Display Ad Measurement Using Observational Data

Srinivas Tunuguntla*

August 31, 2026

Abstract

Estimating the causal effect of digital advertising campaigns using observational data is challenging because exposure is determined by dynamic targeting and bidding decisions in real-time bidding (RTB) auctions. A common empirical strategy is to aggregate data to the consumer level and compare exposed and unexposed users after conditioning on observed characteristics. Recent evidence, however, shows that such methods often produce estimates that differ substantially from experimental benchmarks. We argue that this difficulty reflects a structural mismatch between the sequential auction process that generates ad exposures and the consumer-level representation used in observational analyses. Using auction-level data, we model campaign execution as a sequential decision process and formulate advertising measurement as an off-policy evaluation (OPE) problem. We operationalize a standard doubly robust OPE estimator that combines an exposure model with a continuation-value model for future outcomes, with both components estimated using transformer architectures. We evaluate the approach using randomized holdout experiments from 14 campaigns conducted through a demand-side platform, with advertiser-site visits as the outcome. In this setting, the observational lift estimates have a mean absolute percentage error of 12.3% and a mean percentage error of –3.8% relative to the experimental estimates.

Keywords: Digital advertising; causal inference; off-policy evaluation; real-time bidding; observational data; doubly robust estimation; field experiments

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

1 Introduction

Estimating the causal effect of digital advertising campaigns using observational approaches remains challenging despite the availability of rich, high-dimensional data. The causal effect of a campaign—often referred to as incrementality—is the difference between the outcome realized under the campaign (the factual outcome) and the outcome that would have obtained had the campaign not been run (the counterfactual outcome). The central challenge in advertising measurement is that the counterfactual outcome is not directly observed.

In practice, advertisers often rely on observational approaches that exploit natural variation in ad exposure. Although many consumers may be eligible to receive ads, targeting and bidding algorithms generate heterogeneity in impressions across consumers. A common empirical strategy is to aggregate data to the consumer level over the campaign window, classify consumers as exposed if they receive at least one impression and unexposed otherwise, and compare outcomes across these groups after conditioning on observed consumer characteristics to address selection. Recent evidence, however, shows that consumer-level observational estimates can remain substantially biased relative to experimental benchmarks, even when conditioning on rich covariates using flexible machine learning methods (Gordon et al. 2019; Gordon, Moakler, and Zettelmeyer 2023a).

We argue that this difficulty reflects a structural mismatch between the data-generating process for ad delivery and the representation imposed by consumer-level aggregation. Digital advertising campaigns are executed through real-time bidding (RTB) auctions. For each consumer, the advertiser encounters a sequence of bidding opportunities, where the probability of receiving an impression depends on information available at that moment, including past browsing and purchase behavior as well as prior ad exposures. The effect of an impression is likewise dynamic, depending on its timing and the sequence of prior exposures. As a result, both selection into impressions and the effects of impressions evolve over time, making the relevant counterfactual inherently dynamic. Aggregating this process into a single exposed versus unexposed indicator obscures both the mechanisms governing selection into impressions and the temporal structure of advertising effects. These considerations motivate an auction-level

representation of campaign execution that preserves the sequence of exposure opportunities and outcomes.

Accordingly, we formulate observational advertising measurement in RTB environments as an off-policy evaluation (OPE) problem. We model campaign execution as a sequential decision process in which each auction opportunity is a decision point. During the campaign, consumers receive impressions according to the exposure policy induced by the advertiser's targeting and bidding decisions together with auction outcomes. Our goal is to estimate the outcomes that would arise under a different policy—the no-ads policy—which sets exposure to zero at every auction opportunity. This formulation retains the auction-level structure through which exposure is realized and accommodates both history-dependent selection into impressions and the effects of current exposure on subsequent outcomes. We implement a standard doubly robust OPE estimator (Kallus and Uehara 2020) that combines a model of auction-level exposure probabilities with a continuation-value model of expected future outcomes under the no-ads policy, estimating both using flexible machine learning methods.

Identification requires two sequential analogues of standard assumptions in observational causal inference. First, sequential conditional exogeneity requires that, after conditioning on the observed pre-auction history, residual variation in whether an impression is served be unrelated to the consumer's potential current and future outcomes. In our setting, this condition is plausible because the observed state captures much of the information used by the focal advertiser's bidding system, while residual variation in the auction clearing threshold is generated by competition among many heterogeneous advertisers and may reflect competitor-specific information and shocks that are unrelated to the consumer's response to the focal campaign. Second, sequential overlap requires non-exposure to occur with positive probability at every relevant history. In our setting, this condition is plausible because residual variation in the auction clearing threshold leaves positive probability of losing an auction even after conditioning on the focal advertiser's information. The low auction-level win rates observed in our data provide additional descriptive support for this condition. Under these assumptions and standard regularity conditions, the resulting estimator is consistent and asymptotically normal.

We evaluate the approach using randomized holdout experiments from 14 campaigns conducted in collaboration with a demand-side platform, with advertiser-site visits as the outcome. In each experiment, eligible consumers are randomly assigned to treatment and control groups, with control consumers receiving no ads from the focal campaign. Using only auction-level data from the treatment group, we estimate counterfactual outcomes under the no-ads policy and compare them with outcomes observed in the control group. In this setting, the resulting observational estimates of campaign lift—defined as incrementality expressed as a percentage of the no-ads baseline—have a mean absolute percentage error of 12.3% and a mean percentage error of −3.8% relative to the experimental estimates and substantially smaller errors than the alternative observational approaches considered. Because the control groups provide experimental estimates of outcomes under the no-ads policy, this comparison also assesses the empirical consequences of the identifying assumptions. The relatively small differences between the observational and experimental estimates suggest that potential violations have quantitatively small consequences for estimated lift in this setting.

The rest of the paper is organized as follows. Section 2 reviews the related literature on advertising measurement and sequential causal inference. Section 3 presents the sequential framework for modeling ad delivery and defines the counterfactual quantities of interest. Section 4 describes the field experiments and auction-level data used in our empirical analysis. Section 5 presents the estimation procedure, including the exposure and value-function models and the construction of the doubly robust estimator. Section 6 evaluates the estimator empirically, comparing its performance to experimental benchmarks and to alternative observational approaches. Section 7 concludes.

2 Related Literature

A large literature studies how to measure the incremental effect of digital advertising campaigns. We organize this literature according to the source of information used to identify or predict causal effects: deliberate randomization, quasi-experimental variation generated by advertising-delivery systems, observational adjustment using nonexperimental campaign data, and cross-campaign prediction using prior experimental results.

2.1 Experimental and Quasi-Experimental Approaches

Randomized controlled trials (RCTs) offer a conceptually straightforward approach to measuring incrementality. Eligible consumers are randomly assigned to treatment and control groups; advertising delivery proceeds as usual for treatment-group consumers, while control-group consumers receive no ads from the focal campaign. Despite this simple design, RCTs often require large samples to obtain precise estimates because advertising effects are typically small and outcomes are highly variable (Lewis and Rao 2015; Johnson 2023). Recent work has therefore focused on improving the statistical efficiency of these experiments through designs such as Ghost Ads and extensions for programmatic advertising environments (Johnson, Lewis, and Nubbemeyer 2017b; Barajas and Bhamidipati 2021).

A related literature introduces experimental variation directly into bidding or auction participation. Lewis and Wong (2022) randomize advertiser participation in auctions and use the resulting assignment as an instrument for realized exposure. Waisman, Nair, and Carrion (2025) develop an adaptive bidding procedure that randomizes across bid levels to learn both optimal bids and the associated advertising effects while limiting experimentation costs. These methods address auction-induced selection through deliberate intervention in bidding or auction participation and are therefore especially useful when the platform can modify the delivery system for measurement purposes.

Other approaches exploit naturally occurring or institutionally generated variation rather than deliberately randomizing the focal campaign. Narayanan and Kalyanam (2015) use a regression-discontinuity design around auction-ranking thresholds to estimate the causal effect of position in sponsored-search listings. Hill et al. (2015) exploit variation in whether a served display impression becomes viewable to estimate the causal effect of viewability at the impression level. Gui, Nair, and Niu (2021) use probabilistic auction throttling for budget pacing as an instrument for exposure to identify the average effect of exposure among consumers whose exposure status is affected by throttling.

These quasi-experimental approaches identify causal effects at particular margins of the advertising-delivery process, such as changes in position near an auction-ranking threshold, the

viewability of a served impression, or exposure induced by probabilistic auction participation. However, these local or impression-level effects do not directly recover the overall effect of a campaign. Mapping them into campaign-level incrementality requires additional assumptions about how effects accumulate across repeated exposures and how effects at the identified margin generalize to the campaign's full distribution of exposure opportunities. For example, one may need to assume that impression effects are additive and do not interact through timing, frequency, or prior exposure, or explicitly model these dynamics.

2.2 Observational and Predictive Approaches

A separate literature estimates campaign-level incrementality either by adjusting for nonrandom exposure within the focal campaign or by predicting effects across campaigns using prior experimental results.

A common observational strategy uses variation in realized exposure within a campaign. These approaches aggregate activity over the campaign window, classify consumers according to whether they received one or more impressions, and adjust for observable differences between exposed and unexposed consumers. Gordon et al. (2019) use randomized field experiments to evaluate a broad set of such methods and find that the resulting estimates differ substantially from the corresponding experimental effects. Gordon, Moakler, and Zettelmeyer (2023a) reach a similar conclusion on a much larger scale, showing that even flexible machine learning methods do not reliably reproduce experimental effects.

These findings suggest that the difficulty of nonexperimental advertising measurement may not be resolved solely by increasing the flexibility of the propensity and outcome models. Consumer-level aggregation collapses a sequence of targeting, bidding, and auction outcomes into a single exposed–unexposed indicator. In doing so, it obscures when selection into impressions occurs, what information is available at each exposure opportunity, and how prior exposures affect subsequent delivery and consumer behavior.

An alternative approach is Predicted Incrementality by Experimentation (PIE; Gordon, Moakler, and Zettelmeyer 2023b), which treats prior randomized experiments as labeled observations and learns a campaign-level mapping from pre- and post-campaign features to ex-

perimental estimates of incrementality. PIE and our approach rely on different sources of information and different assumptions. PIE requires a sufficiently large library of comparable prior experiments and stability in the cross-campaign mapping from campaign features to incrementality. Our approach instead uses detailed within-campaign auction histories and relies on sequential conditional exogeneity and overlap. The two approaches therefore apply in different data environments: PIE requires a library of prior experiments, whereas our approach requires detailed auction-level execution data for the focal campaign.[1]

2.3 Dynamic Advertising Response and Policy Evaluation

A substantial literature shows that advertising response can depend on the history, timing, and composition of prior exposures. Braun and Moe (2013) model how the effects of individual impressions depend on the sequence and mix of creatives previously seen by the consumer. Sahni (2015) uses online field experiments to show that the temporal spacing of exposures affects consumer response. Johnson, Lewis, and Nubbemeyer (2017a) document that display-advertising effects can persist after the campaign ends, while Rafieian and Yoganarasimhan (2022) show that the variety of ads previously shown within a mobile session affects response to the next ad. Du et al. (2019) use a recurrent neural network to model how purchase behavior depends on the timing, intensity, and composition of prior exposures. Using randomized bid participation, Lewis and Wong (2022) similarly allow the causal effect of an impression to vary with its timing, attributes, and the consumer's accumulated advertising stock. Together, this literature implies that impressions cannot generally be evaluated independently of the exposure paths in which they occur.

The combination of path-dependent response and history-dependent ad delivery motivates a sequential policy-evaluation formulation. At each auction, the probability of receiving an impression depends on the consumer's evolving history, including prior exposures and behavior. Because these histories also shape current and future response, exposure assignment and outcomes evolve jointly over the campaign. An impression may affect not only the contempo-

[1] When both sources of information are available, the OPE estimate could in principle be included as a post-campaign predictor in a PIE model. Whether it improves prediction beyond other campaign-level features is an empirical question that cannot be evaluated credibly with the 14 campaigns available in our application.

raneous outcome but also subsequent behavior and the histories on which future targeting and bidding decisions are based. Consequently, evaluating the no-ads counterfactual requires accounting jointly for dynamic selection into impressions and the propagation of exposure effects over time.

Estimating the no-ads counterfactual from auction histories generated under the realized campaign is therefore an off-policy evaluation (OPE) problem: data generated under the campaign's exposure policy are used to estimate expected outcomes under a policy that sets exposure to zero. Standard OPE approaches include inverse-probability weighting estimators that reweight observed trajectories toward a target policy (Precup, Sutton, and Singh 2000; Thomas and Brunskill 2016), model-based estimators that approximate the target-policy value function (Fonteneau et al. 2013; Kearns and Singh 2002), and doubly robust estimators that combine weighting and outcome modeling (Dudík, Langford, and Li 2011; Jiang and Li 2016; Kallus and Uehara 2020). Doubly robust estimators are particularly useful when the relevant histories are high-dimensional because their orthogonality allows the nuisance functions to be estimated using flexible machine-learning methods under appropriate rate and regularity conditions (Chernozhukov et al. 2018; Kallus and Uehara 2020).

We contribute to these literatures by formulating advertising measurement in real-time bidding environments as an off-policy evaluation problem. Each auction opportunity is treated as a decision point, the campaign's realized targeting and bidding system induces the behavior policy, and the no-ads counterfactual is represented as a target policy that sets focal-campaign exposure to zero at every opportunity. This formulation targets a campaign-level counterfactual while accounting jointly for history-dependent selection into impressions and the effects of exposures on subsequent outcomes. We operationalize a standard doubly robust OPE estimator (Kallus and Uehara 2020) using auction-level exposure and continuation-value models estimated with transformer architectures. We evaluate the resulting observational estimates against randomized campaign-level benchmarks.

3 Sequential Framework and Identification

3.1 Data Generating Process

We model the execution of a digital advertising campaign as a sequential decision problem. For each consumer, the advertiser participates in a sequence of real-time bidding (RTB) auctions over the campaign horizon, with both exposure and outcomes evolving dynamically as a function of the consumer's observable history.

Let $t = 1, 2, \ldots, T$ index auction opportunities for a given consumer. At each auction t, the advertiser observes a vector of covariates $X_t \in \mathcal{X}$ summarizing information available at that moment. These covariates include auction-level attributes such as page and placement attributes (e.g., ad size, location, and visibility), temporal information (time of day, recency since last visit), and device and browser characteristics. More generally, X_t may also incorporate consumer attributes observed since the previous auction $t-1$, such as interim browsing or purchase behavior, exposures from other advertising channels, or auxiliary data sources; while our empirical application does not observe such information, the framework allows for it.

Based on this information, the advertiser submits a bid $B_t \in \mathbb{R}_+$ at auction t. The auction outcome then determines whether the advertiser serves an impression. Let $D_t \in \{0, 1\}$ denote the exposure indicator, where $D_t = 1$ if an impression is served and $D_t = 0$ otherwise.

