.

We categorize these features into three groups: (1) demographic features, (2) historical features, and (3) session-level features. Demographic and historical features relate to the pre-session state variables (X_i), whereas session-level features relate to the session-level variables ($G_{i,t}$). Figure 11 provides an overview of our feature generation and categorization. In this example, the user is at her fourth exposure in her third session. The features for this particular exposure include the observable demographic features, historical features generated from the prior sessions, and session-level features that are generated from the first three exposures shown in the current session. Clearly, we do not use any information from the future to generate a feature: at any point, we only use the prior history up to that point. In the following sections, we describe all these features in detail.

5.2.1 Demographic Features

This includes the variables that we already observe in our data (see §3.2.1), such as the province, latitude, longitude, smartphone brand, mobile service provider (MSP), and connectivity type. For any session i, we use D_i to denote the set of demographic features. These features do not transition based on the ad that the publisher shows at any time period. As such, we do not use subscript t for them.[8] We include these features because of two reasons. First, these features help predict both users' clicking behavior and app usage. Second, the targeting variables are the main confounding source, and controlling them guarantees that we control for propensity score of ads when estimating the outcomes.

5.2.2 Historical Features

Historical features reflect the user's past activity prior to the current session. While demographic features are available in the data, we need to generate historical features based on the pre-session information. These features are not adaptive, i.e., they remain constant within the session. We follow the approach in Rafieian and Yoganarasimhan (2018b) to generate these features. We refer the reader to that paper for details on why these features help predict user-level outcomes.

Let i, u, t, p, and a denote the session, user, exposure number, app, and ad respectively. Since we only focus on the top app, using the subscript p indicates that we only use impressions in that app to generate a feature. Otherwise, we calculate the feature using all impression. Below, we present the detailed set of our historical features along with their definition:

- $\textit{Imp}_{i,u}$: The total number of impressions user u has seen prior to session i.
- $\textit{Click}_{i,u}$: The total number of clicks user u has made prior to session i.
- $\textit{Imp}_{i,u,p}$: The total number of impressions user u has seen in the top app prior to session i.
- $\textit{Click}_{i,u,p}$: The total number of clicks user u has made in the top app prior to session i.
- $\textit{Imp}_{i,u,t}$: The total number of impressions user u has seen at exposure number t prior to session i.
- $\textit{Click}_{i,u,t}$: The total number of clicks user u has made at exposure number t prior to session i.
- $\textit{Imp}_{i,u,a}$: The total number of impressions of ad a that user u has seen prior to session i.
- $\textit{Click}_{i,u,a}$: The total number of click that user u has made on ad a prior to session i.
- $\textit{Space}_{i,u,a}$: The time space (in minutes) between session i and the last time ad a is shown to user u prior to session i. It takes value zero if there was no prior exposure of ad a.
- $\textit{LastSessionLength}_{i,u}$: The length of last session (in number of exposures) that user u was exposed to prior to session i.
- $\textit{AvgSessionLength}_{i,u}$: The average length of the sessions (in number of exposures) that user u was exposed to prior to session i.
- $\textit{LastFreeTime}_{i,u}$: The free time (in minutes) user u has had between her last session and session i.

[8] One could argue that features such as latitude and longitude may change within the session. While this is possible, it is unlikely to happen as a result of the publisher's ad interventions. Further, the sessions are usually short, and we rarely observe such a change in our data.

- *AvgFreeTime*$_{i,u}$: The average free time (in minutes) user u has had between her sessions prior to session i.
- *Breadth*$_{i,u}$: The total number of distinct ads that user u has seen prior to session i.
- *GiniSimpson*$_{i,u}$: The Gini-Simpson index for ads that user u has seen prior to session i (Simpson, 1949). This metric captures the diversity of prior ad exposures by calculating the probability that two random exposures from the past were of different ads. A higher Gini-Simpson index means that the user has seen a more diverse set of ads. We can write the Gini-Simpson index as follows:

$$\textit{GiniSimpson}_{i,u} = 1 - \sum_{a \in \mathcal{A}} \frac{\textit{Imp}_{i,u,a}(\textit{Imp}_{i,u,a} - 1)}{\textit{Imp}_{i,u}(\textit{Imp}_{i,u} - 1)} \tag{12}$$

For a session i, we denote all these features as H_i. Like demographic features, publisher's actions will not change historical features within the session. However, for the next session that the user participates in, the historical features are updated. Further, it is worth noting that three historical features, *Imp*$_{i,u,a}$, *Click*$_{i,u,a}$, and *Space*$_{i,u,a}$, are ad-specific. It means that these features will change when different ads are selected.

5.2.3 Session-Level Features

The session-level features are key to our analysis as we are interested in optimal sequencing of ads within the session. Further, we allow these features to change by the exposure number. That is, depending on the prior exposures within the session, these features will evolve. Below is a list of session-level temporal features:

- *Imp*$_{i,t}$: The total number of impressions the user has seen in session i prior to exposure number t. For any exposure number t, this feature is $t - 1$.
- *Click*$_{i,t}$: The total number of clicks the user has made in session i prior to exposure number t.
- *Imp*$_{i,a,t}$: The total number of impressions of ad a that user has seen in session i prior to exposure number t.
- *Click*$_{i,a,t}$: The total number of clicks that the user has made on ad a in session i prior to exposure number t.
- *Space*$_{i,a,t}$: The number of exposures between exposure number t and the last time ad a was shown in session i. It takes value 0 when there is no prior exposure of ad a in the session.
- *Breadth*$_{i,t}$: The total number of distinct ads that the user has seen within session i prior to exposure number t. We can define this feature as follows:

$$\textit{Breadth}_{i,t} = \sum_{a \in \mathcal{A}} \mathbb{1}(\textit{Imp}_{i,a,t} > 0) \tag{13}$$

- *Changes*$_{it}$: The total number of consecutive changes of ads prior to the exposure number t within the session i. We can write:

$$\mathit{Change}_{i,t} = \sum_{j=2}^{t-1} \mathbb{1}(A_{i,j} \neq A_{i,j-1}), \tag{14}$$

where $A_{i,j}$ is the ad shown at exposure number j in session i.

- *GiniSimpson*$_{i,t}$: The Gini-Simpson index for the ads shown within session i prior to exposure number t. Following the same logic in Equation (12), we can write:

$$\mathit{GiniSimpson}_{i,t} = 1 - \sum_{a} \frac{\mathit{Imp}_{i,a,t}(\mathit{Imp}_{i,a,t} - 1)}{(t-1)(t-2)} \tag{15}$$

For any session i and exposure number t, we denote all session-level features by $O_{i,t}$. As such, this is the only set of features that has subscript t, indicating that it changes within the session. Therefore, the publisher's actions affect the transition of these features in the session. One could argue that historical features also change within the session as user's history accumulates after each exposure. It is worth noting that we do not update the history within the session because session-level temporal features capture that information. As a result, not updating historical features will not result in any information loss.

5.3 Learning Algorithm

Here we describe the learning algorithm that we use to estimate the click and leave models. Since our goal is to estimate these outcomes as accurately as possible, we need to build a model that is able to capture complex relationships between covariates and outcomes. For both outcomes, we use Extreme Gradient Boosting (XGBoost henceforth) method developed by Chen and Guestrin (2016), which is a fast and scalable version of Boosted Regression Trees (Friedman, 2001). There are some key reasons why we use XGBoost as our main optimization method. First, it has been shown to outperform most existing methods in most prediction contests, especially those related to human decision-making like ours (Chen and Guestrin, 2016). Second, Rafieian and Yoganarasimhan (2018b) show that in the same context, XGBoost achieves the highest predictive accuracy compared to other methods.

There are important implementation details in making an XGBoost model that we summarize as follows. First, following the arguments in Rafieian and Yoganarasimhan (2018b), we use logarithmic loss as our loss function. Second, we split our data into two parts – training and test sets. To tune parameters of XGBoost, we split the training set into two parts and use one as a hold-out validation

Evaluation Metric	Training	Test
RIG	0.2847	0.2768
R^2	0.1899	0.1872
AUC	0.8560	0.8531

Table 2: Performance of the click estimation model on training and test sets

set to prevent the model from over-fitting. Finally, to select the hyper-parameters accurately, we conduct a grid search over a large set of hyper-parameters and select those that give us the best performance on a hold-out validation set. Finally, we evaluate the performance of our model on both training and test set. Note that the test set is not used at any stage to select the model. Thus, the evaluation on the test set demonstrates an out-of-sample predictive accuracy.

5.4 Results

5.4.1 Results from the Click Estimation Model

We use three different evaluation metrics to evaluate the predictive performance of our model:

- *Relative Information Gain (RIG):* This is defined based on the log loss, which is our loss function in the XGBoost model. This metric reflects the percentage improvement in log loss compared to a baseline model that simply predicts the average CTR for all impressions.
- *R-Squared (R^2):* This is the most commonly used metric in marketing and economics, and intuitively calculates the percentage of variance in the outcome that our model can explain.
- *Area Under the Curve (AUC):* It determines how well we can identify *true positives* without identifying *false positives*. This score ranges from 0 to 1 and a higher score indicating better performance.

The results of these three metrics are shown in Table 2. We find that our model achieves over 28.47% and 27.68% RIG on training and test sets, respectively. This predictive accuracy is quite substantial compared to the literature (Yi et al., 2013; Rafieian and Yoganarasimhan, 2018b). Further, the results on R^2 also indicate that our model reaches an excellent predictive performance. Given the inherent noise and variability in clicks, explaining over 18% of the variance in this outcome requires a very powerful model. Finally, the results on the last metric document a very good classification power.

5.4.2 Results from the Leave Estimation Model

We now discuss the results from the leave model. The leave outcome is an important piece of the transition function, since the leave outcome by the user terminates the entire session. We use the same evaluation metrics to evaluate the predictive performance of our leave estimation model. Table 3 summarizes the results in terms of these metrics on both the training and test sets. We find that

Evaluation Metric	Training	Test
RIG	0.0975	0.0949
R^2	0.0942	0.0918
AUC	0.7164	0.7135

Table 3: Performance of the leave estimation model on training and test sets

	Click Model			**Leave Model**		
Model	RIG	R^2	AUC	RIG	R^2	AUC
Full model	0.2768	0.1872	0.8531	0.0949	0.0918	0.7135
Without session-level	0.2245	0.1280	0.8334	0.0881	0.0889	0.7055

Table 4: Performance of models with/without session-level features

our leave estimation model explains over 9% of the variance in the leave outcome on both training and test sets.

Overall, it is worth noting that we should not expect the leave model to to have high predictive ability for a variety of reasons. First, clicking decision is mainly related to the ad shown, whereas leave decision is somewhat independent of ad exposures, especially in the messenger app where people's usage stems from users' messaging behavior that is unobserved to the researcher. Second, we focus on new users for whom we do not have a long panel. Hence, our features reflecting their past usage may not be very informative.

5.4.3 Preliminary Evidence on the Gains from the Adaptive Framework

An adaptive framework allows the publisher to use real-time session-level information to make decisions. Intuitively, the publisher can benefit from an adaptive framework for decision-making if session-level features play an important role in driving the main outcome of interest. In other words, if the user's decision to click on ads is only a function of his demographic or historical features and does not change based on session-level features, using an adaptive framework becomes ineffective. Thus, an intuitive test that can provide some preliminary evidence on the effectiveness of an adaptive framework is to drop session-level features from the main click and leave models and see whether it results in a drop in the predictive performance of these models.

The model without session-level features exclude the following features: $Imp_{i,t}$, $Click_{i,t}$, $Breadth_{i,t}$, $Changes_{i,t}$, $GiniSimpson_{i,t}$, $Imp_{i,a,t}$, $Click_{i,a,t}$, $Space_{i,a,t}$. We compare the performance of this model with the full model for both click and leave estimation models. The results on the performance of these models on the same test set are presented in Table 4. As shown in this table, the performance of the model drops when we exclude session-level features. This finding suggests and publishers can benefit from using adaptive frameworks for ad sequencing.

6 Ad Sequencing Policies

As discussed earlier, our main goal is to use our dynamic framework and develop a sequencing policy that maximizes user engagement in the session. We primarily measure user engagement by the number of clicks per session. To develop sequencing policies, we can use our estimates for the transition and reward functions to solve for the optimal policy in the MDP specified in §4.

In this section, we first present a series of other baseline policies that are commonly used in practice in §6.1. It allows us to compare the performance of the adaptive forward-looking policy developed by our framework with other benchmarks and establish the gains from adopting a dynamic framework. We then discuss the empirical evaluation of all these policies in §6.2. Finally, in §6.4, we present our results on different sequencing policies, explore heterogeneity in gains from the dynamic sequencing policy across sessions, and offer some explanation for the performance of different methods.

6.1 Definition of Sequencing Policies

We present a series of benchmark policies to examine the publisher's gains from adopting a dynamic framework for ad sequencing. We consider benchmark policies that drop modeling components in our dynamic framework and/or reflect the current norm in research and practice. As such, our comparison allows us to pin down how valuable each modeling component is and how much we can improve over the current practice. Starting with the adaptive forward-looking policy, which is based on our dynamic framework, we present the list of policies that we consider as follows:

- **Adaptive forward-looking sequencing policy:** This sequencing policy uses the dynamic framework in §4 with the expected probability of click being the reward function. Using our generic formulation of an MDP in Equation (2), we can write the publisher's optimization problem specific for a click-maximizing objective as follows:

$$\underset{a \in \mathcal{A}_i}{\operatorname{argmax}} \left[R_t(a; S_{i,t}) + \beta \mathbb{E}_{S_{i,t+1}|S_{i,t},a} V_{t+1}(S_{i,t+1}) \right], \tag{16}$$

  where the value function for any state variable at exposure number t can be written as follows:

$$V_t(S_{i,t}) = \max_{a \in \mathcal{A}_i} \left[R_t(a; S_{i,t}) + \beta \mathbb{E}_{S_{i,t+1}|S_{i,t},a} V_{t+1}(S_{i,t+1}) \right] \tag{17}$$

  This policy is both adaptive and forward-looking, i.e., it takes the session-level information (adaptive) and future information (forward-looking) into account. We also call this policy fully dynamic sequencing policy.

- **Adaptive myopic sequencing policy:** This sequencing policy does not take into account the expected future rewards when making the decision at any point. This is equivalent to the adaptive forward-looking sequencing with $\beta = 0$ that turns off the weight on the future rewards. Thus, we can write the objective function for adaptive myopic sequencing as follows:

$$\operatorname*{argmax}_{a \in \mathcal{A}_i} R_t(a; S_{i,t}) \tag{18}$$

In this policy, the publisher selects the ad that maximizes CTR in the current period. It is worth noting that this policy is adaptive, as it uses the session-level information that is time-varying. However, it is myopic in the sense that it ignores future information. This case reflects the common practice of using contextual bandits in the industry.

- **Non-adaptive single-ad policy:** This policy only uses the pre-session information. Since it does not use adaptive information, this policy allocates all the impressions to a single ad that has the highest average CTR. This is similar to the practice of using a fixed ad slot where the whole session is allocated to one ad. The objective in this case is the same as Equation (18) only for $t = 1$.[9]

  This policy provides some insight into the ad sequencing problem because it has two distinct features. First, it captures the potential gains from using a short-lived ad slot as compared to the fixed ad slot. Second, it demonstrates the value of adaptive session-level information.
- **Random sequencing policy:** In this sequencing policy, the publisher randomly selects ads from the ad inventory. While this is a naive policy, it can serve as a benchmark showing how well we can do without any model.

6.2 Empirical Evaluation

We now explain how we use our empirical estimates for both click and leave outcomes to develop and evaluate all the policies presented in §6.1. We start with the solution concept for the baseline policies in §6.2.1. These policies are more straightforward and easier to derive. We then present the solution concept for the adaptive forward-looking sequencing policy in §6.2.2. We finally discuss how we can evaluate the performance of these policies.

6.2.1 Solution Concept for Baseline Policies

We now describe how we obtain the baseline policies described in §6.1: (1) adaptive myopic sequencing policy, (2) non-adaptive single-ad policy, and (3) random sequencing policy. Since none

[9]One could argue that the optimal single-ad that is selected for the entire session may be different from the optimal ad for the first exposure. We acknowledge this issue and check the robustness of our results by using a dynamic optimization constrained by a single ad to be shown for the entire session. In the main text, however, we use the more straightforward approach of allocating the entire session to the ad with the highest CTR in the first exposure.

of these policies incorporate expected future rewards, we need to mainly rely on our estimates for the reward function with the click-maximizing objective, i.e., $\hat{y}_{i,t}(a; S_{i,t})$. We present our empirical approach to solve for these policies below:

- **Adaptive myopic sequencing policy:** As presented in Equation (18), this policy selects the ad with the highest CTR at any point. Therefore, we can write $a_t^m(S_{i,t}) = \text{argmax}_{a \in \mathcal{A}_i} \hat{y}_{i,t}(a; S_{i,t})$, where $a_t^m(S_{i,t})$ indicates the ad to be shown under adaptive myopic sequencing in state $S_{i,t}$. We can characterize the best policy under adaptive myopic sequencing as follows:

$$\hat{\pi}^m(a \mid S_{i,t}) = \begin{cases} 1 & a = a_t^m(S_{i,t}) \\ 0 & a \neq a_t^m(S_{i,t}) \end{cases} \tag{19}$$

- **Non-adaptive single-ad policy:** This case is the solution for the adaptive myopic sequencing policy at $t = 1$. As such, we have $a_t^s(S_{i,t}) = a_t^m(S_{i,1})$ for any t in session i. Thus, the non-adaptive single-ad policy for each session i can be characterized as follows:

$$\hat{\pi}^s(a \mid S_{i,t}) = \begin{cases} 1 & a = a_t^s(S_{i,t}) \\ 0 & a \neq a_t^s(S_{i,t}) \end{cases} \tag{20}$$

- **Random sequencing policy:** This policy basically gives the same chance to all the ads in the inventory. Thus, we can write the random sequencing policy as follows:

$$\hat{\pi}^r(a \mid S_{i,t}) = \begin{cases} \frac{1}{|\mathcal{A}_i|} & a \in \mathcal{A}_i \\ 0 & a \notin \mathcal{A}_i \end{cases}, \tag{21}$$

  where $\mathcal{A}_i$ is the set of ads competing in session i.

6.2.2 Solution Concept for the Adaptive Forward-looking Sequencing Policy

Solving for the best policy in an MDP can become a daunting task when the state space is high dimensional. Given the set of features we use for our estimation tasks, we need to store the full history within the session. For example, to update session-level features such as *Click*$_{i,a,t}$ or *Breadth*$_{i,t}$, we need to know the entire sequence of actions and outcomes within the session up until exposure number t, which can be computationally burdensome for large t.

The session-level history that we need to store for the transition function contains the sequence of all ads and the click outcomes. That is, for exposure number t in session i, this history is defined as $G_{i,t} = \langle A_{i,1}, Y_{i,1}, ..., A_{i,t-1}, Y_{i,t-1} \rangle$. This set can take $(2|\mathcal{A}|)^{t-1}$ unique values for each t, indicating that it grows exponentially in t. Thus, one way to resolve this problem is to consider a finite horizon case with a reasonable T that is sufficiently large to let us exploit the dynamic framework, while reasonably small to help us avoid computational issues. We argue that that $T = 6$ satisfies both these conditions since the user is quite likely not to get to the seventh exposure at all given the results we show in Figure 5a: around 75% of the sessions have shown at most six exposures. Thus, optimizing the exposure after that point has only a marginal effect on the performance of the click-maximizing policy.

To empirically derive the adaptive forward-looking sequencing policy, we need two key estimands – $\hat{y}_{i,t}(a; S_{i,t})$ and $\hat{l}_{i,t}(a; S_{i,t})$. The former affects both the reward function and state transitions, whereas the latter only affects the state transitions. For notational convenience and brevity, let $\tilde{V}_t(a, S_{i,t})$ denote our estimate for the sum of both current period reward and expected future rewards given action a being taken in state $S_{i,t}$. Namely, it is the estimated objective in Equation (16). Using our empirical estimates of expected click and leave probability, we can write $\tilde{V}_t(a, S_{i,t})$ as follows:

$$\begin{aligned}\tilde{V}_t(a, S_{i,t}) =& \hat{y}_{i,t}(a; S_{i,t}) + \left(1 - \hat{l}_{i,t}(a; S_{i,t})\right) \hat{y}_{i,t}(a; S_{i,t}) V_{t+1}\left(\langle S_{i,t}, a, Y_{i,t} = 1\rangle\right) \\ &+ \left(1 - \hat{l}_{i,t}(a; S_{i,t})\right)\left(1 - \hat{y}_{i,t}(a; S_{i,t})\right) V_{t+1}\left(\langle S_{i,t}, a, Y_{i,t} = 0\rangle\right),\end{aligned} \tag{22}$$

where $\langle S_{i,t}, a, Y_{i,t} = 1\rangle$ and $\langle S_{i,t}, a, Y_{i,t} = 0\rangle$ denote the state variables in the next period. The equation above shows that we can easily break the expected future rewards into two deterministic pieces. We now propose a backward induction solution concept for this case as described below:

1. We start from the last period, T. We assume that this is the last exposure number within the session. Hence, the problem is static in that period. We can write:

$$\hat{V}_T(S_{i,T}) = \max_{a \in \mathcal{A}_i} \hat{y}_{i,T}(a; S_{i,T}) \tag{23}$$

$$\hat{a}^d_T(S_{i,T}) = \operatorname*{argmax}_{a \in \mathcal{A}_i} \hat{y}_{i,T}(a; S_{i,T}) \tag{24}$$

   That is the maximum value the publisher can extract from this particular state variable in the last time period.

2. For any $t < T$, we can write:

$$\hat{V}_t(S_{i,t}) = \max_{a \in \mathcal{A}_i} \tilde{V}_{i,t}(a, S_{i,T}) \tag{25}$$

$$\hat{a}^d_t(S_{i,t}) = \operatorname*{argmax}_{a \in \mathcal{A}_i} \tilde{V}_{i,t}(a, S_{i,T}) \tag{26}$$

> Now, if we go backward and solve for the value functions, everything on the right-hand side of Equation (25) is known because we have already solved for the value functions in the next period. Therefore, we can find the value function for all the states in time period and continue this process until exposure number 1.

