Online Appendices

A. Summary of Notations

Notation	Description
TTD	Two-stage target-debt framework
DWO	Debt-weighted offer-set policy
Fluid	Fluid-approximation approach
FV	Fairness-adjusted value
T	Total number of customers
$\mathcal{N}$	Set of ad campaigns
n	Number of ad campaigns
B_i	Total budget of ad campaign i
b_i	Bid price of ad campaign i per click-through
$j(t)$	Type of customer t
$\mathcal{M}$	Set of customer types
m	Number of customer types
p^j	Probability of a customer being type j
$S(t)$	Offer-set displayed to customer t
$\mathfrak{S}^j$	Collection of all possible offer-sets for type-j customers including ad targeting info
$y_i^j(t)$	Number of click-throughs by a type-j customer on ad i in time t
$\bar{y}_i^j$	Per-customer click-throughs of ad i by type-j customers
$\phi_i^j(S)$	Expected value of $y_i^j(t)$ conditioned on $S(t) = S$
$\mathsf{D}_{(j,y)}$	Joint customer type and click-through distribution
$\eta_i^{\mathcal{C}}$	Required click-throughs for customer-type set $\mathcal{C}$ on ad campaign i
$\mathfrak{K}_i$	Set of all $\mathcal{C}$ with $\eta_i^{\mathcal{C}} > 0$ of ad i
r_i^j	Value of each click of ad campaign i by a type-j customer
$F(\boldsymbol{y})$	Fairness metric
$\mathcal{H}_{t-1}$	Realized history until the start of time t
Π	Set of policies
Π_{static}	Set of static policies
Π_{d}	Set of deterministic static policies
$(\mathcal{OP})$	Original stochastic program
$\mathcal{V}^*$	Optimal FV of the original stochastic program
α_j^i	Target for the per-period number of click-throughs of ad i from type-j customers
$\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha})$	FV of the click-through target vector $\boldsymbol{\alpha}$
$(2\mathcal{SSP})$	Two-stage stochastic program
θ_i^j	Dual variable associated with satisfying the click-through target ad i from type-j customers
$(\mathcal{OTP})$	Reformulated optimal target problem
$\boldsymbol{\alpha}^*$	Solution to $(\mathcal{OTP})$
K	Maximum size of an offer-set
DWO-$\boldsymbol{\alpha}$	DWO policy with click-through target vector $\boldsymbol{\alpha}$
$d_i^j(t)$	Debt of the click-throughs from type-j customers on ad i in time t
$\mathcal{OP}(\gamma)$	Family of ad-allocation problems with scaling parameter γ
$\mathcal{V}(\pi\|\gamma)$	Expected FV generated by policy π in $\mathcal{OP}(\gamma)$
$(\mathcal{OP}_{\mathsf{Fluid}})$	Fluid convex program
$\mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z})$	Objective value function of $(\mathcal{OP}_{\mathsf{Fluid}})$

Table 2 Summary of Notations

B. Soft-constraint Formulation

In this section, we consider the original stochastic program with soft constraints as follows:

$$\begin{aligned} \max_{\pi\in\Pi} \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi) + \lambda F(\bar{\boldsymbol{y}}(\pi)) - \nu \sum_{i\in\mathcal{N}}\sum_{\mathcal{C}\in\mathfrak{K}_i}\left(\frac{\eta_i^{\mathcal{C}}}{T} - \sum_{j\in\mathcal{C}}\bar{y}_i^j(\pi)\right)^+\right] \\ \text{s.t. } \frac{1}{T}\sum_{t=1}^{T}\sum_{j\in\mathcal{M}} b_i y_i^j(t|\pi) \le \frac{B_i}{T}, \textit{ almost surely} \text{ for each } i\in\mathcal{N}, \end{aligned} \qquad (\mathcal{OP}_{Soft})$$

where the first term in the objective is the total per-customer-impression value from advertising, which we call the *efficiency* of policy π denoted by $\mathcal{E}(\pi) := \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi)\right]$, the second term is the *fairness* of policy π denoted by $\lambda\cdot\mathcal{F}(\pi) := \lambda\cdot\mathbb{E}[F(\bar{\boldsymbol{y}}(\pi))]$, and the third term in the objective is the soft constraints of the click-through requirements denoted by $\nu\cdot\mathcal{G}(\pi) := -\nu\cdot\mathbb{E}\left[\sum_{i\in\mathcal{N}}\sum_{\mathcal{C}\in\mathfrak{K}_i}\left(\frac{\eta_i^{\mathcal{C}}}{T} - \sum_{j\in\mathcal{C}}\bar{y}_i^j(\pi)\right)^+\right]$ with a parameter $\nu > 0$. Hence, the total objective value denoted by FV_{Soft} under policy π is given by $\mathcal{V}_{Soft}(\pi) := \mathcal{E}(\pi) + \lambda\cdot\mathcal{F}(\pi) + \nu\cdot\mathcal{G}(\pi)$, and we denote the optimal FV_{Soft} as $\mathcal{V}^*_{Soft} = \limsup_{\pi\in\Pi}\mathcal{V}_{Soft}(\pi)$ and the optimal policy (if it exists) as $\pi^* = \arg\max_{\pi\in\Pi}\mathcal{V}_{Soft}(\pi)$. We also remark that the constraint of ($\mathcal{OP}_{Soft}$) refers to the budget constraint of each ad.

We formulate the original stochastic program ($\mathcal{OP}_{Soft}$) as a dynamic program (DP). Specifically, we define $Y_i^j(t) := \sum_{\tau=1}^{t-1} y_i^j(\tau)$ as the accumulative number of click-throughs until the start of time t, and

$$\begin{aligned} \mathbb{V}_t(\boldsymbol{Y}(t)) := \max_{\pi\in\Pi}\mathbb{E}\left[\sum_{\tau=t}^{T}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi) + T\lambda F(\bar{\mathbf{y}}(\pi)) - \nu\sum_{i\in\mathcal{N}}\sum_{\mathcal{C}\in\mathfrak{K}_i}\left(\eta_i^{\mathcal{C}} - \sum_{t=1}^{T}\sum_{j\in\mathcal{C}} y_i^j(t)\right)^+ \middle| \boldsymbol{Y}(t)\right] \\ \text{s.t. } \sum_{\tau=t}^{T}\sum_{j\in\mathcal{M}} b_i y_i^j(t|\pi) \le B_i - \sum_{j\in\mathcal{M}} b_i Y_i^j(t), \textit{ almost surely} \text{ for each } i\in\mathcal{N}. \end{aligned} \qquad (19)$$

Hence, $\mathbb{V}_t(\boldsymbol{Y}(t))$ is the maximum expected FV_{Soft} given that the number of accumulative click-throughs at the beginning of time t is $\boldsymbol{Y}(t) := (Y_i^j(t) : i\in\mathcal{N}, j\in\mathcal{M})$.

To formulate the DP, we first specify the boundary/terminal value function $\mathbb{V}_{T+1}(\boldsymbol{Y}(T+1))$. To this end, we define

$$\mathcal{Y} := \left\{\boldsymbol{Y}(T+1)\in\mathbb{R}_+^{nm} : b_i\sum_{j\in\mathcal{M}} Y_i^j(T+1) \le B_i \text{ for each } i\in\mathcal{N}\right\}$$

as the feasible region for the accumulative number of click-throughs for the entire planning horizon, $\boldsymbol{Y}(T+1)$. The boundary value function is defined as follows:

$$\mathbb{V}_{T+1}(\boldsymbol{Y}(T+1)) = \begin{cases} T\lambda F\left(\frac{\boldsymbol{Y}(T+1)}{T}\right) - \nu\sum_{i\in\mathcal{N}}\sum_{\mathcal{C}\in\mathfrak{K}_i}\left(\eta_i^{\mathcal{C}} - \sum_{j\in\mathcal{C}} Y_i^j(T+1)\right)^+, & \text{if } \boldsymbol{Y}(T+1)\in\mathcal{Y}, \\ -\bar{M}, & \text{otherwise}, \end{cases} \qquad (20)$$

where $\bar{M}$ is a sufficiently large positive number that is far bigger than $\mathcal{V}^*_{Soft}$ (e.g., $\bar{M} := C\cdot\max\{\mathcal{V}^*_{Soft}, 1\}$, where $C > 0$ is a very large positive number).

By the standard backward induction argument, we are now ready to write the Bellman equation to evaluate $\mathbb{V}_t(\boldsymbol{Y}(t))$ in (19):

$$\mathbb{V}_t(\boldsymbol{Y}(t)) = \sum_{j\in\mathcal{M}} p^j \max_{S(t)\in\mathfrak{S}^j}\mathbb{E}_{\boldsymbol{y}(t)}\left[\sum_{i\in\mathcal{N}} r_i^j y_i^j + \mathbb{V}_{t+1}(\boldsymbol{Y}(t) + \boldsymbol{y}(t))\middle| S(t), j(t) = j\right]. \qquad (21)$$

Therefore, the optimal FV_{Soft} for the original problem $(\mathcal{OP}_{Soft})$ is

$$\mathcal{V}^*_{Soft} = \frac{\mathbb{V}_1(\mathbf{0})}{T}, \text{ where } \mathbf{0} := (0, 0, ..., 0)' \in \mathbb{R}^{nm}.$$

Due to the curse of dimensionality, the above DP formulation of $(\mathcal{OP}_{Soft})$ is intractable even when m and n are just moderately large. Therefore, we resort to our TTD framework and the induced DWO algorithm to solve the ad-allocation optimization problem. We can construct a related optimal target problem as follows:

$$\begin{aligned} \max_{\boldsymbol{\alpha}\geq \mathbf{0}} \ & \mathcal{V}_{Soft-\mathsf{CT}}(\boldsymbol{\alpha}) \\ \text{s.t. } & h(\boldsymbol{\alpha}) \geq 0, \\ & b_i \sum_{j\in\mathcal{M}} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i \in \mathcal{N}, \end{aligned} \qquad (\mathcal{OTP}_{Soft})$$

where $\mathcal{V}_{Soft-\mathsf{CT}}(\boldsymbol{\alpha}) := \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j \alpha_i^j + \lambda F(\boldsymbol{\alpha}) + \nu G(\boldsymbol{\alpha})$ and $G(\boldsymbol{\alpha}) := -\sum_{i\in\mathcal{N}}\sum_{\mathcal{C}\in\mathfrak{K}_i}\left(\frac{\eta_i^{\mathcal{C}}}{T} - \sum_{j\in\mathcal{C}}\alpha_i^j\right)^+$. It is straightforward to check that $\mathcal{V}_{Soft-\mathsf{CT}}(\cdot)$ is concave in $\boldsymbol{\alpha}$ and the constraint $h(\boldsymbol{\alpha}) \geq 0$ is equivalent to (3). Solving the convex program $(\mathcal{OTP}_{Soft})$ obtains the optimal targets $\boldsymbol{\alpha}^*_{Soft}$. Then, we apply the DWO-$\boldsymbol{\alpha}^*_{Soft}$ policy to dynamically display the offer-set to each arriving customer. Following the same analysis as our main model with the expected click-through requirements, we can show that the DWO-$\boldsymbol{\alpha}^*_{Soft}$ policy is also asymptotically optimal for $(\mathcal{OP}_{Soft})$. To avoid repetition, we omit the proof details.

C. Metrics of Fairness

In this section, we describe a few commonly adopted metrics of fairness $F(\cdot)$, all of which are concave and can be coherently incorporated into our framework.

Max-min fairness. The recent trend of machine-learning fairness has promoted that minority customers should have sufficient click-throughs in a recommender/advertising system; otherwise, their needs cannot be well taken care of due to data scarcity. A natural choice to accommodate such fairness concern is the max-min fairness metric, which has been extensively studied in the literature of economics (e.g., Young and Isaac 1995), computer science (e.g., Kumar and Kleinberg 2000), and operations research (e.g., Bertsimas et al. 2012). Specifically, we define function $F(\cdot)$ as follows:

$$F(\bar{\boldsymbol{y}}) = \min_{j\in\mathcal{M}}\left\{\sum_{i\in\mathcal{N}} \bar{y}_i^j\right\}. \tag{22}$$

Max-min fairness drives the platform to maximize the minimum per-customer-impression click-throughs from all customer types, ensuring that no customer type has too few click-throughs. We also note that max-min fairness can also be evaluated with respect to advertisers, i.e., $F(\bar{\boldsymbol{y}}) = \min_{i\in\mathcal{N}}\left\{\sum_{j\in\mathcal{M}}\bar{y}_i^j\right\}$, so as to ensure that no advertiser receives too few click-throughs. One may also generalize max-min fairness to the $\alpha-$fairness metric (see, e.g., Bertsimas et al. 2012), i.e., $F(\bar{\boldsymbol{y}}) = \sum_{j\in\mathcal{M}}\frac{1}{1-\alpha}\left(\sum_{i\in\mathcal{N}}\bar{y}_i^j\right)^{1-\alpha}$ if $\alpha \neq 1$ and $F(\bar{\boldsymbol{y}}) = \sum_{j\in\mathcal{M}}\log\left(\sum_{i\in\mathcal{N}}\bar{y}_i^j\right)$ if $\alpha = 1$, which is reduced to the max-min fairness metric if we take $\alpha \to +\infty$.