Following the exposure decision, the consumer generates an outcome $Y_t \in \mathcal{Y}$. This outcome may be binary (e.g., a conversion) or continuous (e.g., purchase value). Outcomes may depend on the entire history of past exposures and covariates, allowing for dynamic effects such as wear-in, wear-out, and carryover.

Define the pre-auction history at time t as

$$H_t = (X_1, D_1, Y_1, \ldots, X_{t-1}, D_{t-1}, Y_{t-1}, X_t)$$

which summarizes all observed information available to the advertiser immediately prior to bidding at auction t.

The joint distribution of the observed data can be written as the following sequence of

conditional distributions:

$$
\begin{aligned}
X_1 &\sim p(x_1), \\
B_t \mid H_t &\sim p(b_t \mid h_t), \\
D_t \mid H_t, B_t &\sim p(d_t \mid h_t, b_t), \\
Y_t \mid H_t, D_t &\sim p(y_t \mid h_t, d_t), \\
X_{t+1} \mid H_t, D_t, Y_t &\sim p(x_{t+1} \mid h_t, d_t, y_t).
\end{aligned}
\tag{1:Sequential DGP}
$$

This representation allows exposure to depend on both the advertiser's bidding policy and the stochastic outcome of the auction, while allowing outcomes and future states to evolve dynamically with the consumer's history.

The bidding policy $p(b_t \mid h_t)$ captures the advertiser's targeting and bidding behavior during the campaign. Although bidding policies are often implemented as deterministic functions of the observed history, we allow for the more general case of stochastic bidding policies. The exposure mechanism $p(d_t \mid h_t, b_t)$ reflects the stochastic outcome of the auction given the submitted bid and unobserved competing bids and auction rules. Finally, the outcome and state transition distributions describe how consumer behavior and subsequent auction attributes evolve following exposure.

We assume that the submitted bid affects outcomes only through its effect on exposure. Formally,

$$
Y_t \perp\!\!\!\perp B_t \,\Big|\, H_t, D_t.
$$

This reflects the institutional structure of display advertising auctions: conditional on whether an impression is delivered, the bid itself does not directly influence the consumer outcome.

3.2 No-Ads Counterfactual and Estimand

The data described in Section 3.1 are generated under the campaign's observed bidding and auction process. Our objective is to estimate the outcomes that would have been realized under

a counterfactual policy in which the advertiser serves no ads.

Although bids are the advertiser's direct decision variable, we formulate the analysis in terms of exposure D_t, which is jointly determined by the advertiser's bid and the auction outcome. The advertiser's bidding policy $p(B_t \mid H_t)$, together with the auction outcome distribution $p(D_t \mid H_t, B_t)$, induces a stochastic exposure policy

$$\pi_b(D_t \mid H_t), \tag{2:Behavior exposure policy}$$

which gives the probability of exposure at auction t conditional on the observed history. We refer to π_b as the *behavior policy*, because it is the exposure policy that generates the observed exposure assignments during the campaign.

More generally, any exposure policy $\pi(D_t \mid H_t)$ induces an expected cumulative outcome

$$\rho_\pi = \mathbb{E}_\pi\left[\sum_{t=1}^{T} Y_t\right], \tag{3:Policy value}$$

which measures the total expected outcome generated over the campaign horizon under that policy, such as the expected number of conversions or total purchase value per consumer.

The expected cumulative outcome under the behavior policy, ρ_b, is directly identified from the observed data and can be estimated by the sample average of realized outcomes.

Our object of interest is the outcome under the *no-ads counterfactual policy* π_0, which sets exposure to zero at every auction opportunity. Formally, $\pi_0(D_t = 0 \mid H_t) = 1$ for all t and all histories H_t. In our setting, this corresponds to bidding zero at every auction, so that the focal advertiser never serves an impression. Let ρ_0 denote the expected cumulative outcome under this no-ads policy, measuring the total conversions or purchase value that would have been realized in the absence of the focal campaign.

The causal effect of the campaign, or incrementality, is therefore

$$\rho_b - \rho_0. \tag{4:Incrementality}$$

This is a campaign-level policy effect averaged over the full population of eligible consumers, including consumers who receive no impressions under the behavior policy. In the randomized holdout design, it corresponds to the intent-to-treat effect of assignment to the campaign: consumers assigned to the campaign are subject to the behavior exposure policy π_b, whereas consumers assigned to the holdout group are subject to the no-ads policy π_0. Equivalently, if assignment to an exposure policy is itself regarded as the treatment, $\rho_b - \rho_0$ is the average treatment effect of assignment to π_b rather than π_0.[2]

Because ρ_b is directly identified from the observed data, the counterfactual quantity that must be estimated is ρ_0. Estimating ρ_0 using data generated under the behavior policy π_b constitutes an off-policy evaluation (OPE) problem.

3.3 Identification Assumptions

We assume that the data generating process described in Section 3.1 satisfies sequential analogues of the standard exogeneity and overlap assumptions in causal inference (Robins 1997). These assumptions are stated with respect to exposure process and the no-ads counterfactual policy defined in Section 3.2.

For each auction t and exposure level $d \in \{0, 1\}$, let $Y_t(d)$ denote the potential outcome at time t if exposure were set to $D_t = d$, and let $X_{t+1}(d)$ denote the corresponding potential next-period state. These quantities are defined conditional on the prior history H_t, with exposure at auction t intervened upon while the remainder of the environment evolves according to the data-generating process.

Assumption 1 (Sequential Conditional Exogeneity)**.** *For each auction $t = 1, 2, \ldots, T$,*

$$D_t \perp\!\!\!\perp \Big(Y_t(0), Y_t(1), X_{t+1}(0), X_{t+1}(1)\Big) \,\Big|\, H_t.$$

Assumption 1 requires that, conditional on the observed pre-auction history, exposure at

[2]Because realized exposure is a potentially history-dependent sequence of impressions rather than a single binary treatment, standard static treatment-effect terminology does not map uniquely to this setting. In particular, defining an ATT requires specifying the relevant exposure regime. ATT-type estimands conditional on realized impressions are distinct from the campaign-level policy effect studied here and would require additional estimation and inference machinery.

auction t is independent of the contemporaneous potential outcomes and the next-period potential states under either exposure condition. Importantly, the assumption does not require exposure to have no effect on outcomes or on the next-period state: $Y_t(0)$ and $Y_t(1)$ may differ, as may $X_{t+1}(0)$ and $X_{t+1}(1)$. Rather, it requires that, after conditioning on H_t, the residual variation determining whether an impression is served is unrelated to these potential outcomes and states. Because the process evolves recursively, different exposure histories may induce different subsequent trajectories of outcomes, states, and auction opportunities; sequential exogeneity is imposed at each realized decision point along those trajectories.

It is worth noting that exposure may affect the consumer's subsequent trajectory and therefore the timing, attributes, and occurrence of future auction opportunities. This is accommodated by the potential next state $X_{t+1}(d)$, which is allowed to differ across exposure conditions. Thus, different exposure histories may generate different sequences of future auctions, including different numbers of auction opportunities over the campaign window. The sequential framework does not hold future auction opportunities fixed; such path dependence is part of the counterfactual process being modeled.

For notational simplicity, we write the horizon as a common T. In practice, consumers may have different—and potentially exposure-dependent—numbers of realized auction opportunities within the fixed campaign window. This can be represented by allowing the transition to X_{t+1} to determine the timing of the next auction and by treating the end of a consumer's auction sequence as an absorbing terminal state, with subsequent outcomes and exposures set to zero. In our empirical implementation, inter-auction times are included in X_t, and sequences of different lengths are handled by padding shorter sequences with a special terminal token that marks the end of the consumer's auction sequence. Thus, the common-T notation does not require the number or timing of auction opportunities to be invariant to exposure.

Advertiser-side rules such as throttling, frequency caps, conversion-based bidding adjustments, and pacing can similarly depend on prior exposures and outcomes. These rules are incorporated in the history-dependent bidding policy $p(B_t|H_t)$, and hence in the induced exposure policy $\pi_b(D_t \mid H_t)$. For example, bids may change after prior exposures or conversions,

both of which are contained in the observed history. Importantly, in our data these rules do not remove auction opportunities from the observed sequence: we observe each bid request even when the DSP responds by submitting a bid of zero. Consequently, such rules affect the bidding and exposure process but do not create policy-dependent missing auction opportunities.

Assumption 2 (Sequential Overlap)**.** *There exists a constant $\varepsilon > 0$ such that, for all t and all histories h_t in the support of H_t,*

$$\pi_b(D_t = 0 \mid H_t = h_t) \;\geq\; \varepsilon.$$

Assumption 2 requires that, conditional on the observed history, non-exposure occurs with positive probability. This condition ensures that the behavior policy induces sufficient variation in exposure to identify counterfactual outcomes under the no-ads policy. In auction environments, this requirement is typically satisfied mechanically, as advertisers lose a substantial fraction of auctions even after submitting positive bids.

Discussion. The two identifying assumptions can be understood more concretely through the auction clearing threshold. Let C_t denote the minimum bid required to win auction t, as determined by competing bids and the auction rules, so that exposure $D_t = \mathbf{1}\{B_t \geq C_t\}$. Given the focal advertiser's bidding policy $p(B_t \mid H_t)$, a useful sufficient condition for sequential exogeneity and overlap is that, conditional on H_t, residual variation in C_t is unrelated to the focal consumer's potential outcomes and states and leaves positive probability that the focal advertiser loses the auction. The first condition provides exposure variation that is conditionally unrelated to consumer responsiveness, while the second ensures support for the no-ads policy.

This condition is plausible in many RTB environments because the clearing threshold is generated by competition among a large number of advertisers whose valuations depend partly on information private to them. Conditioning on the observed pre-auction history removes variation associated with information available to the focal advertiser and with observable consumer and contextual characteristics. The remaining variation in the clearing threshold may then reflect competitor-specific information and idiosyncratic shocks. When the competing

advertisers come from heterogeneous and largely unrelated product categories, such residual variation is less likely to be systematically related to the consumer's potential response to the focal campaign. The presence of many competing advertisers further limits dependence on the information or valuation of any single competitor.

The argument is less compelling in settings with a small number of closely related competitors or where competing advertisers condition their bids on private information that is also informative about outcomes for the focal advertiser. For example, competing firms in the same product category may target consumers using similar latent demand signals, and common-value settings may generate correlation between competitors' private information and the consumer's potential outcomes for the focal advertiser. In such environments, residual variation in the clearing threshold may not satisfy sequential exogeneity, and alternative sources of identifying variation may be required.

This characterization also clarifies which variables should enter the observed history. H_t should contain the pre-auction information used by the focal advertiser's targeting and bidding system, as well as observable variables that may jointly affect auction competition and consumer outcomes. Omitting such information can leave exposure correlated with potential outcomes. At the same time, more conditioning is not necessarily beneficial. In particular, if the realized clearing threshold C_t is observed after the auction, conditioning additionally on C_t would eliminate the auction-side variation on which our identification argument relies. When the focal bid is a deterministic function of H_t, exposure becomes deterministic conditional on H_t and C_t, violating overlap. Thus, observing the clearing threshold does not by itself solve selection on unobservables: our approach requires the residual variation in that threshold, conditional on H_t, to be exogenous. An alternative is deliberate randomization in the focal advertiser's bidding policy. If bids contain independent variation with support on both sides of the clearing threshold, exposure can remain stochastic even after conditioning on auction-side information.

Our use of competing-bid variation differs from regression-discontinuity approaches such as Narayanan and Kalyanam (2015). Those approaches exploit continuity around an auction-ranking threshold to identify a local effect for auctions near the cutoff. We do not use the

clearing threshold as an instrument or restrict attention to auctions near a cutoff. Instead, variation in the clearing threshold motivates the sequential exogeneity and overlap conditions. Under these assumptions across the relevant histories, the sequential framework identifies the value of the no-ads policy and hence the campaign-level policy effect $\rho_b - \rho_0$, rather than a local exposure effect.

These arguments provide reasons why the identifying assumptions may be plausible in our setting. Nevertheless, they remain untestable from observational data alone. We therefore assess their empirical consequences using the randomized holdout experiments in Section 6. The experimental control groups directly identify outcomes under the no-ads policy, allowing us to compare them with the counterfactual outcomes estimated from treatment-group auction histories. The close agreement between the two suggests that violations of the identifying assumptions have quantitatively limited consequences for estimated lift in our empirical setting.

3.4 A Simple Illustration

Before introducing the general off-policy evaluation estimator, we illustrate the sequential structure of the problem using a simple example. We begin with a single auction, where the problem reduces to a familiar static counterfactual comparison, and then introduce a second auction in which exposure can affect subsequent consumer state, future exposure, and future outcomes.

One Auction. Suppose each consumer encounters a single auction opportunity and that all consumers have the same observed state. Let $D_1 \in \{0, 1\}$ indicate whether an impression is served and let $Y_1(d)$ denote the potential outcome when exposure is set to d. Under the behavior policy, define

$$e_b(d) = \Pr_{\pi_b}(D_1 = d), \qquad \mu(d) = \mathbb{E}[Y_1(d)], \qquad d \in \{0, 1\}.$$

Thus, $e_b(d)$ is the probability of exposure level d under the observed campaign, while $\mu(d)$ is the expected potential outcome under that exposure level. Under conditional exogeneity, $\mu(d) = \mathbb{E}[Y_1(d) \mid D_1 = d]$, so these potential-outcome means are identified from the observed outcomes.

The expected outcome under the behavior policy is therefore $\rho_b = e_b(0)\mu(0) + e_b(1)\mu(1)$,

whereas the expected outcome under the no-ads policy is $\rho_0 = \mu(0)$. Campaign incrementality is consequently $\rho_b - \rho_0 = e_b(1)\Big[\mu(1) - \mu(0)\Big]$.

With a single auction, there is no distinction between the sequential problem and a standard static causal-inference problem: estimating the no-ads counterfactual amounts to estimating the expected potential outcome under $D_1 = 0$.

Two Auctions. Now suppose each consumer encounters two auction opportunities. There is no state before the first auction. Between the two auctions, a state $X_2 \in \{H, L\}$ is realized and observed before the second auction, where H and L may represent high- and low-intent browsing states. Importantly, X_2 may depend on the first exposure D_1. Thus, an impression at the first auction may change the consumer's subsequent behavior and, consequently, the state observed at the second auction.

For simplicity, suppose the relevant history at the second auction is summarized by $H_2 = (D_1, X_2)$. The probability of receiving an impression at the second auction may depend on this history, so that $\Pr(D_2 = 1 \mid D_1, X_2)$ can vary with both the first exposure and the state X_2. The outcome may likewise depend on D_1, D_2, and X_2, allowing the first exposure to affect outcomes both through its effect on the subsequent state and through any direct carryover effect.