Once we have estimated the value function and best ad for all the states, we can simply characterize the adaptive forward-looking sequencing policy as follows:

$$\pi^d(a \mid S_{i,t}) = \begin{cases} 1 & a = a_t^d(S_{i,t}) \\ 0 & a \neq a_t^d(S_{i,t}) \end{cases} \tag{27}$$

We use backward induction as our main approach to determine the adaptive forward-looking sequencing policy. The only limitation of this approach is that we have to set an endpoint for the session, which does not allow us to fully exploit exposures with $t > T$. As a robustness check, we also use an infinite horizon with more memory efficient state variables, where only we consider state variables that do not require storing the entire session-level history.

6.3 Evaluation

One of our main goals in this paper is to examine to what extent the dynamic framework helps publishers achieve better user engagement with ads. As such, we need to use an evaluation metric that allows us to evaluate and compare different policies. For any exposure t, we denote a t-step trajectory by g_t and define it as the sequence of states, actions, and rewards in all the steps as follows:

$$g_t = \langle s_1, a_1, r_1^c, \ldots, s_{t-1}, a_{t-1}, r_{t-1}^c, s_t \rangle,$$

where any s_k is determined by the sequence prior to that exposure k and the distribution of transitions, a_k is determined given the policy, and r_k^c is the reward for the pair of s_k and a_k with the click-maximizing objective. Let τ denote the distribution of transitions and π denote any given policy. The joint distribution (τ, π) then determines the probability of each sequence g_t. Let $\rho_T(\pi; S_{i,1})$ denote the expected number of clicks generated in session i for the horizon length T, when using policy π. We can write:

$$\rho_T(\pi; S_{i,1}) = \mathbb{E}_{g_t \sim (\tau, \pi)} \left[\sum_{t=1}^{T} \beta^{t-1} r_t^c \right] \tag{28}$$

Now, we can use our estimates for the distribution of transitions τ and estimate $\rho_T(\pi; S_{i,1})$ for any policy π as follows:

$$\hat{\rho}_T(\pi; S_{i,1}) = \sum_{t=1}^{T} \sum_{g_t \in \mathcal{G}_T} \sum_{a \in \mathcal{A}_i} \pi(a \mid S_{i,t}) \hat{R}_t^c(a; S_{i,t}) P(g_t \mid \tau, \pi), \tag{29}$$

where $\mathcal{G}_T$ is the set of all possible trajectories and g_t denotes the trajectory in the first t periods. The last component in Equation (29) is the probability that a specific trajectory happens given the policy and distribution of transitions.

The main reason why we employ on this approach for evaluation is that it gives us session-level performance metrics on each policy. However, for robustness, we employ other approaches such as importance sampling and doubly robust method.

6.4 Results from the Click-Maximizing Policy

We now present the results on the performance of different sequencing policies when the main objective is to maximize the expected number of clicks per session. In §6.4.1, we illustrate session-level outcomes for different sequencing policies and examine the gains from adopting an adaptive forward-looking policy. In §6.4.2, we explore the heterogeneity in the gains from adopting an adaptive forward-looking sequencing policy across sessions. In §6.4.3, we use the distribution of counterfactual sequencing policies to explain the differences in them.

6.4.1 Gains from the Adaptive Forward-looking Sequencing Policy

We start by demonstrating the gains from the adaptive forward-looking sequencing policy compared to other policies described in §6.1. We use the direct method presented in §6.3 to evaluate the performance of these sequencing policies. As such, we only focus on the first six exposures and do not evaluate sessions after the sixth exposure. We draw a random sample of 1000 users from our test data that gives us 12,136 unique sessions. We estimate the optimal policy under all the sequencing policies defined above and present the results in Table 5.

As expected, all model-based sequencing policies lead to major improvements in the expected number of clicks per session over random allocation of ads. Adopting an adaptive forward-looking sequencing policy results in 0.1772 clicks per session which translates to an 80.36% improvement over the random sequencing, 7.87% over the non-adaptive single-ad sequencing, and 1.50% over the adaptive myopic sequencing. These gains demonstrate the value of each piece of information and modeling paradigm. Our results indicate that both session-level and future information play an important role in ad allocation problem.

Next, we focus on the gap between the adaptive forward-looking and non-adaptive single-ad

	Sequencing Policies			
	Fully Dynamic	*Adaptive Myopic*	*Single-Ad*	*Random*
Expected No. of Clicks	2150.60	2118.81	1993.75	1192.40
Expected No. of Clicks Per Session	0.1772	0.1745	0.1642	0.0983
Expected No. of Impressions	39365.30	39460.26.37	38925.97	38859.06
Expected Session Length	3.24	3.25	3.21	3.20
Expected CTR	5.46%	5.37%	5.12%	3.07%
% Click Increase over Random	80.36%	77.69%	67.21%	0.00%
No. of Users	1000	1000	1000	1000
No. of Sessions	12,136	12,136	12,136	12,136

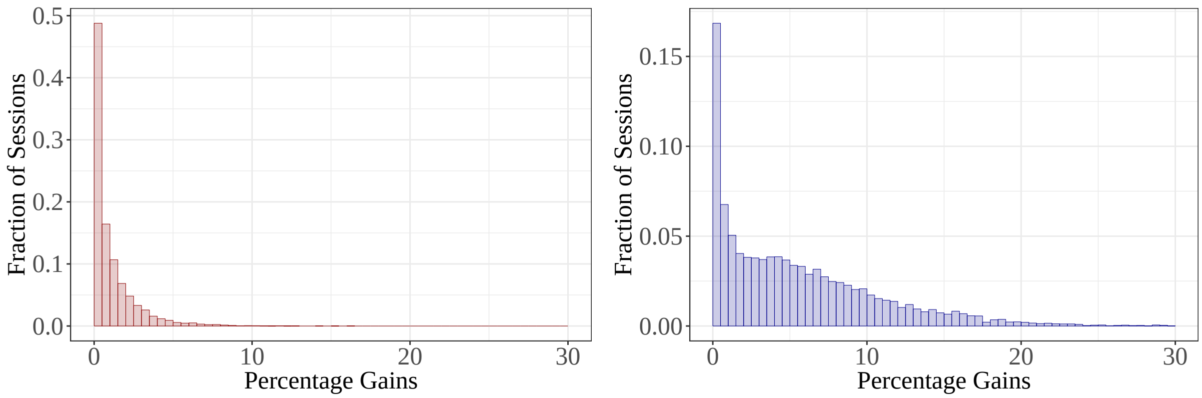
Table 5: Performance of different sequencing policies for a sequence size of 6 with a click-maximizing objective

policy. This comparison relates to a broader question on whether the publisher should use a short-lived ad slot. Our results suggest that the use of dynamic ad slot leads to considerable gains and value creation compared to the fixed ad slot, justifying the current trend of using short-lived ad slots in the industry.[10]

We then examine the performance of adaptive forward-looking sequencing compared to that of adaptive myopic sequencing policy. Our results suggest that there are gains from adopting a forward-looking objective in adaptive ad sequencing. It also provides evidence regarding the interdependence of ads shown within the sequence. It is worth noting that 1.50% improvement is a lower bound for what the publisher can achieve by adopting a forward-looking objective. This is because we only focus on the first six exposures for computational reasons. The sequencing effects likely grow in later exposures. Further, we only focus on a limited ad inventory, with 15 ads. We expect the gains from the adaptive forward-looking ad sequencing to improve as a result of expanding ad inventory.

One likely explanation for the gains from an adaptive forward-looking sequencing policy is that it is the only model that takes users' leave decision into account. As such, one channel through which this policy might improve the outcome is to increase the session length, thereby enhancing the total number of clicks. To see whether this is the channel, we estimate the expected session length for all the policies. We find no significant difference between our policies in terms of the session length. Rather, the adaptive forward-looking sequencing policy generates a higher CTR, which suggests that users are more engaged with ads as a result of adaptive forward-looking ad

[10]However, the caveat here is that our approach in single-ad policy is not the equivalent of the fixed ad slot, as there are very short interruptions in every one minute this ad is being shown, whereas the fixed ad slot keeps the ad immobile throughout the session.



(a) Forward-Looking over Myopic

(b) Forward-Looking over Single-Ad

Figure 12: Distribution of session-level gains from the adaptive forward-looking over both adaptive myopic and single-ad sequencing policies

sequencing.

6.4.2 Heterogeneity in Gains from the Dynamic Framework

While the results in Table 5 establish the average gains in clicks from adopting the adaptive forward-looking sequencing policy, they do not explain how these gains vary across sessions. In this section, we are interested in identifying the sessions for which the adaptive forward-looking sequencing is most helpful. As such, we focus on the session-level gains from the adaptive forward-looking compared to the adaptive myopic and non-adaptive single-ad policies and explore the heterogeneity across sessions. Let $Gain^f_m$ and $Gain^f_s$ denote the percentage gains from the adaptive forward-looking over adaptive myopic and single-ad sequencing respectively. In Figure 12, we show the distribution of these two session-level gains. Both figures show significant heterogeneity in the session-level gains from the adaptive forward-looking sequencing policy over adaptive myopic and single-ad policies.

To further explore this heterogeneity across user history, we regress both $Gain^f_m$ and $Gain^f_s$ on a set of historical variables, controlling for user fixed effects. We also control for the number of ads competing in a session as it is an important factor determining the gains from different sequencing policies. The results are presented in Table 6. The first two columns show the results using the gains from the adaptive forward-looking over the adaptive myopic sequencing policy as the dependent variable. In column 1, we examine how the relative gains from the adaptive forward-looking over adaptive myopic sequencing changes as a user participates in more sessions. On the one hand, as the user participates in more sessions, the publisher has more history and data on him, and it may improve the gains from adopting a forward-looking objective. On the other hand, as the user becomes more experienced in the platform, the relative effectiveness of sequencing and temporal interventions may shrink (Rafieian and Yoganarasimhan, 2018a). As shown in column 1, we find some evidence for the latter: the relative gains from the adaptive forward-looking over adaptive myopic sequencing reduces as the user becomes more experienced.

	Dependent variable			
	$Gain^d_m$	$Gain^d_m$	$Gain^d_s$	$Gain^d_s$
$Session\ Number_{i,u}$	-0.000032***	0.000028**	-0.000447***	-0.000276***
	(-4.51)	(-1.95)	(-17.29)	(-5.41)
$No.\ of\ Competitors_i$	0.000032	-0.000253***	0.004188***	0.003070***
	(0.27)	(-2.17)	(9.90)	(7.31)
$Imp_{i,u}$		0.000019***		0.000056***
		(11.79)		(9.66)
$Breadth_{i,u}$		-0.000401***		-0.001681***
		(-15.61)		(-18.13)
User FE	✓	✓	✓	✓
No. of Obs.	12,136	12,136	12,136	12,136
R^2	0.3021	0.3223	0.2864	0.3097
Adjusted R^2	0.2469	0.2686	0.2300	0.2550
Note:				*p<0.05; **p<0.01; ***p<0.001

Table 6: OLS estimates of session-level gains from the adaptive forward-looking sequencing policy across user history

In column 2, we include two historical variables to further explore which parts of user's prior experience are accountable for the findings in column 1 – total number of impressions ($Imp_{i,u}$), and the total number of distinct ads ($Breadth_{i,u}$) the user has seen prior to the session. With these controls, we find that the number of prior impressions has a positive correlation with the outcome, implying that having more data on a user benefits the adaptive forward-looking sequencing more than the myopic one. However, note that the number of distinct ads the user has seen that makes adaptive forward-looking sequencing less effective. This is probably because the sequencing effects mostly come from users' information processing and differentiation of ads in a sequence. Once the user has processed many different ads in prior sessions, it is less likely that the optimal sequencing of them largely affects her decision.

In columns 3 and 4, we show the results with another dependent variable – gains from the adaptive forward-looking over non-adaptive single-ad sequencing. Again, we find that users' tenure in the platform makes adaptive forward-looking sequencing less effective, and this is because users with more experience have seen more diverse ads, and it is harder to influence their decision by dynamic sequencing of ads. It is worth noting that the results in all columns must be interpreted in the relative terms. That is, a negative coefficient does not mean that adaptive myopic or non-adaptive single-ad policies perform better than the adaptive forward-looking sequencing policy. Rather, it identifies where adaptive forward-looking sequencing performs relatively worse.

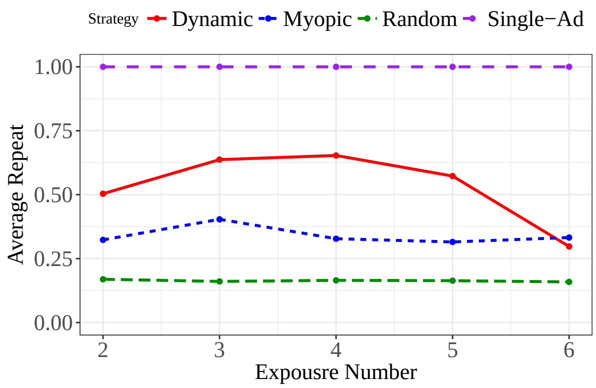
6.4.3 Distribution of Micro-Interventions

In the previous sections, we established the gains from the adaptive forward-looking sequencing policy over other sequencing policies and showed how these gains vary across different sessions based on user-level history. These gains mainly come from the active selection of session-level features through the selection of ads in the adaptive forward-looking sequencing policy. In this section, we first illustrate how these sequencing policies are different in micro-interventions at any point and how that results in different session-level features. We then aim to uncover the reasons behind such differences, given the goal of each sequencing policy.

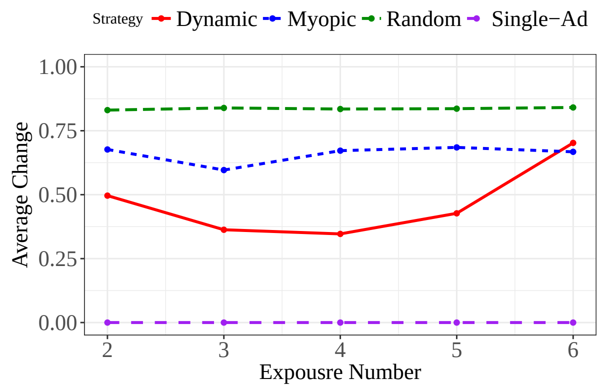
To make sense of how sequences evolve under each policy, we focus on binary micro-interventions as defined in §3.3.3. There are two key points about these micro-interventions. First, these binary micro-interventions together constitute our session-level features. For example, breadth-increasing change at any point results in a higher breadth of variety in a session, whereas repeat increases the number of prior exposures of an ad. Second, as shown in the prior research, all these three micro-interventions can be effective policies from a broader perspective. Recent experimental papers document the positive effects of multiple exposures on sales of an ad (Sahni, 2015a; Johnson et al., 2016b). While confirming the positive effects of multiple exposures of an ad within a session, Rafieian and Yoganarasimhan (2018a) show that both breadth-constant and breadth-increasing changes will increase the likelihood of clicking on the next ad. The fact that these micro-interventions are mutually exclusive makes the decision very challenging from a dynamic point-of-view, as using each comes with trade-offs. Thus, comparing distributions of these micro-interventions under different sequencing policies can help in understanding the decision-making process for each policy.

We present the average values for these micro-interventions at different time periods under each sequencing policy in Figure 13. Since these are binary variables, the y-axis reports the percentage of making any micro-intervention. As shown in these figures, both dynamic and adaptive myopic sequencing policies employ a mix of these micro-interventions. As such, the lines for them lie somewhere between the non-adaptive single-ad and random sequencing policies.

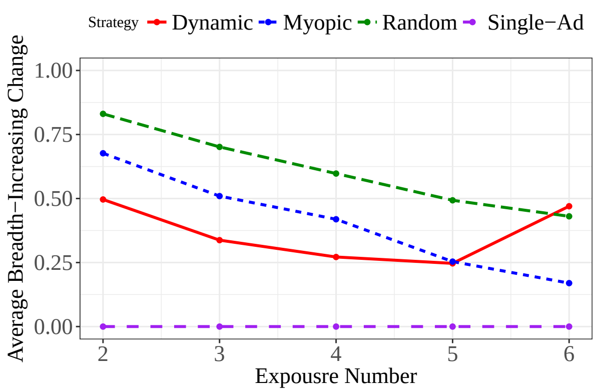
We further explore the difference between the adaptive forward-looking and adaptive myopic sequencing to reflect how these micro-interventions change if the publisher takes expected future rewards into account as in the adaptive forward-looking sequencing. We find that the adaptive



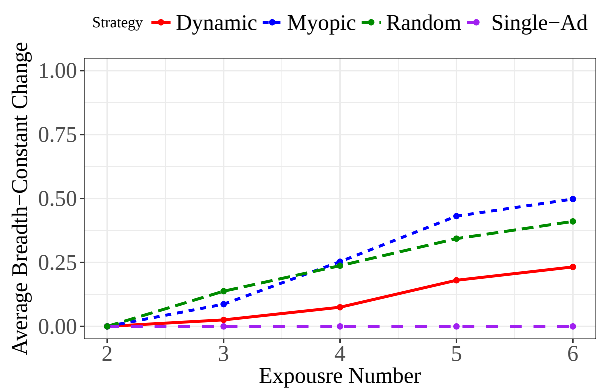
(a) Micro-intervention: repeat



(b) Micro-intervention: change



(c) Micro-intervention: breadth-increasing change



(d) Micro-intervention: breadth-constant change

Figure 13: Distribution of micro-interventions at different time periods under different sequencing policies

forward-looking sequencing policy tends to repeat ads in the earlier time periods, whereas change is a more popular decision in the adaptive myopic sequencing across all time periods. This is likely because repeating the same ad is a sub-optimal decision at the moment, but will help transition to a state where we can reinforce an ad with prior exposures. This is why the adaptive forward-looking sequencing policy takes advantage of repeating in the beginning, while in the last period, it tends to act similarly to the adaptive myopic sequencing.

Next, we look into the difference in the distribution of breadth-increasing and breadth-constant changes at different time periods. Since more prior exposures and changing the last ad helps at the moment, the adaptive myopic sequencing policy employs breadth-constant changes even more than the random sequencing. The adaptive forward-looking sequencing policy, however, seems to have a mix of both micro-interventions.

Overall, Figure 13 illustrates the differences in sequencing policies at a micro level. Our findings demonstrate the trade-offs between these micro-interventions by comparing different sequencing policies. It is generally hard to offer a unifying explanation for the optimal policy in a Markov Decision Process, as it uses an elaborate objective function. Thus, it is important to notice that these findings are descriptive and must be interpreted cautiously.

7 Implications

Our findings in this paper have several implications for marketing practitioners and policy makers. The most direct set of implications is for publishers and ad-networks. We examine the opportunity cost of using a fixed ad slot, which is the current practice in many digital ad platforms. We document considerable loss as a result of running a fixed ad slot, in terms of the value created by making a better match between ads and users. This finding has implications for publishers and ad-networks that want to design their ad format. However, these results must be interpreted with caution.

More importantly, we establish the gains from the adaptive forward-looking sequencing of ads compared to adaptive myopic sequencing. While it significantly adds to the computational complexity of the problem, our findings indicate that inter-temporal trade-offs in ad allocation problem play an important role in optimal policy design. As such, publishers can create value by dynamically sequencing ads.

While advertisers are not the main target of the implications in this paper, our findings offer them some new insights. Given the gains from the adaptive forward-looking sequencing of ads, it can be beneficial for advertisers to incorporate session-level information while targeting their ads. Further, our paper adds to the understanding of short-lived ad formats, which, in turn, can provide some insights for advertisers with regards to their banner design.

More broadly, our general framework can be extended to any context with adaptive interventions. For example, in the context of mobile health, a growing body of work focuses Just-In-Time Adaptive Interventions (JITAI) in mobile apps and studies the impact of them in shaping consumers' health behavior including physical fitness and activity, smoking, alcohol use, and mental illness (Nahum-Shani et al., 2017). Similarly, in any context that adaptive interventions can be used to educate people, we can specify a dynamic framework that helps us achieve better outcomes (Mandel et al., 2014). These showcases can serve as motivation for the public sector to use these tools in cases where collective action is required, such as environmental protection and political participation.

8 Conclusion

Mobile in-app advertising has grown exponentially over the last years. The ability to exploit the time-varying information about a user to personalize ad interventions over time is a key factor in the growth of in-app advertising. Despite the dynamic nature of the information, publishers often use myopic decision-making frameworks to select ads. In this paper, we examine whether a dynamic decision-making framework benefits the publisher in terms of the user engagement with ads, as measured by the number of clicks generated per session. Our dynamic framework has two main components: (1) a theoretical framework that specifies the domain structure such that it captures inter-temporal trade-offs in the decision to show what ad at any time period and (2) an empirical framework that breaks the policy design problem into a combination of machine learning tasks. We apply our framework to large-scale data from the leading in-app ad-network of an Asian country. Our results indicate that the adaptive forward-looking sequencing of ads results in significant gains in the expected number of clicks per session, compared to a set of benchmark policies. Next, we document heterogeneity in gains across sessions and show that adopting an adaptive forward-looking policy is most effective when users are new to the platform. Finally, we illustrate the differences in sequencing policies using a descriptive approach.

Our paper makes three contributions to the literature. First, from a methodological point-of-view, we develop a unified dynamic framework that starts with a theoretical framework that specifies the domain structure in mobile in-app advertising, and an empirical framework that breaks the problem into tasks that can be solved using a combination of machine learning methods and causal inference tools. Second, from a substantive standpoint, we are the first to document the gains from adopting an adaptive forward-looking sequencing policy as compared to the adaptive myopic sequencing policy. This comparison is of particular importance as the adaptive sequencing is the current approach in the industry. Third, we establish the gains from using a short-lived ad slot as compared to a fixed ad slot. The answer to this question informs the publisher's decision to use which type of ad slot.

Nevertheless, there are some limitations in our study that serve as excellent avenues for future research. First, our counterfactual policy evaluation is predicated on the assumption that users do not change their behavior in response to sequencing policies. While we exploit randomization to obtain our counterfactual estimate, it would be important to validate these findings in a field experiment. Further, we use the training data offline to learn counterfactual estimates for click and leave outcomes. Extension of our framework to an online setting that captures exploration/exploitation trade-off is important as offline evaluation may be costly. Finally, we use the entire within-session history to update state variables. Future research can look into more parsimonious frameworks that can be scalable to longer time horizons.