Gini mean difference fairness. Advertisers generally prefer receiving impressions/click-throughs/conversions that are evenly spread across their targeted customer types (e.g., Lejeune and Turner 2019). One way to capture such preference is through Gini mean difference (GMD) fairness. The Gini coefficient/index has long been a canonical measure of income inequality in economics (e.g., Atkinson 1970), and it has recently been studied in the advertising literature to maximize the spreading of impressions across targeted user types (e.g., Lejeune and Turner 2019). Following Lejeune and Turner (2019), given the average click-through vector $\bar{\boldsymbol{y}}_i = (\bar{y}_i^1, \bar{y}_i^2, ..., \bar{y}_i^m)'$ of ad i, we first define the GMD fairness for each ad i:

$$GMD_i(\bar{\boldsymbol{y}}_i) = \frac{2}{(\sum_{j\in\mathcal{M}} p^j)^2} \sum_{j,j'\in\mathcal{M}} p^j p^{j'} \left| \frac{\bar{y}_i^j}{p^j} - \frac{\bar{y}_i^{j'}}{p^{j'}} \right|, \tag{23}$$

where p^j is the proportion of type-j customers, and $\frac{\bar{y}_i^j}{p^j} = \frac{\sum_{t=1}^T y_i^j(t)}{p^j T}$ is the per-type-j customer click-throughs of ad i. Hence, the Gini coefficient of ad i is defined as follows:

$$G_i(\bar{\boldsymbol{y}}_i) = \frac{\left(\sum_{j\in\mathcal{M}} p^j\right) GMD_i(\bar{\boldsymbol{y}}_i)}{2\sum_{j\in\mathcal{M}} \bar{y}_i^j}.$$

We are now ready to define the GMD fairness as a weighted sum of the negative Gini coefficient of each ad:

$$F(\bar{\boldsymbol{y}}) = -\sum_{i\in\mathcal{N}} k_i G_i(\bar{\boldsymbol{y}}_i) = -\sum_{i\in\mathcal{N}} \frac{k_i}{\left(\sum_{j\in\mathcal{M}} \bar{y}_i^j\right)\left(\sum_{j\in\mathcal{M}} p^j\right)} \sum_{j,j'\in\mathcal{M}} |p^{j'}\bar{y}_i^j - p^j \bar{y}_i^{j'}|,$$

where $k_i \geq 0$ is the weight of ad i according to its importance in the GMD fairness metric. Following Lejeune and Turner (2019), we choose $k_i = \sum_{j\in\mathcal{M}} \bar{y}_i^j$ which gives rise to our GMD fairness metric as (24):

$$F(\bar{\boldsymbol{y}}) = -\sum_{i\in\mathcal{N}} \frac{1}{\sum_{j\in\mathcal{M}} p^j} \sum_{j,j'\in\mathcal{M}} |p^{j'}\bar{y}_i^j - p^j \bar{y}_i^{j'}|. \tag{24}$$

It is clear from (24) that the GMD fairness metric prompts the platform to induce click-throughs from each targeted customer type j proportional to its traffic p^j.

Disparate impact. Disparate impact is a widely discussed algorithmic discrimination measure (e.g. Feldman et al. 2015). We show that this measure can be adopted in $F(\cdot)$. Following Feldman et al. (2015), we call there exists disparate impact for type-j customers if

$$\frac{\bar{y}_i^j/p^j}{\bar{y}_i^{j'}/p^{j'}} \leq \tau \text{ for some } i \text{ and } j' \neq j,$$

where τ is a parameter in $[0,1]$. In practice, τ is usually set as 0.8 given the prevailing $80\% - 20\%$ rule (see, e.g. Rubin 1978). Hence, we can define the fairness metric $F(\cdot)$ for eliminating disparate impact as follows:

$$F(\bar{\boldsymbol{y}}) = \min_{i\in\mathcal{N}, j\in\mathcal{J}_i} \left\{ \frac{\bar{y}_i^j}{p^j} - \tau \cdot \max_{j'\in\mathcal{J}_i'} \frac{\bar{y}_i^{j'}}{p^{j'}} \right\}, \tag{25}$$

where $\mathcal{J}_i$ and $\mathcal{J}_i'$ ($\mathcal{J}_i, \mathcal{J}_i' \in \mathcal{M}$) denote the sets of minority and majority types, respectively. It is straightforward to check that (25) is concave, which measures the disparate impact attributed to the ad allocation algorithm.

D. Proof of Statements

We provide the proof of all the technical results in this section.

Proof of Lemma 1

It is evident that any feasible click-through target vector $\boldsymbol{\alpha}$ to (2$\mathcal{SSP}$) must also be feasible to (2), because the feasible region of (2$\mathcal{SSP}$) is a subset of that of (2).

We now show that any feasible click-through target vector $\boldsymbol{\alpha}$ to (2) is also feasible to (2$\mathcal{SSP}$). Consider a feasible $\boldsymbol{\alpha}$ to (2). We have, there exists a feasible policy $\pi \in \Pi$,

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\pi)\right] \geq \alpha_i^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

We define the following probability measure induced by π, $z_{\mathsf{static}}(S|\pi)$: For any $j \in \mathcal{M}$ and $S \in \mathfrak{S}^j$,

$$z^j_{\mathsf{static}}(S|\pi) := \mathbb{P}[\pi_{\tilde{t}}(j(\tilde{t}), \mathcal{H}_{\tilde{t}-1}) = S | j(\tilde{t}) = j],$$

where $\tilde{t}$ is a random variable uniformly distributed on $\{1, 2, ..., T\}$ and independent of everything else. Based on $\boldsymbol{z}_{\mathsf{static}}(\cdot|\pi)$, we construct a static policy π_{static}, which selects offer-set S given each customer type j with probability $z^j_{\mathsf{static}}(S|\pi)$. Straightforward algebraic manipulations and the law of iterated expectations together yield that

$$\mathbb{E}\left[y_i^j(t|\pi_{\mathsf{static}})\right] = \sum_{S \in \mathfrak{S}^j} p^j \phi(j|S) z^j_{\mathsf{static}}(S|\pi) = \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\pi)\right] \geq \alpha_i^j, , \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

Hence, π_{static} and $\boldsymbol{\alpha}$ are feasible to (2$\mathcal{SSP}$). Moreover, because these two problems have the same objective function, any optimal click-through target vector for one problem must be also optimal for another. This completes the proof of Lemma 1. □

Proof of Theorem 1

The proof follows from the discussions before (8). □

Proof of Proposition 1

Before proving Proposition 1, we first state and prove a few auxiliary results. It is sometimes more convenient to use a binary variable representation of a deterministic static offer-set policy $\pi \in \Pi_{\mathsf{d}}$. More specifically, $\pi \in \Pi_{\mathsf{d}}$ can be equivalently represented by an nm-dimensional binary vector $\boldsymbol{x} = (x_i^j \in \{0,1\} : i \in \mathcal{N}, j \in \mathcal{M})$, where $x_i^j = 1$ means that $i \in \pi(j)$, i.e., ad i is included in the offer-set displayed to a type-j customer. With a slight abuse of notation, we denote $\phi_i^j(\boldsymbol{x})$ as the expected click-throughs of a type-j customer for ad i if the offer-set displayed to this customer is $S^j = \{i \in \mathcal{N} : x_i^j = 1\}$. Under the MNL model, we have

$$\phi_i^j(\boldsymbol{x}) = \frac{v_i^j x_i^j}{1 + \sum_{i' \in \mathcal{N}} v_{i'}^j x_{i'}^j}, \tag{26}$$

where $v_i^j > 0$ is the attractiveness of ad i to type-j customers (see, also, (11)). Denote the set of all plausible offer-set representation vectors as $\mathcal{X} \subset \{0,1\}^{nm}$, and the set of plausible offer-set representation vectors displayed to a type-j customer as $\mathcal{X}^j \subset \{0,1\}^n$. Applying Theorems 1 to the MNL choice model (26), we have the following corollary.

Corollary 1. *If customers follow the MNL click-through model (26), a click-through target vector $\boldsymbol{\alpha}$ is single-period feasible if and only if*

$$\max_{\boldsymbol{x}\in\mathcal{X}} \sum_{i\in\mathcal{N},j\in\mathcal{M}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} \geq \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j \theta_i^j \ \textit{for all } \theta_i^j \geq 0 \ (i\in\mathcal{N}, j\in\mathcal{M}). \tag{27}$$

Furthermore, (27) is equivalent to, for each $j\in\mathcal{M}$

$$\max_{\boldsymbol{x}^j\in\mathcal{X}^j} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} \geq \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j \ \textit{for all } \theta_i^j \geq 0 \ (i\in\mathcal{N}). \tag{28}$$

Proof of Corollary 1

Directly applying Theorem 1 to the MNL choice model implies that $\boldsymbol{\alpha}$ is feasible if and only if inequality (27) holds.

We now show that (27) implies (28). If (27) holds for any $\boldsymbol{\theta}\geq\mathbf{0}$, then it also holds for any $\boldsymbol{\theta}\geq\mathbf{0}$ with $\theta_i^{j'}=0$ (for $j'\neq j$ and all $i\in\mathcal{N}$). Therefore, (28) holds.

Finally, we show that if (28) holds for all $j\in\mathcal{M}$, (27) holds as well. Note that the left-hand side of (27) can be decomposed into independent parts as follows:

$$\max_{\boldsymbol{x}\in\mathcal{X}} \sum_{i\in\mathcal{N},j\in\mathcal{M}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} = \max_{\boldsymbol{x}\in\mathcal{X}} \sum_{j\in\mathcal{M}} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} = \sum_{j\in\mathcal{M}} \max_{\boldsymbol{x}^j\in\mathcal{X}^j} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} \geq \sum_{j\in\mathcal{M}} \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j = \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j \theta_i^j,$$

for any $\boldsymbol{\theta}\geq\mathbf{0}$, where the inequality follows from (28). Therefore, that (27) holds is equivalent to that (28) holds for all $j\in\mathcal{M}$. This completes the proof of Corollary 1. □

Leveraging the structural properties of the MNL model, we can give a sharper and simpler characterization for the feasibility condition (as the solution to a linear program). The following lemma characterizes the feasibility condition for a click-through target vector $\boldsymbol{\alpha}$, taking into account the cardinality constraint that the size of an offer-set displayed to any customer is upper bounded by K, i.e., $|S(t)|\leq K$ for any customer t.

Lemma 3. *If customers follow the MNL click-through model (26) and the set of all feasible offer-sets is $\mathfrak{S}^j=\{S\subset\mathcal{N}:|S|\leq K\}$ for each $j\in\mathcal{M}$, we have $\boldsymbol{\alpha}$ is single-period feasible if and only if there exist $\boldsymbol{w}:=(w_i^j:i\in\mathcal{N},j\in\mathcal{M})$ and $\boldsymbol{z}:=(z^j:j\in\mathcal{M})$ that satisfy the following linear constraints*

$$\begin{aligned} &p^j v_i^j w_i^j \geq \alpha_i^j,\ w_i^j \leq z^j,\ w_i^j\geq 0,\ \textit{for each } i\in\mathcal{N}, j\in\mathcal{M},\\ &\sum_{i\in\mathcal{N}} v_i^j w_i^j + z^j = 1,\ \sum_{i\in\mathcal{N}} w_i^j \leq K z^j,\ \textit{for each } j\in\mathcal{M}, \end{aligned} \tag{29}$$

where $z^j := \frac{1}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j}$ *and* $w_i^j := x_i^j z^j = \frac{x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j}$.