Under the no-ads policy, exposure is set to zero at both auctions. By sequential exogeneity, the expected outcome under this policy can be written in terms of observed conditional distributions as

$$\rho_0 = \mathbb{E}[Y_1 \mid D_1 = 0] + \sum_{x \in \{H,L\}} \Pr(X_2 = x \mid D_1 = 0)\mathbb{E}[Y_2 \mid D_1 = 0, X_2 = x, D_2 = 0].$$

The key difference from the one-auction setting is that the no-ads counterfactual requires not only setting exposure to zero at each auction, but also propagating the state distribution generated by those exposure decisions. In particular, the relevant distribution of X_2 is the distribution that arises when $D_1 = 0$, which need not equal the state distribution observed under the behavior policy.

This counterfactual can be expressed recursively using the value functions that appear in the general OPE formulation. Define the second-auction continuation value

$$v_2(d_1, x) = \mathbb{E}\left[Y_2 \mid D_1 = d_1, X_2 = x, D_2 = 0\right].$$

The corresponding value before the first auction is

$$v_1 = \mathbb{E}\left[Y_1 + v_2(0, X_2) \mid D_1 = 0\right],$$

so that $\rho_0 = v_1$. The continuation value v_2 computes the expected future outcome if exposure is turned off from the second auction onward, while v_1 integrates this continuation value over the state distribution generated by turning exposure off at the first auction as well.

A Numerical Illustration. Consider the following parameterization, with relatively large exposure and outcome probabilities chosen to make the mechanics transparent. Unless otherwise noted, all probabilities refer to the behavior policy. At the first auction, $\Pr(D_1 = 1) = 0.5$. The first exposure changes the distribution of the second-period state:

$$\Pr(X_2 = H \mid D_1 = 0) = 0.4, \qquad \Pr(X_2 = H \mid D_1 = 1) = 0.6.$$

At the second auction, exposure probabilities depend on both the prior exposure and the current state:

$$\Pr(D_2 = 1 \mid D_1 = 0, X_2 = H) = 0.8, \quad \Pr(D_2 = 1 \mid D_1 = 0, X_2 = L) = 0.2,$$
$$\Pr(D_2 = 1 \mid D_1 = 1, X_2 = H) = 0.6, \quad \Pr(D_2 = 1 \mid D_1 = 1, X_2 = L) = 0.1.$$

The lower second-auction exposure probabilities following $D_1 = 1$ may, for example, reflect a soft frequency capping.

For simplicity, let $Y_1 = 0$ and suppose that the expected second-period outcome is

$$\mathbb{E}\left[Y_2 \mid D_1 = d_1, X_2 = x, D_2 = d_2\right] = b(x) + 0.04d_1 + 0.03d_2, \tag{5}$$

where $b(H) = 0.20$ and $b(L) = 0.05$. The example therefore contains three sources of dynamics. First, the initial impression changes the distribution of the subsequent consumer state. Second, that state affects the probability of receiving another impression. Third, the initial impression has a carryover effect on the later outcome.

Under this data-generating process, the expected outcome under the no-ads policy is $\rho_0 = 0.11$, while the expected outcome under the behavior policy is $\rho_b = 0.1576$, implying a true campaign lift of 43.3%. A naive consumer-level comparison between consumers who receive at least one impression and those who receive none instead implies a lift of 120.6%. After conditioning on observed consumer characteristics—in this example, just X_2—a consumer-level outcome adjustment implies a lift of 26.1%. Detailed calculations are provided in Web Appendix A.

Both consumer-level approaches fail because they do not reproduce the distribution of consumer states that would arise under the no-ads policy, even though the sequential identifying assumptions hold in the example. The naive exposed–unexposed comparison selects on the realized exposure path: high-intent consumers are more likely to receive an impression at the second auction and are therefore underrepresented among consumers who happen to remain unexposed. Conditioning on X_2 does not fully solve the problem because X_2 is itself affected by the first exposure. A consumer-level adjustment averages conditional outcomes over the observed distribution of X_2, which is partly generated by the behavior policy, rather than over the distribution that would arise if the first exposure were set to zero. Web Appendix A shows these differences explicitly in the numerical example. The sequential formulation instead treats X_2 as part of the history at the second auction while also accounting for how earlier exposure decisions change its distribution.

Connection to Doubly Robust Estimation. The calculations above treat the exposure probabilities and conditional expectations as known. With observational data, these quantities must instead be estimated. The single-auction case illustrates the doubly robust construction used below. Let $\bar{e}(0)$ and $\bar{\mu}(0)$ denote possibly misspecified values of the probability of non-

exposure and the expected outcome under non-exposure. The corresponding score is

$$\psi^{(1)} = \bar{\mu}(0) + \frac{\mathbf{1}\{D_1 = 0\}}{\bar{e}(0)} \left[Y_1 - \bar{\mu}(0) \right].$$

Its bias satisfies

$$\mathbb{E}\left[\psi^{(1)}\right] - \rho_0 = \frac{\bar{e}(0) - e_b(0)}{\bar{e}(0)} \left[\bar{\mu}(0) - \mu(0) \right].$$

Thus, errors in the exposure and outcome models enter multiplicatively: the score remains unbiased if either component is correctly specified, and small estimation errors have no first-order effect around the truth. This is the familiar double-robustness and orthogonality property. In the sequential setting, the outcome regression is replaced by a continuation value and the exposure probability by a sequence of history-dependent exposure probabilities. The next subsection gives the resulting score for an arbitrary number of auctions.

3.5 Efficient Influence Function

Under the identification assumptions in Section 3.3, the estimand ρ_0 admits an efficient influence function representation (Kallus and Uehara 2020). The two-auction score in the preceding example is a special case of this representation. We now define the value functions and density ratios needed for the general setting.

Value and Quality Functions. Consider a generic target exposure policy π. Define the *quality function* (or q-function) as:

$$q_t(H_t, d) \;=\; \mathbb{E}_\pi\left[\sum_{k=t}^{T} Y_k \,\middle|\, H_t, D_t = d \right].$$

Here, exposure at auction t is set to d, subsequent exposures follow the target policy π, and $q_t(H_t, d)$ gives the corresponding expected cumulative outcome from auction t onward.

Similarly, define the *value function* (or v-function) as:

$$v_t(H_t) \;=\; \mathbb{E}_\pi\left[\sum_{k=t}^{T} Y_k \,\middle|\, H_t \right].$$

Equivalently,

$$v_t(H_t) \;=\; \mathbb{E}_{D_t \sim \pi(\cdot \mid H_t)}[q_t(H_t, D_t) \mid H_t]\,.$$

The value function is therefore the expected continuation value immediately before the exposure decision at auction t, generalizing the continuation values introduced in the two-auction example. Under the no-ads policy π_0, which deterministically sets $D_t = 0$, this simplifies to

$$v_t(H_t) = q_t(H_t, 0).$$

Throughout, expectations over future trajectories in the definitions of q_t and v_t are taken under the target policy π, which in our setting is the no-ads policy π_0. Unless explicitly stated otherwise, all other expectations are taken under the observed behavior policy π_b.

Density Ratios. For a target exposure policy π, define the per-auction density ratio relative to the behavior policy π_b as

$$\eta_t(H_t, D_t) = \frac{\pi(D_t \mid H_t)}{\pi_b(D_t \mid H_t)}.$$

The cumulative density ratio up to auction t is then

$$\lambda_t(H_t, D_t) = \prod_{k=1}^{t} \eta_k(H_k, D_k), \qquad \lambda_0 \equiv 1.$$

Because the target policy is the no-ads policy in our setting, the per-auction density ratio simplifies to

$$\eta_t = \frac{\mathbf{1}\{D_t = 0\}}{\pi_b(D_t = 0 \mid H_t)}, \qquad \text{(6:No-ads density ratio)}$$

and the cumulative density ratio λ_t reweights observed histories so that they follow the distribution induced by the no-ads policy. Intuitively, λ_t accumulates inverse-probability weights that correct for history-dependent selection into exposure across auctions.

Efficient Influence Function. Kallus and Uehara (2020) show that the efficient influence

function for ρ_π, the expected cumulative outcome under target policy π, is

$$\phi_\pi(H) = -\rho_\pi + \sum_{t=1}^{T} \lambda_{t-1} \left[\eta_t \Big(Y_t - q_t(H_t, D_t) \Big) + v_t(H_t) \right],$$

where H denotes the full observed trajectory for a consumer.

The structure of the influence function generalizes the doubly robust scores illustrated in the preceding example. At auction t, the cumulative density ratio λ_{t-1} reweights the observed history toward histories generated by the target policy, while η_t adjusts for the difference between the target and behavior exposure probabilities at the current auction. The q- and v-functions provide the corresponding outcome-model adjustment: $q_t(H_t, D_t)$ is the action-specific continuation value after the current exposure decision, whereas $v_t(H_t)$ is the continuation value immediately before that decision. Together, these terms combine inverse-probability weighting with recursive outcome modeling.

We now specialize this expression to the no-ads policy π_0. Under this policy, exposure is zero at every auction, so $v_t(H_t) = q_t(H_t, 0)$. Substituting this relation together with the density ratio in Equation (6:No-ads density ratio) yields

$$\phi(H) = -\rho_0 + \sum_{t=1}^{T} \lambda_{t-1} \left[\frac{\mathbf{1}\{D_t = 0\}}{\pi_b(D_t = 0 \mid H_t)} \Big(Y_t - v_t(H_t) \Big) + v_t(H_t) \right]. \tag{7:No-ads EIF}$$

Equation (7:No-ads EIF) is the general multi-auction counterpart of the one- and two-auction doubly robust scores illustrated in Section 3.4. It accounts simultaneously for history-dependent exposure probabilities and for the propagation of current exposure decisions through future states and outcomes.

Because the efficient influence function has mean zero,

$$\mathbb{E}[\phi(H)] = 0,$$

Equation (7:No-ads EIF) provides an estimating equation for ρ_0. An estimator can therefore be constructed by replacing the population expectation with its sample analogue and unknown

functions with estimates.

The influence function depends on two nuisance components: (i) the exposure model $\pi_b(D_t = 0 \mid H_t)$, which captures history-dependent selection into impressions, and (ii) the value functions v_t, which summarizes expected future outcomes under the no-ads policy. Neither component is of intrinsic interest; they enter the influence function jointly to combine inverse-probability reweighting with recursive outcome adjustment.

A key property of the influence-function moment is *double robustness*. The moment condition $\mathbb{E}[\phi(H)] = 0$ continues to hold if either the exposure model is correctly specified or the sequence of value functions is correctly specified, even if the other component is misspecified. Thus, the resulting estimator remains consistent under partial misspecification of the nuisance components.

A related property is *Neyman orthogonality* (Chernozhukov et al. 2018): small estimation errors in the nuisance functions have no first-order effect on the moment condition. As illustrated in the preceding example, the leading bias instead depends on interactions between errors in the exposure and value-function estimates. This allows the nuisance functions to be estimated using flexible machine-learning methods, subject to appropriate convergence and regularity conditions. When both nuisance components are estimated sufficiently well, the resulting estimator is asymptotically linear and normal and achieves the semiparametric efficiency bound.

Section 5 describes how we estimate these nuisance components in practice and combine them to construct the estimator.

4 Field Experiments and Data

We evaluate our approach using field experiments conducted in collaboration with a demand-side platform (DSP). Across 14 campaigns, eligible consumers were randomly assigned to treatment and holdout control groups, with control consumers receiving no exposure to the focal campaign. These experiments provide a clean benchmark for the no-ads counterfactual defined in Section 3: for control consumers, realized outcomes correspond to the environment in which the focal advertiser does not participate in auctions.

The experimental design therefore allows us to directly observe outcomes under both the

realized campaign policy and the no-ads counterfactual. These experimental benchmarks are used in Section 6 to assess the empirical performance of the observational estimates constructed from auction-level data.

This section describes the RTB institutional environment in which the campaigns were executed, the experimental design, and the structure of the auction-level data used in the empirical analysis.

4.1 RTB Environment

Digital display advertising is implemented through real-time bidding (RTB) auctions, in which advertisers compete for impression opportunities as consumers browse publisher webpages. In practice, advertisers participate in these auctions through demand-side platforms (DSPs) that manage bidding and targeting decisions on their behalf. The campaigns in our study were executed through one such DSP, which operates across multiple ad exchanges and publishers. The DSP's clients span multiple verticals, including retail, travel, finance, and other consumer-facing industries.

Each campaign may include multiple creatives served over the campaign horizon. These creatives typically share a common message and differ modestly in imagery or text. Our object of interest is the incremental effect of the campaign as a whole rather than creative-level heterogeneity. Accordingly, we treat creatives within a campaign as interchangeable and define exposure at the campaign level: an impression is coded as exposure to the focal campaign regardless of the specific creative served.

Consumers are identified using a persistent DSP-level identifier constructed from cookies, device identifiers, and related technologies. This identifier allows us to observe a consumer's auction opportunities and exposures across publishers within the campaign window. While no digital identity system is free from measurement error, the identifier used here is more robust than single-cookie tracking and provides sufficiently stable linkage for our analysis. We do not explicitly adjust for cross-device or cross-browser fragmentation.

Conversions are measured using advertiser-installed tracking pixels that record predefined consumer actions, such as website visits, purchases, or registrations. Campaigns may track

multiple conversion types; in our empirical analysis, the primary metric is website visits. A conversion is therefore defined as a browsing session on the advertiser's website. Following prior research and industry practice (Jansen et al. 2007), we define a session as a continuous period of activity in which successive events are separated by no more than 30 minutes. The DSP observes these conversions for all consumers, irrespective of experimental assignment.

4.2 Experimental Design

For each campaign, the advertiser defines a set of eligible consumers based on pre-specified targeting criteria, such as geography, device type, language, or behavioral segments. Among these eligible consumers, the DSP implements a randomized holdout design: consumers are randomly assigned to either a treatment group or a control group at the outset of the campaign and retain that assignment throughout the campaign horizon.

Consumers assigned to the treatment group are eligible to receive impressions from the focal campaign. For these consumers, the advertiser's targeting and bidding algorithms operate as usual: at each auction opportunity, bids are computed as a function of the consumer's observed history and the contextual attributes of the impression. As a result, exposure within the treatment group is heterogeneous. Some auctions are won and others are lost, generating variation in both the extensive margin (whether a consumer receives any impressions) and the intensive margin (the number and timing of impressions received). Importantly, even within the treatment group, a substantial fraction of consumers receive zero exposure due to auction competition and the advertiser's bidding strategy.

Consumers assigned to the control group are not eligible to receive impressions from the focal campaign. Operationally, this is implemented by submitting a bid of zero for all auction opportunities associated with control consumers. Consequently, control consumers are never exposed to the campaign during the experimental period.

This design directly implements the no-ads counterfactual defined in Section 3. For control consumers, the realized sequence of auction outcomes corresponds to the environment in which the focal advertiser does not participate. If the focal advertiser would have won a given auction absent randomization, the next-highest bidder instead wins the impression for control

consumers. Thus, outcomes for control consumers represent realized behavior under removal of the focal campaign from the auction.