References

A. Ansari and C. F. Mela. E-customization. *Journal of marketing research*, 40(2):131–145, 2003.

N. Arnosti, M. Beck, and P. Milgrom. Adverse selection and auction design for internet display advertising. *American Economic Review*, 106(10):2852–66, 2016.

N. Arora and T. Henderson. Embedded premium promotion: Why it works and how to make it more effective. *Marketing Science*, 26(4):514–531, 2007.

R. Bellman. Dynamic programming. *Science*, 153(3731):34–37, 1966.

T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794. ACM, 2016.

J.-P. Dubé and S. Misra. Scalable price targeting. Technical report, National Bureau of Economic Research, 2017.

D. Dzyabura and H. Yoganarasimhan. Machine Learning and Marketing. In *Handbook of Marketing Analytics*. Edward Elgar Publishing, 2018.

eMarketer. Mobile In-App Ad Spending , 2018. URL `https://forecasts-na1.emarketer.com/584b26021403070290f93a5c/5851918a0626310a2c186a5e`.

eMarketer. Time Spent with Mobile, US , 2019a. URL `https://forecasts-na1.emarketer.com/584b26021403070290f93a96/5851918b0626310a2c186b39`.

eMarketer. Digital Ad Spending 2019 , 2019b. URL `https://content-na1.emarketer.com/us-digital-ad-spending-2019`.

J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.

S. Han, J. Jung, and D. Wetherall. A study of third-party tracking by mobile apps in the wild. *Univ. Washington, Tech. Rep. UW-CSE-12-03-01*, 2012.

J. R. Hauser, O. Toubia, T. Evgeniou, R. Befurt, and D. Dzyabura. Disjunctions of Conjunctions, Cognitive simplicity, and Consideration Sets. *Journal of Marketing Research*, 47(3):485–496, 2010.

G. J. Hitsch and S. Misra. Heterogeneous treatment effects and optimal targeting policy evaluation. *Available at SSRN 3111957*, 2018.

G. A. Johnson, R. A. Lewis, and E. I. Nubbemeyer. The online display ad effectiveness funnel & carryover: A meta-study of predicted ghost ad experiments. 2016a.

G. A. Johnson, R. A. Lewis, and D. Reiley. Location, location, location: repetition and proximity increase advertising effectiveness. 2016b.

W. Kar, V. Swaminathan, and P. Albuquerque. Selection and ordering of linear online video ads. In *Proceedings of the 9th ACM Conference on Recommender Systems*, RecSys '15, pages 203–210, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3692-5. doi: 10.1145/2792838.2800194. URL `http://doi.acm.org/10.1145/2792838.2800194`.

H. Li and P. Kannan. Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research*, 51(1):40–56, 2014.

L. Li, W. Chu, J. Langford, and R. E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 661–670. ACM, 2010.

T. Lu, D. Pál, and M. Pál. Contextual multi-armed bandits. In *Proceedings of the Thirteenth international conference on Artificial Intelligence and Statistics*, pages 485–492, 2010.

P. Manchanda, J.-P. Dubé, K. Y. Goh, and P. K. Chintagunta. The effect of banner advertising on internet purchasing. *Journal of Marketing Research*, 43(1):98–108, 2006.

T. Mandel, Y.-E. Liu, S. Levine, E. Brunskill, and Z. Popovic. Offline policy evaluation across representations with applications to educational games. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 1077–1084. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

S. Mullainathan and J. Spiess. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106, 2017.

I. Nahum-Shani, S. N. Smith, B. J. Spring, L. M. Collins, K. Witkiewitz, A. Tewari, and S. A. Murphy. Just-in-time adaptive interventions (jitais) in mobile health: key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, 52(6):446–462, 2017.

O. Rafieian and H. Yoganarasimhan. How does variety of previous ads influence consumer's ad response? 2018a.

O. Rafieian and H. Yoganarasimhan. Targeting and privacy in mobile advertising, 2018b.

O. J. Rutz and R. E. Bucklin. From generic to branded: A model of spillover in paid search advertising. *Journal of Marketing Research*, 48(1):87–102, 2011.

N. S. Sahni. Effect of temporal spacing between advertising exposures: evidence from online field experiments. *Quantitative Marketing and Economics*, 13(3):203–247, 2015a.

N. S. Sahni. Effect of temporal spacing between advertising exposures: Evidence from online field experiments. *Quantitative Marketing and Economics*, 13(3):203–247, 2015b.

N. S. Sahni. Advertising spillovers: Evidence from online field experiments and implications for returns on advertising. *Journal of Marketing Research*, 53(4):459–478, 2016.

E. M. Schwartz, E. T. Bradlow, and P. S. Fader. Customer acquisition via display advertising using multi-armed bandit experiments. *Marketing Science*, 36(4):500–522, 2017.

E. H. Simpson. Measurement of Diversity. *Nature*, 1949.

Z. Sun, M. Dawande, G. Janakiraman, and V. Mookerjee. Not just a fad: Optimal sequencing in mobile in-app advertising. *Information Systems Research*, 28(3):511–528, 2017.

G. Theocharous, P. S. Thomas, and M. Ghavamzadeh. Personalized ad recommendation systems for life-time value optimization with guarantees. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

O. Toubia, T. Evgeniou, and J. Hauser. Optimization-based and Machine-learning Methods for Conjoint Analysis: Estimation and Question Design. In A. Gustafsso, A. Herrmann, and F. Huber, editors, *Conjoint Measurement: Methods and Applications*, pages 231–258. Springer, 4 edition, 2007.

G. L. Urban, G. Liberali, E. MacDonald, R. Bordley, and J. R. Hauser. Morphing banner advertising. *Marketing Science*, 33(1):27–46, 2013.

J. Yi, Y. Chen, J. Li, S. Sett, and T. W. Yan. Predictive model performance: Offline and online evaluations. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1294–1302. ACM, 2013.

H. Yoganarasimhan. Search personalization using machine learning. *Available at SSRN 2590020*, 2018.

Revenue-Optimal Dynamic Auctions for Adaptive Ad Sequencing

Omid Rafieian*
University of Washington

For Most Updated Version: Click Here

June 30, 2019

*I am grateful to my advisor Hema Yoganarasimhan, and my committee members Arvind Krishnamurthy, Amin Sayedi, Simha Mummalaneni, and Jacques Lawarrée for their guidance and comments. I am also grateful to an anonymous firm for providing the data and to the UW-Foster High Performance Computing Lab for providing me with computing resources. Please address all correspondence to: rafieian@uw.edu.

Abstract

Digital publishers often use real-time auctions to allocate their advertising inventory. These auctions are designed with the assumption that advertising exposures within a user's browsing or app-usage session are independent. Rafieian (2019) empirically documents the interdependence in the sequence of ads in mobile in-app advertising, and shows that dynamic sequencing of ads can improve the match between users and ads. In this paper, we examine the revenue gains from adopting a revenue-optimal dynamic auction to sequence ads. We propose a unified framework with two components – (1) a theoretical framework to derive the revenue-optimal dynamic auction that captures both advertisers' strategic bidding and users' ad response and app usage, and (2) an empirical framework that involves the structural estimation of advertisers' click valuations as well as personalized estimation of users' behavior using machine learning techniques. We apply our framework to large-scale data from the leading in-app ad-network of an Asian country. We document significant revenue gains from using the revenue-optimal dynamic auction compared to the revenue-optimal static auction. These gains stem from the improvement in the match between users and ads in the dynamic auction. The revenue-optimal dynamic auction also improves all key market outcomes, such as the total surplus, average advertisers' surplus, and market concentration.

Keywords: online advertising, dynamic mechanism design, ad sequencing, structural models, optimal auctions, reinforcement learning

1 Introduction

Mobile in-app advertising is now a significant source of revenue for publishers and ad networks. In 2018, over 56% of the total digital ad spend came from in-app advertising (eMarketer, 2018). Like other digital advertising environments, mobile publishers use an auction to determine which ad to show inside an app. An auction is a set of rules that characterizes how to allocate each advertising space and how much each advertiser has to pay, given advertisers' bids. Market outcomes under each auction format can accordingly be different since advertisers can strategically vary their bidding behavior. Thus, auction design plays a central role in the success of the digital advertising ecosystem.

Publishers and ad-networks often use auctions that maximize their revenues.[1] The common practice in this industry is to use a first- or second-price auction with an optimally set reserve price. This is in light of the findings from the seminal paper by Myerson (1981) that has shown these auctions are revenue-optimal for a single item, under the regularity assumption. In an advertising environment, these auctions are revenue-optimal if the publisher can treat each advertising space as a single item: i.e., if advertising spaces are independent, and the auction outcome for one advertising space does not create externalities affecting other advertising spaces.

In the context of mobile in-app advertising, Rafieian (2019) provides empirical evidence on the interdependence of exposures within a session where a user is exposed to multiple short-lived ads. He shows that each ad exposure creates externalities that affect future exposures and documents the publisher's gain from dynamic sequencing of ads, i.e., the policy that captures both the immediate and future outcomes in a session and selects the ad that maximizes the expected number of clicks from that point onward. These findings rule out the independence of advertising spaces within a session, which in turn, imply that first- or second-price auction with an optimally set reserve price is not revenue-optimal in the context of mobile in-app advertising.

While the results in Rafieian (2019) elucidate an opportunity to create value in this market by dynamic sequencing of ads by enhancing consumer engagement and match values, the extent to which the publisher can extract this value as revenue is not clear. Notice that advertisers are strategic agents who can change their bids in response to any change in the allocation mechanism, and thereby appropriate most of this created value. The prior empirical literature on advertising auctions has highlighted cases where publishers cannot necessarily link the improvement in the match to higher revenues in a competitive environment (Athey and Nekipelov, 2010; Rafieian and Yoganarasimhan, 2018b). Thus, when the revenue is the primary outcome of interest, it is crucial

[1] In this paper, we use the publisher, ad-network, and platform interchangeably, when we refer to the agent who designs the auction and makes the ad allocation decision.

for the publisher to incorporate advertisers' bidding behavior as well as users' ad response when designing the ad sequencing policy.

This brings us to the question of optimal auction design, wherein the publisher designs an auction that maximizes her revenues. Our main goal is to theoretically develop the revenue-optimal dynamic auction and compare its outcomes with the revenue-optimal static auction (second-price auction with optimal reserve price). Overall, we aim to answer the following three research questions in this paper:

1. How can we design a revenue-optimal dynamic auction that captures both inter-temporal trade-offs in ad sequencing and advertisers' strategic bidding behavior?

2. How can we build an empirical framework to evaluate the market outcomes such as publisher's revenues and advertisers' surplus under any auction mechanism?

3. What are the gains from using a revenue-optimal dynamic auction as compared to the static one in mobile in-app advertising? How is the advertisers' surplus distributed across advertisers? Do all advertisers benefit when the publisher uses a dynamic auction?

We need to overcome three major challenges to answer these questions. First, to design a revenue-optimal dynamic auction, we need to specify an allocation rule that incorporates inter-temporal trade-offs in ad interventions and a payment rule that governs advertisers' strategic bidding behavior. Second, to empirically evaluate market outcomes (e.g., publisher's revenues) under counterfactual auctions, we need to obtain accurate estimates of both advertisers' and users' behavior that are valid any counterfactual auction. For the former, we need to estimate the distribution of advertisers' click valuations as it is the main structural parameter that governs their bidding behavior in any auction. For the latter, we need to estimate users' behavior: their likelihood of clicking on an ad and leaving the session after seeing an ad under any counterfactual allocation policy. Finally, to measure the gains from both dynamic and static revenue-optimal auctions, we need to first solve for the equilibrium outcome under these counterfactual auctions and then use an evaluation method that estimates the corresponding outcomes for each session.

We present an overview of our approach in Figure 1. This figure illustrates the general framework in the top row and the details specific to our problem in the bottom row. In the general framework, we begin with a theoretical framework that informs our empirical approach regarding how to develop and evaluate optimal auctions. More specifically, we start with designing a revenue-optimal dynamic auction that gives us a combination of allocation and payment rules. This combination captures both the inter-temporal trade-offs and advertisers' bidding strategies. Since our theoretical framework

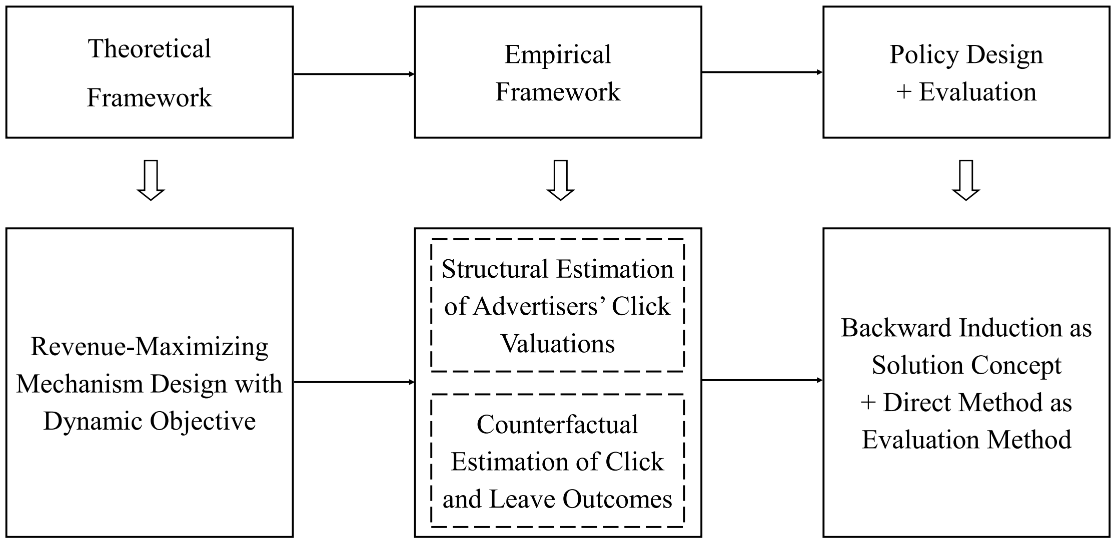


Figure 1: An overview of our approach. The top row presents our general framework and the bottom row shows the specific approach we take in this paper.

involves both advertisers' and users' behavior, our empirical framework requires an estimation procedure that cover both these components. We accordingly break our empirical framework into two separate tasks: (1) estimation of the distribution of advertisers' click valuations since click valuation is the key structural parameter governing advertisers' bidding behavior in any auction, and (2) personalized counterfactual estimation of click and leave outcomes, which are the two user-dependent outcomes that affect the expected revenue per session. The second task is the same as the empirical task in Rafieian (2019). We use the same approach in this paper to estimate users' behavior under counterfactual auctions. Finally, we use all our estimates and numerically derive the optimal policy using backward induction and evaluate this policy using the direct method.

Our theoretical framework directly addresses the first challenge and paves the way to address other challenges. We build our theoretical model on the recent literature on dynamic mechanism design that extends the approach in Myerson (1981) to a dynamic setting. The intuitive idea in this literature is to use the Revelation Principle (Myerson, 1981) and exclusively focus on the case where all bidders report their type truthfully (Kakade et al., 2013; Pavan et al., 2014). We use the modeling framework in Kakade et al. (2013) as their separability assumption is particularly suitable in our context: the value an advertiser extracts from an impression is the product of his private click valuation and the expected probability of click on his ad. It allows us to write down the reward function in the Markov Decision Process (MDP henceforth) in terms of virtual valuations and click probabilities (match valuations), thereby deriving the optimal allocation. Using this allocation function, we can then set the payments such that advertisers will participate in the auction and have no incentive to deviate from truthful reporting (IR and IC constraints). We then show that the auction with these allocation and payment rules is revenue-optimal.

To address our second challenge, we propose a structural framework to estimate the distribution of advertisers' click valuation from their observed bids in the data. The key challenge is that the auction format in the data is a quasi-proportional auction, where truthful bidding is not the equilibrium strategy for advertisers. We first characterize the advertisers' utility function in this setting and then derive the equilibrium properties of this auction. Using the first-order condition, we then write the advertisers' click valuation in terms of their cost and the allocation function used by the ad-network. Since both cost and allocation functions can be estimated from the distribution of observed bids and auction configurations, click valuations are identified under the assumption that advertisers are utility-maximizing. This allows us to estimate the distribution of advertisers' click valuations as well as each advertiser's click valuation.

Next, to develop the ad sequencing policy in the revenue-optimal dynamic auction, we need to solve the MDP for the allocation function. While the transition function is the same as the one in Rafieian (2019), the rewards have an additional multiplicative factor – each advertiser's virtual valuation. We can estimate it using the estimated click valuation for each ad as well as the distribution of click valuations. We plug these estimates into the reward function and solve the dynamic allocation policy using a backward induction solution concept. Finally, like Rafieian (2019), we use a direct method approach for evaluation that directly uses our estimates to simulate a session, equilibrium outcomes, and how it evolves. This method allows us to evaluate the revenue outcome for each session.

We first present the results from our auction estimation framework. We theoretically show that advertisers bid roughly half of their click valuations in the quasi-proportional auction. As such, the distribution of bids alone can approximate the distribution of click valuations. Further, it suggests that the current mechanism (quasi-proportional auction) leads to a substantial loss for the platform in terms of both revenue and efficiency. We then focus on the estimated distribution of click valuations from our structural framework, and empirically show that the regularity assumption is satisfied in our context: the virtual valuations are strictly increasing in click valuations. This is an important requirement for our counterfactual analysis, as the solution to the optimal auction is tractable given this assumption.

Next, we conduct our counterfactual analysis to examine the gains from the revenue-optimal dynamic auction. We set the benchmark as the second-price auction with an optimal reserve price, as it is the revenue-optimal static auction. Our results indicate that the expected revenue per session is 1.60% higher under the revenue-optimal dynamic auction compared to that in the revenue-optimal static auction. This is particularly important because most platforms currently use a version of the static revenue-optimal auction. Thus, our results suggest that publishers and ad-networks can significantly benefit from adopting an optimal dynamic auction. Further, we find that the expected number of clicks per session also improve by 1.80% under the dynamic case, suggesting that the gains in revenues can mostly be attributed to the improved match between users and ads as a result of dynamic sequencing, and not to the greater ability of the publisher to extract rent from advertisers.

We then focus on other market outcomes and show that the optimal dynamic auction achieves better outcomes than the optimal static auction in terms of both total surplus and average advertisers' surplus: the total surplus (efficiency) and the average advertisers' surplus increases by 1.77% and 3.00% under the optimal dynamic auction respectively. Hence, we show that the optimal dynamic auction does not achieve revenue optimality at the expense of efficiency. We then explore the surplus gains across advertisers to see whether the market will become more concentrated as a result of using the revenue-optimal dynamic auction. Using a Herfindahl-Hirschman Index (HHI), we find that the optimal dynamic auction has a lower concentration index than the optimal static auction. This suggests that the reason is that the optimal dynamic auction allocates more to the ad with the second largest surplus, thereby closing the gap between the top two advertisers in an auction.

In sum, our paper makes three key contributions to the literature. First, from a methodological point-of-view, we propose a unified dynamic framework that captures both advertisers' and users' behavior to optimize publisher's revenue. A key contribution of our framework is in illustrating how we can use a theoretical framework to break a complex applied problem into a composite of structural estimation and machine learning tasks. To our knowledge, this is the first paper to empirically examine the revenue gains from dynamic sequencing of ads using an optimal dynamic auction. Second, we present a structural estimation framework to recover the distribution of bidders' private valuations from their observed bidding behavior in a quasi-proportional auction. This is the first paper to propose an estimation procedure for quasi-proportional auctions. Our framework can easily be extended to auctions with non-deterministic allocation rules. Third, from a substantive viewpoint, we establish the revenue gains from adopting a dynamic objective in allocating ads, as opposed to a static objective. This is of particular importance, as the current practice in the industry is to use a static objective. We expect our findings to be of relevance to publishers and ad-networks.

2 Related Literature

First, our paper relates to the growing literature on dynamic mechanism design. While early papers in this literature start in 1980s (Baron and Besanko, 1984; Myerson, 1986; Riordan and Sappington, 1987), most of the major developments in this literature appear more recently, with generic characterizations of both efficient mechanisms (Bergemann and Välimäki, 2010; Athey and Segal, 2013) and revenue-maximizing (optimal) mechanisms (Kakade et al., 2013; Pavan et al., 2014). The majority of applied papers on dynamic mechanism design focus on cases where the inter-temporal trade-offs arise through the dynamics of arrival, departure, or population (Vulcano et al., 2002; Parkes and Singh, 2004; Gallien, 2006; Said, 2012). Our paper adds to this literature by empirically evaluating dynamic mechanisms in a digital advertising context. To our knowledge, this is the first paper to provide an empirical framework to examine the performance of dynamic mechanisms.

Second, our paper relates to the literature on the intersection of mechanism design and online advertising. Early papers in this area examines the theoretical properties of different auctions in sponsored search context (Edelman et al., 2007; Varian, 2007; Lahaie et al., 2007). More specific to our context, a series of work takes externalities in search advertising into account and revisits the question of mechanism design in a context where the higher position of an ad may affect the user's decision to even see lower ranked ads (Ghosh and Mahdian, 2008; Kempe and Mahdian, 2008; Ghosh and Sayedi, 2010). In the context of video ads, Kar et al. (2015) adopt a cascade model similar to Kempe and Mahdian (2008), and provide a mechanism for selection and ordering of video ads. While this stream of work proposes simple mechanisms for allocation, they are only applicable to very basic and unrealistic case where the externality is only imposed through the user's leaving decision. We extend this literature by offering a dynamic framework that captures more complex externalities, under a plausible separability assumption.

Lastly, our paper relates to economics and marketing literature on the estimation of auctions. A significant breakthrough in this literature comes from Guerre et al. (2000) who base their identification strategy on the fact that the equilibrium outcome is achieved when all agents maximize their profits given the distribution of others' behavior. While they study the first-price auction with symmetric independent private valuations and without unobserved heterogeneity, other papers in this literature build on this work and extend it to the cases with affiliated private valuations (Li et al., 2002), asymmetric private valuations (Campo et al., 2003), unobserved heterogeneity (Guerre et al., 2009; Krasnokutskaya, 2011), and also different auctions such as scoring auctions (Bajari et al., 2014) and beauty contest auctions (Yoganarasimhan, 2015). Related to our setting, a few papers study online advertising auctions and propose different empirical approaches for the estimation fi advertising auctions (Athey and Nekipelov, 2010; Yao and Mela, 2011; Choi and Mela, 2016). Our paper adds to this literature by proposing an estimation approach for quasi-proportional auctions that can easily be extended to any randomization-based auction. Further, our counterfactual analysis is the first to consider a dynamic mechanism design in an online advertising context.

