SSM-MTO: A Causal Framework for Session-level Ads Load Optimization

Hui Chen[1] , **Patrick R. Johnstone**[1] , **Taihui Li**[1] , **Shu Wang**[1] , **Chao Cen**[1] , **Qinqin Zhu**[1] and **Jizhe Zhang**[1]

[1]Meta Platforms, Inc.

Abstract

Session-level dynamic ads load optimization seeks to strike a balance between user experience and ads performance score by delivering the right number of ads during online sessions on social networks and e-commerce platforms. Previous approaches struggle with several challenges, including treatment-induced bias, carry-over effects, and business constraints on the trade-off between ads performance score and engagement across different cohorts. To overcome those challenges, we propose to train a session sensitivity model (SSM) as a lever to adjust ads load for each session. Then, we adopt a multi-treatment optimization (MTO) framework by incorporating business constraints to dynamically determine the optimal ads load for each session. The SSM is trained on the data collected from the debiased data collection experiment which randomizes the ads load at the session level to remove the confounding bias caused by ads load treatments. From the offline training data, we showed that the SSM-MTO identifies the efficient sessions for ads load treatment. Furthermore, the SSM-MTO has been put into online A/B tests to serve the online traffic which achieved better efficiency and better trade-off between ads performance score and user experience.

1 Introduction

Ads performance score (APS) and user experience (UE) are two primary goals of interest of a social networking or e-commerce platform [Carrion *et al.*, 2021; Yan *et al.*, 2020]. By personalizing the quantity and pattern of the advertisements that are incorporated into the users' organic consumption journey, ads load optimization has proven to be an effective approach to achieve the optimal trade-off between these two goals [Yan *et al.*, 2020].

In general, there are two common strategies to optimize ads load. The "static" approach personalizes the ads load configuration for each user and applies the same ads load throughout the user trajectory, while the "dynamic" approach optimizes the ads load configuration in real time, e.g., during an online user session [Liu *et al.*, 2025; Liao *et al.*, 2022]. The latter, which is more challenging, is our focus in this paper.

A session generally refers to a period during which a user is actively engaged with the platform. This can include activities such as browsing newsfeeds, posting updates, commenting on other users' posts, sending messages, etc. The ads load of a session could be tuned by certain product features, e.g., by changing the position of the first ad and the minimum gap between two consecutive ads in a typical newsfeed product [Yan *et al.*, 2020]. Increasing ads load is expected to boost short-term APS at the cost of hurting UE, which eventually could cause damage to long-term APS. The key is to control the ads load in real time within an online session to achieve the optimal balance between APS and UE.

For the session-level dynamic ads load optimization, the real-time dynamics pose challenges for causal inference learning. For instance, the confounding bias influences both the treatment and the outcome, resulting in the difficulty to determine the true causal effects [Elwert and Winship, 2014]. Furthermore, the business has lots of constraints for the ads load treatment. We are not allowed to abuse particular cohorts (e.g. high value user's sessions) to deliver more ads to increase APS while leading to long-term dissatisfaction. In this paper, we propose a multi-treatment optimization (MTO) framework to address the challenges. The main contributions of the paper are the following:

- We propose a debiased data collection framework to remove confounding bias from the ads treatments, which is critical to model the causal effects.
- We put the session sensitivity prediction into a multi-treament optimization (MTO) framework to dynamically learn the optimal ads load to each session under the business constraints (for example, the trade-off between APS and UE for particular cohorts).
- Our proposed ads load optimization framework has been put into an A/B test to serve online live traffic. From the A/B test, we have achieved top-line business goals. It shows that the proposed framework has significantly improved our platform's capability to serve both consumers and advertisers effectively.

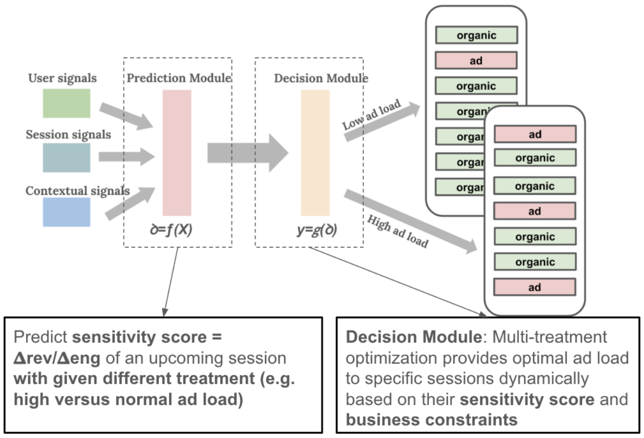

Figure 1: Structure of the session-level dynamic ads load optimization system.