The validity of this interpretation relies on the stability of the auction environment to the removal of the focal bidder. In particular, it requires that other advertisers' bidding strategies do not adjust in response to the focal campaign over the experimental horizon. This condition is likely to hold in our setting. Campaigns are not pre-announced, run over relatively short time periods, and competing advertisers typically observe only whether they win or lose an auction rather than the full bid landscape. Consequently, competitors have limited information and limited ability to condition their strategies on the focal campaign's presence. Under these conditions, the control group provides a credible estimate of outcomes under the no-ads policy.

4.3 Data

Our dataset contains 14 campaigns implemented with randomized treatment and control groups. The sample was constructed from campaigns active on the collaborating DSP during 2025 for which randomized experiments were enabled and the necessary advertiser-specific data-sharing agreements were in place. From this set, we included campaigns that tracked advertiser-site visits as a conversion outcome, ran for at least two weeks, and contained at least 100,000 consumers in the control group. We focus on advertiser-site visits because this outcome is measured more consistently across advertisers than the other conversion metrics available in the DSP data. We did not otherwise screen campaigns based on realized impression delivery, the magnitude or statistical significance of the experimental effect, or the realized precision of the experimental benchmark. The resulting campaigns span multiple industries, including automotive, electronics, travel, retail, e-commerce, finance, food and beverage, and media and entertainment. None of the advertisers are direct product competitors. Although the inclusion criterion required campaigns to run for at least two weeks, the campaigns in the resulting sample each ran for approximately one month.

Table 1 reports summary statistics for each campaign. The studies vary in scale. The smallest campaign includes approximately 0.93 million eligible consumers, while the largest includes approximately 2.83 million. Treatment shares range from 70% to 85%, with the remaining con-

sumers assigned to control. Across campaigns, the total number of impressions varies widely, reflecting differences in bidding intensity, targeting criteria, and auction competitiveness. The outcome metric for all campaigns is website visits, measured using advertiser-installed conversion pixels. A conversion corresponds to a browsing session on the advertiser's website, defined as a sequence of events separated by no more than 30 minutes.

Table 1: Campaign-Level Summary Statistics

Campaign	Consumers	Treatment (%)	Control (%)	Impressions	Conversions
1	2.19M	70	30	3.33M	28,820
2	2.83M	75	25	10.47M	63,221
3	0.93M	70	30	5.84M	8,158
4	1.94M	80	20	2.06M	34,868
5	2.07M	80	20	16.05M	36,271
6	2.33M	70	30	3.31M	19,078
7	1.69M	85	15	3.76M	9,939
8	1.75M	70	30	0.67M	5,202
9	2.02M	70	30	0.82M	35,552
10	2.59M	80	20	2.47M	18,482
11	1.12M	70	30	2.1M	5,414
12	1.66M	70	30	5.83M	19,228
13	2.41M	70	30	2.61M	16,949
14	1.62M	75	25	1.95M	26,945

At the most granular level, the data are organized at the consumer–auction level. For each eligible consumer, we observe the full sequence of auction opportunities over the campaign horizon, together with exposure and outcome realizations.

For each auction, the dataset includes the consumer identifier, a timestamp, and the consumer's experimental assignment (treatment or control). In addition, we observe detailed auction-level covariates, including publisher and contextual attributes such as a domain identifier (with the top 4,095 domains separately identified and the remainder grouped into an "other" category) and a publisher vertical classification defined by the DSP (e.g., news, shopping, travel, finance, or entertainment).

More importantly, for each auction we observe a time-varying 512-dimensional embedding that summarizes the information available to the advertiser at the time of bidding. The

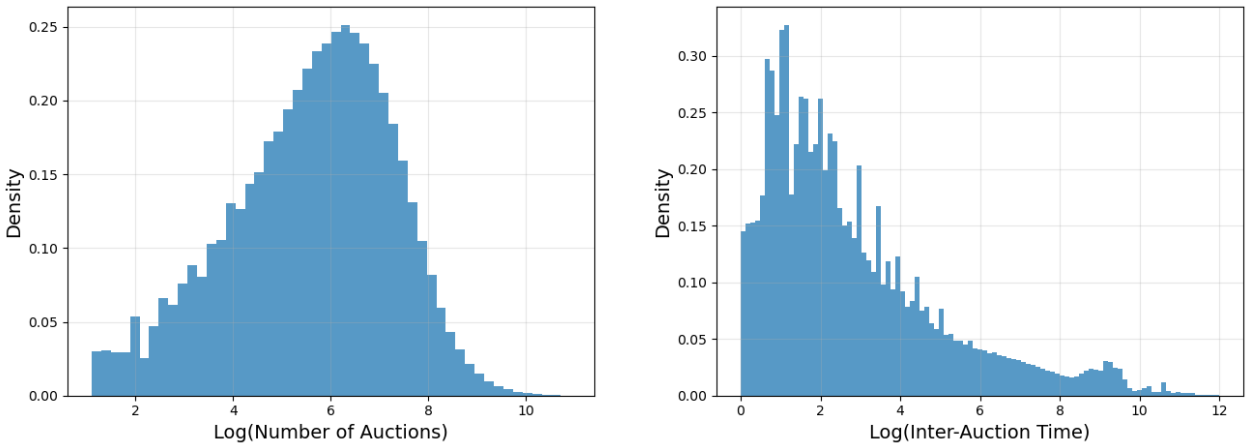
DSP's bidding system constructs this embedding using pre-trained machine learning models. Consumer-level inputs include demographic attributes, inferred interests or affinity groups, and historical engagement measures such as click and conversion rates by vertical, some of which are computed using recent browsing history. Contextual inputs include attributes of the bid request such as time of day, device type, ad slot size and visibility, and webpage content or URL—where URL information is encoded using a pre-trained language model and commonly occurring URLs are pre-parsed. These consumer and contextual representations are combined to form a 512-dimensional embedding representing the observable state at that auction. We refer to this 512-dimensional representation as the DSP auction embedding, denoted $E_t \in \mathbb{R}^{512}$ at auction t.

Although we do not observe each raw input separately, the DSP auction embedding E_t captures the high-dimensional information set used by the advertiser's targeting and bidding algorithm. Together with the auction timestamp, domain, and vertical identifiers, it forms the observable state vector X_t described in Section 3. Because bids are computed as a function of the recent history of these embeddings, together with prior exposures and outcomes, conditioning on this history allows us to model the information governing selection into exposure.

Finally, for each auction t, we observe the bid submitted by the advertiser B_t, whether the advertiser won the impression D_t, and whether a conversion occurred Y_t. The conversion indicator equals one if a conversion occurs after auction t and before the next auction opportunity $t+1$ for the same consumer, and zero otherwise.

These variables provide a detailed auction-level panel for each consumer, containing the observable history used by the advertiser's targeting and bidding policy, the realized exposure sequence, and subsequent outcomes. This structure maps directly to the sequential data generating process formalized in Section 3.

Data Summary. Figure 1 summarizes the distribution of the number of auctions and inter-auction times at the consumer level. Across campaigns, the number of auctions per consumer is highly dispersed—ranging from 1 to 55,109—with a mean of 807 and a median of 322. Inter-auction times also exhibit substantial variation, with many auctions occurring in rapid



(a) Histogram of log auctions

(b) Histogram of inter-auction times

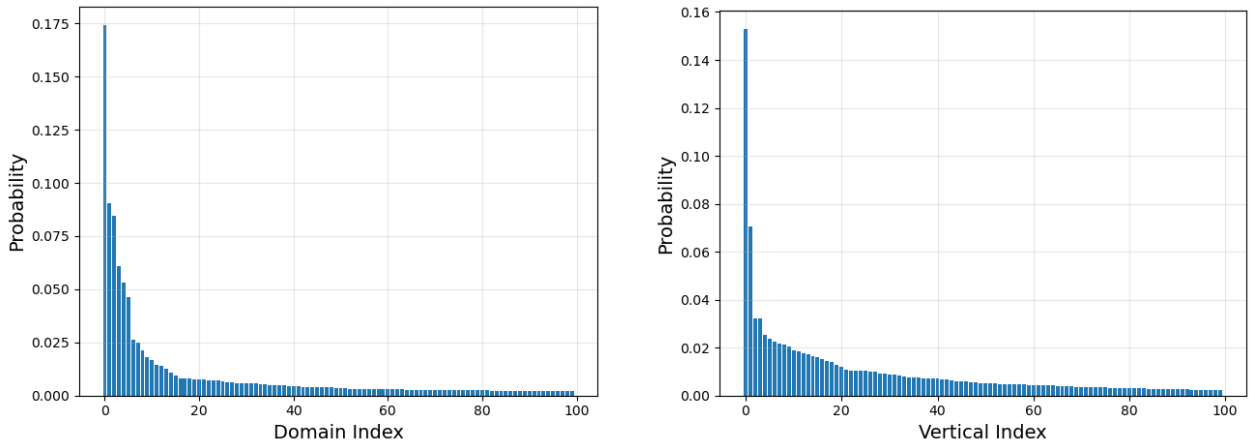
Figure 1: Heterogeneity in consumer browsing activity

succession and a long tail of longer gaps between browsing sessions. These patterns underscore the sequential nature of the data and the heterogeneity in consumers' exposure opportunities over the campaign horizon.

Figure 2 illustrates heterogeneity in the contextual environments in which auctions occur. The distribution of domains is highly concentrated among a relatively small set of frequently visited publishers, with a long tail of less common domains. A similar pattern emerges for publisher verticals (e.g., news, shopping, travel, finance, and entertainment). This dispersion in browsing contexts implies meaningful variation in the observable state X_t, which in turn generates heterogeneity in both bidding decisions and exposure probabilities.

Exposure is rare at the auction level. Across all campaigns, fewer than 1% of auction opportunities result in an impression. Even within the treatment group, more than 50% of consumers receive no impressions over the campaign horizon. This sparsity reflects both auction competition and the advertiser's bidding strategy and generates substantial variation in realized exposure paths.

Conditional on receiving at least one impression, there is considerable heterogeneity in exposure intensity. Figure 3a plots the distribution of the number of impressions per consumer for a representative campaign (conditional on being exposed). The majority of exposed consumers receive only a small number of impressions, while a long right tail captures consumers



(a) Distribution of domains (top 100)

(b) Distribution of verticals (top 100)

Figure 2: Heterogeneity in auction context

who are exposed more frequently. Although the exact shape varies across campaigns, this qualitative pattern is common. Such dispersion motivates modeling exposure as a dynamic, history-dependent process rather than a single binary exposure indicator.

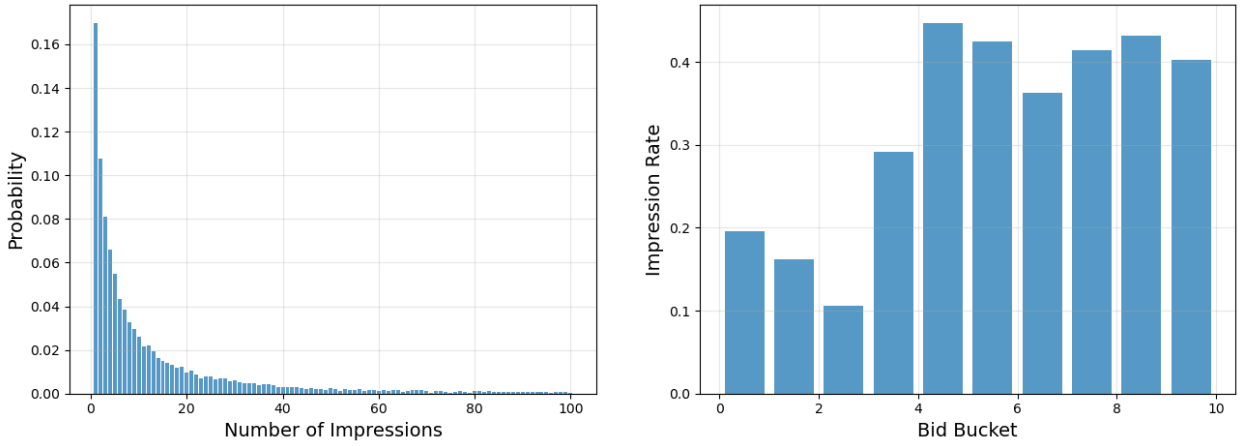
Figure 3b shows the relationship between submitted bids and impression win rates for the same campaign. Win rates increase with bid levels, measured in dollars CPM, but remain strictly between zero and one across all bid buckets. Because bids are functions of the pre-auction history H_t, this pattern, together with the low overall auction-level exposure rate, provides descriptive support for the sequential overlap condition in Section 3, which requires non-exposure to occur with positive probability conditional on the observed history.

Across the full set of observed campaign variables, treatment-control balance tests provide no evidence of systematic differences between the two groups; details are reported in Web Appendix B.

5 Estimation

Our objective is to estimate the expected cumulative outcome under the no-ads counterfactual, ρ_0, for each campaign using only data from the treatment group.

Section 3 shows that ρ_0 admits the efficient influence function in Equation (7:No-ads EIF). Because the influence function has mean zero, $\mathbb{E}[\phi(H)] = 0$, this moment condition provides the



(a) Distribution of impressions per consumer (conditional on exposure)

(b) Impression win rate by bid bucket

Figure 3: Exposure intensity and auction win probabilities for a representative campaign.

basis for estimating ρ_0. We replace the population expectation with its empirical analogue over consumers in the treatment group and replace the unknown nuisance functions with estimated counterparts.

Solving the resulting empirical estimating equation yields an estimate of ρ_0. This section describes the estimation of the two nuisance components, the construction of the estimator, and the procedure used to obtain standard errors and confidence intervals.

5.1 Nuisance Components

The estimating equation depends on two nuisance components: (i) the exposure model $\pi_b(D_t = 0 \mid H_t)$, which characterizes auction-level exposure assignment given the observed history, and (ii) the value function $v_t(H_t)$, which maps histories into expected future outcomes under the no-ads policy. Together, these components account for history-dependent selection into exposure and the evolution of outcomes under the no-ads counterfactual.

We estimate both nuisance components using flexible machine learning methods. As discussed in Section 3.5, the moment condition is Neyman orthogonal, so sufficiently small estimation errors in the nuisance functions have no first-order effect on the estimator. Under appropriate convergence-rate and regularity conditions, the resulting second-order remainder is asymptotically negligible, allowing valid inference.

5.1.1 Exposure Model. The exposure model estimates the auction-level propensity

$$\pi_b(D_t = 0 \mid H_t),$$

the probability that the focal advertiser does not win auction t conditional on the observed pre-auction history H_t.

Because $H_t = (X_1, D_1, Y_1, \ldots, X_t)$ grows with t, directly modeling the full history is computationally infeasible. We therefore approximate H_t using a truncated history of fixed length L, consisting of the most recent L auction opportunities prior to and including t. Given the sequential and history-dependent nature of exposure, we use a transformer-based sequence model to map

$$(D_{t-L}, Y_{t-L}, X_{t-L+1}, D_{t-L+1}, Y_{t-L+1}, \ldots, X_t)$$

to the probability of non-exposure at auction t.