3 Optimal Auctions

The main results in Rafieian (2019) establish the gains from the dynamic sequencing of ads in terms of the expected number of clicks generated per session. Intuitively, we expect the increase in the number of clicks to be linked to higher publisher revenues, as clicks are the key revenue-generating source for publishers. However, the fact that advertisers can respond to the change in the allocation by changing their bids in an auction environment makes it unclear how the publisher can extract more revenues from dynamic sequencing of ads. For example, if advertisers decrease their bid as a response to better allocation by dynamic sequencing in the equilibrium, the publisher may end up selling more clicks at a lower price. Thus, it is crucial to take advertisers' bidding behavior into account if the publisher wants to maximize her revenues through adaptive ad sequencing.

To examine revenue gains from adopting a dynamic framework, we incorporate advertisers' utility model into our framework and design revenue-optimal auctions with both dynamic and static objectives. Our framework, in turn, captures both users' behavior and advertisers' strategic bidding. The publisher's problem is then to design an auction with certain allocation and payment rules that maximizes her revenues, i.e., the total payments made by advertisers. To find the revenue-optimal auctions under each objective, we build on the seminal paper by Myerson (1981) and design the allocation and payment rules in the auction to maximize publisher revenues.

This section proceeds as follows: We first define our model environment and assumptions in §3.1. We then introduce the case where the publisher's objective is static and discuss the optimal auction in this case in §3.2. In §3.3, we focus on the optimal auction with a dynamic objective and present the allocation and payments rules in this case.

3.1 Auction Environment and Assumptions

We now describe the auction environment in this problem. For each session i, there are $\mathcal{A}_i$ risk-neutral bidders competing for impressions in this session. Each advertiser a has a private click valuation x_a which is drawn from the distribution F_a with support $[\underline{x}_a, \bar{x}_a]$. This parameter is a private signal that reflects how much the advertiser values a click. We assume that each ad's private valuation is independent of other ads' private valuation and does not vary across impressions.[2] The total value generated from showing ad a at exposure t of session i will then be the product of that ad's click valuation and the probability of click on that ad, which can be written as follows:

$$w_{i,a,t}(x_a; S_{i,t}) = x_a P(Y_{i,t} \mid a, S_{i,t}), \tag{1}$$

[2] We can extend our theoretical analysis to the cases where click valuations vary across sessions and exposures. However, in our empirical analysis, we can only identify one value for each ad. This is why we restrict our attention to this case.

where $w_{i,a,t}(x_a; S_{i,t})$ is the total value ad a receives from being shown at exposure t of session i, and $S_{i,t}$ is the state variable at that point which captures the prior history of actions and outcomes within the session, as well as the pre-session information (Please see §4.2.1 in Rafieian (2019) for more details on state variables).

The publisher's problem is to design a mechanism/auction that maximizes her revenues. Each mechanism is characterized by two components – (1) allocation rule, and (2) payment rule. We can define any mechanism as follows:

Definition 1. *A mechanism $M(q, e)$ is defined as a combination of an allocation rule $q(\cdot)$ and a payment rule $e(\cdot)$. Given a profile of reported bids $b = (b_1, b_2, \ldots, b_{\mathcal{A}})$, we can characterize the allocation and payment rules for each ad in the exposure number t in session i as follows:*

- *The allocation rule $q^M_{i,a,t}(b; S_{i,t})$ determines the probability that the item is allocated to a.*
- *The payment rule $e^M_{i,a,t}(b; S_{i,t})$ determines what ad a pays in expectation.*

We assume that the publisher has full commitment power, i.e., the players believe that the publisher follows the rules. Now, under any mechanism $M(q, e)$, we can characterize advertiser's utility function. The only decision variable for any advertiser is to submit a bid that reflects their willingness to pay for a click on their ad. Given advertiser a's bid, we can characterize their utility in exposure t of session i as follows:

$$u^M_{i,a,t}(b_a; x_a, S_{i,t}) = \mathbb{E}_{b_{-a}}\left[w_{i,a,t}(x_a; S_{i,t}) q^M_{i,a,t}(b; S_{i,t}) - e^M_{i,a,t}(b; S_{i,t})\right], \tag{2}$$

where b_{-a} is a bid profile of all ads except a, and b is the profile of all bids. We assume that advertisers maximize their utility.

3.2 Warm-Up Case: Optimal Static Auction

We begin by describing the case where the publisher's objective is static. In this case, the publisher only considers the current period rewards. As such, at any point, the goal is to sell the slot to the ad that maximizes the publisher's revenues. The analysis of this case is almost identical to that of the seminal paper by Myerson (1981) on optimal auctions. However, we present this case as a warm-up example for our main goal – deriving the optimal dynamic auction.

In general, the choice of the optimal auction may seem impossible as there is no bound on the set of feasible auctions. However, in light of the Revelation Principle and without loss of generality, we can only focus on direct-revelation mechanisms wherein advertisers truthfully bid their click valuations (Myerson, 1981). A direct revelation mechanism is feasible if it satisfies:

1. *Plausibility:* For any profile of reported click valuations x, we have $\sum_{a \in \mathcal{A}_i} q_{i,a,t}(x; S_{i,t}) \leq 1$

and $q_{i,a,t}(x; S_{i,t}) \geq 1$ for all $a \in \mathcal{A}_i$. This condition guarantees that each impression is allocated to at most one ad.

2. *Individual Rationality (IR):* Given reported click valuations, all advertisers receive non-negative utility, i.e., $u^M_{i,a,t}(x_a; x_a, S_{i,t}) \geq 0$ for all $a \in \mathcal{A}_i$. This condition guarantees that all competing ads for a session will participate in the auction.

3. *Incentive Compatibility (IC):* No advertiser has incentive to deviate from bidding truthfully, given that everyone else bids truthfully. Therefore, we have:

$$u^M_{i,a,t}(x_a; x_a, S_{i,t}) \geq u^M_{i,a,t}(b_a; x_a, S_{i,t}) \tag{3}$$

for any $b_a \in [\underline{x}_a, \bar{x}_a]$. This condition guarantees that reporting truthfully is a Bayesian Nash Equilibrium for all advertisers.

The Revelation Principle helps us reduce the set of all auction to the set of feasible direct revelation mechanism denoted by $\mathcal{M}_{\text{direct}}$. As such, our search is over a more structured set with clear constraints. We can write the publisher's optimization problem as follows:

$$\max_{M \in \mathcal{M}_{\text{direct}}} \mathbb{E}_x \left[\sum_{a \in \mathcal{A}_i} e^M_{i,a,t}(x; S_{i,t}) \right] \tag{4}$$

In this optimization, $M \in \mathcal{M}_{\text{direct}}$ implies that the mechanism must satisfy all three constraints presented above. While focusing only on feasible direct revelation mechanisms helps constrain the problem, we still need further transformations to find the optimal solution. One key transformation is to use envelope condition instead of the IC constraint. The following lemma shows the link between these two:

Lemma 1. *If the mechanism $M(q, e)$ is IC, we have:*

$$\begin{aligned} u^M_{i,a,t}(x_a; x_a, S_{i,t}) = \ & u^M_{i,a,t}(x'_a; x'_a, S_{i,t}) \\ & + P(Y_{i,t} \mid a, S_{i,t}) \int_{x'_a}^{x_a} \mathbb{E}_{x_{-a}} \left[q^M_{i,a,t}(b_a, x_{-a}; S_{i,t}) \right] db_a \end{aligned} \tag{5}$$

Now, we can use this lemma to derive the publisher's expected revenue under any IC mechanism as follows:

Lemma 2. *If the mechanism $M(q,e)$ is IC, the publisher's expected revenue can be written as follows:*

$$\mathbb{E}_x\left[\sum_{a\in\mathcal{A}_i}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x;S_{i,t})\right]-\sum_{a\in\mathcal{A}_i}u^M_{i,a,t}(\underline{x}_a;\underline{x}_a,S_{i,t}) \tag{6}$$

The result of Lemma 2 is transformative in finding the optimal auction, as for any given environment, the expected revenue of an auction only depends on the allocation at the equilibrium and advertisers' expected utility of their lowest click valuation. We can now use Equation (6) as the objective function and maximize it subject to the *Plausibility* and *Individual Rationality (IR)* constraints and obtain the optimal auction.

Another important feature of writing the objective function in Equation (6) is that it additively separates allocation and payment functions: the first component is independent of the payment function. Thus, one candidate for the optimal auction is to find a plausible allocation function that maximizes the first component and a payment function that minimizes the second component. This brings us to the following proposition:

Proposition 1. *The mechanism $M(q,e)$ is optimal if q maximizes*

$$\mathbb{E}_x\left[\sum_{a\in\mathcal{A}_i}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x;S_{i,t})\right] \tag{7}$$

subject to q being plausible and $\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]$ increasing in x_a, and the payment function e is

$$e^M_{i,a,t}(x;S_{i,t})=w_{i,a,t}(x_a;S_{i,t})q^M_{i,a,t}(x;S_{i,t})-P(Y_{i,t}\mid a,S_{i,t})\int_{\underline{x}_a}^{x_a}q^M_{i,a,t}(b_a,x_{-a};S_{i,t})db_a \tag{8}$$

As shown in Proposition 1, we can solve for the optimal allocation and payment functions. In next sections, we discuss the details of each component.

3.2.1 Allocation Rule

Proposition 1 offers a constrained optimization to find the allocation function under the static case. Without the constraint on $\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]$ being increasing in x_a, we can simply maximize the objective function by allocating to the ad with the highest $\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}\right)P(Y_{i,t}\mid a,S_{i,t})$. However, there is no guarantee whether the constraint is satisfied unless we impose the following assumption on the distribution F_a for all ads:

Assumption 1. *Distribution F_a is regular, i.e., the function $c_a(x_a) = x_a - \frac{1-F_a(x_a)}{f_a(x_a)}$ is strictly increasing in x_a.*

It is worth noting that this is not an unrealistic assumption, since most familiar distributions satisfy this condition. Under this assumption, the resulting allocation rule will allocate the item to ad with the highest non-negative $c_a(x_a)P(Y_{i,t} \mid a, S_{i,t})$ for any exposure t in session i. If all values are negative, the publisher does not sell the item. It is easy to show that under this allocation, the constraint on $\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x_a, x_{-a}; S_{i,t})\right]$ being increasing in x_a is satisfied: an increase in x_a will increase the expected probability of winning the item for ad a, since $c_a(x_a)$ is increasing in x_a.

3.2.2 Payments

Following Equation (8) in the second part of Proposition 1, we can now determine the optimal payment functions, consistent with the allocation function presented in §3.2.1. We can show that in this case, losing ads will not pay anything: the first component in Equation (8) is zero, and it is easy to show that the integral is also zero. The payment for the winning ad, however, can be calculated as follows:

$$
\begin{aligned}
e^M_{i,a,t}(x; S_{i,t}) &= w_{i,a,t}(x_a; S_{i,t})q^M_{i,a,t}(x; S_{i,t}) - P(Y_{i,t} \mid a, S_{i,t}) \int_{\underline{x}_a}^{x_a} q^M_{i,a,t}(b_a, x_{-a}; S_{i,t}) db_a \\
&= w_{i,a,t}(x_a; S_{i,t}) - P(Y_{i,t} \mid a, S_{i,t}) \int_{x^*_a}^{x_a} db_a \\
&= x^*_a P(Y_{i,t} \mid a, S_{i,t}),
\end{aligned} \tag{9}
$$

where x^*_a is the minimum bid that still wins the impression for the winning ad. The simple allocation rule in the static case helps us find the analytical solution to the integral in Equation (8). With this payment rule, it is easy to check the incentive compatibility of the proposed optimal mechanism.

Finally, if we consider the case of symmetric click valuations, we can simplify the optimal auction to a greater extent. In this case, instead of having ad-specific distributions F_a for each ad, we have one distribution F from which all click valuations are drawn independently. We can show the following corollary for this specific case:

Corollary 1. *With symmetric independent click valuations, the optimal auction is a second-price auction with a reserve price of $c^{-1}(0)$. The auction allocates the item to the ad with the highest $w_{i,a,t}(x_a; S_{i,t})$ and that ad pays the second-highest $w_{i,a,t}(x_a; S_{i,t})$ in expectation.*

Thus, the optimal auction either allocates the impression to the highest valuation advertiser or does not allocate it at all, meaning that it never allocates an impression to an advertiser who does not have the highest valuation for that impression.

3.3 Optimal Dynamic Auction

We now focus on the optimal auction with the dynamic objective. In this case, the publisher wants to incorporate the expected future revenues as well as expected revenues from the current period. This is in contrast with optimal static auction where the publisher only cares about her expected revenues in the current period. The publisher's goal in this dynamic environment is to design a mechanism $M(q, e)$ that maximizes her expected revenues from the session, i.e., $\mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1} \left(\sum_{a \in \mathcal{A}_i} e^M_{i,a,t}(b; S_{i,t})\right)\right]$.

This change in the publisher's objective clearly changes critical aspects of our environment. The most important change is that the publisher no more sells each exposure t in session i, but rather auctions off the entire session i. As such, the optimal auction in the static case is no more optimal under the current objective. An important factor that helps us simplify the dynamic problem is that only one piece is private to each advertiser in the entire session: their click valuation x_a. That is, advertisers' click valuation will not change within the session, and anything that changes their overall valuation of each impression (e.g., the probability of click in different states) is common knowledge. Thus, we can treat this problem as a case of static auction where advertisers just submit only one bid and the publisher decides how to allocate the exposures within the session, using a dynamic objective.

In line with the change in the publisher's objective, we must re-define advertisers' utility function for the session as follows:

$$U^M_{i,a}(b_a; x_a, S_i) = \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1} \left(w_{i,a,t}(x_a; S_{i,t}) q^M_{i,a,t}(b; S_{i,t}) - e^M_{i,a,t}(b; S_{i,t})\right)\right], \tag{10}$$

where the expectation is taken over bidding strategies and the mechanism, and S_i denotes the pre-session information.

Similar to the static case, without loss of generality, we only focus on dynamic direct revelation mechanisms to find the optimal auction (Myerson, 1986). As such, for the dynamic case, we re-write the requirements for a feasible direct revelation mechanism as follows:

1. *Plausibility:* This condition states that at each time period, the exposure is allocated to at most one ad, given that advertisers report their click valuations truthfully. That is, we have $\sum_{a \in \mathcal{A}} q_{i,a,t}(x; S_{i,t}) \leq 1$ and $q_{i,a,t}(x; S_{i,t}) \geq 1$ for all $a \in \mathcal{A}$ and for any t in session i.

2. *Individual Rationality (IR):* Given truthful reporting of click valuations, the advertiser's expected utility over the session is non-negative, i.e., $U^M_{i,a}(x_a; x_a, S_i) \geq 0$ for all $a \in \mathcal{A}_i$.

3. *Incentive Compatibility (IC):* No advertiser has incentive to deviate from bidding truthfully, given that everyone else reports their bid truthfully. Hence, we can write:

$$U_{i,a}^{M}(x_a; x_a, S_i) \geq U_{i,a}^{M}(x'_a; x_a, S_i) \tag{11}$$

where the expectation is taken over other advertisers' click valuations and the mechanism. This constraint guarantees that truth-telling is a Bayesian Nash Equilibrium for all advertisers.

While this restriction to the direct revelation mechanisms reduces the set of auctions the publisher considers, we still need some transformations in these constraints to be able to solve for the optimal mechanism. Like the static case, we show the resulting envelope condition from the IC constraint in the dynamic case as follows:

Lemma 3. *If the mechanism $M(q, e)$ is IC, we have:*

$$\begin{aligned} U_{i,a}^{M}(x_a; x_a, S_i) = \; & U_{i,a}^{M}(x'_a; x'_a, S_i) \\ & + \int_{x'_a}^{x_a} \mathbb{E}_{x_{-a}} \left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q_{i,a,t}^{M}(b_a, x_{-a}; S_{i,t}) \right] db_a \end{aligned} \tag{12}$$

Now, with the dynamic version of the envelope condition, we can show that the publisher's expected revenues under any IC mechanism can be written as follows:

Lemma 4. *If the mechanism $M(q, e)$ is IC, the publisher's expected revenue can be written as follows:*

$$\begin{aligned} \mathbb{E}_x \Bigg[& \sum_{t=1}^{\infty} \beta^{t-1} \left(\sum_{a \in \mathcal{A}_i} \left(x_a - \frac{1 - F_a(x_a)}{f_a(x_a)} \right) P(Y_{i,t} \mid a, S_{i,t}) q_{i,a,t}^{M}(x_a; S_{i,t}) \right) \\ & - \sum_{a \in \mathcal{A}_i} U_{i,a}^{M}(\underline{x}_a; \underline{x}_a, S_i) \Bigg] \end{aligned} \tag{13}$$

Lemma 4 is the equivalent of Lemma 2 for the dynamic case. Now, if we optimize this new objective subject to both *Plausibility* and *Individual Rationality*, we can find the optimal auction. Further, an important finding of this lemma is that the publisher's revenues cannot exceed the first component in Equation (13). Thus, roughly speaking, if we find a dynamic allocation policy that maximizes the first component and payments are such that the second component is zero, the corresponding mechanism is optimal. More precisely, we can write the following proposition on the optimal auction with dynamic objective as follows:

Proposition 2. *The mechanism $M(q, e)$ is optimal if q maximizes*

$$\mathbb{E}_x \left[\sum_{t=1}^{\infty} \beta^{t-1} \left(\sum_{a \in \mathcal{A}_i} \left(x_a - \frac{1 - F_a(x_a)}{f_a(x_a)} \right) P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(x_a; S_{i,t}) \right) \right] \tag{14}$$

subject to q being plausible and $\mathbb{E}_{x_{-a}} \left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(x_a, x_{-a}; S_{i,t}) \right]$ increasing in x_a, and the payment function e is

$$\begin{aligned} e^M_{i,a}(x; S_i) =& \mathbb{E} \left[\sum_{t=1}^{\infty} \beta^{t-1} \left(w_{i,a,t}(x_a; S_{i,t}) q^M_{i,a,t}(x; S_{i,t}) \right) \right] \\ & - \int_{\underline{x}_a}^{x_a} \mathbb{E} \left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t}) \right] db_a, \end{aligned} \tag{15}$$

where the expectation is taken over the stochasticity induced by the dynamic process and state transitions, and not over other advertisers' click valuations.

This proposition states that if an allocation mechanism that maximizes the first part of Equation (13) and sets the payment to make the second part zero, this mechanism is optimal if the allocation mechanism satisfies both plausibility and monotonicity conditions. We explain the details of both allocation and payment components in next sections.

3.3.1 Allocation Rule

We start by finding the optimal allocation rule. Again, we impose the assumption that the distribution F_a is regular for each ad a, i.e., $c_a(x_a) = x_a - \frac{1-F_a(x_a)}{f_a(x_a)}$ is increasing in x_a. In the static case, we show that under this assumption, the optimal ad at any time period is simply the one that maximizes $c_a(x_a)P(Y_{i,t} \mid a, S_{i,t})$. In the dynamic case, however, we want to design a dynamic allocation policy that maximizes the expected revenues for the entire session.

An important result of Lemma 4 is that we can re-write the reward function that is independent of payments. We present a generic definition of reward function with revenue-maximizing objective as follows:

$$R^r_t(a; S_{i,t}) = \left(x_a - \frac{1 - F_a(x_a)}{f_a(x_a)} \right) P(Y_{i,t} \mid a, S_{i,t}) \tag{16}$$

Now, we can use the reward function in Equation (16) to write down the publisher's optimization problem and find the optimal allocation. The following lemma characterizes the optimal allocation function:

Lemma 5. *If the distribution F_a is regular for each ad a (i.e., $c_a(x_a) = x_a - \frac{1-F_a(x_a)}{f_a(x_a)}$ is increasing* *in x_a), then the optimal allocation is the solution to the following Markov Decision Process:*

$$\underset{a}{\operatorname{argmax}}\, R_t^r(a; S_{i,t}) + \beta \mathbb{E}_{G_{i,t+1}|S_{i,t},a}\left[V_{t+1}^r(G_{i,t+1})\right], \tag{17}$$

where the value function is defined for each exposure number as follows:

$$V_{t+1}^r(S_{i,t}) = \max_a R_t^r(a; S_{i,t}) + \beta \mathbb{E}_{G_{i,t+1}|S_{i,t},a}\left[V_{t+1}^r(G_{i,t+1})\right] \tag{18}$$

Depending on the context of our problem, we can use various approaches to design the optimal dynamic allocation policy. Based on this dynamic policy, we can then easily define the allocation function $q_{i,a,t}(x, S_{i,t})$. This lemma guarantees that under the assumption that click valuations come from a regular distribution, the chosen q based on Equation (17) and Equation (18) satisfies both plausibility and monotonicity constraints.

3.3.2 Payments

The intuition behind the payment function in the dynamic case is the same as that in the static case. Each advertiser pays the expected valuation they received from the session, minus an informational rent which is determined by integration of their allocation over the set of lower possible values. This payment function guarantees that the IR constraint is satisfied and the second component in Equation (13) will be zero.

While the informational rent has an analytical solution in the static case as shown in Equation (9), it is harder to derive it analytically in the dynamic case, since the allocation function contains more elaborate rules. The first component in Equation (15) is the expected valuation advertiser a receives from the session, and the second component is the informational rent. To calculate the amount of this rent, we basically need to move down from the true click valuation and see how the number of impressions allocated to advertiser a shrinks. Integrating over this function over the possible values will then give us the amount of rent advertiser a is able to extract. Intuitively, it is the total leeway advertiser a has in this optimal auction.