Proof of Lemma 3

A standard result in fractional programming postulates that the left-hand side of (28) is quasi-convex in $\boldsymbol{x}^j$ for all $j\in\mathcal{M}$, so there always exists a maximizer on the boundary of the feasible region. Thus, we can relax the binary constraint $x_i^j\in\{0,1\}$ to $x_i^j\in[0,1]$ in (28), which is therefore equivalent to

$$\max_{\boldsymbol{x}^j\in[0,1]^n,\ \sum\limits_{i\in\mathcal{N}} x_i^j\leq K} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} \geq \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j \ \text{for all } \boldsymbol{\theta}^j\geq\mathbf{0} \text{ and } j\in\mathcal{M}. \tag{30}$$

We change the decision variable and define, for all $j \in \mathcal{M}$,

$$z^j := \frac{1}{1+\sum_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} \text{ and } w_i^j := x_i^j z^j = \frac{x_i^j}{1+\sum_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j}.$$

Then, we can rewrite (30) as, for any $j \in \mathcal{M}$,

$$\begin{aligned} \min_{\boldsymbol{\theta}^j \geq \mathbf{0}} & \left(\max_{w_i^j, z^j} \sum_{i\in\mathcal{N}} p^j v_i^j w_i^j \theta_i^j - \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j \right) \geq 0 \\ \text{s.t. } & \sum_{i\in\mathcal{N}} v_i^j w_i^j + z^j = 1, \\ & \sum_{i\in\mathcal{N}} w_i^j \leq K z^j, \\ & 0 \leq w_i^j \leq z^j, \text{ for each } i \in \mathcal{N}. \end{aligned} \tag{31}$$

By Sion's minimax theorem, we can exchange the maximization and minimization operators so that (31) is equivalent to, for any $j \in \mathcal{M}$:

$$\begin{aligned} \max_{\boldsymbol{w}^j, z^j} \min_{\boldsymbol{\theta}^j \geq 0} & \sum_{i\in\mathcal{N}} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \geq 0, \\ \text{s.t. } & \sum_{i\in\mathcal{N}} v_i^j w_i^j + z^j = 1, \\ & \sum_{i\in\mathcal{N}} w_i^j \leq K z^j, \\ & 0 \leq w_i^j \leq z^j, \text{ for each } i \in \mathcal{N}. \end{aligned} \tag{32}$$

Therefore, (32) holds if and only if there exist $\boldsymbol{w}^j$ and z^j such that all the constraints in (32) hold and $\sum_{i\in\mathcal{N}} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \geq 0$ holds for all $\boldsymbol{\theta}^j \geq \mathbf{0}$, which is equivalent to $p^j v_i^j w_i^j - \alpha_i^j \geq 0$ for all $i \in \mathcal{N}$. Therefore, (32) is equivalent to that, for any $j \in \mathcal{M}$,

$$\begin{aligned} & p^j v_i^j w_i^j - \alpha_i^j \geq 0, \text{ for each } i \in \mathcal{N}, \\ & \sum_{i\in\mathcal{N}} v_i^j w_i^j + z^j = 1, \\ & \sum_{i\in\mathcal{N}} w_i^j \leq K z^j, \\ & 0 \leq w_i^j \leq z^j, \text{ for each } i \in \mathcal{N}. \end{aligned} \tag{33}$$

That (33) holds for all $j \in \mathcal{M}$ is equivalent to that (29) holds. This completes the proof of Lemma 3. □

We now prove Proposition 1 itself. It suffices to show that, taking into account the cardinality constraint $|S| \leq K$, the (first-stage) feasible region for the first-stage click-through target vector $\boldsymbol{\alpha}$ is given by the following linear constraints:

$$\mathcal{A}_{MNL} := \left\{ \boldsymbol{\alpha} \in \mathbb{R}_+^{nm} : \sum_{i'\in\mathcal{N}} \alpha_{i'}^j + \frac{\alpha_i^j}{v_i^j} \leq p^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}, \text{ and } \sum_{i\in\mathcal{N}} \alpha_i^j + \frac{1}{K} \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j} \leq p^j, \text{ for each } j \in \mathcal{M} \right\}. \tag{34}$$

We first show that if (29) holds, then $\boldsymbol{\alpha} \in \mathcal{A}_{MNL}$. By the first inequality of (29), we have $v_i^j w_i^j \geq \frac{\alpha_i^j}{p^j}$ for all $i \in \mathcal{N}$ and $j \in \mathcal{M}$. Plugging this into the first equality of (29), we have

$$1 - z^j = \sum_{i\in\mathcal{N}} v_i^j w_i^j \geq \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j}, \text{ for each } j \in \mathcal{M}.$$

Thus, by the first and second inequalities of (29), we have

$$\sum_{i'\in\mathcal{N}} \frac{\alpha_{i'}^j}{p^j} \le 1 - z^j \le 1 - w_i^j \le 1 - \frac{\alpha_i^j}{p^j v_i^j} \text{ for each } i\in\mathcal{N}, j\in\mathcal{M}.$$

Rearranging the terms, we have

$$p^j \ge \sum_{i'\in\mathcal{N}} \alpha_{i'}^j + \frac{\alpha_i^j}{v_i^j} \text{ for each } i\in\mathcal{N}, j\in\mathcal{M}.$$

The first, second, and fourth inequalities and the first equality of (29) imply that

$$\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j v_i^j} \le \sum_{i\in\mathcal{N}} w_i^j \le K z^j = K(1 - \sum_{i\in\mathcal{N}} v_i^j w_i^j) \le K\left(1 - \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j}\right) \text{ for each } j\in\mathcal{M}.$$

Rearranging the terms, we have

$$p^j \ge \sum_{i\in\mathcal{N}} \alpha_i^j + \frac{1}{K}\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j} \text{ for each } j\in\mathcal{M}.$$

Therefore, if (29) holds, we have $\boldsymbol{\alpha}\in\mathcal{A}_{MNL}$.

Next, we show that if $\boldsymbol{\alpha}\in\mathcal{A}_{MNL}$, (29) holds. Given $\boldsymbol{\alpha}\in\mathcal{A}_{MNL}$, define

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \text{ for each } i\in\mathcal{N} \text{ and } j\in\mathcal{M}, \text{ and } z^j = 1 - \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j} \text{ for each } j\in\mathcal{M}.$$

To show (29), it suffices to show the first, second and fourth inequalities hold because the rest of the constraints hold trivially.

Since $p^j \ge \sum_{i\in\mathcal{N}} \alpha_i^j + \frac{1}{K}\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j}$ for each $j\in\mathcal{M}$, we have

$$\sum_{i\in\mathcal{N}} w_i^j = \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j v_i^j} = \frac{1}{p^j}\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j} \le K\left(1 - \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j}\right) = K z^j \text{ for each } j\in\mathcal{M}.$$

Hence, the second inequality of (29) holds. Since $p^j \ge \sum_{i'\in\mathcal{N}} \alpha_{i'}^j + \frac{\alpha_i^j}{v_i^j}$ for each $i\in\mathcal{N}, j\in\mathcal{M}$, we have

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \le 1 - \sum_{i'\in\mathcal{N}} \frac{\alpha_{i'}^j}{p^j} = z^j \text{ for each } i\in\mathcal{N}, j\in\mathcal{M}.$$

Therefore, (29) holds. Hence, the first-stage feasible region of $\boldsymbol{\alpha}$ is characterized by (34). This completes the proof of Proposition 1. □

Proof of Theorem 2

Let us consider a problem identical to $\mathcal{OP}(\gamma)$ but without budget constraints (i.e., $B_i(\gamma) = +\infty$ for all $i\in\mathcal{N}$ and $\gamma > 0$), which we denote as $\mathcal{OP}_*(\gamma)$. By definition, in $\mathcal{OP}_*(\gamma)$, any ad i will not run out of budget throughout the planning horizon. Throughout the proof of Theorem 2, we write $y_i^j(t) = y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}))$ whenever there is no confusion.

- *Step 1.* For problem $\mathcal{OP}_*(\gamma)$, if $\boldsymbol{\alpha}$ is single-period feasible, it holds that

$$\liminf_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{t=1}^{T(\gamma)} y_i^j(t) \ge \alpha_i^j \text{ for all } i\in\mathcal{N} \text{ and } j\in\mathcal{M}. \tag{35}$$

Under the DWO-$\boldsymbol{\alpha}$ algorithm, we have that

$$t\alpha_i^j - \sum_{\tau=1}^{t} y_i^j(\tau) = d_i^j(t+1) \le \left(d_i^j(t+1)\right)^+.$$

Therefore, it suffices to show that, if (8) holds,

$$\limsup_{t\to+\infty} \frac{1}{t}\cdot d_i^j(t+1) \le 0, \text{ almost surely in problem } \mathcal{OP}_*(+\infty).$$

For a vector $\boldsymbol{x}\in\mathbb{R}^n$, we use $\boldsymbol{x}^+$ to denote the component-wise positive part of $\boldsymbol{x}$. Also note that, for any $A,B\in\mathbb{R}$, $((A+B)^+)^2 \le (A^+ + B)^2$. We have

$$\begin{aligned}\mathbb{E}||\left(\boldsymbol{d}(t+1)\right)^+||_2^2 =& \mathbb{E}||\left(\boldsymbol{d}(t)+\boldsymbol{\alpha}-\boldsymbol{y}(t)\right)^+||_2^2 \le \mathbb{E}||(\boldsymbol{d}(t))^+ + \boldsymbol{\alpha} - \boldsymbol{y}(t)||_2^2 \\ =& \mathbb{E}||(\boldsymbol{d}(t))^+||_2^2 + \mathbb{E}||\boldsymbol{\alpha}-\boldsymbol{y}(t)||_2^2 + 2\mathbb{E}\left[\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}(d_i^j(t))^+\cdot\alpha_i^j - \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}(d_i^j(t))^+\cdot y_i^j(t)\right],\end{aligned}$$

where $||\cdot||_2$ denotes the ℓ_2-norm in a Euclidean space. Since $(d_i^j(t))^+\ge 0$ for all $i\in\mathcal{N}$ and $j\in\mathcal{M}$, inequality (8) implies that

$$\mathbb{E}\left[\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}(d_i^j(t))^+\cdot\alpha_i^j - \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}(d_i^j(t))^+\cdot y_i^j(t)\right]\le 0.$$

Furthermore,

$$\mathbb{E}||\boldsymbol{\alpha}-\boldsymbol{y}(t)||_2^2 \le n\cdot m\cdot \max_{i\in\mathcal{N},j\in\mathcal{M}}\left\{\mathbb{E}[(y_i^j(t))^2]+(\alpha_i^j)^2\right\}\le n\cdot m\cdot C, \text{ where } C := \max_{i\in\mathcal{N},j\in\mathcal{M}}(\alpha_i^j)^2+1\le 2.$$

Therefore,

$$\mathbb{E}||(\boldsymbol{d}(t+1))^+||_2^2 \le ||(\boldsymbol{d}(1))^+||_2^2 + tnmC \text{ for all } t\ge 1. \tag{36}$$

By Jensen's inequality and that $||\cdot||_2^2$ is convex,

$$||\mathbb{E}[(\boldsymbol{d}(t+1))^+]||_2^2 \le \mathbb{E}||(\boldsymbol{d}(t+1))^+||_2^2 \le tnmC \text{ for all } t\ge 1. \tag{37}$$

Therefore,

$$0\le \frac{1}{t}||\mathbb{E}[(\boldsymbol{d}(t+1))^+]||_2 \le \sqrt{\frac{nmC}{t}}, \text{ which implies that } \limsup_{t\to+\infty}\frac{1}{t}||\mathbb{E}[(\boldsymbol{d}(t+1))^+]||_2 = 0.$$

Hence

$$\limsup_{t\to+\infty}\frac{1}{t}\cdot(d_i^j(t+1))^+ = 0, \text{ almost surely for each } i\in\mathcal{N} \text{ and } j\in\mathcal{M}.$$

Inequality (35) then follows immediately.