We set the context length to $L = 512$. This choice was guided by discussions with the collaborating DSP, which indicated that most predictive signal for bidding decisions lies within this horizon. Empirically, increasing the context size to 1,024 yields negligible improvements in out-of-sample predictive performance and does not materially affect the resulting estimates.

Per-Auction Representation. At each auction s within the context window, we construct a $d_{\text{model}} = 512$ dimensional representation h_s. The auction-level covariates X_s include (i) the time elapsed since the previous auction, (ii) a 512-dimensional DSP auction embedding summarizing the observable bidding state, and (iii) categorical identifiers for the publisher domain and publisher vertical. In addition, we incorporate lagged exposure and outcome indicators, D_{s-1} and Y_{s-1}, which summarize the most recent realized interaction.

Categorical variables are represented using learned embedding vectors estimated jointly with the model. Let u_s denote the domain index (0–4095) and v_s the vertical index (0–511). We first embed domains into 64-dimensional vectors and verticals into 32-dimensional vectors. These are then linearly projected to dimension $d_{\text{model}} = 512$ to match the transformer input dimension.

Inter-auction times are highly dispersed (Section 4.3). Rather than use raw time differences,

we discretize the elapsed time into 128 logarithmically spaced buckets and represent the bucket index using a learned 32-dimensional embedding, again projected to dimension 512. Logarithmic binning preserves resolution for short intervals—where most auctions occur—while controlling dimensionality in the long tail. Let w_s be the time bucket index for auction s.

Binary exposure and conversion indicators are embedded directly into 512-dimensional vectors. For the first auction in a sequence, we set $D_0 = 0$ and $Y_0 = 0$.

Formally, the representation at auction s is

$$\begin{aligned} h_s = E_s &+ \mathrm{Emb}_{\mathrm{dom}}(u_s)W_{\mathrm{dom}} + \mathrm{Emb}_{\mathrm{vert}}(v_s)W_{\mathrm{vert}} + \mathrm{Emb}_{\mathrm{time}}(w_s)W_{\mathrm{time}} \\ &+ \mathrm{Emb}_D(D_{s-1}) + \mathrm{Emb}_Y(Y_{s-1}), \end{aligned}$$

where $E_s \in \mathbb{R}^{512}$ is the DSP auction embedding, $\mathrm{Emb}_{(\cdot)}(\cdot)$ denotes learned embedding maps, and $W_{(\cdot)}$ are linear projection matrices mapping lower-dimensional embeddings to $\mathbb{R}^{512}$. Because attention mechanisms are permutation-invariant, we add learned positional embeddings to each h_s to encode the temporal order within the context window.

Consistent with the identification discussion in Section 3.3, the model inputs are restricted to information observed prior to the focal auction. We include the observed information used by the DSP's targeting and bidding system, together with observed contextual variables and prior exposures and outcomes. We do not condition on realized auction-side variables such as the clearing threshold, which would absorb the residual auction variation underlying our identification argument. In our data, the clearing threshold, competing bids, and other competitor-side information are not observed.

As in any observational implementation, the recorded state may not perfectly reproduce every input or real-time update used by the DSP's bidding system; identification therefore requires that any omitted information not induce residual dependence between exposure and potential outcomes after conditioning on the observed history.

Model Architecture. The sequence $(h_{t-L+1}, \ldots, h_t)$ is processed using a transformer model (Vaswani et al. 2017), which captures history-dependent patterns in exposure through self-

attention over the sequence.

We use 8 transformer layers, each with 8 attention heads and a feedforward network of inner dimension 2048. The attention mechanism allows the model to capture interactions across different points in the history, while the feedforward component applies a nonlinear transformation at each position. Dropout with rate 0.1 is used for regularization.

After processing the sequence, we extract the final hidden representation corresponding to auction t and pass it through a feedforward prediction head consisting of two fully connected layers: a hidden layer with 512 units and GELU activation, followed by a linear output layer and a sigmoid transformation mapping the result to a scalar in $(0, 1)$. This scalar is interpreted as the estimated probability of non-exposure, $\hat{\pi}_b(D_t = 0 \mid H_t)$.

Training Procedure. The model is trained to minimize binary cross-entropy loss:

$$\mathcal{L} = -\frac{1}{N} \sum_{i,t} \left[(1 - D_{it}) \log \hat{\pi}_b(D_{it} = 0 \mid H_{it}) + D_{it} \log(1 - \hat{\pi}_b(D_{it} = 0 \mid H_{it})) \right].$$

Optimization is performed using the Adam optimizer with learning rate 3×10^{-4} and batch size 64. We use an 80/20 split of consumers into training and validation sets, monitor validation loss, and apply early stopping if validation performance does not improve for two consecutive epochs.

Common Architecture Across Campaigns. The architecture and training protocol are held fixed across all campaigns. Hyperparameters are selected ex ante based on standard scaling heuristics for transformer models and are not tuned separately by campaign. Because the exposure model enters only as a nuisance component in the moment condition, we prioritize stability and generalization over aggressive hyperparameter optimization. Empirically, modest variations in architectural and optimization parameters have minimal impact on the resulting estimates, consistent with the orthogonality properties of the influence-function approach.

Overall, this implementation provides a flexible estimator of $\pi_b(D_t = 0 \mid H_t)$ that captures history-dependent selection into impressions.

5.1.2 Value Function. The second nuisance component is the value function $v_t(H_t)$, which maps histories to expected future outcomes under the no-ads policy. As defined in Section 3.5, under the target policy π_0 that sets $D_t = 0$ for all t, the value function satisfies

$$v_t(H_t) = \mathbb{E}_{\pi_0}\left[\sum_{k=t}^{T} Y_k \,\middle|\, H_t\right].$$

Model Architecture. We estimate $v_t(H_t)$ using a transformer-based sequence model analogous to the exposure model described in Section 5.1.1. The input is the same truncated history of length $L = 512$, with the same per-auction representation h_s for $s \in \{t - L + 1, \ldots, t\}$. The transformer backbone—including the number of layers, attention heads, hidden dimension, and feedforward width—is identical to that used for the exposure model.

The prediction head differs only in the output layer. After processing the sequence, we extract the hidden representation corresponding to auction t and pass it through the same two-layer feedforward head. Instead of applying a sigmoid transformation, we use a linear output layer that produces a scalar prediction $\hat{v}_t(H_t)$, interpreted as the expected cumulative future outcome under the no-ads policy conditional on the truncated history.

Bellman Recursion and Pseudo-Outcomes. The value function is not directly observed in the data. However, under the no-ads policy, it satisfies a Bellman recursion that characterizes its dynamic structure. We exploit this recursion to construct training targets for estimation. Because π_0 deterministically sets $D_t = 0$, the value function satisfies the recursion

$$v_t(H_t) = \mathbb{E}\left[Y_t + v_{t+1}(H_{t+1}) \,\middle|\, H_t,\, D_t = 0\right].$$

This recursion implies that, for auctions with $D_t = 0$, the quantity

$$Y_t + v_{t+1}(H_{t+1})$$

serves as a target for $v_t(H_t)$. Under sequential exogeneity and overlap, the continuation value is identified from observed non-exposure transitions conditional on H_t. Accordingly, we restrict

value-function training to auctions with $D_t = 0$. For auctions with $D_t = 1$, the transition that would have occurred under non-exposure is not observed, and these observations therefore do not contribute directly to the Bellman target.

At the terminal auction for a consumer, the pseudo-outcome reduces to the observed outcome Y_T for non-exposed terminal auctions, as the recursion simplifies to $v_T(H_T) = \mathbb{E}\left[Y_T \mid H_T,\ D_T = 0\right]$.

Training Objective and Target Network. The Bellman recursion implies that, for auctions with $D_t = 0$, the quantity $Y_t + v_{t+1}(H_{t+1})$ serves as a regression target for $v_t(H_t)$. Replacing the unknown continuation value with its estimate yields the empirical target

$$\tilde{Y}_t = Y_t + \hat{v}_{t+1}(H_{t+1}).$$

A natural estimator therefore minimizes the squared Bellman residual over non-exposed auctions:

$$\mathcal{L} = \frac{1}{N} \sum_{i,t:D_{it}=0} \left(\hat{v}_t(H_{it}) - \tilde{Y}_{it}\right)^2.$$

However, directly using the same network to compute both $\hat{v}_t(H_t)$ and the continuation term $\hat{v}_{t+1}(H_{t+1})$ can lead to instability and convergence issues (Mnih et al. 2015), because the regression target depends on the parameters currently being updated during training. This creates a moving-target problem during optimization.

To stabilize training, we adopt a target-network approach commonly used in dynamic programming methods (Mnih et al. 2015). Specifically, we maintain a separate copy of the value network, denoted $\hat{v}^{\text{target}}$, which is used exclusively to compute the continuation term in the pseudo-outcome:

$$\tilde{Y}_t = Y_t + \hat{v}_{t+1}^{\text{target}}(H_{t+1}).$$

The parameters of the target network are updated every 1,000 gradient steps to match those of the main value network. This separation mitigates the moving-target problem inherent in recursive Bellman updates and improves the stability of the estimation procedure (Mnih et al. 2015).

The architecture and training protocol for the value function mirror those used for the exposure model and are held fixed across all campaigns. As with the exposure model, hyperparameters are selected ex ante and not tuned separately by campaign.

Overall, this procedure yields a flexible estimator of $v_t(H_t)$ that captures the recursive structure of counterfactual outcomes under the no-ads policy.

Sources of Identifying Variation. Both nuisance components are estimated using auction level data and require sufficient support for non-exposure events to be well identified. In our setting, such non-exposure events are abundant. As documented in Section 4.3, fewer than 1% of auction opportunities result in an impression, and even within the treatment group a substantial fraction of consumers receive no exposure over the campaign horizon. This implies that, at the auction level, the probability of non-exposure is high across a wide range of histories, providing rich support for estimation of $\pi_b(D_t = 0 \mid H_t)$ and for the Bellman recursion used to estimate $v_t(H_t)$, both of which rely on observed $D_t = 0$ transitions. This feature is particularly favorable for evaluating the no-ads policy. In settings with substantially higher exposure probabilities, the no-ads policy would be farther from the observed behavior policy, resulting in less favorable overlap and potentially larger cumulative importance weights. Counterfactual estimates may therefore be substantially less precise in such settings. Neyman orthogonality mitigates the first-order effect of estimation errors in the nuisance functions, but does not eliminate this underlying overlap problem.

5.2 Estimation and Inference

Building on the estimating equation implied by the efficient influence function, we construct an estimator of ρ_0 by replacing the population expectation with its empirical analogue and the unknown nuisance functions with their estimates $\hat{\pi}_b(D_t = 0 \mid H_t)$ and $\hat{v}_t(H_t)$.

For each consumer $i \in \mathcal{T}$ in the treatment group and auction t, we compute the estimated per-auction density ratio

$$\hat{\eta}_{it} = \frac{\mathbf{1}\{D_{it} = 0\}}{\hat{\pi}_b(D_{it} = 0 \mid H_{it})},$$

and the corresponding cumulative density ratio

$$\hat{\lambda}_{it} = \prod_{k=1}^{t} \hat{\eta}_{ik}, \qquad \hat{\lambda}_{i0} \equiv 1.$$

Substituting the estimated nuisance functions into the estimating equation yields the doubly robust off-policy estimator

$$\hat{\rho}_0^{\text{DR-OPE}} = \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \sum_{t=1}^{T_i} \hat{\lambda}_{i,t-1} \left[\hat{\eta}_{it} \Big(Y_{it} - \hat{v}_t(H_{it}) \Big) + \hat{v}_t(H_{it}) \right],$$

where $n_{\mathcal{T}}$ is the number of consumers in the treatment group.

Cross-Fitting. To mitigate overfitting and facilitate valid inference with flexibly estimated nuisance functions, we implement two-fold cross-fitting at the consumer level (Chernozhukov et al. 2018). Consumers are randomly partitioned into two folds. The nuisance functions are estimated on one fold and evaluated on the other when constructing the estimator, and the roles of the folds are then reversed. The final estimate averages contributions across both folds. This procedure ensures that nuisance predictions used in the estimating equation are obtained from models trained on independent data.

The number of folds can matter in smaller samples because it determines the amount of data available for nuisance estimation. Under K-fold cross-fitting, each nuisance model is estimated using a fraction $(K-1)/K$ of the sample. Thus, increasing the number of folds can improve finite-sample performance when nuisance estimation is limited by sample size, although the choice of a fixed K does not affect the first-order asymptotic distribution of the estimator. In our application, because of the large sample size, we use two-fold cross-fitting to limit the computational cost of repeatedly estimating the nuisance models.

Variance Estimation and Inference. The influence function satisfies $\mathbb{E}[\phi(H)] = 0$, and the estimator admits an asymptotically linear representation with influence function $\phi(H)$. As a result,

$$\sqrt{n_{\mathcal{T}}} \Big(\hat{\rho}_0^{\text{DR-OPE}} - \rho_0 \Big) \;\Rightarrow\; \mathcal{N}(0,\, \mathbb{V}[\phi(H)]).$$

Thus, the asymptotic variance of the estimator is $\mathbb{V}[\phi(H)]/n_{\mathcal{T}}$.

Let

$$\hat{\phi}_i = \sum_{t=1}^{T_i} \hat{\lambda}_{i,t-1} \left[\hat{\eta}_{it}\Big(Y_{it} - \hat{v}_t(H_{it})\Big) + \hat{v}_t(H_{it}) \right] - \hat{\rho}_0^{\text{DR-OPE}}$$

denote the estimated influence-function contribution for consumer i. We estimate $\mathbb{V}[\phi(H)]$ by the sample variance of $\hat{\phi}_i$:

$$\widehat{\mathbb{V}}[\phi(H)] = \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \hat{\phi}_i^2.$$

The variance of the estimator is therefore

$$\widehat{\text{Var}}(\hat{\rho}_0^{\text{DR-OPE}}) = \frac{1}{n_{\mathcal{T}}} \widehat{\mathbb{V}}[\phi(H)] = \frac{1}{n_{\mathcal{T}}} \cdot \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \hat{\phi}_i^2.$$

Standard errors and confidence intervals for ρ_0 are constructed using the normal approximation implied by asymptotic linearity.

6 Empirical Validation

We evaluate the DR-OPE estimator by comparing its estimates of the no-ads counterfactual with experimental benchmarks from the holdout control groups. We then compare its performance with alternative observational approaches that progressively incorporate auction-level information, history-dependent exposure, and continuation values.

6.1 Comparison to RCT Estimates

We begin by comparing the DR-OPE estimate of the no-ads counterfactual, ρ_0, with its experimental counterpart. Section 5 constructs the observational estimate using data from the treatment group alone. The randomized holdout design provides a direct experimental benchmark using outcomes from the control group.