4 Empirical Strategy

In light of the theoretical results in Proposition 1 and 2, we can characterize the reward function at any point for both static and dynamic cases as follows:

$$R_t^r(a; S_{i,t}) = \left(x_a - \frac{1 - F_a(x_a)}{f_a(x_a)}\right) P(Y_{i,t} \mid a, S_{i,t}), \tag{19}$$

where $S_{i,t}$ is the set of state variables in session i at exposure t. Since the reward specification is the same at any point, we can present a unifying specification of the publisher's optimization in both static and dynamic cases as follows:

$$\underset{a}{\operatorname{argmax}}\, R_t^r(a; S_{i,t}) + \beta \mathbb{E}_{S_{i,t+1}|S_{i,t},a}\left[V_{t+1}^r(S_{i,t+1})\right], \tag{20}$$

where $\beta = 0$ when the objective is static and the value function is defined as the maximum future rewards at a given state.

Given the unified optimization problem in Equation (20), we need to know three key elements to determine the outcomes: (1) distribution of click valuations, (2) match values or click probabilities, and (3) distribution of transitions. The first two elements are required for both static and dynamic objectives, whereas the distribution of transitions is only required for the dynamic case. To conduct an empirical analysis of both static and dynamic optimal auctions presented in §3, we need to obtain estimates for all three unknown elements in Equation (20). Thus, we can identify three empirical tasks as follows:

Task 1: For any ad a in the data, we want to estimate the click valuations x_a based on their observed bidding behavior in the data, under the quasi-proportional auction.

Task 2: For any set of state variables observed in the data, we want to accurately estimate the click probability for all ads if shown in that impression. That is:

$$\hat{y}_{i,t}(a; S_{i,t}) = P(Y_{i,t} \mid a, S_{i,t}), \forall a \tag{21}$$

Task 3: For any set of state variables observed in the data, we want to accurately estimate the leave probability for all ads if shown in that impression. That is:

$$\hat{l}_{i,t}(a; S_{i,t}) = P(L_{i,t} \mid a, S_{i,t}), \forall a \tag{22}$$

The empirical strategy for Tasks 2 and 3 is presented in Rafieian (2019) in great details. We use the same approach for these tasks. We present our empirical strategy to estimate the click valuations in this section.

4.1 Setting

The setting of this problem is the same as that of Rafieian (2019). However, since we are interested in estimation of click valuations, we present more details on the auction environment and advertisers' decisions in this section.

4.1.1 Auction Mechanism

For any exposure that is recognized, the ad-network runs an auction to serve an ad. Unlike the common practice in this industry, the ad-network runs a quasi-proportional auction to select the ad for each exposure. The most notable feature of this auction is in the probabilistic allocation rule that is in contrast with the commonly used mechanisms such as first- or second-price auctions.

1. **Reserve Price:** There is a reserve price b_0 that advertisers' bid must exceed to participate in the auction.

2. **Allocation Function:** For any exposure i and any set of participating ads $\mathcal{A}_i$ with bidding profile $(b_1, b_2, \ldots, b_{|\mathcal{A}_i|})$, ad a has the following probability to win each exposure t within session i:
$$q^p_{i,a,t}(b;z) = \frac{b_a z_a}{\sum_{j \in \mathcal{A}_i} b_j z_j}, \tag{23}$$
where z_a is ad a's quality score which is a measure reflecting the profitability of ad a. The rationale behind using such quality score adjustments is to run the auction based on the expected revenue the ad-network can extract from the ad, rather than their willingness to pay per click. For example, there may be an ad with a very high bid but no chance of getting a click. So despite its high willingness to pay per click, the ad-network cannot actually extract much from it as it will not get many clicks.

   While quality scores can technically vary across auctions, the ad-network does not update them regularly. They only take value zero when the ad is not available for the auction due to budget exhaustion or targeting decisions. Likewise, bids do not change across t, since advertisers cannot choose their bids per auction[3]. Further, changing bids is unlikely as bids reflect advertisers' structural parameters that are stable across auctions.

3. **Payment Scheme:** The ad-network employs a cost-per-click (CPC henceforth) payment scheme wherein advertisers are only when charged when a user clicks on their ad. The amount an ad is charged per click is determined by a next-price rule similar to that of Google's sponsored search auctions. That is, ads are first ranked based on their product of bid and quality score, and each ad pays the minimum amount that guarantees their rank, if a click

[3]There are over 500 auctions run in a second and the information shared with advertisers is minimal about them, disabling them from distinguishing auctions.

happens on their ad:

$$e^p_{i,a,t}(b;z) = \inf \left\{ b' \mid \sum_{j\in\mathcal{A}_i, j\neq a} \mathbb{1}(b'z_a \leq b_j z_j) = \sum_{j\in\mathcal{A}_i, j\neq a} \mathbb{1}(b_a z_a \leq b_j z_j) \right\}, \tag{24}$$

where $\sum_{j\in\mathcal{A}_i, j\neq a} \mathbb{1}(b_a q_a \leq b_j q_j)$ indicates the rank of advertiser a, and the payment b' is the minimum amount of bid that guarantees the same rank for ad a. For example, if there are three bidders with bids 1, 2, and 3, and quality scores 0.1, 0.2, and 0.3, the scores will be 0.1, 0.4, and 0.9. Now, if the second-ranked bidder gets a click, she will pay the price that would have guaranteed her second rank. That is, she only needs to pay $\frac{1\times 0.1}{0.2} = 0.5$, as it guarantees her score to be higher than the third-ranked bidder. Since the item being sold is a click, advertisers' bids reflect their willingness to pay per click. Our goal is to use advertisers' observed bid to estimate their click valuations.

4.1.2 Advertisers' Decisions

Advertisers can make the following decisions:

- **Bid:** Advertisers can set their bid indicating how much they are willing to pay per click.
- **Targeting:** Advertisers can specify their targeting decisions on the following variables: (1) app category, (2) province, (3) hour of the day, (4) smartphone brand, (5) connectivity type, and (6) mobile service provider (MSP). As such, they can exclude some categories from the variables listed above. It will guarantee that their ad will not be shown in the excluded categories.
- **Budget:** Advertisers become unavailable if they do not have enough budget in their account. They can refill their budget and make their campaigns available again.
- **Design of the Banner:** They can design a small banner for their ad.

While we mainly focus on their bidding behavior, it is important to take into account what other decisions they can make. Further, it is important to notice what information they receive from the auction. Each advertiser has a profile in which she can track some key performance metrics on an hourly basis such as the number of impressions, number of clicks, and the total cost of clicks. As such, they do not have granular access to exposure-level information and can only incorporate this information at an aggregate level.

4.2 Data

We use data from a leading in-app ad-network in a large Asian country for over a one week time period from October 22 to 30 in 2015. The analysis sample is the same as the one in Rafieian (2019). However, for estimation of auction, we use all the impressions as it adds to the precision of our estimates. The original data are at the impression-level, indicating the characteristics of each impression, the ad shown, and the final clicking decision by the user. Please see § 3.2.1 in Rafieian (2019) for the description of impression-level data.

The data required for the estimation of auction are not readily available, but we can construct that from the impression-level data. Auction information in the impression-level data is limited to the winning ad and the potential CPC for that particular ad. However, for estimation of auction, we need more information on each auction (impression) regarding the losing ads as well their bids. We recover these two pieces of information using the following approach:

- **Set of Competing Ads:** This task is equivalent to the task of identifying ads that could have been shown in a session. Thus, we use the filtering strategy presented in §5.1.1 in Rafieian (2019).
- **Inferring Bid Amounts:** Actual bids are not directly observed from the original data. The amount reported in the data is the one presented in Equation (24). Since the CPC is always lower than or equal to the actual bid, the maximum observed CPC can be the best approximation for the actual bid. We can write:

$$b_a \approx \sup_i \inf \left\{ b' \mid \sum_{j \in \mathcal{A}_i, j \neq a} \mathbb{1}(b' z_a \leq b_a z_a) = \sum_{j \in \mathcal{A}_i, j \neq a} \mathbb{1}(b_a z_a \leq b_a z_a) \right\} \tag{25}$$

  Again, being in a data-rich environment with extensive variation in the next-price enormously helps. One intuitive validation for this method would be the ending digits of approximated bids as bidders are more likely to round up their bids. We find that this is the case for most approximated bids.

Overall, it provides us with 547,626,756 impressions and 398 distinct ads competing to win the impressions.

4.2.1 Summary Statistics

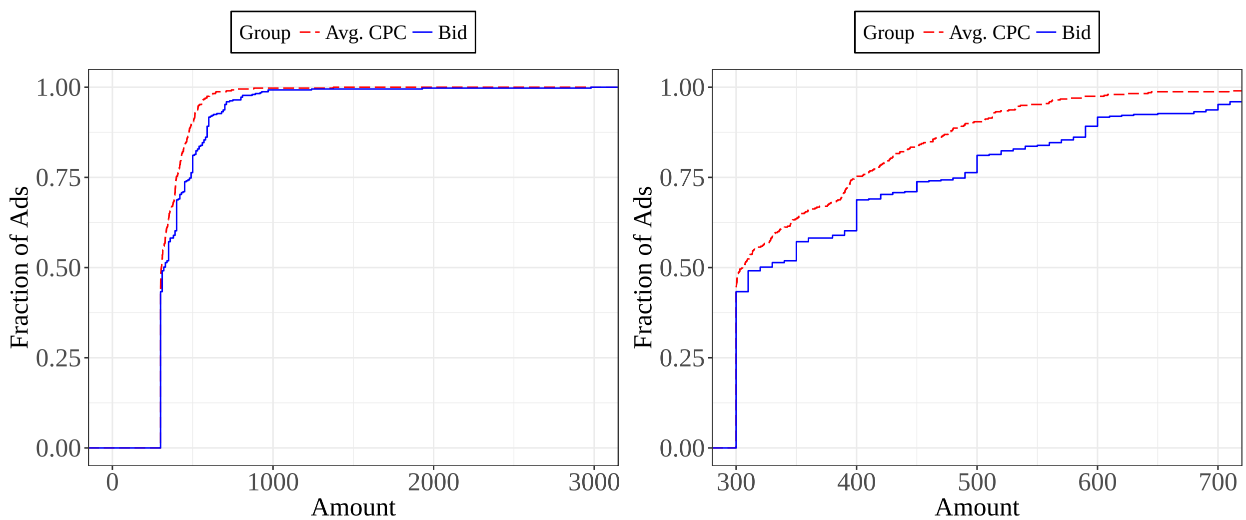
Here we present some summary statistics of the data. Overall, there are 398 ads participating in the timeline of the study. Table 1 shows mean, standard deviation, minimum and maximum of key ad-level variables. All these variables are defined by ads and reflect an important aspect of advertiser's decision, including bidding, targeting, and budget.

As shown in this table, there is substantial variation in ads participating in this auction, ranging from smaller ads with a very short lifetime and few clicks, to larger ads with permanent availability and significant expenditure. Further, Table 1 presents number of categories targeted out of all 86 targeting categories, indicating that most ads are not targeting at all.

Variable	Mean	Std. Dev	Min	Max
Bid	411.13	208.25	300.00	2976.00
Avg. CPC	363.30	103.98	300.00	1375.84
Total Hours of Availability	48	70	1	217
No. of Impressions	1,379,411	5,402,346	7	66,228,977
No. of Clicks	12,619	55,208	0	656,680
Click-Through Rate	0.0109	0.0112	0.0000	0.1429
Total Expenditure	4,880,013	21,032,815	0	216,434,237

Table 1: Summary statistics of key ad-level variables

Figure 2 shows the the empirical CDF of ads' bids and their average CPC. The figure on the left (Figure 2b) shows the distribution for all values of bids. As shown in this figure, there is a reserve price 300 that censors the left side of the graph. With no reserve price, there could have been bids lower than 300. This raises an important identification challenge that we address in the estimation approach, especially because a substantial portion of ads are reserve bidders.



(a) Zoomed Out

(b) Zoomed In

Figure 2: Empirical CDF of ads' bids and average CPCs.

Figure 2b zooms into the bids for a shorter interval of amounts that capture over 90% of all ads. While there are discontinuities in bids especially at round numbers, average CPCs show a smoother pattern. Overall, both figures demonstrate a first-order stochastic dominance relationship between the empirical CDF of average CPC and bids which emphasizes the fact an ad's CPC cannot exceed her bid.

4.3 Estimation of Auction

We now present our approach to estimate the distribution of click valuations from the observed auction data. We first develop advertisers' utility model and provide and equilibrium analysis in §4.3.1. We then state the set of assumptions required for the identification of the distribution of click valuations in §4.3.2. Finally, in §4.3.3, we propose our estimation approach.

4.3.1 Advertisers' Model

We begin by characterizing advertisers' utility model in the context of our data. Given the allocation and payment functions defined in §4.1.1, we know that advertiser a receives the following utility from exposure t in session i:

$$\mathbb{1}(A_{i,t} = a, Y_{i,t} = 1) \left(x_a - e^p_{i,a,t}(b)\right),$$

where the indicator function takes value one if ad a is selected through the allocation mechanism and got clicked in that impression. Here we explicitly assume that advertisers only receive utility from clicks and there is no utility from an impression alone.[4]

Mirrokni et al. (2010) study this auction with the case of identity cost function, i.e., $e^p_{i,a,t}(b) = b_a$. They show that advertisers' optimal bidding strategy depends on their expected probability of winning. While this expected probability varies across auctions in our case, we do not observe bid-changing behavior. This is possibly because the effect of the expected probability of winning on their bid is very small in a competitive market and it does not exceed the bid-changing cost.[5]

The observation that advertisers do not change their bids informs us in modeling advertisers' utility function. We can characterize the expected utility of advertiser a from quasi-proportional auction p as follows:

$$u^p_a(b_a; x_a) = m_a \left(x_a \tilde{q}^p_a(b_a) - \tilde{e}^p_a(b_a)\right), \tag{26}$$

where m_a is the expected probability of click on ad a conditional on winning. Since we treat this probability as independent to other parts of the expected utility, we separate that in the specification. Further, functions $\tilde{q}^p_a$ and $\tilde{e}^p_a$ are advertiser a's expected allocation and payment functions. Before defining these two functions, we define an important distribution that we use in our estimation. Let $\mathcal{C}_a$ denote the joint distribution of auctions that advertiser a participates in. Each draw C_a from this distribution contains the information about the distribution of bids and quality scores, the impression

[4] We believe this assumption is valid in our context, because almost all ads are mobile apps whose objective is more app installs. This assumption may be violated in the presence of brand ads whose objective is more reach.

[5] Please notice that this is just a behavioral cost and the platform does not charge them for bid-changing. As shown in Rafieian and Yoganarasimhan (2018b), the marginal effect of the expected winning probability α is $\frac{1}{1-\alpha}$. Hence, in a market with many homogeneous competitors, α has a very small effect on advertisers' equilibrium bidding strategy.

characteristics (session i and exposure t), and the number of ads competing in that auction. Using this distribution, we can now characterize the allocation function as follows:

$$\tilde{q}_a^p(b_a) = \mathbb{E}_{\mathcal{C}_a}\left[q_{i,a,t}^p(b_a, b_{-a}^{C_a})\right], \tag{27}$$

where the expectation is taken over all the configurations C_a. This function essentially returns the expected probability of winning by advertiser a in a random impression. Given the Continuous Mapping Theorem, we know that $\tilde{q}_a^p(b_a) \xrightarrow{p} \frac{b_a z_a}{b_a z_a + \zeta_a}$, where $\zeta_a = \mathbb{E}_{\mathcal{C}_a}\left[\sum_{j \in \mathcal{A}_i, j \neq a} b_j q_j\right]$. Therefore, we characterize the functional form for the function $\tilde{q}$ as follows:

$$\tilde{q}_a^p(b_a) = \frac{b_a z_a}{b_a z_a + \zeta_a} \tag{28}$$

Similarly, we can define the payment function as follows:

$$\tilde{e}_a^p(b_a) = \mathbb{E}_{\mathcal{C}_a}\left[\mathbb{1}\left(A_{i,t} = a; C_a\right) e_{i,a,t}^p\left(b_a, b_{-a}^{C_a}; z_a, z_{-a}^{C_a}\right)\right] \tag{29}$$

The term $\mathbb{1}\left(A_{i,t} = a; C_a\right)$ in Equation (29) shows whether ad a wins the impressions and the second term basically computes the cost-per-click using Equation (24). Now, using the utility specification in Equation (26), we can write the first-order condition as follows:

$$x_a = \frac{\partial \tilde{e}_a^p(b_a)}{\partial b_a}\left(\frac{\partial \tilde{q}_a^p(b_a)}{\partial b_a}\right)^{-1} \tag{30}$$

This equation lays out our estimation approach, as we need to empirically estimate both $\tilde{q}$ and $\tilde{e}$ using the data at hand. In the next section, we discuss the assumptions required for identification and overall identification strategy.

4.3.2 Assumptions and Identification

We make a series of assumptions required for our estimation task. Some of these are commonly used in the context of auction estimation, whereas some other assumptions are more specific to the context of quasi-proportional auction in our data. While the former is necessary for identification, we impose the latter mostly for the ease of estimation. For robustness check, we relax those specific assumptions and show the results will not change.

Our first assumption characterizes how advertisers make decisions regarding their own bidding strategy:

Assumption 2. ***[Profit-Maximizing Advertisers]*** *Advertisers are profit-maximizing, i.e., they choose* *the bid that maximizes their profit.*

In light of Assumption 2, we can treat observed bids in the data as equilibrium bids and use Lemma 6 to estimate click valuations. As shown in Equation (26), advertisers choose their optimal bids given their own private click valuations and their belief about other advertisers. The following assumption characterizes other advertisers' private click valuations:

Assumption 3. ***[Independent Private Values (IPV)]*** *Advertisers' private click valuations are drawn independently from a distribution $F(\cdot)$ with a continuous density.*

The assumption of Independent Private Values (IPV) is an assumption used in most auction settings (Guerre et al., 2000; Athey and Haile, 2007). This assumption hints a straightforward approach to estimate both $\tilde{q}$ and $\tilde{e}$ in Equation (26), by simulating the distribution of $\mathcal{C}_a$ for each ad.

Now, we make assumptions more specific to the case of quasi-proportional auctions. Some of these assumptions are necessary to use the first-order condition in Lemma 6. However, we impose some of these assumptions to provide an analytically simpler solution. We start with the following assumption:

Assumption 4. ***[Zero Impression Value]*** *Advertisers' valuation from an unclicked impression is zero.*

This assumption is widely used in both theoretical and empirical literature on cost-per-click auctions (Edelman et al., 2007; Varian, 2007; Athey and Nekipelov, 2010). In our case, we believe the value from impressions is negligible as most ads are mobile apps whose objective is to get more app installs and can be considered as performance ads. In the next assumption, we make an assumption about the role of budget in advertisers bidding behavior:

Assumption 5. ***[Budget Independent Bidding]*** *Advertisers' bidding behavior is independent of their budget.*

In principle the equilibrium bidding behavior may change for budget-constrained advertisers (Borgs et al., 2007; Asadpour et al., 2014). We make this assumption in our case for two reasons. First, we observe quite a few advertisers who have run out of budget during our study and refilled it later. However, there is no difference in their bidding behavior. The second reason is more empirical, as we do not observe the exact budget in our data. As such, we need to approximate this budget which can be quite noisy.

The assumptions so far are enough to establish the identification of the distribution of click valuations for bidders who bid above the reserve price. However, we make an additional assumption that helps us derive a simpler analytical solution for this auction.

Assumption 6. *[Separability of Allocation] We can separate the allocation function from the payment function as follows:*

$$\tilde{e}^p_a(b_a) = \tilde{q}^p_a(b_a)\mathbb{E}_{\mathcal{C}_a}\left[e^p_{i,a,t}\left(b_a, b^{C_a}_{-a}; z_a, z^{C_a}_{-a}\right)\right]$$

This assumption allows us to simplify the first-order condition presented in Equation (30). For brevity, let $\epsilon^p_a(b_a) = \mathbb{E}_{\mathcal{C}_a}\left[e^p_{i,a,t}\left(b_a, b^{C_a}_{-a}; z_a, z^{C_a}_{-a}\right)\right]$. It is easy to show that the function ϵ^p_a is monotonic for each advertiser a. However, we need to make the following assumption about this function:

Assumption 7. *[Twice Differentiability of Payment] The expenditure function ϵ^p_a is twice differentiable in b_a.*

Now, we can use all these assumptions and show the following lemma which is the key idea behind our identification:

Lemma 6. *If advertiser a's equilibrium bid is greater than the reserve price ($b_a > b_0$) and the function $b^2 \frac{\partial \epsilon^p_a(b)}{\partial b}$ is increasing in the local neighbourhood around b_a, her click valuation x_a can be written in terms of equilibrium bids as follows:*

$$x_a = \epsilon^p_a(b_a) + \frac{b_a \frac{\partial \epsilon^p_a(b)}{\partial b}\Big|_{b=b_a}}{1 - \tilde{q}_a(b_a)} \tag{31}$$

Lemma 6 shows the first-order condition for the case where the advertiser's equilibrium bid is greater than the reserve price. We need $b^2 \frac{\partial \epsilon^p_a(b)}{\partial b}$ to be increasing at b_a to satisfy the second-order condition. Intuitively, if the function e is too concave, this assumption may not hold. However, it is a testable assumption as we can empirically test it for all the bidders.

Now, we focus on the advertisers whose bid is equal to the reserve price. For these advertisers, we cannot use the inverse bidding equation in Equation (31), as their optimal bid could have been lower than the reserve price, if they had been allowed to bid lower. Figure **??** illustrates this point by plotting an advertiser's utility against her bid. As shown in this figure, while she submitted the reserve price b_0, her first-order condition is satisfied in point $b^* < b_0$. Instead of the first-order condition, we have other conditions for reserve price bidders. First, we know that their participation constraint is satisfied, i.e., $x_a \geq b_0$. Second, we know that their first-derivative at b_0 is lower than or equal to zero, as the utility must be decreasing at the truncation point. Together, we can write the following lemma to characterize the link between the click valuation and reserve bidding behavior:

Lemma 7. *If advertiser a's equilibrium bid is equal the reserve price ($b_a = b_0$), we can obtain lower and upper bounds for her click valuation x_a as follows:*

$$b_0 \leq x_a \leq \epsilon_a^p(b_0) + \frac{b_0 \frac{\partial \epsilon_a^p(b)}{\partial b} \mid_{b=b_0}}{1 - \tilde{q}_a(b_0)} \tag{32}$$

In light of Lemma 32, we cannot point-identify click valuations for the reserve bidders. However, we can use lower and upper bounds in Equation (32). To complete our identification for the distribution of valuations for all participating advertisers, we need one more assumption on the click valuations of the reserve bidders:

Assumption 8. ***[Uniformity of Reserve Bidders' Valuations]*** *The click valuation x_a for any reserve price bidder is drawn from a uniform distribution with the following bounds:*

$$x_a \sim \mathcal{U}\left(b_0, \epsilon_a^p(b_0) + \frac{b_0 \frac{\partial \epsilon_a^p(b)}{\partial b} \mid_{b=b_0}}{1 - \tilde{q}_a(b_0)}\right) \tag{33}$$

With this assumption, we can now state our identification proposition:

Proposition 3. *If all Assumptions 2 to 8 hold, then the distribution of advertisers' private click valuations are non-parametrically identified.*

The proof for this part is similar to the identification proof for most auction models with independent private values (Guerre et al., 2000; Athey and Haile, 2007). We can directly observe all the elements in Lemma 6 and 7. Thus, we can estimate advertisers' click valuations and form the distribution F.