- *Step 2.* For problem $\mathcal{OP}_*(\gamma)$, if $\boldsymbol{\alpha}$ is single-period feasible, it holds that

$$\limsup_{\gamma\uparrow+\infty}\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_i^j(t)\le\alpha_i^j \text{ almost surely for all } i\in\mathcal{N} \text{ and } j\in\mathcal{M}. \tag{38}$$

Assume that, to the contrary, there exists (i_0,j_0) such that

$$\limsup_{\gamma\uparrow+\infty}\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_{i_0}^{j_0}(t) > \alpha_{i_0}^{j_0}.$$

Hence, there exists some $\Delta>0$, such that

$$\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_{i_0}^{j_0}(t) > \alpha_{i_0}^{j_0}+\Delta \text{ for infinitely many } \gamma. \tag{39}$$

Denote the set of γ's that satisfy (39) as Γ. Note that $\frac{1}{T(\gamma)}(\sum_{t=1}^{s} y_{i_0}^{j_0}(t))$ increases by at most $1/(T(\gamma))$ as s increases by 1. Hence, for all $\gamma \in \Gamma$ and $\gamma > 3/(T\Delta)$, $\frac{1}{T(\gamma)}(\sum_{t=1}^{s} y_{i_0}^{j_0}(t))$ increases by no more than $\Delta/3$ if s increases by 1. Therefore, for all $\gamma \in \Gamma$ and $\gamma > 3/(T\Delta)$, there exists a $s(\gamma) < T(\gamma)$, such that

$$\alpha_{i_0}^{j_0} + \frac{\Delta}{3} < \frac{1}{T(\gamma)} \sum_{t=1}^{s(\gamma)} y_{i_0}^{j_0}(t) < \alpha_{i_0}^{j_0} + \frac{2\Delta}{3} \tag{40}$$

By (40), we have that, for infinitely many γ,

$$\sum_{t=1}^{s(\gamma)} y_{i_0}^{j_0}(t) > T(\gamma)\left(\alpha_{i_0}^{j_0} + \frac{\Delta}{3}\right).$$

Hence, for infinitely many γ,

$$(d_{i_0}^{j_0}(t))^+ = \left((t-1)\alpha_{i_0}^{j_0} - \sum_{\tau=1}^{t-1} y_{i_0}^{j_0}(\tau)\right)^+ = 0 \text{ for all } t \geq s(\gamma)+1,$$

where the equality follows from

$$\sum_{\tau=1}^{t-1} y_{i_0}^{j_0}(\tau) \geq \sum_{\tau=1}^{s(\gamma)} y_{i_0}^{j_0}(\tau) > T(\gamma)\alpha_{i_0}^{j_0} > (t-1)\alpha_{i_0}^{j_0}.$$

Therefore, ad i_0 will not be offered to customer type j_0 for all $t \geq s(\gamma)+1$. Hence, $y_{i_0}^{j_0}(t) = 0$ for all $t \geq s(\gamma)+1$ and $t \leq T(\gamma)$. By (40), we have

$$\frac{\sum_{t=1}^{T(\gamma)} y_{i_0}^{j_0}(t)}{T(\gamma)} = \frac{\sum_{t=1}^{s(\gamma)} y_{i_0}^{j_0}(t)}{T(\gamma)} < \alpha_{i_0}^{j_0} + \frac{2\Delta}{3} \text{ for } \gamma \in \Gamma \text{ and } \gamma > \frac{3}{T\Delta},$$

which contradicts inequality (39). Therefore, for the system of $\mathcal{OP}_*(\gamma)$, we have inequality (38) holds.

- *Step 3.* For problem $\mathcal{OP}(\gamma)$, if $\boldsymbol{\alpha}$ is first-stage feasible, then (13) holds.

Inequalities (35) and (38) together imply that (13) holds for problem $\mathcal{OP}_*(\gamma)$. In particular, for problem $\mathcal{OP}_*(\gamma)$, we have no stock-out occurs for any ad asymptotically, i.e.,

$$\lim_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) = \sum_{j\in\mathcal{M}} \alpha_i^j \leq \frac{B_i(\gamma)}{b_i T(\gamma)} = \frac{B_i}{b_i T} \text{ for all } i \in \mathcal{N}. \tag{41}$$

Furthermore, by construction, the click-through process of $\mathcal{OP}_*(\gamma)$ is *identical* to that of $\mathcal{OP}(\gamma)$ before stock-out occurs in $\mathcal{OP}(\gamma)$.

We now show that the stock-out probability of any ad's budget converges to 0 for $\mathcal{OP}(\gamma)$ as $\gamma \uparrow +\infty$. If stock-out occurs in $\mathcal{OP}(\gamma)$, there exists some $i \in \mathcal{N}$, such that $b_i \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) > B_i(\gamma)$ for $\mathcal{OP}_*(\gamma)$. We have $\frac{1}{T(\gamma)} \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) > \frac{B_i(\gamma)}{b_i T(\gamma)} = \frac{B_i}{b_i T} \geq \sum_{j\in\mathcal{M}} \alpha_i^j$ by the feasibility of $\boldsymbol{\alpha}$. Hence, for $\mathcal{OP}_*(\gamma)$,

$$\mathbb{P}\left\{\limsup_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) > \frac{B_i}{b_i T}\right\} \leq \mathbb{P}\left\{\limsup_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) > \sum_{j\in\mathcal{M}} \alpha_i^j\right\} = 0, \tag{42}$$

where the equality follows from (41). Because $\mathcal{OP}_*(\gamma)$ and $\mathcal{OP}(\gamma)$ are equivalent before stock-out occurs, (42) implies that stock-out occurs with probability 0 as $\gamma \uparrow +\infty$ for $\mathcal{OP}(\gamma)$. Therefore, $\mathcal{OP}_*(\gamma)$ and $\mathcal{OP}(\gamma)$ are equivalent with probability 1 as $\gamma \uparrow +\infty$. Since (13) holds for problem $\mathcal{OP}_*(\gamma)$, a standard coupling argument implies that (13) holds for $\mathcal{OP}(\gamma)$ as well. This completes the proof of Theorem 2. □

Before the proof of Theorem 3, we prove Lemma 2 and Proposition 2 first.

Proof of Lemma 2

We first check the feasibility of $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})$ by directly plugging $\hat{\alpha}_i^j(\boldsymbol{z})$ into the constraints of $(\mathcal{OTP})$. Since $\boldsymbol{z}$ is feasible to $(\mathcal{OP}_{\mathsf{Fluid}})$, we have, for all $i\in\mathcal{N}$,

$$b_i\sum_{j\in\mathcal{M}}\hat{\alpha}_i^j(\boldsymbol{z})=b_i\sum_{j\in\mathcal{M},S\in\mathfrak{S}^j}p^j\phi_i^j(S)z^j(S)\le\frac{B_i}{T};$$

and, for all $i\in\mathcal{N}$ and $\mathcal{C}\in\mathfrak{K}_i$,

$$\sum_{j\in\mathcal{C}}\hat{\alpha}_i^j(\boldsymbol{z})=\sum_{j\in\mathcal{C},S\in\mathfrak{S}^j}p^j\phi_i^j(S)z^j(S)\ge\frac{\eta_i^{\mathcal{C}}}{T}.$$

In addition,

$$\mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}))=\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}r_i^j\sum_{S\in\mathfrak{S}^j}\hat{\alpha}_i^j(\boldsymbol{z})+\lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}))=\sum_{i\in\mathcal{N},j\in\mathcal{M},S\in\mathfrak{S}^j}r_i^jp^j\phi_i^j(S)z^j(S)+\lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}))=\mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}).$$

Therefore, it remains to show that

$$\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}))]=\hat{\alpha}_i^j(\boldsymbol{z}).\tag{43}$$

Applying the law of total probability, we directly evaluate that

$$\begin{aligned}\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}))]=&\mathbb{P}[j(t)=j]\sum_{S\in\mathfrak{S}^j}\mathbb{P}[\pi_{\mathsf{Fluid}}(j)=S|j(t)=j]\mathbb{E}[y_i^j(t)|\pi_{\mathsf{Fluid}}(j)=S,j(t)=j]\\=&\mathbb{P}[j(t)=j]\sum_{S\in\mathfrak{S}^j}\mathbb{P}[\pi_{\mathsf{Fluid}}(j)=S]\mathbb{E}[y_i^j(t)|\pi_{\mathsf{Fluid}}(j)=S,j(t)=j]\\=&p^j\sum_{S\in\mathfrak{S}^j}z^j(S)\phi_i^j(S)\\=&\hat{\alpha}_i^j(\boldsymbol{z}),\end{aligned}$$

i.e., (43) holds. Therefore, the click-through target vector $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})$ is first-stage feasible. In particular, $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is first-stage feasible with $\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))]=\hat{\alpha}_i^j(\boldsymbol{z}^*)$. We defer the proof of $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$'s optimality for $(\mathcal{OTP})$ to the proof of Theorem 3. □

Before proving Theorem 3, we present the proof of Proposition 2 first.

Proof of Proposition 2

We prove (17) by showing each individual equality or inequality thereof.

- *Step 1.* The FV generated by $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ in $\mathcal{OP}(\gamma)$ is asymptotically identical to the optimal FV in $(\mathcal{OP}_{\mathsf{Fluid}})$, i.e.,

$$\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma)=\mathcal{V}^*_{\mathsf{Fluid}}.\tag{44}$$

We first show that $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ is asymptotically *feasible* for $\mathcal{OP}(\gamma)$. As $\gamma\uparrow+\infty$, we have

$$\lim_{\gamma\uparrow+\infty}\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{j\in\mathcal{M}}b_iy_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))=\mathbb{E}\left[\sum_{j\in\mathcal{M}}b_iy_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))\right]=\sum_{j\in\mathcal{M},S\in\mathfrak{S}^j}b_ip^j\phi_i^j(S)z^{j*}(S)\le\frac{B_i(\gamma)}{T(\gamma)},\tag{45}$$

where the first equality follows from the strong law of large numbers, the second from the definition of $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$, and the inequality follows from $\boldsymbol{z}^*$ is feasible for $(\mathcal{OP}_{\mathsf{Fluid}})$. Similarly, we have, under the policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$,

$$\lim_{\gamma\uparrow+\infty}\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{j\in\mathcal{C}}y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))\right]=\sum_{j\in\mathcal{C},S\in\mathfrak{S}^j}p^j\phi_i^j(S)z^j(S)\ge\frac{\eta_i^{\mathcal{C}}(\gamma)}{T(\gamma)},\tag{46}$$

where the equality follows from the definition of $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$, and the inequality follows from $\boldsymbol{z}^*$ is feasible for $(\mathcal{OP}_{\mathsf{Fluid}})$. Inequalities (45) and (46) together imply that $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ is asymptotically feasible for $\mathcal{OP}(\gamma)$ as $\gamma \uparrow +\infty$.

Then, we evaluate the FV of the Fluid-$\boldsymbol{z}^*$ policy as follows:

$$\begin{aligned}\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma) &= \lim_{\gamma\uparrow+\infty}\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)) + \lambda F(\bar{\boldsymbol{y}}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)))\right] \\ &= \sum_{i\in\mathcal{N},j\in\mathcal{M},S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^{j*}(S) + \lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)) \\ &= \mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}^*) \\ &= \mathcal{V}^*_{\mathsf{Fluid}},\end{aligned} \tag{47}$$

where the second equality follows from the law of large numbers and the dominated convergence theorem. Therefore, (44) follows from (47).

- *Step 2.* The following inequality holds:

$$\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma) \geq \lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma). \tag{48}$$

Since $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ is asymptotically feasible, the optimality of $\mathcal{V}^*(\gamma)$ implies that $\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma) \geq \lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma)$.