2 Related Work

There is a series of literature on related ads allocation problems in social networking and e-commerce applications in industry [Carrion *et al.*, 2021; Liao *et al.*, 2022; Sagtani *et al.*, 2024; Wang *et al.*, 2022; Yan *et al.*, 2020], focusing mainly on how to place a fixed number of ads and organic content locations. Several approaches have been proposed to find optimal positions via constrained optimization problems [Yan *et al.*, 2020], multi-objective optimization [Carrion *et al.*, 2021], and end-to-end RL [Liao *et al.*, 2022; Rafieian, 2023; Wang *et al.*, 2022].

However, our ads load optimization problem distinguishes ads allocation from three perspectives. Firstly, ads load optimization is built on top of the existing mechanism to make further personalization and optimization, e.g., the highest position of the first ad or the minimum distance between two consecutive ads. As a comparison, Yan et al. [Yan *et al.*, 2020] adopt a fixed highest position and minimum distance as a fixed rule, and some others do not consider ads load directly. Secondly, many existing works on ads allocation seem to have overlooked the strong carry-over effect, and are therefore incapable of capturing how historical treatment affects future observations, including state representations and treatment outcomes. Thirdly, many works rely on the complex deep learning models to optimize the ads load, which are hard to explain why those models work in production. Our proposed causal models depend on the LightGBM [Ke *et al.*, 2017], a tree-based learning algorithm, which makes it easier to understand how the model reaches a prediction.

3 Method

The proposed session-level dynamic ads load optimization system was shown in Figure 1. It consists of two main modules: the prediction module and the decision module. The prediction module, implemented as a session sensitivity causal model, predicts the sensitivity of each session to the ad load treatment; the decision module, implemented in a multi-treatment optimization framework, provides the optimal ad load for each session to maximize the ads performance score (APS) under the business constraints.

3.1 Debiased Data Collection Framework

In order to learn the causal effects of the treatment, we need to set up random control trials to collect the data to train the session level sensitivity models.

We propose the debiased data collection experiment framework randomizes ads load at the session level, which was shown in Figure 2. Different sessions can be exposed to different ads load treatments in a fully randomized way. The framework can measure the treatment effects at the session level by attributing the metrics (e.g., APS and UE) to each session and comparing the impact across different ads load treatments. Furthermore, the framework can also remove feature bias from the ads load treatment since the randomization of logic guarantees that the impact of different ads load treatment is homogenous to all sessions.

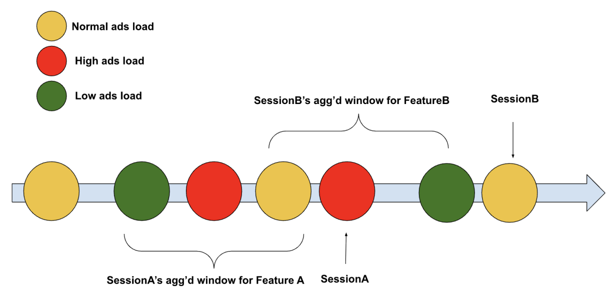

Figure 2: Illustration of the debiased data collection.

By running the above debiased data collection, we collect the session-level features. We then ran the Kolmogorov–Smirnov test, Mann-Whitney U test, and T-test under the treatment and control groups, and identified a few features that are significantly different between the two groups. Those features were removed from training the session causal models.

3.2 Session Sensitivity Model

The architecture of the proposed session sensitivity model (SSM) was shown in Figure 3. In the context of ads supply optimization, we leverage the causal learning model framework [Künzel *et al.*, 2019; Liu *et al.*, 2023; Nie and Wager, 2021; Wu *et al.*, 2023] to learn the conditional average treatment effect (CATE) of each session to predict the session sensitivity in terms of the ads load change.