Let $Y_i = \sum_{t=1}^{T_i} Y_{it}$ denote the cumulative outcome for consumer i over the campaign horizon. Because control-group consumers receive no exposure to the focal campaign, their realized outcomes correspond to the no-ads policy. Under random assignment, the experimental estimator

Table 2: Comparison of Experimental and Observational Estimates of ρ_0

Campaign	$\hat{\rho}_0^{\text{RCT}}$	95% CI	$\hat{\rho}_0^{\text{DR-OPE}}$	95% CI
1	2.74	[2.61, 2.86]	3.00	[2.87, 3.12]
2	7.83	[7.63, 8.04]	8.17	[8.01, 8.33]
3	2.88	[2.68, 3.08]	3.32	[3.06, 3.58]
4	12.14	[11.79, 12.48]	11.52	[10.03, 13.01]
5	6.32	[6.08, 6.56]	6.37	[5.87, 6.87]
6	2.25	[2.14, 2.36]	2.70	[2.56, 2.84]
7	1.31	[1.17, 1.45]	1.40	[1.36, 1.44]
8	0.90	[0.82, 0.98]	1.01	[0.98, 1.05]
9	6.02	[5.82, 6.21]	6.40	[6.11, 6.69]
10	2.59	[2.45, 2.72]	2.83	[2.74, 2.93]
11	1.16	[1.04, 1.27]	0.90	[0.85, 0.95]
12	3.29	[3.13, 3.45]	3.68	[3.39, 3.97]
13	3.01	[2.89, 3.14]	2.92	[2.60, 3.23]
14	7.07	[6.82, 7.33]	6.86	[6.43, 7.29]

Notes: All values of ρ_0 are reported in units of 10^{-3}. Confidence intervals are computed as $\hat{\rho} \pm 1.96 \times SE$.

of ρ_0 is therefore

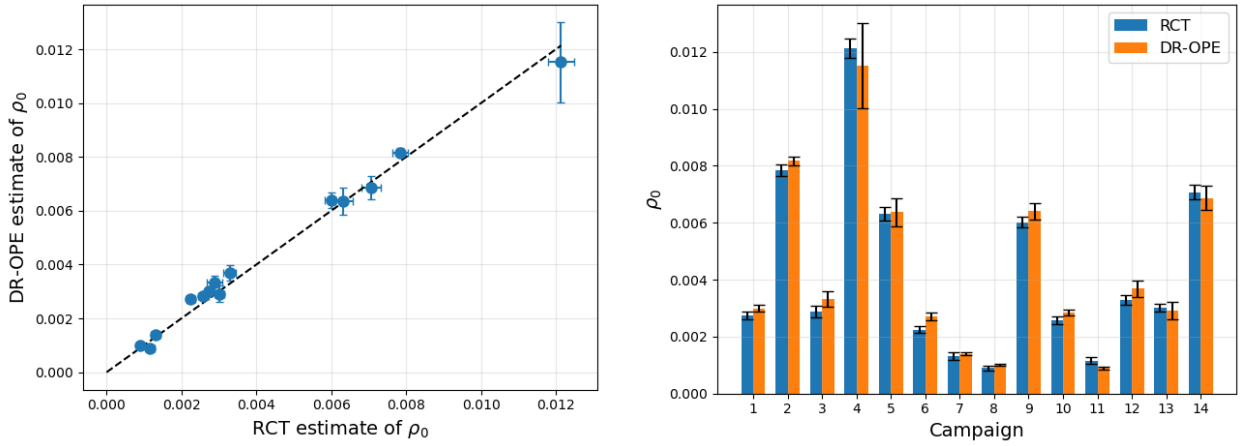
$$\hat{\rho}_0^{\text{RCT}} = \frac{1}{n_{\mathcal{C}}} \sum_{i \in \mathcal{C}} Y_i,$$

where $\mathcal{C}$ denotes the set of control consumers and $n_{\mathcal{C}}$ is the corresponding sample size.

Let $s_{\mathcal{C}}^2$ denote the sample variance of Y_i in the control group. The estimated variance of $\hat{\rho}_0^{\text{RCT}}$ is $s_{\mathcal{C}}^2/n_{\mathcal{C}}$, with the corresponding standard error $s_{\mathcal{C}}/\sqrt{n_{\mathcal{C}}}$.

Table 2 reports $\hat{\rho}_0^{\text{RCT}}$ and $\hat{\rho}_0^{\text{DR-OPE}}$ for each campaign, together with their 95% confidence intervals. All values of ρ_0 are reported in units of 10^{-3}. Because the underlying data are confidential, the reported estimates and confidence intervals are scaled by a random constant. This masking preserves proportional differences across estimates while preventing recovery of the underlying conversion rates.

Figure 4 provides a graphical comparison of the observational and experimental estimates. Panel (a) plots $\hat{\rho}_0^{\text{DR-OPE}}$ against $\hat{\rho}_0^{\text{RCT}}$ across campaigns, together with a 45-degree line that represents perfect agreement. Panel (b) shows the campaign-level estimates from both methods with their associated 95% confidence intervals.



(a) Observational versus experimental estimates of ρ_0

(b) Campaign-level estimates of ρ_0 with 95% confidence intervals

Figure 4: Comparison of observational and experimental estimates of ρ_0 across campaigns.

To summarize the discrepancies across campaigns, define the error for campaign j as

$$\Delta_j = \hat{\rho}_{0j}^{\text{DR-OPE}} - \hat{\rho}_{0j}^{\text{RCT}}.$$

We first measure these errors relative to the campaign-specific experimental benchmarks. The mean percentage error (MPE) and mean absolute percentage error (MAPE) are defined as

$$\text{MPE} = \frac{100}{J} \sum_{j=1}^{J} \frac{\Delta_j}{\hat{\rho}_{0j}^{\text{RCT}}},$$

and

$$\text{MAPE} = \frac{100}{J} \sum_{j=1}^{J} \left| \frac{\Delta_j}{\hat{\rho}_{0j}^{\text{RCT}}} \right|.$$

Across the 14 campaigns, the MPE is 4.5%, with a 95% confidence interval of [2.0%, 7.0%], and the MAPE is 9.3%, with a 95% confidence interval of [6.9%, 11.8%]. Confidence intervals are constructed using the delta method and account for estimation uncertainty in the campaign-level observational and experimental benchmarks.

We next decompose the errors into systematic bias and cross-campaign variation. Let $\bar{\Delta} = \frac{1}{J}\sum_{j=1}^{J} \Delta_j$ denote the mean error across campaigns. The mean squared error (MSE) can then

be written as

$$\text{MSE} = \frac{1}{J}\sum_{j=1}^{J}\Delta_j^2 = \bar{\Delta}^2 + \frac{1}{J}\sum_{j=1}^{J}\left(\Delta_j - \bar{\Delta}\right)^2.$$

The first component is squared mean bias, while the second is the cross-campaign variance of the errors. In our setting, squared mean bias accounts for 12.0% of the MSE, while cross-campaign variance accounts for the remaining 88.0%. Thus, most of the discrepancy between the observational and experimental estimates reflects variation across campaigns rather than a common directional bias.

It is also useful to express the magnitude of these errors relative to the experimental baseline. Let $\bar{\rho}_0^{\text{RCT}} = \frac{1}{J}\sum_{j=1}^{J}\hat{\rho}_{0j}^{\text{RCT}}$ denote the mean RCT estimate of ρ_0 across campaigns. Relative to this benchmark, the root mean squared error (RMSE) is 7.6%, the mean bias is 2.6%, and the cross-campaign standard deviation of the errors is 7.2%. Taken together, these results show that the errors are small relative to the experimental benchmark and that most of the remaining error reflects cross-campaign variation rather than systematic bias.[3]

We next examine campaign lift, defined as incrementality expressed as a percentage of the no-ads baseline:

$$\ell = 100 \times \frac{\rho_b - \rho_0}{\rho_0}.$$

We report lift rather than the absolute difference $\rho_b - \rho_0$ because percentage effects are easier to interpret and are the standard metric used in industry practice. Because ρ_b is identified directly from the treatment group and estimated as the sample average of cumulative outcomes, it is common to both the experimental benchmark and the observational estimator. Differences in estimated lift therefore arise from differences in the estimated no-ads counterfactual, $\hat{\rho}_0$. Standard errors for lift are obtained using the delta method applied to the estimates of ρ_b and ρ_0.

Table 3 reports the experimental and DR-OPE estimates of lift for each campaign, together with their 95% confidence intervals. As with the ρ_0 results, the reported lift values are scaled by

[3]The 95% confidence interval for the MPE excludes zero, indicating a statistically detectable average proportional discrepancy between the DR-OPE and RCT estimates. An Hausman–Wald test of zero mean discrepancy likewise rejects equality between the two estimates ($|Z| = 4.84$, $p < 0.001$). Thus, although the estimates are statistically distinguishable, the magnitude of the discrepancy is relatively small.

Table 3: Comparison of Experimental and Observational Estimates of Lift

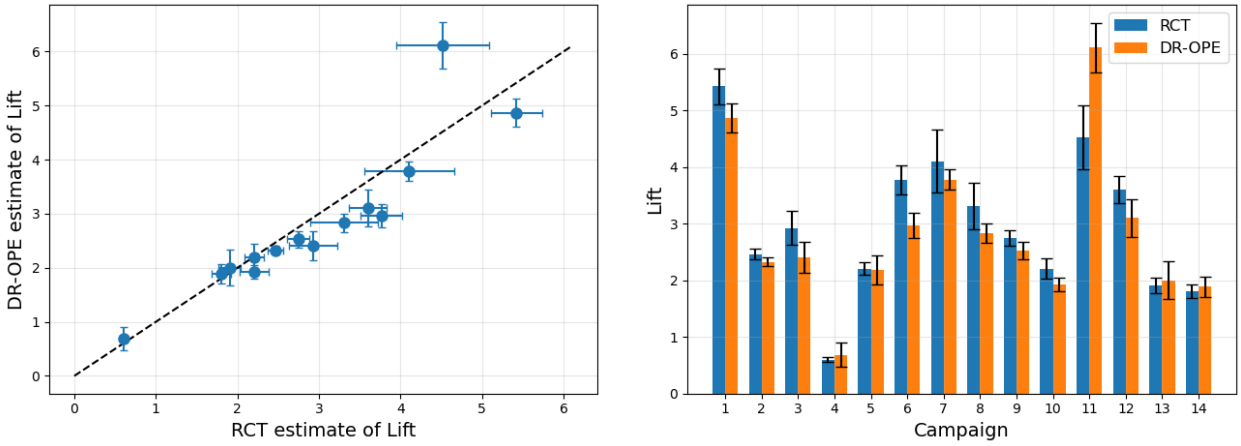
Campaign	$\hat{\ell}^{\text{RCT}}$	95% CI	$\hat{\ell}^{\text{DR-OPE}}$	95% CI
1	542.22	[511.66, 572.78]	486.72	[461.58, 511.86]
2	246.50	[236.99, 256.01]	232.29	[225.23, 239.35]
3	292.54	[263.99, 321.09]	240.53	[212.49, 268.57]
4	60.07	[55.19, 64.95]	68.66	[46.82, 90.50]
5	220.82	[208.12, 233.52]	218.50	[193.17, 243.83]
6	376.93	[352.34, 401.52]	296.86	[275.33, 318.39]
7	410.55	[354.74, 466.36]	378.02	[360.99, 395.05]
8	330.70	[289.93, 371.47]	282.99	[265.75, 300.23]
9	274.94	[262.14, 287.74]	252.56	[236.01, 269.11]
10	220.34	[202.55, 238.13]	192.41	[181.40, 203.42]
11	452.16	[395.32, 509.00]	610.80	[567.33, 654.27]
12	359.92	[336.69, 383.15]	310.46	[277.78, 343.14]
13	190.34	[177.30, 203.38]	200.00	[167.17, 232.83]
14	180.15	[169.35, 190.95]	188.93	[170.66, 207.20]

Notes: Confidence intervals are computed as $\hat{\ell} \pm 1.96 \times SE$.

a random constant to preserve confidentiality while maintaining comparability across methods.

Figure 5 provides a graphical comparison of the experimental and observational lift estimates. Panel (a) plots $\hat{\ell}^{\text{DR-OPE}}$ against $\hat{\ell}^{\text{RCT}}$ across campaigns with a 45-degree reference line. Panel (b) shows campaign-level lift estimates from both methods with their associated confidence intervals.

The campaign-level lift estimates exhibit the same broad pattern as the estimates of ρ_0: the DR-OPE estimates are generally similar in magnitude to the experimental benchmarks, although discrepancies are larger for some campaigns. Because lift is a nonlinear function of ρ_0, differences in the estimated no-ads outcome can be amplified on the lift scale, particularly when the baseline ρ_0 is small. In the next subsection, we use lift as a common metric to compare the DR-OPE estimator with alternative observational approaches and examine their errors in terms of both systematic bias and cross-campaign variation.



(a) Observational versus experimental estimates of lift

(b) Campaign-level estimates of lift with 95% confidence intervals

Figure 5: Comparison of observational and experimental estimates of lift across campaigns.

6.2 Comparison to Alternative Observational Methods

Section 6.1 compares the DR-OPE estimates with the experimental benchmarks. We now compare DR-OPE with a sequence of alternative observational approaches to examine how estimator performance changes as the analysis moves from consumer-level aggregation to richer representations of the auction process.

The first two methods operate at the consumer level, as is standard in much of the advertising measurement literature: consumers are aggregated over the campaign horizon and classified as exposed if they receive at least one impression and unexposed otherwise. The remaining methods operate at the auction level and progressively incorporate richer information about the exposure process and future outcomes. For each method, we construct an implied estimate of campaign lift that is directly comparable to the experimental benchmark. Detailed definitions and implementation of the alternative estimators are provided in Appendix C.

Method 1 (EU): Consumer-level exposed/unexposed comparison. The first benchmark compares average outcomes between treatment-group consumers who receive at least one impression and those who receive none. It does not adjust for observable differences between exposed and unexposed consumers and is included as a naive observational benchmark.

Method 2 (DML-C): Consumer-level double machine learning. The second benchmark applies double machine learning at the consumer level, adjusting for selection into exposure using flexible propensity and outcome models. The covariates consist of consumer-level summaries of browsing activity, auction contexts, and the DSP auction embeddings over the campaign horizon. This method allows rich adjustment for observed heterogeneity but abstracts from the sequential structure of exposure.

Method 3 (A-Static): Auction-level estimator without history. The third method moves the analysis to the auction level but conditions only on contemporaneous auction covariates. It uses an auction-level exposure model together with a contemporaneous outcome regression, thereby incorporating auction-level variation while abstracting from history-dependent exposure and future outcomes.