- *Step 3.* The optimal FV of $(\mathcal{OP}_{\mathsf{Fluid}})$ dominates that of $\mathcal{OP}(\gamma)$, i.e.,

$$\mathcal{V}^*_{\mathsf{Fluid}} \geq \mathcal{V}^*(\gamma) \text{ for any } \gamma > 0. \tag{49}$$

Consider an arbitrary policy $\pi\in\Pi$ feasible for $\mathcal{OP}(\gamma)$. We first define the following probability measure induced by π, $\boldsymbol{z}_{\mathsf{Fluid}}(\pi)$ for $(\mathcal{OP}_{\mathsf{Fluid}})$: For $j\in\mathcal{M}$ and $S\in\mathfrak{S}^j$,

$$z^j_{\mathsf{Fluid}}(S|\pi) := \mathbb{P}\left[\pi_{\tilde{t}}(j(\tilde{t}),\mathcal{H}_{\tilde{t}-1}) = S | j(\tilde{t}) = j\right], \tag{50}$$

where $\tilde{t}$ is a random variable uniformly distributed on $\{1,2,...,T(\gamma)\}$ and independent of everything else. Because π is feasible for $\mathcal{OP}(\gamma)$, all the constraints of $\mathcal{OP}(\gamma)$ will also be satisfied in the expected sense as well, i.e.,

$$\frac{1}{T(\gamma)}\mathbb{E}\left[\sum_{t=1}^{T(\gamma)}\sum_{j\in\mathcal{M}} b_i y_i^j(t|\pi)\right] \leq \frac{B_i(\gamma)}{T(\gamma)}, \text{ for each } i\in\mathcal{N},$$

and

$$\frac{1}{T(\gamma)}\mathbb{E}\left[\sum_{t=1}^{T}\sum_{j\in\mathcal{C}} y_i^j(t|\pi)\right] \geq \frac{\eta_i^{\mathcal{C}}(\gamma)}{T(\gamma)}, \text{ for each } i\in\mathcal{N} \text{ and } \mathcal{C}\in\mathfrak{K}_i,$$

where the expectations are taken with respect to $j(t)$, π, and $\boldsymbol{y}$. By (50), straightforward algebraic manipulation and the law of iterated expectations together yield that

$$\mathbb{E}[\bar{y}_i^j] = \mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_i^j(t|\pi)\right] = \sum_{S\in\mathfrak{S}^j} p^j\phi_i^j(S) z^j_{\mathsf{Fluid}}(S|\pi) = \hat{\alpha}_i^j(\boldsymbol{z}_{\mathsf{Fluid}}(S|\pi)).$$

Plugging this identity into the constraints of $(\mathcal{OP}_{\mathsf{Fluid}})$, we have

$$\sum_{j\in\mathcal{M},S\in\mathfrak{S}^j} b_i p^j \phi_i^j(S) z^j_{\mathsf{Fluid}}(S|\pi) \leq \frac{B_i(\gamma)}{T(\gamma)} \text{ for each } i\in\mathcal{N}$$

and that

$$\sum_{j\in\mathcal{C},S\in\mathfrak{S}^j} p^j\phi_i^j(S)z_{\mathsf{Fluid}}^j(S|\pi)\geq\frac{\eta_i^{\mathcal{C}}(\gamma)}{T(\gamma)}\text{ for each } i\in\mathcal{N}\text{ and }\mathcal{C}\in\mathfrak{K}_i.$$

Therefore, $\boldsymbol{z}_{\mathsf{Fluid}}(\pi)$ is feasible for $(\mathcal{OP}_{\mathsf{Fluid}})$ and, hence,

$$\mathcal{V}^*_{\mathsf{Fluid}}\geq\mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}_{\mathsf{Fluid}}(\pi)).\tag{51}$$

By Jensen's inequality and the concavity of the fairness metric $F(\cdot)$,

$$\begin{aligned}\mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}_{\mathsf{Fluid}}(\pi)) &= \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}\sum_{S\in\mathfrak{S}^j} r_i^jp^j\phi_i^j(S)z_{\mathsf{Fluid}}^j(S|\pi)+\lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}_{\mathsf{Fluid}}(S|\pi)))\\ &\geq\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^jy_i^j(t|\pi)+\lambda F(\bar{\boldsymbol{y}}(\pi))\right]\\ &=\mathcal{V}(\pi|\gamma)\end{aligned}\tag{52}$$

Since π is arbitrary, inequalities (51) and (52) together imply that (49) holds.

Therefore, putting the (in)equalities (44), (48), and (49) together, we have (17) holds, which completes the proof of Proposition 2. □

We are now ready to prove Theorem 3.

Proof of Theorem 3

We prove (14) by the following three steps.

- *Step 1.* The following inequality holds:

$$\mathcal{V}^*_{\mathsf{CT}}\geq\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma).\tag{53}$$

By Proposition 2, there exists one optimal static policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ of $(\mathcal{OP})$ in the asymptotic regime. By Lemma 2, we have

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)=(\hat{\alpha}_i^j(\boldsymbol{z}^*),i\in\mathcal{N},j\in\mathcal{M})\in\mathbb{R}_+^{(nm)}.$$

Therefore, there exists a static policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$, such that

$$\begin{aligned}&\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))]\geq\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*),\text{ for each } i\in\mathcal{N},j\in\mathcal{M},\\ &b_i\sum_{j\in\mathcal{M}}\hat{\alpha}_i^j(\boldsymbol{z}^*)\leq\frac{B_i}{T},\text{ for each } i\in\mathcal{N},\\ &\sum_{j\in\mathcal{C}}\hat{\alpha}_i^j(\boldsymbol{z}^*)\geq\frac{\eta_i^{\mathcal{C}}}{T},\text{ for each } i\in\mathcal{N}\text{ and }\mathcal{C}\in\mathfrak{K}_i,\end{aligned}$$

hence $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is a feasible solution of the first-stage click-through target optimization problem $(\mathcal{OTP})$. By the law of large numbers and the dominated convergence theorem, we obtain $\mathcal{V}^*_{\mathsf{CT}}\geq\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma)$ as follows:

$$\begin{aligned}\mathcal{V}^*_{\mathsf{CT}}&\geq\mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*))\\ &=\sum_{i\in\mathcal{N},j\in\mathcal{M}} r_i^j\hat{\alpha}_i^j(\boldsymbol{z}^*)+\lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*))\\ &=\lim_{\gamma\uparrow+\infty}\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^jy_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))\right]+\lambda F\left(\mathbb{E}\left[\bar{\boldsymbol{y}}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))\right]\right)\\ &=\lim_{\gamma\uparrow+\infty}\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^jy_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))+\lambda F(\bar{\boldsymbol{y}}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)))\right]\\ &=\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma).\end{aligned}\tag{54}$$

Since $\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma)=\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma)$ by Proposition 2, (53) follows.

- *Step 2.* The following inequality holds:

$$\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma)\geq\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma). \tag{55}$$

The optimality of $\mathcal{V}^*(\gamma)$ implies that $\mathcal{V}^*(\gamma)\geq\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)$ for all $\gamma>0$.

- *Step 3.* The DWO-$\boldsymbol{\alpha}^*$ policy generates the same asymptotic FV in $(\mathcal{OP})$ as the optimal FV in $(\mathcal{OTP})$, i.e.,

$$\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)=\mathcal{V}^*_{\mathsf{CT}} \tag{56}$$

By Theorem 2, the DWO-$\boldsymbol{\alpha}^*$ policy is asymptotically feasible for the original problem $(\mathcal{OP})$ with (13) holding true. We now evaluate $\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)$ in the asymptotic regime:

$$\begin{aligned}\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)&=\lim_{\gamma\uparrow+\infty}\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}r_i^jy_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))+\lambda F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))\right]\\&=\sum_{i\in\mathcal{N},j\in\mathcal{M}}r_i^j\alpha_i^{j*}+\lambda F(\boldsymbol{\alpha}^*)\\&=\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}^*)\\&=\mathcal{V}^*_{\mathsf{CT}},\end{aligned} \tag{57}$$

where the second equality follows from equality (13) and the dominated convergence theorem. Hence, equality (56) follows from equality (57). Therefore, putting the (in)equalities (53), (55), and (56) together, we have (14) holds. As a by-product of the proof, (14) also implies

$$\mathcal{V}^*_{\mathsf{CT}}=\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma)=\mathcal{V}^*_{\mathsf{Fluid}}=\mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)).$$

By (49) in the proof of Proposition 2, we have

$$\mathcal{V}^*_{\mathsf{CT}}=\mathcal{V}^*_{\mathsf{Fluid}}\geq\mathcal{V}^*(\gamma)\text{ for any }\gamma>0.$$

Hence, we have (15) holds. Moreover, $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is optimal for $(\mathcal{OTP})$, which completes the proof of Lemma 2.

To complete the proof, we now show (16). Because (15) holds for any $\gamma>0$, we have

$$\begin{aligned}&\mathcal{V}^*(\gamma)-\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)\leq\mathcal{V}^*_{\mathsf{CT}}-\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)\\&=\left[\sum_{i\in\mathcal{N},j\in\mathcal{M}}r_i^j\alpha_i^{j*}+\lambda F(\boldsymbol{\alpha}^*)\right]-\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}r_i^jy_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))+\lambda F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))\right]\\&=\sum_{i\in\mathcal{N},j\in\mathcal{M}}r_i^j\mathbb{E}\left[\alpha_i^{j*}-\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))\right]+\lambda\left(F(\boldsymbol{\alpha}^*)-\mathbb{E}[F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))]\right)\end{aligned} \tag{58}$$

By the definition of the debt vector $\boldsymbol{d}(t)$, we can bound the first term of (58) as follows

$$\begin{aligned}\sum_{i\in\mathcal{N},j\in\mathcal{M}}r_i^j\mathbb{E}\left[\alpha_i^{j*}-\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))\right]&=\frac{1}{T(\gamma)}\sum_{i\in\mathcal{N},j\in\mathcal{M}}r_i^j\mathbb{E}\left[d_i^j(T(\gamma)+1)\right]\\&\leq\frac{1}{T(\gamma)}\sum_{i\in\mathcal{N},j\in\mathcal{M}}r_i^j\mathbb{E}\left[(d_i^j(T(\gamma)+1))^+\right]\\&\leq\frac{1}{T(\gamma)}||\boldsymbol{r}||_2\cdot||\mathbb{E}\left[(\boldsymbol{d}(T(\gamma)+1))^+\right]||_2\\&\leq\frac{\mathcal{C}_1}{\sqrt{\gamma}}\end{aligned} \tag{59}$$

where the constant $\mathcal{C}_1 := ||\boldsymbol{r}||_2 \cdot \sqrt{\frac{2mn}{T}}$, the second inequality follows from the Cauchy–Schwarz inequality, and the last from (37).

Since $F(\cdot)$ is a concave function, it has subgradient $\boldsymbol{f}(\cdot) := (f_i^j(\cdot) : i \in \mathcal{N}, j \in \mathcal{M}) \in \mathbb{R}^{nm}$. Define $f_{\max} := \max_{i\in\mathcal{N}, j\in\mathcal{M}, \boldsymbol{y}\in[0,1]^{nm}} |f_i^j(\boldsymbol{y})|$ and $F_{\max} := \max_{\boldsymbol{\alpha}\in\in[0,1]^{nm}} |F(\boldsymbol{\alpha})|$. We have

$$F(\boldsymbol{\alpha}^*) \le F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))) + \boldsymbol{f}(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))^\intercal(\boldsymbol{\alpha}^* - \bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))).$$

Hence, we can bound the second term of (58) as follows, for $\gamma \ge 1$,

$$\begin{aligned} F(\boldsymbol{\alpha}^*) - \mathbb{E}[F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))] &\le \mathbb{E}[\boldsymbol{f}(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))^\intercal(\boldsymbol{\alpha}^* - \bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))] \\ &\le \frac{1}{T(\gamma)}\sqrt{\mathbb{E}||\boldsymbol{f}(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))||_2^2} \cdot \sqrt{\mathbb{E}||\boldsymbol{d}(T(\gamma)+1)||_2^2} \\ &\le \frac{1}{T(\gamma)}\sqrt{\mathbb{E}||\boldsymbol{f}(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))||_2^2} \cdot \sqrt{\mathbb{E}||(\boldsymbol{d}(T(\gamma)+1))^+||_2^2 + mn} \\ &\le \frac{\sqrt{nm}\cdot f_{\max}}{T(\gamma)} \cdot \sqrt{2nmT(\gamma)+mn} \\ &\le \frac{\mathcal{C}_2}{\sqrt{\gamma}}, \end{aligned} \tag{60}$$

where the constant $\mathcal{C}_2 := \max\{\frac{\sqrt{nm}\cdot f_{\max}}{T} \cdot \sqrt{2nmT+mn}, F_{\max}\}$, the second inequality follows from the Cauchy–Schwarz inequality, the third from the fact that

$$d_i^j(T(\gamma)+1))^- = \left(T(\gamma)\alpha_i^{j*} - \sum_{t=1}^{T(\gamma)} y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))\right)^- \le 1, \tag{61}$$

and the fourth from (36). Inequality (61) holds because, under the DWO-$\boldsymbol{\alpha}^*$ policy, once $\sum_{t=1}^{s} y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)) \ge T(\gamma)\alpha_i^{j*}$ for some $s \le T(\gamma)$, ad i will not be offered to type j customer for all $t \ge s+1$. For $\gamma < 1$, it holds that

$$F(\boldsymbol{\alpha}^*) - \mathbb{E}[F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))] \le \frac{F_{\max}}{\sqrt{\gamma}} \le \frac{\mathcal{C}_2}{\sqrt{\gamma}}. \tag{62}$$

Combining (59), (60), and (62) yields that

$$\mathcal{V}^*(\gamma) - \mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma) \le \frac{\mathcal{C}_1 + \lambda\cdot\mathcal{C}_2}{\sqrt{\gamma}}.$$

Hence, we have (16) holds with $\mathcal{C} := \mathcal{C}_1 + \lambda \cdot \mathcal{C}_2 > 0$. This completes the proof of Theorem 3. □

Proof of Proposition 3

The proof follows from the same argument as *Step 3* in the proof of Theorem 3 by replacing $\boldsymbol{\alpha}^*$ with any feasible $\boldsymbol{\alpha}$. To avoid repetition, we omit the proof details. □

E. Feasible Click-Through Targets Under the MNL Choice Model

Proposition 1 characterizes the feasible region of the click-through targets $\mathcal{A}_{MNL}$ if customers follow the MNL choice model. This section seeks to deliver additional insights on when the click-through targets are feasible. We observe that (34) is equivalent to

$$p^j \ge \sum_{i\in\mathcal{N}} \alpha_i^j + \max\left\{\frac{1}{K}\sum_{i\in\mathcal{N}}\frac{\alpha_i^j}{v_i^j}, \max_{i\in\mathcal{N}}\left\{\frac{\alpha_i^j}{v_i^j}\right\}\right\}, \text{ for each } j \in \mathcal{M}.$$

Here, p^j is the expected (per-user) traffic of type-j customers in each period. Clearly, $\sum_{i\in\mathcal{N}} \alpha_i^j$ is the total required traffic for type-j customers if a customer will click one of the ad in the offer set with probability 1. In practice, however, a customer may end up not choosing any ad from the offer set, so we need some buffer traffic for type-j customers that accounts for the non-click circumstance.