Since the user sessions' behaviors are heterogeneous, we can apply personalized ads load based on the predicted sensitivity to each session to win a cumulative gain. To achieve this goal, we need to address the following key question: how much APS will be increased and UE will be lost due to the high ads load treatment. We take advantage of the meta-learner error canceling (EC) X-learner framework to address this question. The notations used for EC X-learner are defined in the following:

- $\boldsymbol{X}$: the covariate or feature vector

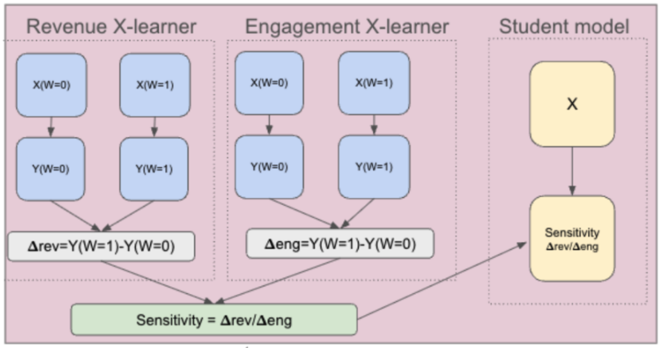

Figure 3: The proposed structure of session sensitivity model.

- $\boldsymbol{W} \in \{0, 1\}$: treatment indicator (0: control, 1: treatment)
- $\boldsymbol{Y}(0)$: potential outcome under control
- $\boldsymbol{Y}(1)$: potential outcome under treatment
- $\tau(x)$: $E[\boldsymbol{Y}(1) - \boldsymbol{Y}(0)|\boldsymbol{X} = x]$ (CATE)

The challenge of meta-learner roots from the counterfactuals for each individual since one individual can only be present in either the treatment or control group. For example, when an individual is assigned to the test group, one's outcome under control $\boldsymbol{Y}(0)$ becomes unmeasurable, which we refer to as the counterfactual. The methodology of meta-learner is to use dedicated models to predict counterfactuals—$\boldsymbol{Y}(1)$ or $\boldsymbol{Y}(0)$ in the CATE function.

The meta-learner has multiple variants, and we use the EC X-learner framework (see Figure 4) to learn the treatment effects of APS and UE.

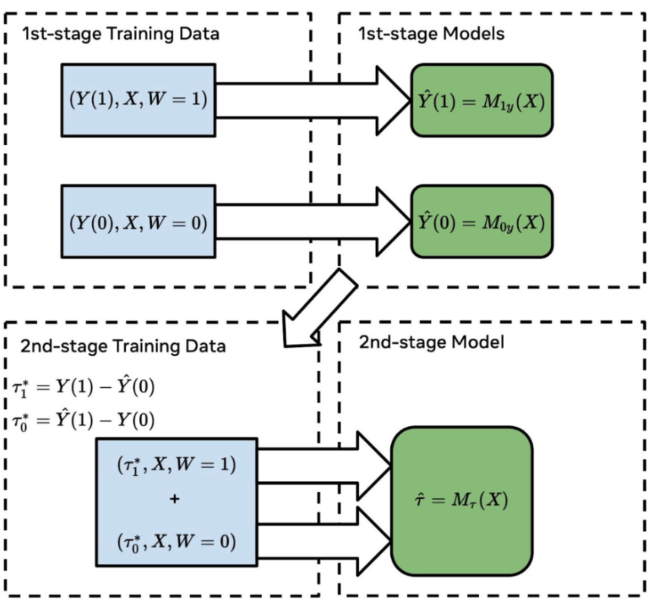

Figure 4: The structure of error canceling (EC) X-learner.

The EC X-learner is an extension of T-learner (T means two) and it consists of two stages, with the 1st-stage usually being T-learner. The 1st-stage base learners provide the counterfactual predictions that complement the absences of outcomes—$\boldsymbol{Y}(0)$ for treatment group and $\boldsymbol{Y}(1)$ for control group—in the CATE function. The EC X-learner pools the 2nd-stage training data from the treatment and control groups to train a single regressor to predict the treatment effects for APS and UE. We calculate the sensitivity and train a student regression model to approximate the sensitivity. All the learners used in this framework are based on LightGBM [Ke *et al.*, 2017].