4.3.3 Estimation Method

Our estimation approach relies on the findings in §4.3.2. We first estimate the joint distribution of configurations $\mathcal{C}_a$ for all ads. This involves the estimation of quality scores, distribution of observed bids, and observed impressions. We then use this distribution to form both allocation and payment functions, which in turn, allows us to estimate the distribution of click valuations.

Before describing the estimation procedure, we need to define a time period η at which advertisers update their bids. While advertisers do not change their bid in our data, we set this time period for two main reasons. First, advertisers can technically change their bids if they want. Therefore, it is more reasonable to assume that they revise their decision every once in a while. Second, from an empirical point-of-view, it allows us to capture the variance in the set of advertisers competing. Thus, following the arguments in Rafieian and Yoganarasimhan (2018a), we set an hourly time period such that each η distinguish a different hour-day combination. The denote the last time period by L.

We provide a step-by-step procedure for our estimation as follows:

- **Step 1:** We estimate all the quality scores $z_{a,\eta}$ for all ads and time periods. We use the proportional nature of the allocation to identify quality scores. If ads a and a' both participate in the auction for exposure t in session i, the denominator of their winning probability is the same. Thus, the odds ratio of these two ads can be written as follows:

$$\frac{\Pr(A_{i,t} = a)}{\Pr(A_{i,t} = a')} = \frac{b_a z_a}{b_{a'} z_{a'}}$$

  Since we observe their bids, we can easily estimate the ratio $\frac{z_a}{z_{a'}}$ by calculating the number of impressions awarded to each ad over the set of auctions where they both have participated. Let $\mathcal{I}^{\eta}_{a,a'}$ denote the set of impressions wherein both ads a and a' participate in time period η. We can write:

$$\frac{\hat{z}_{a,\eta}}{\hat{z}_{a',\eta}} = \frac{b_{a'} \sum_{(i,t)\in\mathcal{I}^{\eta}_{a,a'}} \mathbb{1}(A_{i,t} = a)}{b_a \sum_{(i,t)\in\mathcal{I}^{\eta}_{a,a'}} \mathbb{1}(A_{i,t} = a')} \tag{34}$$

  As such, we can estimate all the ratios. For ad a^* that has participated in all the impressions, we set $\hat{z}_{a^*,\eta} = 1$ for all time periods. This allows us to estimate all quality scores $\hat{z}_{a,\eta}$.
- **Step 2:** We empirically estimate the joint distribution $\mathcal{C}_{a,\eta}$ for all ads in all time periods. This distribution contains the information about all the impressions that ad a has participated in, number of bidders in corresponding impressions, and the joint distribution of bids and quality scores. We call the estimated distribution $\hat{\mathcal{C}}_{a,\eta}$.
- **Step 3:** Given all estimated configurations $\hat{\mathcal{C}}_{a,\eta}$, we can estimate the allocation probability for all ads over all time periods as follows:

$$\hat{q}_{a,\eta}(b_a) = \frac{1}{N_a} \sum_{(i,t)\sim\hat{\mathcal{C}}_{a,\eta}}^{N_a} \mathbb{1}(A_{i,t} = a), \tag{35}$$

  where N_a is the number of draws we get from the distribution $\hat{\mathcal{C}}_{a,\eta}$. It is worth noting that we do not need to fully estimate the allocation function as the relationship in Equation (30) only depends on the final allocation probability given the observed bid.
- **Step 4:** Again, we use the distribution estimate configurations $\hat{\mathcal{C}}_{a,\eta}$ to estimate our cost-per-click

function $\hat{\epsilon}_{a,\eta}(b)$. For any value b, we can estimate the cost-per-click function as follows:

$$\hat{\epsilon}_{a,\eta}(b) = \frac{1}{N_a} \sum_{(i,t)\sim\hat{\mathcal{C}}_{a,\eta}}^{N_a} \inf \left\{ b' \mid \sum_{j\in\mathcal{A}_i, j\neq a} \mathbb{1}(b' z_a \leq b_j z_j) - \mathbb{1}(b z_a \leq b_j z_j) = 0 \right\} \tag{36}$$

This gives us the estimate for the cost-per-click function. We can then take the numerical derivatives of $\hat{\epsilon}_{a,\eta}(b)$ and estimate $\hat{\epsilon'}_{a,\eta}(b)$ for all values of b.

- **Step 5:** Using the estimates for the allocation and cost-per-click functions, we can now identify click valuations $x_{a,\eta}$ for any combination of a and η as follows:

$$\hat{x}_{a,\eta} = \begin{cases} \hat{\epsilon}_{a,\eta}(b_a) + \frac{b_a \hat{\epsilon'}_{a,\eta}(b_a)}{1-\hat{q}_{a,\eta}(b_a)} & b_a > b_0 \\ x_0 \sim \mathcal{U}\left(b_0, \hat{\epsilon}_{a,\eta}(b_a) + \frac{b_a \hat{\epsilon'}_{a,\eta}(b_a)}{1-\hat{q}_{a,\eta}(b_a)}\right) & b_a = b_0 \end{cases} \tag{37}$$

We also obtain single estimates for each x_a without using the time periods. The procedure is the same as what described above. We only need to use $\hat{\mathcal{C}}_a$ instead of this distribution for all time periods. We then use these estimates throughout for click valuations in our analysis.

- **Step 6:** Finally, we use the estimates in Equation (37) to form the distribution of click valuations. We can write:

$$\hat{F}(x) = \frac{1}{L}\sum_{\eta=1}^{L} \frac{1}{|\mathcal{A}_\eta|} \sum_{a\in\mathcal{A}_\eta} \mathbb{1}(x_{a,\eta} \leq x) \tag{38}$$

$$\hat{f}(x) = \frac{1}{L}\sum_{\eta=1}^{L} \frac{1}{|\mathcal{A}_\eta|} \sum_{a\in\mathcal{A}_\eta} \frac{1}{h}\mathcal{K}\left(\frac{x - \hat{x}_{a,\eta}}{h}\right), \tag{39}$$

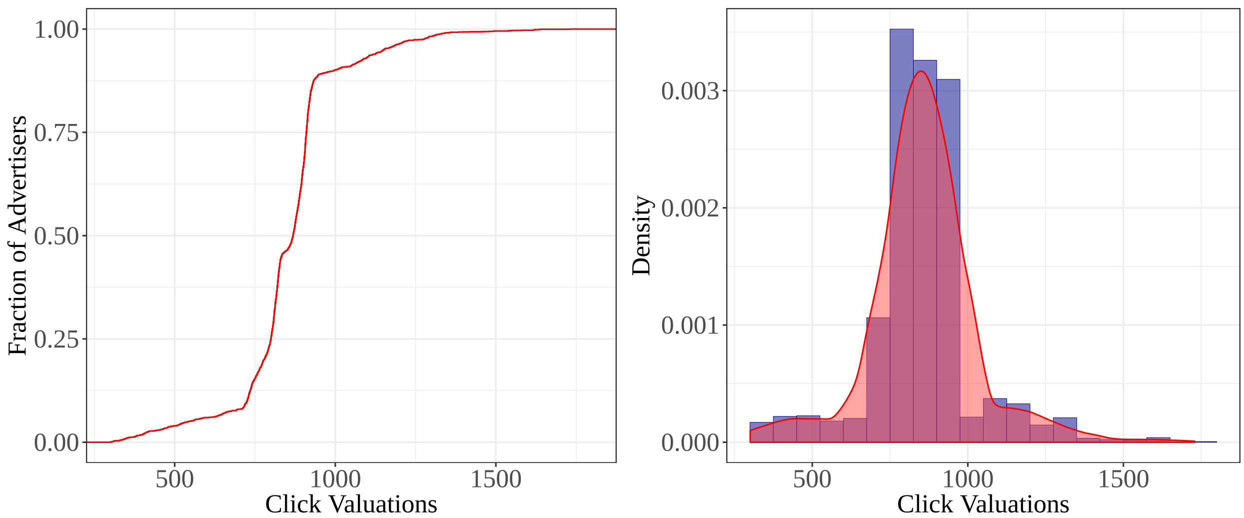
where $\mathcal{K}$ is the kernel function. In this case, we use Epanechnikov kernel, i.e., $\mathcal{K}(u) = \frac{3}{4}(1-u^2)\mathbb{1}(|u| \leq 1)$.

4.4 Results

We now present our results on the estimation of click valuations. We first show the main findings on the distribution of click valuations in §4.4.1. Next, in §4.4.2, we discuss the validity of our assumption that the distribution of click valuations is regular.

4.4.1 Estimated Distribution of Click Valuations

We first show the estimated distribution of click valuations for top ads. We only focus on the top ads, as it is the set of ads that we use for the counterfactual evaluation of the optimal auctions. However, we can technically recover the distribution for any set of ads. Figure 3 show the empirical CDF



(a) CDF

(b) Density

Figure 3: Estimated distribution of click valuations for top 15 ads.

and estimated density of the distribution. As shown in Figure 3a, the range between 700 to 1000 constitutes the vast majority of click valuations. Over 80% of click valuations lie in this range.

Figure 3b shows the estimated density for the distribution of click valuations. We use Epanechnikov kernel with bandwidth size of 75. The shape of density is similar to a bell curve with a low variance. The mean and standard deviation for this distribution are 854.91 and 164.73 respectively.

4.4.2 On the Regularity of the Distribution of Click Valuations

Now, we use the results shown in Figure 3 and check whether the estimated distribution is regular. As described in Assumption 1, the estimated distribution $\hat{F}$ is regular if $x - \frac{1-\hat{F}(x)}{\hat{f}(x)}$ is strictly increasing in x. Using our estimates for $\hat{F}$ and $\hat{f}$, we plot the virtual valuations in Figure 4. As demonstrated in this figure, virtual values are monotonic in click valuations.

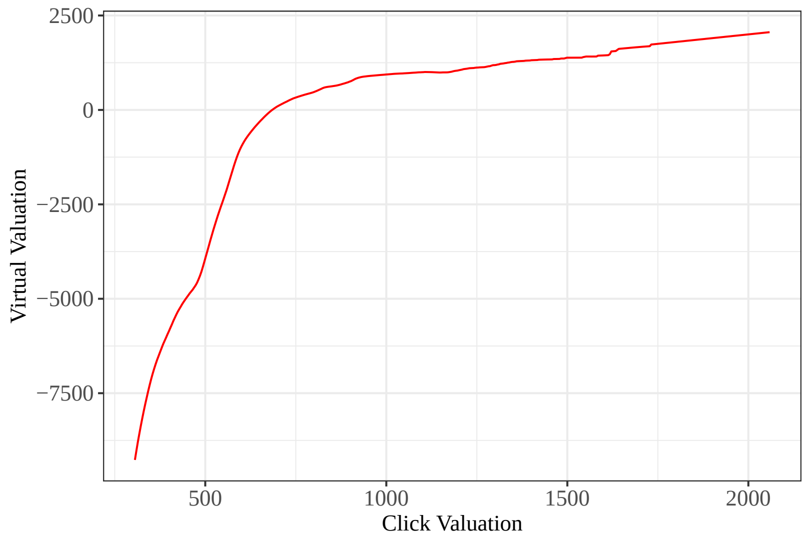


Figure 4: Virtual valuations against click valuations

5 Counterfactual Evaluation of Optimal Auctions

In this section, we take our theoretical models in §3.2 and §3.3 to the data and present our approach to evaluate the counterfactual outcomes under optimal auctions. In both cases, we need to use the reward function that is presented Equation (16) in a generic way. More specific to our context, we can write it as follows:

$$R_t^r(a; S_{i,t}) = \left(x_a - \frac{1 - F_a(x_a)}{f_a(x_a)}\right) P(Y_{i,t} \mid a, S_{i,t}), \tag{40}$$

Now, we want to estimate the reward function in this case, given our data. We use estimates for both click probability at any state (i.e., $P(Y_{i,t} \mid a, S_{i,t})$) as well as click valuations x_a for all ads from Rafieian (2019). We also impose symmetry assumption on the distribution of click valuations and estimate $\hat{F}$ and $\hat{f}$, based on the estimated click valuations. Given these estimates, we can now estimate the reward function in Equation (41) and re-write it as follows:

$$\hat{R}_t^r(a; S_{i,t}) = \left(\hat{x}_a - \frac{1 - \hat{F}(\hat{x}_a)}{\hat{f}(\hat{x}_a)}\right) \hat{y}_{i,t}(a; S_{i,t}) \tag{41}$$

We use our estimates for the reward function in this case to obtain optimal auctions in both static and dynamic cases and then present our how we evaluate the resulting auctions.

5.1 Solution Concept for the Optimal Static Auction

We start with the simpler case where the publisher's objective is static. In §3.2.1 and §3.2.2, we show how we can derive the optimal mechanism $M^m(q^m, e^m)$. More specifically, we show that when $\hat{F}$ is symmetric, the optimal auction will be a second-price auction with an optimal reserve price. We can estimate the reserve price $\hat{e}_0$ as follows:

$$\hat{e}_0 = \hat{c}^{-1}(0),$$

where $\hat{c}(x) = \frac{1-\hat{F}(x)}{\hat{f}(x)}$ and we can find $\hat{e}_0$ by solving for $\hat{e}_0 = \frac{1-\hat{F}(\hat{e}_0)}{\hat{f}(\hat{e}_0)}$. Using this reserve price, we can then derive the optimal allocation q^m as follows:

$$q^m_{i,a,t}(\hat{x}; S_{i,t}) = \begin{cases} 1 & \hat{R}^r_t(a; S_{i,t}) > \max_{a' \in \mathcal{A}_i \backslash a} \left(\hat{R}^r_t(a'; S_{i,t}), \hat{e}_0 \right) \\ 0 & \text{otherwise} \end{cases} \tag{42}$$

This allocation rule indicates that the highest reward ad (which is the ad with the highest expected valuation) will win the impression.[6] As specified in §3.2.2, the payments are determined as follows:

$$e^m_{i,a,t}(\hat{x}; S_{i,t}) = \begin{cases} \max_{a' \in \mathcal{A}_i \backslash a} \left(\hat{R}^r_t(a'; S_{i,t}), \hat{e}_0 \right) & q^m_{i,a,t}(x \mid S_{i,t}) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{43}$$

We can now use both Equation (42) and Equation (43) to find the optimal mechanism and apply it to each session.[7]

5.2 Solution Concept for the Optimal Dynamic Auction

Now, we focus on the optimal auction with dynamic objective and show how we can use our estimates and determine the optimal mechanism $M^d(q^d, e^d)$. As shown in §3.3.1, we can find the optimal allocation by solving a Markov Decision Process that incorporates expected future rewards as well as the current period reward. Like Rafieian (2019), we consider a finite case with and solve for the optimal allocation using backward induction. For notational convenience, we first define the

[6]It is worth noting that in a case where there are multiple highest reward ads, we randomly allocate the item to one of the highest reward ads. However, since this a very rare event in an empirical setting, for brevity, we exclude that from Equation (42).

[7]It is important to notice that we can recover the distribution of click valuations specific to each session. For simplicity, we focus on the case one global distribution for all sessions. As a robustness check, we also consider the case where distributions are session-specific.

function $\tilde{V}_t^r(a, S_{i,t})$ for a pair of action and state as follows:

$$\begin{aligned}\tilde{V}_t^r(a, S_{i,t}) = &\left(\hat{x}_a - \frac{1 - \hat{F}(\hat{x}_a)}{\hat{f}(\hat{x}_a)}\right) \hat{y}_{i,t}(a; S_{i,t}) \\ &+ \left(1 - \hat{l}_{i,t}(a; S_{i,t})\right) \hat{y}_{i,t}(a; S_{i,t}) V_{t+1}^r\left(\langle S_{i,t}, a, Y_{i,t} = 1\rangle\right) \\ &+ \left(1 - \hat{l}_{i,t}(a; S_{i,t})\right) \left(1 - \hat{y}_{i,t}(a; S_{i,t})\right) V_{t+1}^r\left(\langle S_{i,t}, a, Y_{i,t} = 0\rangle\right),\end{aligned} \tag{44}$$

where $\hat{y}_{i,t}(a; S_{i,t})$ and $\hat{l}_{i,t}(a; S_{i,t})$ are estimated leave and click probabilities respectively. We can use Equation (44) and describe our backward induction solution concept as follows:

1. We begin from the last period T. Since there is no expected future at that point, the value function can be estimated as follows:

$$\hat{V}_T^c(S_{i,T}) = \max_{a \in \mathcal{A}_i} \hat{R}_t^r(a, S_{i,t}) \tag{45}$$

2. For any $t < T$, we can determine the value function as follows:

$$\hat{V}_t^c(S_{i,t}) = \max_{a \in \mathcal{A}_i} \tilde{V}_t^r(a, S_{i,T}) \tag{46}$$

   We can easily estimate the value function for any $t < T$ if we have all the value functions for the next periods. We can satisfy that by going backward and find the value function for all states at any t and continue this process until exposure number 1.

Once we have identified the value function for all the states, we can find the optimal allocation with dynamic objective as follows:

$$q_{i,a,t}^d(\hat{x}; S_{i,t}) = \begin{cases} 1 & a = \text{argmax}_{a \in \mathcal{A}_i} \tilde{V}_t^r(a, S_{i,t}), t < T \\ 1 & a = \text{argmax}_{a \in \mathcal{A}_i} \hat{R}_t^r(a, S_{i,t}), t = T \\ 0 & \text{otherwise} \end{cases} \tag{47}$$

Now, given the allocation, we can determine the payments using Equation (15). It is important to notice that the payment is determined in expectation over the entire session for each ad. For any ad a, the first component in Equation (15) is the average value ad a would get given the allocation, and the second component is the integral of the expected value ad a would get if she reduces her bid, taken over all possible bids that she could submit. The first component is easier to calculate as it is an expected sum of the total value each ad receives given each sequence. The second part, however, is more computationally intensive as it involves a numerical integration. To do that, we first need to estimate the function inside the integral. Let $\hat{Q}_{i,a}(b_a; S_{i,1})$ denote the expectation inside the integral in Equation (15). This will be the expected number of clicks ad a would get for any bid, given other players report their click valuations. We interpolate this function by finding its values for a set of points on the interval $[\underline{x}_a, x_a]$. Since the maximum number of expected number of clicks an ad could get is 6 when $T = 6$, we split this interval into 6 equally length intervals and find the function values for the points splitting the interval. We operationalize that with $h_a = \frac{\hat{x}_a - \hat{\underline{x}}_a}{T}$, that indicates the length of each interval for ad a. As such, if $b_a \in (\hat{\underline{x}}_a + (i-1)h_a, \hat{\underline{x}}_a + ih_a]$, we can estimate this function as follows:

$$\hat{Q}_{i,a}(b_a, \hat{x}_{-a}; S_{i,1}) = \sum_{t=1}^{T} \sum_{g_t \in \mathcal{G}_T} q^d_{i,a,t}(\hat{x}_a + ih_a, \hat{x}_{-a}; S_{i,t}) P(g_t \mid \tau, \pi), \tag{48}$$

where g_t contains the states and ads prior to period t. In Equation (48), we approximate the function $\mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t})\right]$ with a step function, which makes the integral computation significantly easier. Now, we can determine the payments as follows:

$$\begin{aligned} e^d_{i,a}(x; S_{i,1}) = &\sum_{t=1}^{T} \sum_{g_t \in \mathcal{G}_T} \hat{x}_a \hat{y}_{i,t}(a; S_{i,t}) q^d_{i,a,t}(x; S_{i,t}) P(g_t \mid \tau, \pi) \\ &- h_a \sum_{t=1}^{T} \hat{Q}_{i,a}(\hat{x}_a + th_a, \hat{x}_{-a}; S_{i,1}), \end{aligned} \tag{49}$$

where the first component is the total value ad a receives from the session given the allocation, and the second component is the estimate of advertiser's surplus.

5.3 Evaluation

To evaluate the performance of each auction, we implement the direct method with the reward function with the revenue-maximizing objective. As such, we can define the expected revenue from session i with the horizon length T as $\rho^r_T(M; S_{i,1})$ for any mechanism M as follows:

$$\rho_T(M; S_{i,1}) = \mathbb{E}_{g_t \sim (\tau, M)} \left[\sum_{t=1}^{T} \beta^{t-1} r^r_t \right], \tag{50}$$

where the sequence g_t is determined by the joint distribution of transitions τ, and the allocation rule in mechanism M. Further, r^r_t denotes the reward function with the revenue-maximizing objective for the pair of state and action shown in the corresponding g_t. We can use our estimates for the distribution of transitions and evaluate $\rho_T(M; S_{i,1})$ for any mechanism M as follows:

$$\hat{\rho}_T(M; S_{i,1}) = \sum_{t=1}^{T} \sum_{g_t \in \mathcal{G}_T} \sum_{a \in \mathcal{A}_i} q_{i,a,t}^M(\hat{x}; S_{i,t}) \hat{R}_t^r(a; S_{i,t}) P(g_t \mid \tau, M) \tag{51}$$

This gives us the expected revenue that the publisher can extract from session i in the first T periods, when using the mechanism M.

6 Results

6.1 Gains from the Optimal Dynamic Auction

We start by presenting different session-level outcomes under both dynamic and static optimal auctions as well as the actual outcomes under the current auction – quasi-proportional auction. We evaluate the counterfactual outcomes under optimal auctions using our approach in §5. The sample of sessions that we use is the same as the one in Rafieian (2019). In our evaluation, we only focus on the first six exposures in any session, as we set $T = 6$ as the length of the horizon. We present our results in Table 2.