More specifically, let $\mathfrak{S}_i$ denote the collection of all offer sets containing ad i. Since the offer-set policy may be random, we define $\mu_j(S)$ as the probability of displaying offer-set $S\subseteq\mathcal{N}$ to type-j customers. Thus, the desired click-through goal for ad i and type-j customer is

$$\sum_{S\in\mathfrak{S}_i} \mu_j(S)\cdot \frac{v_i^j}{1+\sum_{i'\in S} v_{i'}^j} \geq \alpha_i^j.$$

Thus, the non-click probability of the ads for a type-j customer when ad i ($i\in\mathcal{N}$) is offered satisfies that

$$\alpha_i^j(o) := \sum_{S\in\mathfrak{S}_i} \mu_j(S)\cdot \frac{1}{1+\sum_{i'\in S} v_{i'}^j} \geq \frac{\alpha_i^j}{v_i^j}.$$

Therefore, to ensure the click-through goal of type-j customers and ad i, the traffic of customer type j must satisfy $p^j \geq \sum_{i'\in\mathcal{N}} \alpha_{i'}^j + \alpha_i^j(o) \geq \sum_{i'\in\mathcal{N}} \alpha_{i'}^j + \frac{\alpha_i^j}{v_i^j}$ for all $i\in\mathcal{N}$.

The cardinality constraint for the offer set size would impose an additional bound on the non-click probability of type-j customers. Specifically, let $\mathfrak{S} := \bigcup_{i=1}^n \mathfrak{S}_i$ be the set of all offer sets displayed to a customer. Because $|S|\leq K$ for any $S\in\mathfrak{S}$, $|\{i\in\mathcal{N} : S\in\mathfrak{S}_i\}|\leq K$ for all S. We have, given customer type j,

$$K\sum_{S\in\mathfrak{S}} \mu_j(S)\cdot \frac{1}{1+\sum_{i\in S} v_i^j} \geq \sum_{i\in\mathcal{N}}\sum_{S\in\mathfrak{S}_i} \mu_j(S)\cdot \frac{1}{1+\sum_{i\in S} v_i^j} \geq \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j}.$$

Thus, the non-click probability of all ads for type-j customer satisfies that

$$\alpha_o^j := \sum_{S\in\mathfrak{S}} \mu_j(S)\cdot \frac{1}{1+\sum_{i\in S} v_i^j} \geq \frac{1}{K}\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j}.$$

Therefore, given the cardinality constraint of an offer set, to ensure the click-through targets of type-j customers with respect to all ads, the traffic of customer type j must satisfy $p^j \geq \sum_{i\in\mathcal{N}} \alpha_i^j + \alpha_o^j \geq \sum_{i\in\mathcal{N}} \alpha_i^j + \frac{1}{K}\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j}$. In summary, the characterization for the feasibility of $\boldsymbol{\alpha}$ demonstrates that, to meet the click-through targets, we should also account for the *non-click* cases.

F. Optimal Target Convex Program Formulation for Specific Choice Models

In this section, we introduce the characterization of first-stage feasible region $\mathcal{A} := \{\boldsymbol{\alpha}\in[0,1]^{nm} : h(\boldsymbol{\alpha})\geq 0\}$ for independent and generalized attraction choice models. Similar to the characterization of $\mathcal{A}$ if the customers follow the MNL model, we use a binary variable representation of a deterministic static offer-set policy $\pi\in\Pi_{\mathsf{d}}$ (see, also, the proof of Proposition 1). More specifically, $\pi\in\Pi_{\mathsf{d}}$ can be equivalently represented by an $nm-$dimensional binary vector $\boldsymbol{x} = (x_i^j\in\{0,1\} : i\in\mathcal{N}, j\in\mathcal{M})$, where $x_i^j = 1$ means that $i\in\pi(j)$, i.e., ad i is included in the offer-set displayed to a type-j customer. We denote $\phi_i^j(\boldsymbol{x})$ as the expected click-throughs of a type-j customer for ad i if the offer-set displayed to this customer is $S^j = \{i\in\mathcal{N} : x_i^j = 1\}$. Denote the set of all plausible offer-set representation vectors as $\mathcal{X}\subset\{0,1\}^{nm}$, and the set of plausible offer-set representation vectors displayed to a type-j customer as $\mathcal{X}^j\subset\{0,1\}^n$.

F.1. Independent Choice Model

If customers follow the independent choice model, the click-throughs only depend on the customer type j and ad i, but not on the offer-set S^j displayed to the customer. This is actually the most commonly adopted choice models in practical advertising (see, e.g., Feldman et al. 2022), where the click-through rate (CTR) prediction algorithm of the platform outputs:

$$\phi_i^j(\boldsymbol{x}) = c_i^j x_i^j, \tag{63}$$

where $c_i^j > 0$ is the CTR of ad i with respect to a type-j customer. It is evident from (63) that the number of click-throughs is independent of the ads in S^j other than ad i. The following proposition is a counterpart of Proposition 1 with the independent choice model.

Proposition 4. *If customers follow the independent click-through model* (63)*, the first-stage feasible region $\mathcal{A}_{IND}$ is given by the following linear constraints:*

$$\mathcal{A}_{IND} = \left\{\boldsymbol{\alpha} \in [0,1]^{nm} : p^j c_i^j \geq \alpha_i^j, \textit{ for each } i \in \mathcal{N}, j \in \mathcal{M}\right\}. \tag{64}$$

It is expected that, under the independent choice model, we have an independent feasibility condition for each ad-customer type pair. With Proposition 4, we we can simplify the optimal target problem $(\mathcal{OTP})$ under the independent choice model as the following convex program with linear constraints only:

$$\begin{aligned}
\max_{\boldsymbol{\alpha} \geq \mathbf{0}} \ & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\
\text{s.t. } & p^j c_i^j \geq \alpha_i^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M} \\
& b_i \sum_{j \in \mathcal{M}} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i \in \mathcal{N}, \\
& \sum_{j \in \mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \text{ for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i.
\end{aligned} \tag{$\mathcal{OTP}-\mathcal{IND}$}$$

F.2. Generalized Attraction Model

The generalized attraction model (GAM) is a generalization of MNL accounts for the possibility that a customer may look for a product outside the offer-set (see, e.g., Gallego et al. 2015). Under the GAM, the expected number of click-throughs of ad i by a type-j customer is given by

$$\phi_i^j(\boldsymbol{x}) = \frac{v_i^j x_i^j}{1 + \sum_{i' \in \mathcal{N}} \omega_{i'}^j (1 - x_{i'}^j) + \sum_{i \in \mathcal{N}} v_{i'}^j x_{i'}^j}, \tag{65}$$

where $v_i^j > 0$ is the attraction value of ad i to a type-j customer, and $w_i^j \in [0, v_i^j]$ is the shadow attraction value of ad i to a type-j customer, capturing the customer's looking for a product outside the offer-set. Hence, by defining $\tilde{v}^j := 1 + \sum_{i \in \mathcal{N}} \omega_i^j > 0$ and $\tilde{v}_i^j := v_i^j - \omega_i^j \geq 0$, we have, under the GAM,

$$\phi_i^j(\boldsymbol{x}) = \frac{v_i^j x_i^j}{\tilde{v}^j + \sum_{i' \in \mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j}. \tag{66}$$

As Corollary 1, we first rewrite the necessary and sufficient condition for $\boldsymbol{\alpha}$ under the GAM.

COROLLARY 2. *If customers follow the GAM click-through model (66), a click-through target vector $\boldsymbol{\alpha}$ is single-period feasible if and only if, for each $j \in \mathcal{M}$*

$$\max_{\boldsymbol{x}^j \in \mathcal{X}^j} \sum_{i \in \mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{\tilde{v}^j + \sum_{i' \in \mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j} \geq \sum_{i \in \mathcal{N}} \alpha_i^j \theta_i^j \text{ for any } \theta_i^j \geq 0 \ (i \in \mathcal{N}). \tag{67}$$

Similar to Proposition 1, we can characterize the first-stage feasible region of the click-through target vector under the GAM, $\mathcal{A}_{GAM} := \{\boldsymbol{\alpha} : (67) \text{ holds for each } j \in \mathcal{M}\}$, using linear constraints. The following proposition characterizes $\mathcal{A}_{GAM}$ and accounts for the offer-set cardinality constraint.

PROPOSITION 5. *If customers follow the GAM (66) and the set of all feasible offer-sets is $\mathfrak{S}^j = \{S \subset \mathcal{N} : |S| \leq K\}$ for each $j \in \mathcal{M}$, the first-stage feasible region $\mathcal{A}_{GAM}$ is given by the following linear constraints:*

$$\mathcal{A}_{GAM} := \left\{ \boldsymbol{\alpha} \in \mathbb{R}_+^{nm} : \sum_{i' \in \mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{\tilde{v}^j \alpha_i^j}{v_i^j} \leq p^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}, \text{ and } \sum_{i \in \mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j}{K} \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{v_i^j} \leq p^j, \text{ for each } j \in \mathcal{M} \right\}. \tag{68}$$

With Proposition 5, we can simplify the optimal target problem $(\mathcal{OTP})$ under the generalized attraction model as the following convex program with linear constraints only:

$$\begin{aligned}
\max_{\boldsymbol{\alpha} \geq \mathbf{0}} \ & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\
\text{s.t.} \ & \sum_{i' \in \mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{\tilde{v}^j \alpha_i^j}{v_i^j} \leq p^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}, \\
& \sum_{i \in \mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j}{K} \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{v_i^j} \leq p^j, \text{ for each } j \\
& b_i \sum_{j \in \mathcal{M}} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i \in \mathcal{N}, \\
& \sum_{j \in \mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \text{ for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i.
\end{aligned} \tag{$\mathcal{OTP}-\mathcal{GAM}$}$$

F.3. Proofs

In this subsection, we give proofs of the technical results presented in Appendix F.

Proof of Proposition 4

Directly applying Theorem 1 to the independent choice model implies that $\boldsymbol{\alpha}$ is feasible if and only if the following inequality holds.

$$\max_{\boldsymbol{x} \in \mathcal{X}} \sum_{i \in \mathcal{N}, j \in \mathcal{M}} p^j c_i^j \theta_i^j x_i^j \geq \sum_{i \in \mathcal{N}, j \in \mathcal{M}} \alpha_i^j \theta_i^j \text{ for any } \boldsymbol{\theta} \geq \mathbf{0}. \tag{69}$$

Setting $\theta_i^j = 1$ and all other θ's equal to zero in (69) immediately implies that if $\boldsymbol{\alpha}$ is single-period feasible, then (64) holds. Reversely, if (64) holds, then for any $\boldsymbol{\theta} \geq \mathbf{0}$, we have

$$\max_{\boldsymbol{x} \in \mathcal{X}} \sum_{i \in \mathcal{N}, j \in \mathcal{M}} p^j c_i^j \theta_i^j x_i^j = \sum_{i \in \mathcal{N}, j \in \mathcal{M}} p^j c_i^j \theta_i^j \geq \sum_{i \in \mathcal{N}, j \in \mathcal{M}} \alpha_i^j \theta_i^j,$$

i.e., (69) holds and, therefore, $\boldsymbol{\alpha}$ is single-period feasible. This concludes the proof of Proposition 4. □

Proof of Corollary 2

The proof follows from exactly the same argument as that of Corollary 1. We omit the details to avoid repetition. □

Proof of Proposition 5

As in the proof of Proposition 1, we first state and prove the following auxiliary lemma.