3.3 Multi-Treatment Optimization

We integrate the session sensitivity predictions into the multi-treatment optimization (MTO) framework to estimate the optimal ads load for each session to maximize the objective (e.g., APS) under the given constraints (e.g., UE). We formulate it as a constrained optimization problem and solve it with the primal-dual approach. Given the objective

$$\begin{aligned} \max_{\boldsymbol{\theta}_c} \quad & \Sigma_{c \in C} R(\boldsymbol{\theta}_c) \\ \text{s.t.} \quad & \Sigma_{c \in C} I(\boldsymbol{\theta}_c) \leq A, \end{aligned} \tag{1}$$

Where C is the set of session cohorts, $\boldsymbol{\theta}_c$ is the AAP scalar for cohort c, $R(\boldsymbol{\theta}_c)$ is the objective (e.g., APS) function for cohort c, $I(\boldsymbol{\theta}_c)$ is the constraint (e.g., UE) function for cohort c, and A is the total constraints. We convert it to a dual problem using Lagrangian duality:

$$L(\boldsymbol{\theta}_c, \lambda) = \Sigma_{c \in C}(R(\boldsymbol{\theta}_c) - \lambda(I(\theta_c) - A)). \tag{2}$$

The dual problem becomes $min_{\lambda \geq 0} max_{\boldsymbol{\theta}_c} L(\boldsymbol{\theta}_c, \lambda)$ and it could be solved by various algorithms, such as LP, SQP, and SLSQP, depending on the particular use cases.

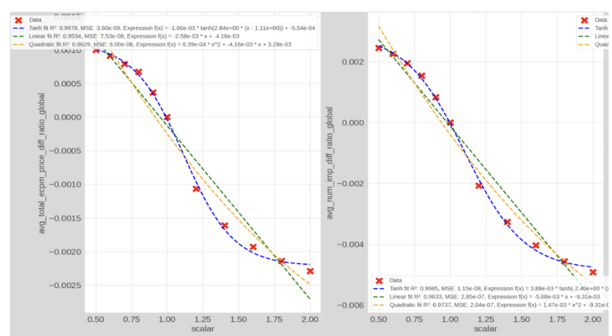

Figure 5: The non-linear curve ($tanh$) fitting of the APS (left) and UE (right) w.r.t the ads load (red cross is the historical data point).

In the application of MTO to the collected data, how APS and UE change with respect to the ads load, we performed a nonlinear curve fitting to ensure the optimization smoothness, numerical stability and computational efficiency for the optimization solver (see Figure 5). After $tanh()$ non-linear curve fitting, the optimization problem formulation would be

$$\begin{aligned} \max \quad & \Sigma_{i=1}^{n} a_i \cdot tanh(k_i \cdot (x_i - x_{0,i})) + b_i \\ \text{s.t.} \quad & \Sigma_{i=1}^{n} c_i \cdot tanh(m_i \cdot (x_i - x_{c,i})) + d_i \leq A \\ & 0.5 \leq x_i \leq 2.0 \quad \forall_i \in \{1, 2, \cdots, n\}, \end{aligned} \tag{3}$$

where a, k, x_0, b are parameters of $tanh()$ for the objective (e.g., APS), and c, m, x_c, d are the ones for cost (e.g., UE), A is the constraint threshold (e.g., 0.0001), x is continuous with a bound of 0.5, 2.0 that represent the ads load, n is the number of cohorts. Then we follow the steps below to solve the problems:

- Maximization over Scalars $\boldsymbol{\theta}_c$. For a fixed λ, solve the $max()$ part of the above problem using numerical optimization methods (e.g. SLSQP given problem is non-convex);
- Minimization over λ. Update λ using gradient ascent $\lambda^{k+1} = \lambda^k + \eta(\Sigma_{c \in C} I(\boldsymbol{\theta}_c^{(k)}) - A)$;
- Stopping Criterion: iterate until the constraint violation is minimal and objective stabilizes.

4 Experiments

4.1 Offline Model Training

We set up the debiased data collection to serve the online traffic and collect training data for a period of 3 months. The features we used for training are mainly the real-time signals. For example, we count how many ads click and how many organic contents consumed in different time windows (15m, 1h and 6h). Those features are fed to the meta learner to estimate the sensitivity of sessions.

We evaluate the performance of lift using the Area Under the Uplift Curve (AUUC), which measures the percentage of total uplift achieved by targeting the specific cohorts. A higher AUUC score indicates better model performance, which suggests the model can efficiently identify the cohorts for treatments.