Method 4 (A-Dynamic): Auction-level estimator with history-dependent exposure. The fourth method allows the probability of exposure to depend on the consumer's observed auction history, using the same history-dependent exposure model as DR-OPE. It continues, however, to use a myopic outcome regression for the contemporaneous outcome rather than a continuation value for future outcomes.

Method 5 (DR-OPE): Full auction-level estimator. The final method is the DR-OPE estimator developed in Section 5. It combines the history-dependent exposure model with the continuation-value model for outcomes under the no-ads policy.

Results. We next compare the performance of these methods using common measures of error relative to the experimental lift estimates. For method m and campaign j, define the error as

$$\Delta_j^m = \hat{\ell}_j^m - \hat{\ell}_j^{\mathrm{RCT}}.$$

We use the same error measures introduced in Section 6.1, applied here to lift. The mean percentage error (MPE) and mean absolute percentage error (MAPE) measure proportional discrepancies relative to the campaign-specific RCT estimates. We also examine errors on the level scale using the root mean squared error (RMSE), mean bias, and cross-campaign standard

Table 4: Accuracy of Observational Lift Estimates Relative to RCT Benchmarks

	EU	DML-C	A-Static	A-Dynamic	DR-OPE
Panel A: Proportional Error (%)					
MPE	258.3	123.0	104.1	25.8	-3.8
	[248.4, 268.2]	[115.7, 130.2]	[98.2, 109.9]	[20.3, 31.3]	[-7.9, 0.2]
MAPE	268.3	176.1	112.1	48.9	12.3
	[258.3, 278.2]	[168.8, 183.3]	[106.2, 117.9]	[43.4, 54.5]	[8.2, 16.4]
Panel B: Error Relative to Mean RCT Lift (%)					
RMSE	333.8	251.2	151.4	61.4	19.1
Bias	261.3	133.2	106.5	31.4	-4.8
SD	207.8	212.9	107.6	52.8	18.5
Panel C: MSE Decomposition (%)					
Bias^2	61.2	28.1	49.5	26.2	6.3
Variance	38.8	71.9	50.5	73.8	93.7

Notes: The table reports errors in observational lift estimates relative to the corresponding RCT estimates across the 14 campaigns. MPE denotes mean percentage error and MAPE denotes mean absolute percentage error. Brackets report 95% confidence intervals constructed using the delta method. RMSE, mean bias, and cross-campaign standard deviation (SD) are normalized by the mean RCT lift across campaigns. In Panel C, squared mean bias and cross-campaign variance are reported as percentages of MSE and sum to 100%.

deviation, each normalized by the mean RCT lift across campaigns,

$$\bar{\ell}^{\text{RCT}} = \frac{1}{J}\sum_{j=1}^{J} \hat{\ell}_j^{\text{RCT}}.$$

Finally, we decompose the MSE for each method into squared mean bias and cross-campaign variance, as in Section 6.1. Table 4 reports these measures for all five observational approaches. Campaign-level lift estimates for all methods are reported in Appendix C.

The results show a clear improvement in performance as the analysis moves from consumer level aggregation to richer auction-level representations. Both consumer-level methods exhibit large positive errors relative to the experimental benchmarks. Flexible adjustment through DML-C reduces this bias, but substantial error and cross-campaign variation remain. Moving to the auction level improves performance further, and allowing exposure probabilities to depend on the consumer's auction history produces another substantial reduction in both systematic

bias and overall error. These patterns indicate that the level of aggregation and the information used to model exposure are empirically important in our setting.

The full DR-OPE estimator performs substantially better than the alternative approaches along each measure in Table 4. Its MPE is -3.8% and its MAPE is 12.3%; its normalized RMSE and cross-campaign standard deviation are 19.1% and 18.5%, respectively, the lowest among the methods considered. Moreover, only a small share (6.3%) of its MSE reflects systematic bias, indicating that the remaining error is primarily cross-campaign variation rather than systematic over- or underestimation.[4]

The comparison between A-Dynamic and DR-OPE also illustrates the practical value of the doubly robust, orthogonal construction. A-Dynamic uses the same history-dependent exposure model as DR-OPE but replaces the continuation value with a myopic outcome regression. If the exposure probabilities were known, or estimated sufficiently accurately, weighting alone would be sufficient to recover the no-ads counterfactual. In practice, however, these probabilities must themselves be estimated from long, high-dimensional histories. This makes accurate estimation of the exposure model difficult in finite samples, so relying primarily on that component can leave non-negligible estimation error. The DR-OPE estimator additionally estimates the continuation value, and Neyman orthogonality implies that the leading effect of nuisance-estimation error is second order, depending on the interaction between errors in the exposure and continuation-value models. The substantially better finite-sample performance of DR-OPE therefore shows that the continuation-value component is empirically valuable in our setting.

7 Conclusion

This paper studies how to measure the causal effect of digital advertising campaigns using observational data. Using auction-level data, we model campaign execution as a sequential decision process and formulate advertising measurement as an off-policy evaluation problem. We operationalize a standard doubly robust OPE estimator using auction-level exposure and continuation-value models estimated with flexible machine learning methods. This formulation

[4] As with the ρ_0 comparison, a Hausman–Wald test rejects equality between the DR-OPE and RCT lift estimates ($|Z| = 4.50$, $p < 0.001$). Thus, the observational and experimental lift estimates remain statistically distinguishable despite the substantially smaller magnitude of the DR-OPE errors relative to the alternative observational methods.

preserves the history-dependent structure of ad delivery while allowing valid statistical inference under sequential conditional exogeneity, overlap, and standard regularity conditions.

We evaluate the approach using randomized holdout experiments from 14 campaigns conducted through a demand-side platform, with advertiser-site visits as the outcome. In this setting, the DR-OPE estimates are substantially closer to the experimental benchmarks than the alternative observational approaches considered. The remaining discrepancies are modest relative to the experimental benchmarks and are driven primarily by variation across campaigns rather than systematic bias. The comparison with alternative methods also shows substantial gains from moving beyond consumer-level aggregation and incorporating richer auction histories.

At the same time, our approach relies on detailed auction-level data, which may not always be readily available in practice. Storing and processing such data can be computationally intensive, and access may be limited by platform constraints or privacy considerations. Developing methods that approximate the benefits of auction-level modeling using more aggregated or privacy-preserving data representations remains an important direction for future research.

Several additional avenues merit further exploration. First, while this paper focuses on average campaign-level effects, the framework naturally extends to estimating heterogeneous effects, such as impression-level or consumer-level conditional treatment effects. Such estimates could provide finer-grained insights into when and for whom advertising is most effective. Second, the sequential formulation of the problem opens the door to optimizing targeting and bidding policies using observational data. Extending the current approach to policy learning—designing optimal exposure strategies based on estimated value functions—represents a promising direction at the intersection of advertising and reinforcement learning.

More broadly, the results suggest that the level at which observational advertising data are represented can materially affect empirical performance. When sufficiently detailed auction histories are available and the relevant identifying assumptions are credible, aligning the empirical analysis with the sequential process that generates exposure can substantially improve observational measurement of campaign effectiveness.

References

Barajas, Joel and Narayan Bhamidipati (2021), "Incrementality Testing in Programmatic Advertising: Enhanced Precision with Double-Blind Designs," *Proceedings of the Web Conference 2021*.

Braun, Michael and Wendy W Moe (2013), "Online Display Advertising: Modeling the Effects of Multiple Creatives and Individual Impression Histories," *Marketing Science*, 32 (5), 753–767.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins (2018), "Double/debiased Machine Learning for Treatment and Structural Parameters," *The Econometrics Journal*.

Du, Ruihuan, Yu Zhong, Harikesh Nair, Bo Cui, and Ruyang Shou (2019), "Causally Driven Incremental Multi Touch Attribution Using a Recurrent Neural Network," *arXiv preprint arXiv:1902.00215*.

Dudík, Miroslav, John Langford, and Lihong Li (2011), "Doubly Robust Policy Evaluation and Learning," *arXiv*.

Fonteneau, Raphael, Susan A Murphy, Louis Wehenkel, and Damien Ernst (2013), "Batch Mode Reinforcement Learning Based on the Synthesis of Artificial Trajectories," *Annals of Operations Research*, 208 (1), 383–416.

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

Gordon, Brett R, Robert Moakler, and Florian Zettelmeyer (2023b), "Predicted Incrementality by Experimentation (PIE) for Ad Measurement," *arXiv preprint arXiv:2304.06828*.

Gordon, Brett R, Florian Zettelmeyer, Neha Bhargava, and Dan Chapsky (2019), "A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook," *Marketing Science*, 38 (2), 193–225.

Gui, George, Harikesh Nair, and Fengshi Niu (2021), "Auction Throttling and Causal Inference of Online Advertising Effects," *arXiv preprint arXiv:2112.15155*.

Hill, Daniel N, Robert Moakler, Alan E Hubbard, Vadim Tsemekhman, Foster Provost, and Kiril Tsemekhman (2015), "Measuring Causal Impact of Online Actions via Natural Experiments: Application to Display Advertising," *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Jansen, Bernard J, Amanda Spink, Chris Blakely, and Sherry Koshman (2007), "Defining a session on Web search engines," *Journal of the American Society for Information Science and Technology*, 58 (6), 862–871.

Jiang, Nan and Lihong Li (2016), "Doubly Robust Off-policy Value Evaluation for Reinforcement Learning," *International conference on machine learning*.

Johnson, Garrett, Randall A Lewis, and Elmar Nubbemeyer (2017a), "The Online Display Ad Effectiveness Funnel & Carryover: Lessons from 432 Field Experiments," *Available at SSRN*.

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

Johnson, Garrett A, Randall A Lewis, and Elmar I Nubbemeyer (2017b), "Ghost Ads: Improving the Economics of Measuring Online Ad Effectiveness," *Journal of Marketing Research*, 54 (6), 867–884.

Kallus, Nathan and Masatoshi Uehara (2020), "Double Reinforcement Learning for Efficient Off-policy Evaluation in Markov Decision Processes," *Journal of Machine Learning Research*, 21 (167), 1–63.

Kearns, Michael and Satinder Singh (2002), "Near-Optimal Reinforcement Learning in Polynomial Time," *Machine Learning*, 49 (2), 209–232.

Lewis, Randall and Jeffrey Wong (2022), "Incrementality Bidding and Attribution," *arXiv preprint arXiv:2208.12809*.

Lewis, Randall A and Justin M Rao (2015), "The Unfavorable Economics of Measuring the Returns to Advertising," *The Quarterly Journal of Economics*, 130 (4), 1941–1973.

Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski et al. (2015), "Human-level Control Through Deep Reinforcement Learning," *Nature*, 518 (7540), 529–533.

Narayanan, Sridhar and Kirthi Kalyanam (2015), "Position Effects in Search Advertising and their Moderators: A Regression Discontinuity Approach," *Marketing Science*, 34 (3), 388–407.

Precup, Doina, Richard S Sutton, and Satinder Singh (2000), "Eligibility traces for off-policy policy evaluation," *Working Paper*.

Rafieian, Omid and Hema Yoganarasimhan (2022), "Variety Effects in Mobile Advertising," *Journal of Marketing Research*, 59 (4), 718–738.

Robins, James M (1997), "Causal Inference from Complex Longitudinal Data," *Latent Variable Modeling and Applications to Causality*.

Sahni, Navdeep S (2015), "Effect of Temporal Spacing Between Advertising Exposures: Evidence from Online Field Experiments," *Quantitative Marketing and Economics*, 13 (3), 203–247.

Thomas, Philip and Emma Brunskill (2016), "Data-Efficient Off-Policy Policy Evaluation for Reinforcement Learning," *International conference on machine learning*.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin (2017), "Attention Is All You Need," *Advances in Neural Information Processing Systems*, 30.

Waisman, Caio, Harikesh S Nair, and Carlos Carrion (2025), "Online Causal Inference for Advertising in Real-Time Bidding Auctions," *Marketing Science*, 44 (1), 176–195.

Web Appendix

Web Appendix A: Numerical Illustration

This appendix provides the calculations underlying the numerical illustration in Section 3.4.

True Campaign Lift. Under the no-ads policy, $D_1 = D_2 = 0$, so the relevant distribution of X_2 is the distribution generated by $D_1 = 0$. The second-auction continuation values along the no-ads path are

$$v_2(0, H) = 0.20, \qquad v_2(0, L) = 0.05.$$

Consequently,

$$\rho_0 = 0.4(0.20) + 0.6(0.05) = 0.11.$$

Under the behavior policy,

$$\Pr(X_2 = H) = 0.5(0.4) + 0.5(0.6) = 0.5,$$

and

$$\begin{aligned}\Pr(D_2 = 1) &= 0.5\{0.4(0.8) + 0.6(0.2)\} \\ &\quad + 0.5\{0.6(0.6) + 0.4(0.1)\} \\ &= 0.42.\end{aligned}$$

Following Equation 5, the expected outcome under the behavior policy is therefore

$$\begin{aligned}\rho_b &= 0.5(0.20) + 0.5(0.05) + 0.5(0.04) + 0.42(0.03) \\ &= 0.1576.\end{aligned}$$

The resulting campaign lift is

$$100 \times \frac{0.1576 - 0.11}{0.11} = 43.3\%.$$

Consumer-Level Observational Approaches. Consider first a consumer-level analysis that collapses the two-auction exposure sequence into a single indicator

$$W = \mathbf{1}\{D_1 + D_2 > 0\},$$

where $W = 1$ if the consumer receives at least one impression and $W = 0$ otherwise. A naive exposed–unexposed comparison would use the mean outcome among unexposed consumers as an estimate of the no-ads counterfactual:

$$\rho_0^{EU} = \mathbb{E}[Y_2 \mid W = 0].$$

In the numerical example,

$$\begin{aligned}\Pr(W = 0) &= \Pr(D_1 = 0, D_2 = 0) \\ &= 0.5\Big[0.4(0.2) + 0.6(0.8)\Big] \\ &= 0.28.\end{aligned}$$

Among consumers who receive no impressions,

$$\Pr(X_2 = H \mid W = 0) = \frac{0.5(0.4)(0.2)}{0.28} = \frac{1}{7}.$$

It follows that

$$\begin{aligned}\rho_0^{EU} &= \mathbb{E}[Y_2 \mid W = 0] \\ &= \frac{1}{7}(0.20) + \frac{6}{7}(0.05) \\ &= 0.0714.\end{aligned}$$

Using $\rho_b = 0.1576$, the corresponding estimate of campaign lift is

$$100 \times \frac{0.1576 - 0.0714}{0.0714} = 120.6\%,$$

substantially above the true no-ads lift of 43.3%.

This discrepancy can arise even when the sequential identifying assumptions hold at each auction. Aggregating the exposure sequence into W discards the temporal ordering of exposure and the evolving consumer state. In the example, consumers in the high-intent state are more likely to receive an impression at the second auction and are therefore underrepresented among consumers who happen to remain unexposed. Consequently,

$$\Pr(X_2 = H \mid W = 0) = \frac{1}{7},$$

whereas under the no-ads policy,

$$\Pr(X_2 = H \mid \pi_0) = \Pr(X_2 = H \mid D_1 = 0) = 0.4.$$

Thus, the consumers who happen to receive no impressions under the behavior policy do not reproduce the state distribution that would arise if the no-ads policy were imposed on the full population.