LEMMA 4. *If customers follow the GAM click-through model (66) and the set of all feasible offer-sets is $\mathfrak{S}^j = \{S \subset \mathcal{N} : |S| \le K\}$ for each $j \in \mathcal{M}$, we have $\boldsymbol{\alpha}$ is single-period feasible if and only if there exist $\boldsymbol{w} := (w_i^j : i \in \mathcal{N}, j \in \mathcal{M})$ and $\boldsymbol{z} := (z^j : j \in \mathcal{M})$ that satisfy the following linear constraints*

$$\begin{aligned} &p^j v_i^j w_i^j \ge \alpha_i^j,\ w_i^j \le z^j,\ w_i^j \ge 0,\ \text{for each } i \in \mathcal{N}, j \in \mathcal{M},\\ &\sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1,\ \sum_{i\in\mathcal{N}} w_i^j \le K z^j,\ \text{for each } j \in \mathcal{M}, \end{aligned} \tag{70}$$

where $z^j := \frac{1}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j}$ and $w_i^j := x_i^j z^j = \frac{x_i^j}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j}$.

Proof of Lemma 4 Similar to the proof of Lemma 3, it is straightforward to check that the left-hand side of (67) is quasi-convex in $\boldsymbol{x}^j$ for all j, so there always exists a maximizer on the boundary of the feasible region. Thus, we can relax the binary constraint $x_i^j \in \{0,1\}$ to $x_i^j \in [0,1]$ in (67), which is therefore equivalent to

$$\max_{\boldsymbol{x}^j \in [0,1]^n,\ \sum_{i\in\mathcal{N}} x_i^j \le K} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j} \ge \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j \ \text{for all } \boldsymbol{\theta}^j \ge \mathbf{0} \text{ and } j \in \mathcal{M}. \tag{71}$$

We change the decision variable and define, for all $j \in \mathcal{M}$,

$$z^j := \frac{1}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j} \text{ and } w_i^j := x_i^j z^j = \frac{x_i^j}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j}.$$

Then, we can rewrite (71) as, for any j,

$$\begin{aligned} &\min_{\boldsymbol{\theta}^j \ge \mathbf{0}} \left(\max_{w_i^j, z^j} \sum_{i\in\mathcal{N}} p^j v_i^j w_i^j \theta_i^j - \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j \right) \ge 0\\ &\text{s.t. } \sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1,\\ &\qquad \sum_{i\in\mathcal{N}} w_i^j \le K z^j,\\ &\qquad 0 \le w_i^j \le z^j \text{ for each } i \in \mathcal{N}. \end{aligned} \tag{72}$$

By Sion's minimax theorem, we can exchange the maximization and minimization operators so that (72) is equivalent to, for any $j \in \mathcal{M}$:

$$\begin{aligned} &\max_{\boldsymbol{w}^j, z^j} \min_{\boldsymbol{\theta}^j \ge 0} \sum_{i\in\mathcal{N}} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \ge 0,\\ &\qquad \text{s.t. } \sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1,\\ &\qquad \sum_{i\in\mathcal{N}} w_i^j \le K z^j,\\ &\qquad 0 \le w_i^j \le z^j,\ \text{for each } i \in \mathcal{N}. \end{aligned} \tag{73}$$

Therefore, (73) holds if and only if there exist $\boldsymbol{w}^j$ and z^j such that all the constraints in (73) hold and $\sum_{i\in\mathcal{N}} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \ge 0$ holds for all $\boldsymbol{\theta}^j \ge \mathbf{0}$, which is equivalent to $p^j v_i^j w_i^j - \alpha_i^j \ge 0$ for all $i \in \mathcal{N}$. Therefore, (73) is equivalent to that, for any $j \in \mathcal{M}$,

$$\begin{aligned} & p^j v_i^j w_i^j - \alpha_i^j \ge 0, \text{ for each } i \in \mathcal{N}, \\ & \sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1, \\ & \sum_{i\in\mathcal{N}} w_i^j \le K z^j, \\ & 0 \le w_i^j \le z^j, \text{ for each } i \in \mathcal{N}. \end{aligned} \tag{74}$$

That (74) holds for all $j \in \mathcal{M}$ is equivalent to that (70) holds. This completes the proof of Lemma 4. □

We now prove Proposition 5 itself. We first show that if (70) holds, then $\boldsymbol{\alpha} \in \mathcal{A}_{GAM}$. By the first inequality of (70), we have $w_i^j \ge \frac{\alpha_i^j}{p^j v_i^j}$ for all $i \in \mathcal{N}$ and $j \in \mathcal{M}$. Plugging this into the first equality of (70), we have

$$1 - \tilde{v}^j z^j = \sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j \ge \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j} \text{ for each } j \in \mathcal{M}.$$

Thus, by the first and second inequalities of (70), we have

$$\sum_{i'\in\mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{p^j v_{i'}^j} \le 1 - \tilde{v}^j z^j \le 1 - \tilde{v}^j w_i^j \le 1 - \frac{\tilde{v}^j \alpha_i^j}{p^j v_i^j} \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

Rearranging the terms, we have

$$p^j \ge \sum_{i'\in\mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{\tilde{v}^j \alpha_i^j}{v_i^j} \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

The first, second, and fourth inequalities and the first equality of (70) imply that

$$\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j v_i^j} \le \sum_{i\in\mathcal{N}} w_i^j \le K z^j = K \frac{1 - \sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j}{\tilde{v}^j} \le \frac{K}{\tilde{v}^j} \left(1 - \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right) \text{ for each } j \in \mathcal{M}.$$

Rearranging the terms, we have

$$p^j \ge \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j}{K} \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j} \text{ for each } j \in \mathcal{M}.$$

Therefore, if (70) holds, we have $\boldsymbol{\alpha} \in \mathcal{A}_{GAM}$.

Next, we show that if $\boldsymbol{\alpha} \in \mathcal{A}_{GAM}$, then (70) holds. Given $\boldsymbol{\alpha} \in \mathcal{A}_{GAM}$, define

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \text{ for each } i \in \mathcal{N} \text{ and } j \in \mathcal{M}, \text{ and } z^j = \frac{1}{\tilde{v}^j}\left(1 - \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right), \text{ for each } j \in \mathcal{M}.$$

To show (70), it suffices to show the first, second and fourth inequalities hold because the rest of the constraints hold trivially.

Since $p^j \ge \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j}{K} \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j}$ for all $j \in \mathcal{M}$, we have

$$\sum_{i\in\mathcal{N}} w_i^j = \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j v_i^j} = \frac{1}{p^j} \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j} \le \frac{K}{\tilde{v}^j}\left(1 - \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right) = K z^j \text{ for each } j \in \mathcal{M}.$$

Hence, the second inequality of (70) holds. Since $p^j \ge \sum_{i'\in\mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{\tilde{v}^j \alpha_i^j}{v_i^j}$ for each $i \in \mathcal{N}, j \in \mathcal{M}$, we have

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \le \frac{1}{\tilde{v}^j}\left(1 - \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right) = z^j \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

Therefore, (70) holds. Hence, the feasible region of $\boldsymbol{\alpha}$ is characterized by (68). This completes the proof. □

G. Resource Allocation and Other Ad-Allocation Problems

For an arbitrary resource-allocation problem, we can reformulate the program $(\mathcal{OTP})$ by setting the cardinality of offer set to 1, and the choice probability ϕ_i^j to 1 when i is offered, otherwise 0, and changing b_i to b_i^j, as follows:

$$\begin{aligned}
&\max_{\boldsymbol{\alpha}\geq\mathbf{0}} \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\
&\text{s.t.} \sum_{j\in\mathcal{M}} p^j \max_{i\in\mathcal{N}} \theta_i^j - \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j\theta_i^j \geq 0, \text{ for each } \boldsymbol{\theta}\geq\mathbf{0} \\
&\qquad \sum_{j\in\mathcal{M}} b_i^j\alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i\in\mathcal{N}.
\end{aligned} \tag{75}$$

Setting $\theta_i^j = 1$ for some j and all other $\boldsymbol{\theta}$'s equal to zero in the first group of constraints of (75), immediately implies that if $\boldsymbol{\alpha}$ is single-period feasible, we have

$$\sum_{i\in\mathcal{N}} \alpha_i^j \leq p^j, \text{ for each } j\in\mathcal{M}. \tag{76}$$

Reversely, if (76) holds, then for any $\boldsymbol{\theta}\geq\mathbf{0}$, we have

$$\sum_{j\in\mathcal{M}} p^j \max_{i\in\mathcal{N}} \theta_i^j \geq \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j \max_{i\in\mathcal{N}} \theta_i^j \geq \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j\theta_i^j,$$

i.e. the first group of constraints of (75) holds and it is equivalent (76).

Especially, for the AdWords problem, each vertex $i\in\mathcal{N}$ has budget B_i, and edge (i,j) has a bid b_i^j. When a vertex $j\in\mathcal{M}$ arrivals, we have to match it to a vertex $i\in\mathcal{N}$ who has not yet spent all its budget. After the matching, b_i^j is depleted from B_i. The goal is to maximize the total bid spent. So the formulation is as follows:

$$\begin{aligned}
&\max_{\boldsymbol{\alpha}\geq\mathbf{0}} \sum_{i\in\mathcal{N},j\in\mathcal{M}} b_i^j\alpha_i^j \\
&\text{s.t.} \sum_{i\in\mathcal{N}} \alpha_i^j \leq p^j, \text{ for each } j\in\mathcal{M}, \\
&\qquad \sum_{j\in\mathcal{M}} b_i^j\alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i\in\mathcal{N},
\end{aligned} \tag{77}$$

and the Display Ads problem, each edge (i,j) has a weight w_i^j, and each vertex $i\in\mathcal{N}$ has capacity c_i. When a vertex $j\in\mathcal{M}$ arrivals, we have to match it to a vertex $i\in\mathcal{N}$ who can be matched at most c_i times. The goal is to maximize the total weight of the matched edges. So the formulation is as follows:

$$\begin{aligned}
&\max_{\boldsymbol{\alpha}\geq\mathbf{0}} \sum_{i\in\mathcal{N},j\in\mathcal{M}} w_i^j\alpha_i^j \\
&\text{s.t.} \sum_{i\in\mathcal{N}} \alpha_i^j \leq p^j, \text{ for each } j\in\mathcal{M}, \\
&\qquad \sum_{j\in\mathcal{M}} \alpha_i^j \leq \frac{c_i}{T}, \text{ for each } i\in\mathcal{N}.
\end{aligned} \tag{78}$$

After solving the above problems and then obtaining the optimal $\boldsymbol{\alpha}$, we can apply the DWO policy to get the allocation. Hence, our DWO policy can be applied to these resource-allocation problems.