As expected, both optimal auctions generate substantial gains over the current auction in terms of all session-level outcomes, except the expected advertisers' surplus. Three key components of the current auction explain its performance relative to the optimal auctions. First, as discussed before, advertisers bid roughly half of their click valuation in this auction, lowering the revenues the publisher is able to extract. This, in turn, explains higher advertisers' surplus in the current auction as advertisers can extract huge informational rent by bidding half of their valuations. Second, the allocation mechanism is probabilistic in the current auction, which enables advertisers with low valuations to win impressions and clicks. This significantly reduces the efficiency or total surplus generated in the current auction compared to optimal auctions. Finally, as mentioned before, the ad-network does not incorporate sophisticated customization tools to show more relevant ads to users. This also greatly contributes to the gap between the current auction and other optimal auctions.

Next, we focus on our main goal in this paper, and compare the session-level outcomes under dynamic and static optimal auctions. We find that using an optimal dynamic auction leads to 1.60% increase in the expected revenue per session, compared to the optimal static auction. This is of particular importance to the publishers and ad-networks as they usually use optimal static auctions (e.g., second-price auction).

While revenue is the main outcome most publishers and ad-networks care about, they do not usually want to achieve better revenue outcomes at the expense of efficiency and advertisers' surplus.

	Auctions		
	Dynamic	*Static*	*Current*
Expected Publisher's Revenue Per Session	126.49	124.49	31.63
Expected Total Surplus Per Session	143.45	140.96	66.60
Expected Advertisers' Surplus Per Session	16.96	16.47	34.97
Expected No. of Clicks Per Session	0.1577	0.1549	0.0823
Expected Session Length	3.24	3.25	3.12
Expected CTR	4.86%	4.76%	2.64%
No. of Users	1000	1000	1000
No. of Sessions	12,136	12,136	12,136

Table 2: Market outcomes under different auctions for a sequence size of 6

Our results show that the optimal dynamic auction yields higher total and advertisers' surplus than the optimal static auction: the expected total surplus and advertisers' surplus grow by 1.77% and 3.00% respectively. Thus, advertisers also benefit from the revenue-optimal dynamic auction, as compared to the revenue-optimal static auction. Finally, we focus on the user-level outcomes – expected number of clicks per session and the session length. While the difference in the session length is very small, we find 1.83% increase in the expected number of clicks. This finding is in line with the findings in Rafieian (2019). More importantly, it suggest that the revenue gains from the optimal dynamic auction likely come from the improvement in the match between ads and users, and not from the greater ability to extract informational rent from advertisers.

6.2 Distribution of Advertisers' Surplus

Our results in Table 2 indicate that the average advertisers' surplus increases by 3.00% in the optimal dynamic auction, as compared to the optimal static auction. We now explore how this surplus is distributed across advertisers. We use two main approaches to examine the distribution of advertisers' surplus: (1) number of advertisers who benefit from the optimal dynamic auction compared to the optimal static auction, and (2) the Herfindahl-Hirschman Index (HHI) which is a well-known measure to study market concentration.

We first focus on the distribution of advertisers' surplus under both auctions. We compute the log of each advertiser's surplus over all 12,136 sessions and present it in Figure 5. Interestingly, we find that only 3 out of 15 advertisers benefit from the optimal dynamic auction compared to the one with the optimal static auction. In contrast, 9 out of 15 advertisers prefer the optimal static auction. However, it is worth noting that these advertisers have a very small surplus in both cases and their gains are small in magnitude. Therefore, the average advertisers' surplus is higher under

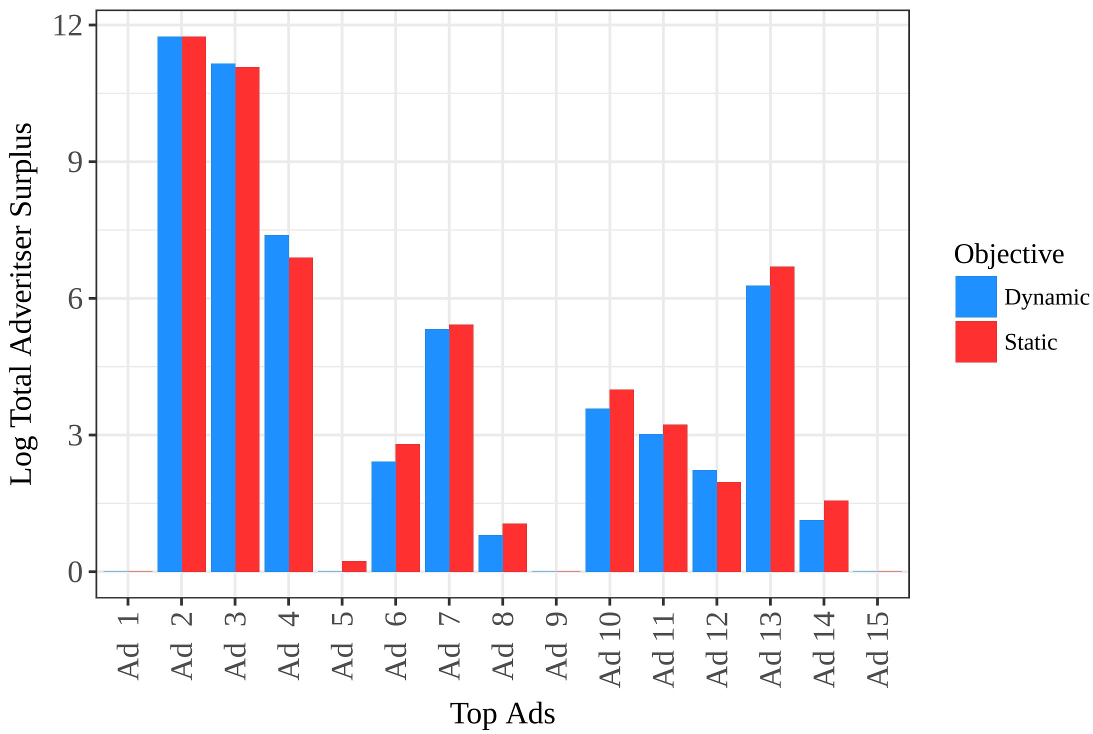


Figure 5: Log advertiser surplus under auctions with both dynamic and static objectives. The values are computed over all 12,136 session.

the optimal dynamic auction.

Next, we focus on the concentration of advertiser surplus in the market and calculate the Herfindahl-Hirschman Index (HHI) under both auctions. While more advertisers prefer the optimal static auction, the concentration in the optimal dynamic auction is lower: the Herfindahl-Hirschman Index (HHI) for the optimal dynamic auction is 0.5285, whereas it is 0.5398 for the optimal static auction. Although both auctions seem quite concentrated as they both allocate most impressions to one ad, optimal dynamic auction achieves a lower concentration by allocating more to the second-largest ad, thereby closing the gap between the two largest advertisers.

7 Conclusion

Mobile in-app advertising has grown exponentially over the last years. One important reason contributing to this growth is the publishers' ability to make adaptive interventions – using the time-varying information about the users within the session to personalize ad interventions for these users. Rafieian (2019) focuses on the match between ads and users as the main outcome of interest and shows that dynamic sequencing of ads improves the match outcome per session, compared static sequencing of ads. While dynamic sequencing of ads leads to improvements in the match outcome, it is not clear whether it can be linked to higher revenues, as advertisers can change their bids in response to the change in the allocation. In this paper, we investigate the revenue gains from adopting a dynamic sequencing framework in a competitive environment, as opposed to a static sequencing framework. We propose a unified framework that contains two key components: (1) a theoretical framework that solves for the revenue-optimal auction with the dynamic objective, and (2) an empirical framework that estimates the counterfactual market outcomes under this auction. Our empirical framework comprises of structural estimation of advertisers' private valuations as well as personalized estimation of the click outcome given any pair of user-ad using machine learning methods. We apply our framework to large-scale data from the leading in-app ad-network of an Asian country. We demonstrate that adopting a dynamic sequencing framework increases the expected revenue by 1.60% compared to the static sequencing framework. We then show that the improved match outcome is the key factor in achieving these gains. Further, we explore other outcomes such as the total surplus (efficiency) and advertisers' surplus and document gains from the dynamic framework over the static framework. Thus, our optimal dynamic auction leads to improvement in all primary market outcomes.

Our paper makes three key contributions to the literature. First, from a methodological standpoint, we present a unified dynamic framework that incorporates both advertisers' and users' behavior and examines the market outcomes under the optimal dynamic auction. To our knowledge, this is the first paper to empirically study the revenue gains from adopting an optimal dynamic auction. Second, we propose a methodological framework for structural estimation of quasi-proportional auction. Our framework adds to the literature on the non-parametric estimation of auctions and can easily be extended to any auction that employs a randomized allocation rule. Finally, from a substantive viewpoint, we show that dynamic sequencing of ads can lead to considerable gains in the publisher's revenue, over the existing auctions. This is particularly important, because the current practice in marketing ignores the gains from using a dynamic framework.

Our findings in this paper have several implications for marketing practitioners. First, we propose a framework that helps publishers evaluate market outcomes under counterfactual auctions. Second, our substantive finding shows that adopting an optimal dynamic auction results in considerable gains in terms of three key market outcomes – revenue, total surplus, and advertisers' surplus. Thus, we expect this finding to inform the publishers' decision as to whether use an optimal dynamic auction. Further, we highlight that adaptive interventions require more careful consideration of players' incentives, when there are strategic players whose actions can affect market outcomes. Thus, our framework has implications for the practitioners and policy makers that increasingly use adaptive interventions in the presence of strategic players.

Nevertheless, there remains some limitations in our study that serve as excellent avenues for future research. First, our optimal dynamic auction involves a bit complex allocation and payment rules. As a result, it may be possible that advertisers will not behave in the expected way. Finding auctions with easier rules that achieve approximately the same outcomes would be an interesting avenue for future study. Second, our structural estimation framework assumes that advertisers are risk-neutral and aware of how their bid affects their probability of winning. However, advertisers' behavior would be different if they are risk-averse. Future research can study risk aversion in probabilistic auctions and empirically identify the risk elements in advertisers' utility function. Finally, while our paper provides counterfactual estimates optimal auctions, we do not run these auctions in the field. An interesting future research is to test these auction in a field setting and examine market outcomes.

References

A. Asadpour, M. H. Bateni, K. Bhawalkar, and V. Mirrokni. Concise bid optimization strategies with multiple budget constraints. In *International Conference on Web and Internet Economics*, pages 263–276. Springer, 2014.

S. Athey and P. A. Haile. Nonparametric approaches to auctions. *Handbook of econometrics*, 6:3847–3965, 2007.

S. Athey and D. Nekipelov. A structural model of sponsored search advertising auctions. In *Sixth Ad Auctions Workshop*, volume 15, 2010.

S. Athey and I. Segal. An efficient dynamic mechanism. *Econometrica*, 81(6):2463–2485, 2013.

P. Bajari, S. Houghton, and S. Tadelis. Bidding for incomplete contracts: An empirical analysis of adaptation costs. *American Economic Review*, 104(4):1288–1319, 2014.

D. P. Baron and D. Besanko. Regulation and information in a continuing relationship. *Information Economics and policy*, 1(3):267–302, 1984.

D. Bergemann and J. Välimäki. The dynamic pivot mechanism. *Econometrica*, 78(2):771–789, 2010.

C. Borgs, J. Chayes, N. Immorlica, K. Jain, O. Etesami, and M. Mahdian. Dynamics of bid optimization in online advertisement auctions. In *Proceedings of the 16th international conference on World Wide Web*, pages 531–540. ACM, 2007.

S. Campo, I. Perrigne, and Q. Vuong. Asymmetry in first-price auctions with affiliated private values. *Journal of Applied Econometrics*, 18(2):179–207, 2003.

H. Choi and C. F. Mela. Online marketplace advertising. *Available at SSRN*, 2016.

B. Edelman, M. Ostrovsky, and M. Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American economic review*, 97(1):242–259, 2007.

eMarketer. Mobile In-App Ad Spending , 2018. URL `https://forecasts-na1.emarketer.com/584b26021403070290f93a5c/5851918a0626310a2c186a5e`.

J. Gallien. Dynamic mechanism design for online commerce. *Operations Research*, 54(2):291–310, 2006.

A. Ghosh and M. Mahdian. Externalities in online advertising. In *Proceedings of the 17th international conference on World Wide Web*, pages 161–168. ACM, 2008.

A. Ghosh and A. Sayedi. Expressive auctions for externalities in online advertising. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 371–380, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-799-8. doi: 10.1145/1772690.1772729. URL `http://doi.acm.org/10.1145/1772690.1772729`.

E. Guerre, I. Perrigne, and Q. Vuong. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574, 2000.

E. Guerre, I. Perrigne, and Q. Vuong. Nonparametric identification of risk aversion in first-price auctions under exclusion restrictions. *Econometrica*, 77(4):1193–1227, 2009.

S. M. Kakade, I. Lobel, and H. Nazerzadeh. Optimal dynamic mechanism design and the virtual-pivot mechanism. *Operations Research*, 61(4):837–854, 2013.

W. Kar, V. Swaminathan, and P. Albuquerque. Selection and ordering of linear online video ads. In *Proceedings of the 9th ACM Conference on Recommender Systems*, RecSys '15, pages 203–210, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3692-5. doi: 10.1145/2792838.2800194. URL `http://doi.acm.org/10.1145/2792838.2800194`.

D. Kempe and M. Mahdian. A cascade model for externalities in sponsored search. In *International Workshop on Internet and Network Economics*, pages 585–596. Springer, 2008.

E. Krasnokutskaya. Identification and estimation of auction models with unobserved heterogeneity. *The Review of Economic Studies*, 78(1):293–327, 2011.

S. Lahaie, D. M. Pennock, A. Saberi, and R. V. Vohra. Sponsored search auctions. *Algorithmic game theory*, pages 699–716, 2007.

H. Li, S. M. Edwards, and J.-H. Lee. Measuring the intrusiveness of advertisements: Scale development and validation. *Journal of advertising*, 31(2):37–47, 2002.

V. Mirrokni, S. Muthukrishnan, and U. Nadav. Quasi-proportional mechanisms: Prior-free revenue maximization. In *Latin American Symposium on Theoretical Informatics*, pages 565–576. Springer, 2010.

R. B. Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

R. B. Myerson. Multistage games with communication. *Econometrica: Journal of the Econometric Society*, pages 323–358, 1986.

D. C. Parkes and S. P. Singh. An mdp-based approach to online mechanism design. In *Advances in neural information processing systems*, pages 791–798, 2004.

A. Pavan, I. Segal, and J. Toikka. Dynamic mechanism design: A myersonian approach. *Econometrica*, 82 (2):601–653, 2014.

O. Rafieian. Optimizing user engagement through adaptive ad sequencing. Technical report, Working paper, 2019.

O. Rafieian and H. Yoganarasimhan. How does variety of previous ads influence consumer's ad response? 2018a.

O. Rafieian and H. Yoganarasimhan. Targeting and privacy in mobile advertising, 2018b.

M. H. Riordan and D. E. Sappington. Information, incentives, and organizational mode. *The Quarterly Journal of Economics*, 102(2):243–263, 1987.

M. Said. Auctions with dynamic populations: Efficiency and revenue maximization. *Journal of Economic Theory*, 147(6):2419–2438, 2012.

H. R. Varian. Position auctions. *international Journal of industrial Organization*, 25(6):1163–1178, 2007.

G. Vulcano, G. Van Ryzin, and C. Maglaras. Optimal dynamic auctions for revenue management. *Management Science*, 48(11):1388–1407, 2002.

S. Yao and C. F. Mela. A dynamic model of sponsored search advertising. *Marketing Science*, 30(3):447–468, 2011.

H. Yoganarasimhan. Estimation of beauty contest auctions. *Marketing Science*, 35(1):27–54, 2015.

Appendices

A Proofs

Proof of Lemma 1. In a mechanism M, we can write the maximum utility advertiser a can receive as follows:

$$\max_{b_a} u^M_{i,a,t}(b_a; x_a, S_{i,t}) = \max_{b_a} \mathbb{E}_{x_{-a}} \left[w_{i,a,t}(x_a; S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t}) - e^M_{i,a,t}(b_a, x_{-a}; S_{i,t}) \right], \tag{52}$$

where the expectation is over other advertisers' click valuations, as we have IC. The first derivative of advertiser a with respect to her click valuation x_a as follows:

$$\frac{\partial u^M_{i,a,t}(b_a; x_a, S_{i,t})}{\partial x_a} = \mathbb{E}_{x_{-a}} \left[P(Y_{i,a} \mid a, S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t}) \right] \tag{53}$$

Given IC constraint, we know that $u^M_{i,a,t}(x_a; x_a, S_{i,t}) = \max_{b_a} u^M_{i,a,t}(b_a; x_a, S_{i,t})$. Therefore, based on envelope theorem, we have:

$$\frac{\partial u^M_{i,a,t}(b_a; x_a, S_{i,t})}{\partial x_a} \Big|_{b_a = x_a} = \mathbb{E}_{x_{-a}} \left[P(Y_{i,a} \mid a, S_{i,t}) q^M_{i,a,t}(x; S_{i,t}) \right] \tag{54}$$

Now, since u is differentiable, we can write:

$$\begin{aligned} u^M_{i,a,t}(x_a; x_a, S_{i,t}) - u^M_{i,a,t}(x'_a; x'_a, S_{i,t}) &= \int_{x'_a}^{x_a} \frac{\partial u^M_{i,a,t}(b_a; x_a, S_{i,t})}{\partial x_a} db_a \\ &= P(Y_{i,a} \mid a, S_{i,t}) \int_{x'_a}^{x_a} \mathbb{E}_{x_{-a}} \left[q^M_{i,a,t}(b_a, x_{-a}; S_{i,t}) \right] db_a \end{aligned} \tag{55}$$

This directly implies Equation (5) and completes the proof for Lemma 1. □

Proof of Lemma 2. We can write the publisher's revenues as follows:

$$\begin{aligned} \mathbb{E}_x \left[\sum_{a \in \mathcal{A}_i} e^M_{i,a,t}(x; S_{i,t}) \right] &= \sum_{a \in \mathcal{A}_i} \mathbb{E}_x \left[w_{i,a,t}(x_a; S_{i,t}) q^M_{i,a,t}(x; S_{i,t}) \right] \\ &\quad - \sum_{a \in \mathcal{A}_i} \mathbb{E}_x \left[w_{i,a,t}(x_a; S_{i,t}) q^M_{i,a,t}(x; S_{i,t}) - e^M_{i,a,t}(x; S_{i,t}) \right], \end{aligned} \tag{56}$$

where each component of the second term is advertiser a's surplus and allows us to write it as follows:

$$\mathbb{E}_x \left[\sum_{a \in \mathcal{A}_i} e^M_{i,a,t}(x; S_{i,t}) \right] = \sum_{a \in \mathcal{A}_i} \mathbb{E}_x \left[w_{i,a,t}(x_a; S_{i,t}) q^M_{i,a,t}(x; S_{i,t}) \right] - \sum_{a \in \mathcal{A}_i} \mathbb{E}_{x_a} \left[u^M_{i,a,t}(x_a; x_a, S_{i,t}) \right] \tag{57}$$

Using calculus theorems and Lemma 1, we can transform each element of the second term. For brevity and with slight abuse of notation, we drop $S_{i,t}$ from the function specification, and define $p_{i,a} = P(Y_{i,a} \mid a, S_{i,t})$.

We can write:

$$
\begin{aligned}
\mathbb{E}_{x_a}\left[u^M_{i,a,t}(x_a;x_a)\right] &= \int_{\underline{x}_a}^{\bar{x}_a} u^M_{i,a,t}(x_a) f_a(x_a) dx_a \\
&= \int_{\underline{x}_a}^{\bar{x}_a}\left(u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) + p_{i,a}\int_{\underline{x}_a}^{x_a}\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(b_a,x_{-a})\right] db_a\right) f_a(x_a) dx_a \\
&= u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) + p_{i,a}\int_{\underline{x}_a}^{\bar{x}_a}\int_{\underline{x}_a}^{x_a}\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(b_a,x_{-a})\right] f_a(x_a) db_a dx_a \\
&= u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) + p_{i,a}\int_{\underline{x}_a}^{\bar{x}_a}\int_{b_a}^{\bar{x}_a} f_a(x_a)\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(b_a,x_{-a})\right] dx_a db_a \\
&= u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) + p_{i,a}\int_{\underline{x}_a}^{\bar{x}_a}\left(1-F_a(b_a)\right)\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(b_a,x_{-a})\right] db_a \\
&= u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) + p_{i,a}\int_{\underline{x}_a}^{\bar{x}_a}\left(1-F_a(b_a)\right)\int_{x_{-a}} q^M_{i,a,t}(b_a,x_{-a}) f_{-a}(x_{-a}) dx_{-a} db_a \\
&= u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) + p_{i,a}\int_{\underline{x}_a}^{\bar{x}_a}\frac{1-F_a(b_a)}{f_a(b_a)}\int_{x_{-a}} q^M_{i,a,t}(b_a,x_{-a}) f_{-a}(x_{-a}) dx_{-a} f_a(b_a) db_a \\
&= u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) + p_{i,a}\int_{x}\frac{1-F_a(b_a)}{f_a(b_a)} q^M_{i,a,t}(b_a,x_{-a}) f(b_a,x_{-a}) dx_{-a} db_a \\
&= u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) + p_{i,a}\int_{x}\frac{1-F_a(x_a)}{f_a(x_a)} q^M_{i,a,t}(x) f(x) dx \\
&= u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) + p_{i,a}\mathbb{E}_x\left[\frac{1-F_a(x_a)}{f_a(x_a)} q^M_{i,a,t}(x)\right]
\end{aligned}
\tag{58}
$$

Now we can transform the publisher's revenues in Equation (57) using Equation (58) to complete the proof:

$$
\begin{aligned}
\mathbb{E}_x\left[\sum_{a\in\mathcal{A}_i} e^M_{i,a,t}(x;S_{i,t})\right] &= \sum_{a\in\mathcal{A}_i}\mathbb{E}_x\left[w_{i,a,t}(x_a;S_{i,t}) q^M_{i,a,t}(x;S_{i,t})\right] \\
&\quad - \sum_{a\in\mathcal{A}_i} u^M_{i,a,t}(\underline{x}_a;\underline{x}_a) - \sum_{a\in\mathcal{A}_i} p_{i,a}\mathbb{E}_x\left[\frac{1-F_a(x_a)}{f_a(x_a)} q^M_{i,a,t}(x)\right] \\
&= \mathbb{E}_x\left[\sum_{a\in\mathcal{A}_i}\left(x_a - \frac{1-F_a(x_a)}{f_a(x_a)}\right) p_{i,a} q^M_{i,a,t}(x;S_{i,t})\right] - \sum_{a\in\mathcal{A}_i} u^M_{i,a,t}(\underline{x}_a;\underline{x}_a,S_{i,t})
\end{aligned}
$$

□

Proof of Proposition 1. Given the payments in Equation (8), it is easy to check that $u^M_{i,a,t}(\underline{x}_a;\underline{x}_a,S_{i,t}) = 0$.