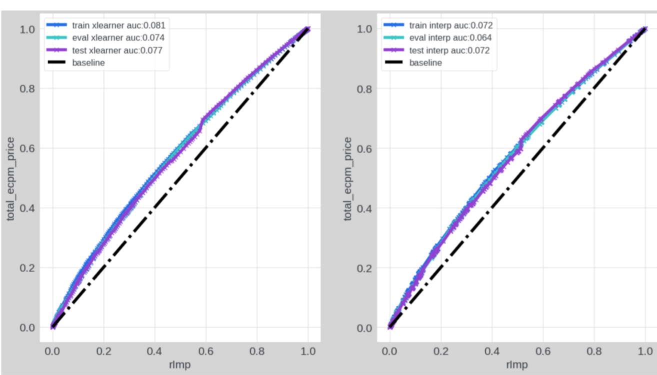

Figure 6: The train/eval/test AUUC of EC X-learner and the student model (on the right).

The train, evaluation and test AUUC plot was shown in Figure 6. From Figure 6, we can see that the model has very similar AUUC performance on the train/evaluation/test dataset, and the student model achieves similar performance as the EC X-learner. We set up a A/B test to compare the model with the random treatment baseline and we can achieve much better efficiency over the random baseline in terms of Δ_{APS} over the Δ_{UE} listed in the Table 1.

Table 1: The uplift improvement by the SSM over different cohorts.

Cohort	Efficiency↑
Global	4.35x
US and Canada	4.17x
UDV9	7.14x
UDV8	1.52x
UDV7	3.68x

4.2 SSM vs. MTO

Following the steps described in Section 3.3, we can find a global solution to maximize the overall total APS under the constraints of per-UDV cohort UE within a certain threshold. With MTO, we can optimize both global and per-UDV efficiency and mitigate the udv-bias (high UDV cohorts get higher ads load). Figure 7 shows the ads load per-UDV cohort for MTO (on the top) and the SSM (on the bottom). We clearly observe that MTO helps to reduce UDV-bias, since SSM itself tends to deliver more ads to high UDV cohorts.

Figure 7: The ads load per UDV cohort for MTO and SSM.

4.3 Online Results

We put the MTO into an online A/B test. To measure the UE impact of an ad load treatment, we set up a control group by applying the ad load treatment to randomly assigned sessions. We report online outcomes using UE and APS to evaluate the treatment efficiency, and time spent (i.e., the total duration a user engages with the entire session) as the engagement-related metric. These metrics were chosen due to their strong long-term correlation with our core business goals (e.g., time spent is highly correlated to long-term engagement metrics such as daily active users and session counts).

From Table 2, we clearly see that the MTO has an efficiency of APS/UE 1.83, which is around $8\times$ more efficient than the random treatment. The random treatment tends to increase more ads impression at the cost of the user engagement (-0.028% time spent in the sessions), while MTO treatment increases the users' time spent in the sessions, suggesting that our proposed ads optimization framework is effective in optimizing APS and UE tradeoff.

Table 2: Online results of MTO and the random treatment baseline.

Online Metrics	MTO	Random
UE	0.300%	1.670%
APS	0.550%	0.39%
Time Spent	0.056%	−0.028%
Efficiency	1.830	0.230

5 Conclusion

In this paper, we proposed an SSM-MTO framework for session-level ads load optimization. It first leverages a session sensitivity model (SSM) as a lever to adjust ads load for each session; then it uses multi-treatment optimization (MTO) with the incorporated business constraints to dynamically determine the optimal ads load for each session. The SSM is trained on the data collected from the debiased data collection which randomizes the ads load at the session level to remove the feature bias caused by ads load treatments. From the offline training data, we showed that the SSM-MTO identifies the efficient sessions for ads load treatment. We also set up an A/B test for the SSM-MTO; the online results showed that the SSM-MTO has $8\times$ more efficiency than the baseline and presents a better trade-off between ads performance score and user experience.

References

[Carrion *et al.*, 2021] Carlos Carrion, Zenan Wang, Harikesh S. Nair, Xianghong Luo, Yulin Lei, Xiliang Lin, Wenlong Chen, Qiyu Hu, Changping Peng, Yongjun Bao, and Weipeng Yan. Blending advertising with organic content in e-commerce: A virtual bids optimization approach. *CoRR*, abs/2105.13556, 2021.