A natural response is to control for the observed state X_2 in the consumer-level analysis. To isolate whether this resolves the problem, suppose the analyst knows the conditional outcome regression for unexposed consumers without estimation error:

$$m_C(x) = \mathbb{E}[Y_2 \mid W = 0, X_2 = x].$$

In this example,

$$m_C(H) = 0.20, \qquad m_C(L) = 0.05.$$

A standard consumer-level outcome adjustment would average these conditional means over the observed population distribution of X_2. Because

$$\Pr(X_2 = H) = \Pr(X_2 = L) = 0.5,$$

the resulting estimate of the no-ads counterfactual is

$$\begin{aligned}\rho_0^C &= \sum_{x\in\{H,L\}} \Pr(X_2 = x) m_C(x) \\ &= 0.5(0.20) + 0.5(0.05) \\ &= 0.125.\end{aligned}$$

The corresponding estimate of campaign lift is

$$100 \times \frac{0.1576 - 0.125}{0.125} = 26.1\%,$$

which also differs from the true lift of 43.3%.

The discrepancy remains even though X_2 is observed and the outcome regression is known without error. The reason is that X_2 is not a pre-exposure characteristic: its distribution is itself affected by the first exposure. Consequently, the observed distribution $\Pr(X_2)$ is partly generated by the behavior policy and is not the state distribution relevant for the no-ads counterfactual. The latter requires averaging outcomes over

$$\Pr(X_2 \mid D_1 = 0),$$

the distribution generated when the first exposure is set to zero. The sequential analysis therefore uses X_2 as part of the history at the second auction while also accounting for how earlier exposure decisions change its distribution. Simply including X_2 as another consumer-level control does not make this adjustment.

Web Appendix B: Randomization Checks

To verify that randomization was implemented correctly, we compare means across treatment and control groups for each campaign and for each variable. These variables include the embedding components, counts of domains and verticals, the number of auctions, and inter-auction times. Across all campaigns and variables, this yields 71,708 hypothesis tests. Under the null hypothesis of equal means, the resulting p-values should be uniformly distributed on the unit interval. The empirical distribution closely matches this benchmark: 10.41% of p-values are below 0.10, 5.28% are below 0.05, and 1.06% are below 0.01, all near their nominal levels. Figure 6 plots the distribution of p-values across all tests. A Kolmogorov–Smirnov test fails to reject the null that the p-values are uniformly distributed (p-value = 0.79). Taken together, these results provide no evidence of systematic imbalance between treatment and control groups, supporting the validity of the experimental benchmark.

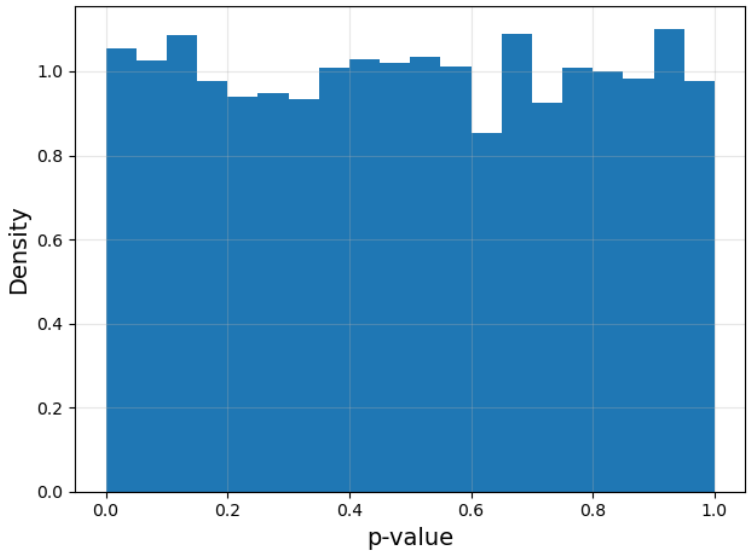


Figure 6: Randomization check: distribution of p-values across all campaigns and variables

Web Appendix C: Alternative Observational Methods

This appendix provides additional details on the alternative observational methods summarized in Section 6.2 and reports the resulting campaign-level lift estimates for all methods.

C.1 Alternative Estimator Definitions

Method 1 (EU): Consumer-level exposed/unexposed difference in means. The simplest observational benchmark aggregates outcomes to the consumer level and classifies treatment-group consumers as exposed or unexposed based on whether they receive at least one impression during the campaign.

Let

$$Y_i = \sum_{t=1}^{T_i} Y_{it}$$

denote cumulative outcomes for consumer i, and define the exposure indicator

$$W_i = \mathbf{1}\Big\{ \sum_{t=1}^{T_i} D_{it} > 0 \Big\},$$

which equals one if consumer i receives at least one impression and zero otherwise.

The EU estimator compares average outcomes across exposed and unexposed consumers in the treatment group:

$$\hat{\tau}^{\,\mathrm{EU}} = \frac{1}{\sum_{i\in\mathcal{T}} W_i} \sum_{i\in\mathcal{T}} W_i Y_i - \frac{1}{\sum_{i\in\mathcal{T}} (1-W_i)} \sum_{i\in\mathcal{T}} (1-W_i) Y_i.$$

This estimator ignores all observable differences between exposed and unexposed consumers and therefore does not adjust for selection into exposure. It is included as a naive benchmark.

To map this estimate to the same scale as the experimental benchmark, we express it at the level of the treatment-group average outcome. Let

$$\hat{p}_W = \frac{1}{n_{\mathcal{T}}} \sum_{i\in\mathcal{T}} W_i$$

denote the fraction of treatment-group consumers who receive at least one impression. We then

scale the exposed/unexposed difference by $\hat{p}_W$ to obtain the implied incremental effect at the treatment-group level:

$$\hat{\tau}^{\mathrm{EU}}_{\mathrm{scaled}} = \hat{p}_W \, \hat{\tau}^{\mathrm{EU}}.$$

Let

$$\hat{\rho}_b = \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} Y_i$$

denote the average cumulative outcome in the treatment group. The implied no-ads counterfactual is then

$$\hat{\rho}_0^{\mathrm{EU}} = \hat{\rho}_b - \hat{\tau}^{\mathrm{EU}}_{\mathrm{scaled}},$$

which yields the corresponding lift estimate

$$\hat{\ell}^{\mathrm{EU}} = 100 \times \frac{\hat{\rho}_b - \hat{\rho}_0^{\mathrm{EU}}}{\hat{\rho}_0^{\mathrm{EU}}}.$$

This adjustment is necessary because exposed/unexposed comparisons operate at the level of consumers who receive impressions, whereas the experimental benchmark is defined relative to the treatment-group average outcome. Similar rescaling approaches are standard in the advertising measurement literature when comparing observational estimates to experimental benchmarks (Gordon et al. 2019; Gordon, Moakler, and Zettelmeyer 2023a).

Method 2 (DML-C): Consumer-level double machine learning. The second benchmark applies a double machine learning (DML) procedure at the consumer level. As in Method 1, treatment is defined by the exposure indicator W_i and the outcome is the cumulative campaign-level outcome Y_i.

The estimator uses flexible machine learning methods to adjust for selection into exposure based on consumer-level covariates. In particular, it estimates both a propensity model for the probability of exposure and an outcome regression for the expected outcome conditional on exposure and covariates, and combines these components to account for systematic differences between exposed and unexposed consumers (Chernozhukov et al. 2018; Gordon, Moakler, and Zettelmeyer 2023a).

The covariates consist of summary statistics constructed from each consumer's campaign history, including measures of browsing activity (e.g., number of auctions, number of browsing sessions, and moments of inter-auction times), distributional features of the DSP auction embeddings (mean and standard deviation), and exposure to different contexts (shares of the top 100 domains and verticals). These features capture rich consumer-level heterogeneity while abstracting from the sequential structure of the data.[5]

Let $\hat{\tau}^{\text{DML-C}}$ denote the resulting estimate of the difference in outcomes between exposed and unexposed consumers after adjusting for observed covariates. As in Method 1, we rescale this estimate to the treatment-group level using the exposure rate $\hat{p}_W$:

$$\hat{\tau}^{\text{DML-C}}_{\text{scaled}} = \hat{p}_W \, \hat{\tau}^{\text{DML-C}}.$$

The implied no-ads counterfactual is then

$$\hat{\rho}_0^{\text{DML-C}} = \hat{\rho}_b - \hat{\tau}^{\text{DML-C}}_{\text{scaled}},$$

which yields the corresponding lift estimate

$$\hat{\ell}^{\text{DML-C}} = 100 \times \frac{\hat{\rho}_b - \hat{\rho}_0^{\text{DML-C}}}{\hat{\rho}_0^{\text{DML-C}}}.$$

This estimator represents the natural comparison to the dominant observational approach in the literature, which aggregates the campaign to the consumer level and adjusts for selection using flexible machine learning methods.

Method 3 (A-Static): Auction-level estimator without history. The first auction-level benchmark moves from consumer-level aggregation to the auction level but restricts the information set to contemporaneous covariates. Specifically, it replaces the history-dependent propensity score $\pi_b(D_{it} = 0 \mid H_{it})$ with a static propensity model based only on current-auction

[5] We estimate the propensity score and outcome regression using feedforward neural networks with three hidden layers, ReLU activations, and dropout regularization. The propensity model is trained using cross-entropy loss and the outcome model using squared error loss. Estimation is performed using cross-fitting to mitigate overfitting and preserve orthogonality.

covariates:

$$\pi_X(D_{it} = 0 \mid X_{it}).$$

It also replaces the continuation value $v_t(H_{it})$ with a myopic outcome regression that models only the contemporaneous outcome:

$$m_t(X_{it}) = \mathbb{E}[Y_{it} \mid X_{it}, D_{it} = 0].$$

Substituting these components into the DR-OPE structure yields the estimator[6]

$$\hat{\rho}_0^{\text{A-static}} = \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \sum_{t=1}^{T_i} \hat{\lambda}_{i,t-1}^X \left[\hat{\eta}_{it}^X \Big(Y_{it} - \hat{m}_t(X_{it}) \Big) + \hat{m}_t(X_{it}) \right],$$

where

$$\hat{\eta}_{it}^X = \frac{\mathbf{1}\{D_{it} = 0\}}{\hat{\pi}_X(D_{it} = 0 \mid X_{it})}, \qquad \hat{\lambda}_{i,t-1}^X = \prod_{s=1}^{t-1} \hat{\eta}_{is}^X.$$

This estimator departs from the data-generating process in two key ways. First, by conditioning only on X_{it}, it ignores the dependence of exposure on the full history, thereby failing to capture dynamic selection into impressions. Second, by modeling only the contemporaneous outcome Y_{it}, it ignores the effect of current exposure on future outcomes, omitting dynamic treatment effects such as carryover and timing.

Despite these restrictions, the estimator yields an implied estimate of the no-ads counterfactual $\hat{\rho}_0^{\text{A-static}}$ and a directly comparable lift estimate.

The A-Static estimator isolates the role of auction-level variation alone, abstracting from both dynamic selection and dynamic treatment effects.

Method 4 (A-Dynamic): Auction-level estimator with history-dependent propensity scores. The second auction-level benchmark extends Method 3 by incorporating history dependent selection into exposure while continuing to abstract from dynamic treatment effects.

[6]The propensity model $\pi_X(D_{it} = 0 \mid X_{it})$ and outcome regression $m_t(X_{it})$ are estimated using feedforward neural networks with nonlinear activations and dropout regularization. The propensity model is trained using cross-entropy loss and the outcome model using squared error loss. Estimation is performed using cross-fitting.

Specifically, it uses a propensity model that conditions on the full observed history,

$$\pi_b(D_{it} = 0 \mid H_{it}),$$

but replaces the continuation value with a myopic outcome regression that models only the contemporaneous outcome:

$$m_t(H_{it}) = \mathbb{E}[Y_{it} \mid H_{it}, D_{it} = 0].$$

Substituting these components into the DR-OPE structure yields the estimator

$$\hat{\rho}_0^{\text{A-dynamic}} = \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \sum_{t=1}^{T_i} \hat{\lambda}_{i,t-1} \left[\hat{\eta}_{it} \Big(Y_{it} - \hat{m}_t(H_{it}) \Big) + \hat{m}_t(H_{it}) \right],$$

where

$$\hat{\eta}_{it} = \frac{\mathbf{1}\{D_{it} = 0\}}{\hat{\pi}_b(D_{it} = 0 \mid H_{it})}, \qquad \hat{\lambda}_{i,t-1} = \prod_{s=1}^{t-1} \hat{\eta}_{is}.$$

Relative to Method 3, this estimator captures dynamic selection into exposure by conditioning on the full history H_{it}. However, by modeling only the contemporaneous outcome Y_{it}, it continues to ignore the effect of current exposure on future outcomes, omitting dynamic treatment effects such as carryover and timing.

The A-Dynamic estimator therefore isolates the role of modeling dynamic selection while abstracting from dynamic treatment effects. If exposure mainly affects outcomes contemporaneously, then replacing the continuation value with a myopic regression should have limited impact. If, instead, impressions influence future outcomes through path-dependent effects, this estimator should perform materially worse than the full DR-OPE estimator.

C.2 Campaign-Level Lift Estimates

Table 5 reports the campaign-level lift estimates for the four alternative observational methods, DR-OPE, and the experimental benchmark. Standard errors are similar in magnitude across methods; Table 6 summarizes their distributions across campaigns.

Table 5: Campaign-Level Lift Estimates (%)

Campaign	EU	DML-C	A-Static	A-Dynamic	DR-OPE	RCT
1	1829	430	807	529	487	542
2	989	1120	741	8	232	247
3	716	521	234	384	241	293
4	251	43	115	68	69	60
5	1384	572	619	335	219	221
6	1266	1561	868	720	297	377
7	2372	937	1055	585	378	411
8	1249	316	568	397	283	331
9	1231	1329	836	518	253	275
10	174	-178	198	333	192	220
11	2229	2313	1574	812	611	452
12	185	-136	267	496	310	360
13	353	611	380	71	200	190
14	794	259	326	210	189	180

Notes: The table reports campaign-level lift estimates, expressed as percentages of the estimated no-ads baseline and rounded to the nearest percentage point.

Table 6: Summary of Standard Errors for Campaign-Level Lift Estimates

	EU	DML-C	A-Static	A-Dynamic	DR-OPE	RCT
Mean	16	17	14	15	12	13
Median	15	14	13	15	11	11
Min	6	5	6	5	4	2
Max	44	56	42	29	22	29

Notes: The table summarizes standard errors across the 14 campaigns. Standard errors are reported in percentage points and rounded to the nearest integer.