We can write:

$$
\begin{aligned}
u^M_{i,a,t}(\underline{x}_a;\underline{x}_a,S_{i,t}) &= \mathbb{E}_{x_{-a}}\left[w_{i,a,t}(\underline{x}_a;S_{i,t})q^M_{i,a,t}(\underline{x}_a,x_{-a};S_{i,t}) - e^M_{i,a,t}(\underline{x}_a,x_{-a};S_{i,t})\right] \\
&= \mathbb{E}_{x_{-a}}\left[P(Y_{i,t}\mid a,S_{i,t})\int_{\underline{x}_a}^{\underline{x}_a} q^M_{i,a,t}(b_a,x_{-a};S_{i,t})db_a\right] \\
&= 0
\end{aligned} \tag{59}
$$

This implies that the second part of the publisher's revenues in Equation (6) is zero, i.e., $\sum_{a\in\mathcal{A}_i} u^M_{i,a,t}(\underline{x}_a;\underline{x}_a,S_{i,t}) = 0$. Thus, since the q is chosen to maximize the first part of Equation (6) given some constraints, we know that mechanism M is optimal given those constraint. It is now sufficient to show the following two statements: 1) mechanism M is a direct revelation mechanism, and 2) any direct mechanism satisfies the constraints: q is plausible and $\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]$ is increasing in x_a.

We start by showing that M is a direct revelation mechanism. The plausibility is satisfied by definition. We only need to show both IR and IC. Given the payment function, we can write the utility function for advertiser a as follows:

$$
\begin{aligned}
u^M_{i,a,t}(x_a;x_a,S_{i,t}) &= \mathbb{E}_{x_{-a}}\left[w_{i,a,t}(x_a;S_{i,t})q^M_{i,a,t}(x;S_{i,t}) - e^M_{i,a,t}(x;S_{i,t})\right] \\
&= \mathbb{E}_{x_{-a}}\left[P(Y_{i,t}\mid a,S_{i,t})\int_{\underline{x}_a}^{x_a} q^M_{i,a,t}(b_a,x_{-a};S_{i,t})db_a\right] \\
&= P(Y_{i,t}\mid a,S_{i,t})\int_{\underline{x}_a}^{x_a} \mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(b_a,x_{-a};S_{i,t})\right]db_a
\end{aligned} \tag{60}
$$

Given $u^M_{i,a,t}(\underline{x}_a;\underline{x}_a,S_{i,t}) = 0$, this is equivalent to the envelope condition. Since the integral on the RHS is always non-negative, $u_{i,a,t}$ is increasing, which combined with $u^M_{i,a,t}(\underline{x}_a;\underline{x}_a,S_{i,t}) = 0$ imply IR constraint. We now need to show IC constraint is also satisfied. We prove that by contradiction. Suppose that there is an x'_a that gives a higher utility to advertiser a than truthful reporting, given everyone else bidding their true click valuations. Let γ denote the gains advertiser a receives by reporting x'_a instead of x_a. We can write:

$$
\begin{aligned}
\gamma &= u^M_{i,a,t}(x'_a;x_a,S_{i,t}) - u^M_{i,a,t}(x_a;x_a,S_{i,t}) \\
&= u^M_{i,a,t}(x'_a;x'_a,S_{i,t}) - \left(u^M_{i,a,t}(x'_a;x'_a,S_{i,t}) - u^M_{i,a,t}(x'_a;x_a,S_{i,t})\right) - u^M_{i,a,t}(x_a;x_a,S_{i,t}) \\
&= u^M_{i,a,t}(x'_a;x'_a,S_{i,t}) - u^M_{i,a,t}(x_a;x_a,S_{i,t}) - (x'_a - x_a)P(Y_{i,t}\mid a,S_{i,t})\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x'_a,x_{-a};S_{i,t})\right] \\
&= P(Y_{i,t}\mid a,S_{i,t})\left(\int_{x_a}^{x'_a}\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(b_a,x_{-a};S_{i,t})\right]db_a - (x'_a - x_a)\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x'_a,x_{-a};S_{i,t})\right]\right) \\
&= P(Y_{i,t}\mid a,S_{i,t})\int_{x_a}^{x'_a}\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(b_a,x_{-a};S_{i,t}) - q^M_{i,a,t}(x'_a,x_{-a};S_{i,t})\right]db_a
\end{aligned} \tag{61}
$$

Now, given that $\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]$ is increasing in x_a, it is easy to check $\gamma \le 0$ regardless of whether $x'_a > x_a$ or not. This completes the proof of part 1: mechanism M is a direct revelation mechanism.

Now, we show the second part: any direct revelation mechanism satisfies the constraints: q is plausible and $\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]$ is increasing in x_a. A direct revelation mechanism M satisfies IC constraints.

Hence, for $x'_a > x_a$ we can write:

$$u^M_{i,a,t}(x_a; x_a, S_{i,t}) \geq u^M_{i,a,t}(x'_a; x_a, S_{i,t}) \tag{62}$$
$$u^M_{i,a,t}(x'_a; x'_a, S_{i,t}) \geq u^M_{i,a,t}(x_a; x'_a, S_{i,t}) \tag{63}$$

Subtracting these two equations, we have:

$$u^M_{i,a,t}(x_a; x_a, S_{i,t}) - u^M_{i,a,t}(x_a; x'_a, S_{i,t}) \geq u^M_{i,a,t}(x'_a; x_a, S_{i,t}) - u^M_{i,a,t}(x'_a; x'_a, S_{i,t}) \tag{64}$$

Simplifying Equation (74) gives us:

$$(x'_a - x_a)\left(\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x'_a, x_{-a}; S_{i,t})\right] - \mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x_a, x_{-a}; S_{i,t})\right]\right) \geq 0 \tag{65}$$

Since $x'_a > x_a$, we can show that $\mathbb{E}_{x_{-a}}\left[q^M_{i,a,t}(x_a, x_{-a}; S_{i,t})\right]$ is increasing in x_a. This completes the proof for the second part. □

Proof of Lemma 3. The proof for this lemma is almost identical to the proof for Lemma 1. We start by writing down the maximization problem advertiser a faces in mechanism M:

$$\max_{b_a} U^M_{i,a}(b_a; x_a, S_i) = \max_{b_a} \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1}\left(w_{i,a,t}(x_a; S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t}) - e^M_{i,a,t}(b_a, x_{-a}; S_{i,t})\right)\right], \tag{66}$$

where the expectation is over other advertisers' click valuations as well as the stochasticity induce by the dynamic process. The reason we take the expectation over other advertisers' click valuation is the main condition in the lemma: IC constraint. We can write the first derivative of advertiser a with respect to her click valuation x_a as follows:

$$\frac{\partial U^M_{i,a}(b_a; x_a, S_i)}{\partial x_a} = \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t})\right], \tag{67}$$

where, again, the expectation is over other advertisers' click valuations as well as the stochasticity induce by the dynamic process. IC constraint implies that $U^M_{i,a}(x_a; x_a, S_i) = \max_{b_a} U^M_{i,a}(b_a; x_a, S_i)$. We can now apply the envelope theorem as follows:

$$\frac{\partial u^M_{i,a,t}(b_a; x_a, S_{i,t})}{\partial x_a} |_{b_a = x_a} = \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t})\right] \tag{68}$$

Now, due to the first-differentiability of U, we have:

$$\begin{aligned} U^M_{i,a}(x_a; x_a, S_i) - U^M_{i,a}(x'_a; x'_a, S_{i,t}) &= \int_{x'_a}^{x_a} \frac{\partial U^M_{i,a}(b_a; x_a, S_i)}{\partial x_a} db_a \\ &= \int_{x'_a}^{x_a} \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t})\right] db_a, \end{aligned} \tag{69}$$

which is equivalent to Equation (12) and completes the proof for Lemma 3. □

Proof of Lemma 4. The steps for this proof is almost identical to the proof for Lemma 2. We start by writing the publisher's objective function:

$$
\begin{aligned}
\mathbb{E}\left[\sum_{a\in\mathcal{A}_i} e^M_{i,a}(x;S_i)\right] &= \sum_{a\in\mathcal{A}_i}\mathbb{E}_x\left[\sum_{t=1}^{\infty}\beta^{t-1}w_{i,a,t}(x_a;S_{i,t})q^M_{i,a,t}(x;S_{i,t})\right] \\
&\quad \sum_{a\in\mathcal{A}_i}\mathbb{E}_x\left[\sum_{t=1}^{\infty}\beta^{t-1}w_{i,a,t}(x_a;S_{i,t})q^M_{i,a,t}(x;S_{i,t}) - e^M_{i,a}(x;S_i)\right] \\
&= \sum_{a\in\mathcal{A}_i}\mathbb{E}_x\left[\sum_{t=1}^{\infty}\beta^{t-1}w_{i,a,t}(x_a;S_{i,t})q^M_{i,a,t}(x;S_{i,t})\right] - \sum_{a\in\mathcal{A}_i}\mathbb{E}_{x_a}\left[U^M_{i,a}(x_a;x_a,S_i)\right],
\end{aligned}
\tag{70}
$$

where all the expectations are over the specified click valuations and the stochasticity induced by the dynamic process. Now, we can transform each element of the second term. For brevity and with slight abuse of notation, we drop S_i and $S_{i,t}$ from the function specification, and define $p_{i,a,t} = P(Y_{i,a} \mid a, S_{i,t})$. We can write:

$$
\begin{aligned}
\mathbb{E}_{x_a}\left[U^M_{i,a}(x_a;x_a)\right] &= \int_{\underline{x}_a}^{\bar{x}_a} U^M_{i,a}(x_a)f_a(x_a)dx_a \\
&= \int_{\underline{x}_a}^{\bar{x}_a}\left(U^M_{i,a}(\underline{x}_a;\underline{x}_a) + \int_{\underline{x}_a}^{x_a}\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(b_a,x_{-a})\right]db_a\right)f_a(x_a)dx_a \\
&= U^M_{i,a}(\underline{x}_a;\underline{x}_a) + \int_{\underline{x}_a}^{\bar{x}_a}\int_{\underline{x}_a}^{x_a}\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(b_a,x_{-a})\right]f_a(x_a)db_a dx_a \\
&= U^M_{i,a}(\underline{x}_a;\underline{x}_a) + \int_{\underline{x}_a}^{\bar{x}_a}\int_{b_a}^{\bar{x}_a} f_a(x_a)\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(b_a,x_{-a})\right]dx_a db_a \\
&= U^M_{i,a}(\underline{x}_a;\underline{x}_a) + \int_{\underline{x}_a}^{\bar{x}_a}\left(1-F_a(b_a)\right)\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(b_a,x_{-a})\right]db_a \\
&= U^M_{i,a}(\underline{x}_a;\underline{x}_a) + p_{i,a}\int_{\underline{x}_a}^{\bar{x}_a}(1-F_a(b_a))\int_{x_{-a}}\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(b_a,x_{-a})f_{-a}(x_{-a})dx_{-a}db_a \\
&= U^M_{i,a}(\underline{x}_a;\underline{x}_a) + \int_{\underline{x}_a}^{\bar{x}_a}\frac{1-F_a(b_a)}{f_a(b_a)}\int_{x_{-a}}\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(b_a,x_{-a})f_{-a}(x_{-a})dx_{-a}f_a(b_a)db_a \\
&= U^M_{i,a}(\underline{x}_a;\underline{x}_a) + \int_{x}\frac{1-F_a(b_a)}{f_a(b_a)}\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(b_a,x_{-a})f(b_a,x_{-a})dx_{-a}db_a \\
&= U^M_{i,a}(\underline{x}_a;\underline{x}_a) + \int_{x}\frac{1-F_a(x_a)}{f_a(x_a)}\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(x)f(x)dx \\
&= U^M_{i,a}(\underline{x}_a;\underline{x}_a) + p_{i,a}\mathbb{E}_x\left[\frac{1-F_a(x_a)}{f_a(x_a)}\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(x)\right]
\end{aligned}
\tag{71}
$$

Now we can re-write Equation (70) using Equation (71) to complete the proof:

$$
\begin{aligned}
\mathbb{E}_x\left[\sum_{a\in\mathcal{A}_i} e^M_{i,a}(x;S_i)\right] &= \sum_{a\in\mathcal{A}_i}\mathbb{E}_x\left[\sum_{t=1}^{\infty}\beta^{t-1}w_{i,a,t}(x_a;S_{i,t})q^M_{i,a,t}(x;S_{i,t})\right] \\
&\quad -\sum_{a\in\mathcal{A}_i}U^M_{i,a}(\underline{x}_a;\underline{x}_a)-\sum_{a\in\mathcal{A}_i}\mathbb{E}_x\left[\frac{1-F_a(x_a)}{f_a(x_a)}\sum_{t=1}^{\infty}\beta^{t-1}p_{i,a,t}q^M_{i,a,t}(x)\right] \\
&= \mathbb{E}_x\left[\sum_{a\in\mathcal{A}_i}\sum_{t=1}^{\infty}\beta^{t-1}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}\right)p_{i,a,t}q^M_{i,a,t}(x;S_{i,t})\right] \\
&\quad -\sum_{a\in\mathcal{A}_i}U^M_{i,a}(\underline{x}_a;\underline{x}_a,S_i) \\
&= \sum_{a\in\mathcal{A}_i}\mathbb{E}_x\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}\right)p_{i,a,t}q^M_{i,a,t}(x;S_{i,t})\right] \\
&\quad -\sum_{a\in\mathcal{A}_i}U^M_{i,a}(\underline{x}_a;\underline{x}_a,S_i)
\end{aligned}
$$

□

Proof of Proposition 2. The proof for this proposition is very similar to the one for Proposition 1. We begin by providing the necessary and sufficient conditions for the IC constraint. We first show that a mechanism M is IC if and only if $\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty}\beta^{t-1}P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]$ is increasing in x_a and we have the envelope condition as presented in Equation (12).

We start our proof by the only if part. We want to show that if the mechanism M is IC, then $\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty}\beta^{t-1}P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]$ is increasing in x_a and we have the envelope condition as presented in Equation (12). The latter is the result of Lemma 3. So we only need to show the that IC implies monotonicity. Given the IC, for any $x'_a > x_a$ we can write:

$$U^M_{i,a}(x_a;x_a,S_i)\geq U^M_{i,a}(x'_a;x_a,S_i) \tag{72}$$
$$U^M_{i,a}(x'_a;x'_a,S_i)\geq U^M_{i,a}(x_a;x'_a,S_i) \tag{73}$$

Subtracting these two equations, we have:

$$U^M_{i,a}(x_a;x_a,S_i)-U^M_{i,a}(x_a;x'_a,S_i)\geq U^M_{i,a}(x'_a;x_a,S_i)-U^M_{i,a}(x'_a;x'_a,S_i) \tag{74}$$

Simplifying Equation (74) gives us:

$$(x'_a-x_a)\left(\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty}\beta^{t-1}P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]\right)\geq 0 \tag{75}$$

Since $x'_a > x_a$, we can show that $\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty}\beta^{t-1}P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]$ is increasing in x_a. This completes the *only if* part of the statement.

Now, we show the *if* part of the statement. If we have plausibility and monotonicity conditions as above, then we IC constraint is satisfied. We assume that the IC is not satisfied and then show contradiction. If IC is not satisfied, there exists an x'_a that gives a higher utility to advertiser a than truthful reporting, given IC for other advertisers. We denote the gains from deviating by γ. We can write:

$$\begin{aligned}
\gamma &= U^M_{i,a}(x'_a; x_a, S_i) - U^M_{i,a}(x_a; x_a, S_i) \\
&= U^M_{i,a}(x'_a; x'_a, S_i) - \big(U^M_{i,a}(x'_a; x'_a, S_i) - U^M_{i,a}(x'_a; x_a, S_i)\big) - U^M_{i,a}(x_a; x_a, S_i) \\
&= U^M_{i,a}(x'_a; x'_a, S_i) - U^M_{i,a}(x_a; x_a, S_i) \\
&\quad - (x'_a - x_a)\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(x'_a, x_{-a}; S_{i,t})\right] \\
&= \int_{x_a}^{x'_a} \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t})\right] db_a \\
&\quad - (x'_a - x_a)\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(x'_a, x_{-a}; S_{i,t})\right] \\
&= \int_{x_a}^{x'_a} \mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t}) \right. \\
&\quad \left. - \sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(x'_a, x_{-a}; S_{i,t})\right] db_a
\end{aligned} \tag{76}$$

Now, given that $\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(x_a, x_{-a}; S_{i,t})\right]$ is increasing in x_a, it is easy to check $\gamma \le 0$ independent of the relationship between x'_a and x_a. This contradicts the assumption that $\gamma > 0$ and shows that mechanism M is IC.

Now we show that the proposed mechanism is optimal. This mechanism maximizes the first component in Equation (13) subject to the plausibility and monotonicity conditions, while setting the payment such that the second component is zero. Given the payments, we can write:

$$\begin{aligned}
U^M_{i,a}(x_a; x_a, S_i) &= \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1} \left(w_{i,a,t}(x_a; S_{i,t}) q^M_{i,a,t}(x_a, x_{-a}; S_{i,t}) - e^M_{i,a,t}(x_a, x_{-a}; S_{i,t})\right)\right] \\
&= \int_{\underline{x}_a}^{x_a} \mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(b_a, x_{-a}; S_{i,t})\right] db_a
\end{aligned} \tag{77}$$

Equation Equation (77) shows that $U^M_{i,a}(\underline{x}_a; \underline{x}_a, S_i) = 0$ for all a, which implies that we have the envelope condition as in Equation (12). Together, these imply IR constraint. Further, the envelope condition and the monotonicity constraint are equivalent to the IC constraint. Therefore, IC is also satisfied in our case.

Now we only need to show the optimality of the mechanism M. Since $\sum_{a \in \mathcal{A}_i} U^M_{i,a}(\underline{x}_a; \underline{x}_a, S_i) = 0$, the choice of q maximizes the publisher's objective given the plausibility and monotonicity constraint. Since these two constraints are necessary for any direct revelation mechanism, the mechanism M is optimal.

This equivalence proves two statements: 1) IC is satisfied, and 2) any direct revelation mechanism must satisfy monotonicity and plausibility. $\square$

Proof of Lemma 5. We only need to show that $\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty} \beta^{t-1} P(Y_{i,t} \mid a, S_{i,t}) q^M_{i,a,t}(x_a, x_{-a}; S_{i,t})\right]$ increasing in x_a. Let q^M and $q^{M'}$ denote the optimal allocation functions derived by Equation (17) and Equation (18) for click valuation profiles x and x' respectively. Since these mechanisms are optimal for corresponding cases, we can write the following two inequalities:

$$\begin{aligned} &\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(\sum_{a\in\mathcal{A}_i}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x_a;S_{i,t})\right)\right] \\ \geq\ &\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(\sum_{a\in\mathcal{A}_i}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^{M'}_{i,a,t}(x'_a;S_{i,t})\right)\right], \end{aligned} \tag{78}$$

and

$$\begin{aligned} &\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(\sum_{a\in\mathcal{A}_i}\left(x'_a-\frac{1-F_a(x'_a)}{f_a(x'_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^{M'}_{i,a,t}(x'_a;S_{i,t})\right)\right] \\ \geq\ &\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(\sum_{a\in\mathcal{A}_i}\left(x'_a-\frac{1-F_a(x'_a)}{f_a(x'_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^{M}_{i,a,t}(x_a;S_{i,t})\right)\right] \end{aligned} \tag{79}$$

Subtracting these two inequalities will give us the following inequality:

$$\begin{aligned} &\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(\sum_{a\in\mathcal{A}_i}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}-x'_a+\frac{1-F_a(x'_a)}{f_a(x'_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x_a;S_{i,t})\right)\right] \\ \geq\ &\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(\sum_{a\in\mathcal{A}_i}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}-x'_a+\frac{1-F_a(x'_a)}{f_a(x'_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^{M'}_{i,a,t}(x'_a;S_{i,t})\right)\right] \end{aligned} \tag{80}$$

Now, suppose that x and x' are the same at each element except the a-th element, i.e., $x_a \neq x'_a$ and $x_j = x'_j$ for all $j \neq a$. Further, suppose that $x_a > x'_a$. We can then write:

$$\begin{aligned} &\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}-x'_a+\frac{1-F_a(x'_a)}{f_a(x'_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x_a;S_{i,t})\right] \\ \geq\ &\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(x_a-\frac{1-F_a(x_a)}{f_a(x_a)}-x'_a+\frac{1-F_a(x'_a)}{f_a(x'_a)}\right)P(Y_{i,t}\mid a,S_{i,t})q^{M'}_{i,a,t}(x'_a;S_{i,t})\right] \end{aligned} \tag{81}$$

Now, since the distribution F_a is regular, we have:

$$\mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x_a;S_{i,t})\right] \geq \mathbb{E}\left[\sum_{t=1}^{\infty}\beta^{t-1}P(Y_{i,t}\mid a,S_{i,t})q^{M'}_{i,a,t}(x'_a;S_{i,t})\right] \tag{82}$$

The last inequality directly implies $\mathbb{E}_{x_{-a}}\left[\sum_{t=1}^{\infty}\beta^{t-1}P(Y_{i,t}\mid a,S_{i,t})q^M_{i,a,t}(x_a,x_{-a};S_{i,t})\right]$ increasing in x_a and completes the proof. $\square$