[Elwert and Winship, 2014] Felix Elwert and Christopher Winship. Endogenous selection bias: The problem of conditioning on a collider variable. *Annual review of sociology*, 40(1):31–53, 2014.

[Ke *et al.*, 2017] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.

[Künzel *et al.*, 2019] Sören R Künzel, Jasjeet S Sekhon, Peter J Bickel, and Bin Yu. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the national academy of sciences*, 116(10):4156–4165, 2019.

[Liao *et al.*, 2022] Guogang Liao, Ze Wang, Xiaoxu Wu, Xiaowen Shi, Chuheng Zhang, Yongkang Wang, Xingxing Wang, and Dong Wang. Cross DQN: cross deep Q network for ads allocation in feed. In Frédérique Laforest, Raphaël Troncy, Elena Simperl, Deepak Agarwal, Aristides Gionis, Ivan Herman, and Lionel Médini, editors, *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, pages 401–409. ACM, 2022.

[Liu *et al.*, 2023] Dugang Liu, Xing Tang, Han Gao, Fuyuan Lyu, and Xiuqiang He. Explicit feature interaction-aware uplift network for online marketing. In Ambuj K. Singh, Yizhou Sun, Leman Akoglu, Dimitrios Gunopulos, Xifeng Yan, Ravi Kumar, Fatma Ozcan, and Jieping Ye, editors, *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2023, Long Beach, CA, USA, August 6-10, 2023*, pages 4507–4515. ACM, 2023.

[Liu *et al.*, 2025] Tao Liu, Qi Xu, Wei Shi, Zhigang Hua, and Shuang Yang. Session-level dynamic ad load optimization using offline robust reinforcement learning. In Yizhou Sun, Flavio Chierichetti, Hady W. Lauw, Claudia Perlich, Wee Hyong Tok, and Andrew Tomkins, editors, *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining, V.1, KDD 2025, Toronto, ON, Canada, August 3-7, 2025*, pages 2458–2468. ACM, 2025.

[Nie and Wager, 2021] Xinkun Nie and Stefan Wager. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108(2):299–319, 2021.

[Rafieian, 2023] Omid Rafieian. Optimizing user engagement through adaptive ad sequencing. *Mark. Sci.*, 42(5):910–933, 2023.

[Sagtani *et al.*, 2024] Hitesh Sagtani, Madan Gopal Jhawar, Rishabh Mehrotra, and Olivier Jeunen. Ad-load balancing via off-policy learning in a content marketplace. In Luz Angelica Caudillo-Mata, Silvio Lattanzi, Andrés Muñoz Medina, Leman Akoglu, Aristides Gionis, and Sergei Vassilvitskii, editors, *Proceedings of the 17th ACM International Conference on Web Search and Data Mining, WSDM 2024, Merida, Mexico, March 4-8, 2024*, pages 586–595. ACM, 2024.

[Wang *et al.*, 2022] Ze Wang, Guogang Liao, Xiaowen Shi, Xiaoxu Wu, Chuheng Zhang, Bingqi Zhu, Yongkang Wang, Xingxing Wang, and Dong Wang. Hybrid transfer in deep reinforcement learning for ads allocation. In Mohammad Al Hasan and Li Xiong, editors, *Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, October 17-21, 2022*, pages 4560–4564. ACM, 2022.

[Wu *et al.*, 2023] Guojun Wu, Ge Song, Xiaoxiang Lv, Shikai Luo, Chengchun Shi, and Hongtu Zhu. Dnet: Distributional network for distributional individualized treatment effects. In Ambuj K. Singh, Yizhou Sun, Leman Akoglu, Dimitrios Gunopulos, Xifeng Yan, Ravi Kumar, Fatma Ozcan, and Jieping Ye, editors, *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2023, Long Beach, CA, USA, August 6-10, 2023*, pages 5215–5224. ACM, 2023.

[Yan *et al.*, 2020] Jinyun Yan, Zhiyuan Xu, Birjodh Tiwana, and Shaunak Chatterjee. Ads allocation in feed via constrained optimization. In Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash, editors, *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020*, pages 3386–3394. ACM, 2020.