H. Implementation Details of the Re-Solving Benchmarks

In this section, we provide the implementation details of the re-solving benchmarks: the Fluid-R, Fluid-I-R, and Fluid-E-R policies. For all policies, we preset a set of re-solving epochs $\mathcal{T} := \{t_u : u = 0, 1, 2, ..., U\}$ in which the algorithm re-solves the Fluid convex program with updated ad budgets and click-through requirements, where $t_0 = 1$ refers to solving ($\mathcal{OP}_{\mathsf{Fluid}}$) at the beginning of the planning horizon. At each re-solving epoch t_u $(1 \le u \le U)$, an Fluid-R, Fluid-I-R, or Fluid-E-R policy will re-solve the following convex program with budget and click-through requirement updates (similar to Appendix B, we define $Y_i^j(t) := \sum_{\tau=1}^{t-1} y_i^j(\tau)$ as the cumulative click-throughs until the beginning of time t):

$$
\begin{aligned}
\max_{\boldsymbol{z}} \ & \mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}|u) := \sum_{i\in\mathcal{N}, j\in\mathcal{M}, S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^j(S) + \lambda F(\boldsymbol{\zeta}) \\
\text{s.t.} \ & \sum_{j\in\mathcal{M}, S\in\mathfrak{S}^j} b_i p^j \phi_i^j(S) z^j(S) \le \frac{B_i - b_i \sum_{j\in\mathcal{M}} Y_i^j(t_u)}{T - t_u + 1} \text{ for each } i \in \mathcal{N} \\
& \sum_{j\in\mathcal{C}, S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \ge \frac{\eta_i^{\mathcal{C}} - \sum_{j\in\mathcal{C}} Y_i^j(t_u)}{T - t_u + 1} \text{ for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i \\
& \sum_{S\in\mathfrak{S}^j} z^j(S) \le 1 \text{ for each } j \in \mathcal{M} \\
& z^j(S) \ge 0 \text{ for each } j \in \mathcal{M}, S \in \mathfrak{S}^j \\
& \boldsymbol{\zeta} \in \mathbb{R}^{nm}, \text{ with } \zeta_i^j = \frac{1}{T} \cdot Y_i^j(t_u) + \frac{T - t_u + 1}{T} \cdot \sum_{S\in\mathfrak{S}^j} p^j z^j(S) \phi_i^j(S).
\end{aligned}
\tag{79}
$$

It is clear from the formulation that (79) is reduced to ($\mathcal{OP}_{\mathsf{Fluid}}$) with $u = 0$. We denote the solution to (79) in re-solving epoch u as $\boldsymbol{z}^*(u)$. Therefore, the Fluid-R, Fluid-I-R, and Fluid-E-R policies re-solve (79) for U times at the re-solving epochs $\mathcal{T}$, and follows the static policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*(u))$ from time t_u to time $t_{u+1} - 1$ for $u = 1, 2, ..., U$, where we adopt the convention $t_{U+1} = T + 1$.

The only difference between the Fluid-R, Fluid-I-R, and Fluid-E-R policies is the pattern of the re-solving epochs $\mathcal{T}$. More specifically, for the Fluid-R policy (see, also, Jasin and Kumar 2012), $\mathcal{T}$ is evenly spread across all time, i.e., $t_u = \lceil \frac{Tu}{U} \rceil$, where $\lceil \cdot \rceil$ refers to the ceiling function. For the Fluid-I-R policy (see, also, Bumpensanti and Wang 2020), $\mathcal{T}$ is sparser at the beginning of the planning horizon and denser at the end. Specifically, following Bumpensanti and Wang (2020), we set the re-solving epoch $t_u = \lceil T - T^{(5/6)^u} \rceil + 1$ for all $u \in \{1, 2, ..., U\}$, where $U = \lceil \frac{\log(\log(T))}{\log(6/5)} \rceil$ (in which the logarithm base is 10). We remark that, in our numerical experiments, we set $T = 1,000$ and, therefore, $U = 7$ for the Fluid-I-R policy. To make the benchmark policies more comparable, we set $U = 7$ for the Fluid-R policy as well. Finally, for the Fluid-E-R policy, it re-solves (79) in each period, i.e., $U = 999$ and $t_u = u + 1$ for each u (see, also, Balseiro et al. 2023).

I. Settings and Parameters in Numerical Experiments

We consider an ad-allocation problem with $T = 1,000$ customers of five types and 50 ads. The customer-type distribution $(p^1, p^2, ..., p^5)$, where $\mathbb{P}[j(t) = j] = p^j$ and $\sum_{j=1}^{5} p^j = 1$, is generated from a five-dimensional Dirichlet distribution. We sample the per-click value r_i^j for each ad i by type j independently from a uniform distribution on the interval $[10, 50]$. We model the click-through behavior of the customers using MNL, i.e., for each $i \in S$ and $j \in \mathcal{M}$, $\phi_i^j(S)$ is given by (11). Each ad/customer-type pair is associated with an attraction index v_i^j. For ad i and customer type j, let $v_i^j := \exp(u_i^j)$, where u_i^j is independently sampled from the uniform distribution on the interval $[0,5]$. We set the cardinality constraint such that the maximum size of an offer-set is 2, i.e., $|S(t)| \leq 2$ for each customer t. The fairness metric we use in the numerical studies is the GMD fairness (24) (in Appendix C) with $\lambda = 10$.

Our first set of numerical studies is based on problem instances generated by systematically varying two focal parameters: (a) the concentration parameter (CP) associated with the proportion of each customer type, and (b) the loading factor (LF), defined as the ratio of total expected demand to total supply. Specifically, the concentration parameters are determined by the parameters of the Dirichlet distribution that we use to generate $(p^1, p^2, \cdots, p^5)$. We vary CP to change the uniformness of proportions $p^1, ..., p^m$ of m customer types, which are generated by a m-dimension Dirichlet distribution. The m-dimension Dirichlet distribution has m concentration parameters $\beta_1, ..., \beta_m$. In our experiments, we set $CP := \beta_0 = \beta_1 = \beta_2 = ... = \beta_m$. Note that, for all j, $\mathbb{E}[p^j] = \beta_j / \sum_{k=1}^{m} \beta_k = \frac{1}{m}$ and $Var(p^j) = \frac{m-1}{m^2(m\beta_0+1)}$, which is decreasing in β_0. For $j \neq k$, the covariance between p^j and p^k is $-\frac{1}{m^2(m\beta_0+1)}$, which is increasing in β_0. Hence, if β_0 is larger, the sampled customer type distribution will be close to the uniform distribution on $\{1, 2, ..., m\}$. In contrast, if $CP = \beta_0$ is small, the customer type distribution is more likely to be concentrated on a subset of $\{1, 2, ..., m\}$. In other words, the higher the CP, the more uniform the generated distribution of customer types. The loading factor is the ratio of total user traffic to total affordable traffic with the budgets of the advertisers, namely $LF = T / \sum_{i \in \mathcal{N}} \frac{B_i}{b_i}$. In our experiments, we set $b_i = 1$ and B_i equal to the value of rounded $T/(n \times LF)$ for all i. As is clear from their definitions, CP measures the uniformity of the customer-type distribution, while LF measures the tightness of the ad budget. The higher the CP, the more uniform the distribution of customer types; the higher the LF, the tighter the budget constraints for the ad campaigns.

In our experiments, we vary CP in the set $\{0.1, 1, 10, 100\}$ and LF in the set $\{0.5, 0.75, 1, 1.25, 1.5\}$. For each problem instance, we solve the problem ($\mathcal{OTP} - \mathcal{MNL}$) with different per-click values sampled from the same distribution as r_i^j's and without click-through requirements to obtain the solution-optimal targets $\boldsymbol{\alpha}^*$. A click-through requirement $\eta_i^{\{j\}}$, where $\{j\}$ is a singleton set of customer type j, is generated by the product of α_i^{j*} and a random number independently sampled from the $[0,1]$ uniform distribution for each ad i and each customer type j.

In Figure 2, we pick an ad with index $i = 1$ from 50 ads to show the quantiles of the click-through sample-paths for the five approaches. Because both the per-click value r_i^j and the attraction index v_i^j for each ad i by type j are generated independently from two distributions respectively, the case of the ad with index $i = 1$ is typical as well as any other ads. Hence, the smooth budget depletion of our proposed algorithm could happen to all ads generally.

J. Mean-Reverting Behavior of the DWO Policy

To highlight the mean-reverting property of our proposed DWO algorithm, we examine the intertemporal correlation between of the click-through $y_i^j(t)$ of ad i by type-j customers in period t and the *per-period debt*, defined as $\Delta_i^j(t) := d_i^j(t)/t$ where $d_i^j(t)$ is the debt of ad i for customer segment j at the beginning of time t (as defined in Algorithm 1). Recall that debt measures the gap between the click-through target set by the algorithm and the realized click-throughs. Therefore, if the correlation between $y_i^j(t)$ and $\Delta_i^j(t)$ is larger, it implies that the algorithm "pays back" the "debt" faster and, therefore, the mean-reversion of the click-through process is stronger.

For each of the 5 algorithms studied in our numerical experiments, we regress the click-through on the per-period debt using the following model specification with 30 million randomly drawn samples for each policy studied:

$$y_i^j(t) = a_0 + a_1 \Delta_i^j(t) + \epsilon$$

The regression results are reported in Table 3. If we instead regress the click-through on the total debt $d_i^j(t)$, the results will be similar because the per-period debt is a constant multiplication of the total debt.

Policy	Coefficient	Estimation	Standard Error	t-statistics	p-value
Fluid	a_0	0.0036905	1.1071e-05	333.35	0
	a_1	-0.00042435	0.0015441	-0.27481	0.78346
Fluid-R	a_0	0.0037289	1.1128e-05	335.09	0
	a_1	0.044064	0.0016192	27.213	4.505e-163
Fluid-I-R	a_0	0.0037046	1.1092e-05	334	0
	a_1	-0.0051239	0.0015941	-3.2144	0.0013071
Fluid-E-R	a_0	0.0036979	1.1084e-05	333.63	0
	a_1	0.061641	0.0016548	37.25	1.0702e-303
DWO	a_0	0.0038891	1.1218e-05	346.67	0
	a_1	0.48795	0.0020018	243.76	0

Table 3 The Regression Results of the Intertemporal Correlations Between Click-Throughs and Per-Period Debts

Table 3 shows that our DWO algorithm clearly drives the mean-reverting pattern for the click-through process, captured by the fact that the estimate $\hat{a}_1 = 0.48795$ is positive, large and statistically significant. This is expected given that the DWO policy gives a higher weight for the ad/customer pair with a larger debt at each time t. An important observation from our regression results is that, the estimated coefficient $\hat{a}_1$ of our DWO algorithm (0.48795) is about one order of magnitude larger than that of the Fluid-based benchmarks. Such an observation delivers an intriguing insight that our debt-based algorithm drives the click-through process toward its mean (i.e., the target set by the first-stage optimization) and, as a consequence, result in a more stable budget depletion process for the ads. Finally, we remark that, because of the budget constraints of the ads, the Fluid-based benchmarks also exhibit certain mean-reverting property weaker than our DWO algorithm.

K. Efficiency-Fairness Trade-off

We demonstrate the efficiency-fairness trade-off by varying the parameter λ in our setting. Specifically, we consider the problem instance with $\lambda \in \{10^{i_\lambda} : i_\lambda = -1 + 0.1 \times (i-1), i = 1, 2, ..., 31\} \cup \{0\}$ (under the problem instance $LF = 1$ and $CP = 100$). We plot the relationship between the efficiency and Gini fairness in Figure 4 for different values of λ, where the x-axis (resp. y-axis) is the ratio between the expected efficiency (resp. expected GMD fairness) with respect to λ and that with respect to $\lambda = 0$ (i.e., the system is purely efficiency-driven). Our numerical results reveal the trade-off between efficiency and fairness. Importantly, we find that introducing the fairness term in the objective function could substantially reduce the algorithmic bias without much compromising the advertising efficiency. For example, a 1% (resp. 5%) optimality gap in efficiency could reduce about 50% (resp. 90%) of the algorithmic bias.

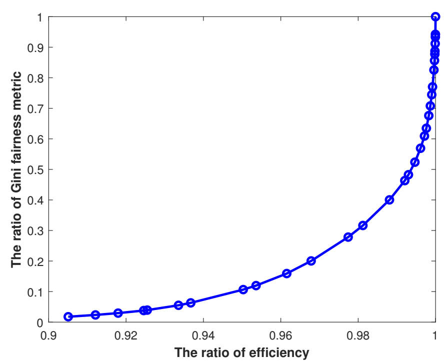

Figure 4 **The Trade-Off Between Optimal Efficiency and Gini Fairness**

L. Comparison With the Inventory-Balancing Policies

In this section, we compare our DWO algorithm against another family of benchmarks called the inventory balancing (IB) policies. Specifically, Golrezaei et al. (2014) propose two IB algorithms which implement real-time personalized offer-set optimization with an exponential penalty function (the EIB policy) and a linear penalty function (the LIB policy), respectively, to reweight the value of each ad. Upon the arrival of customer t, the IB policies solve a single-period offer-set optimization problem with a discounted value $r_i\Phi(B_i(t-1)/B_i)$, where $\Phi(\cdot)$ is an increasing discount function and $B_i(t-1)$ is the budget of ad i at the end of time $t-1$. The discount function is $\Phi(x) = (e/(e-1)) \cdot (1-e^{-x})$ under the EIB policy and is $\Phi(x) = x$ under the LIB policy. It is hard, if not impossible, to incorporate the non linear fairness metric and the click-through requirements into the IB policies. To account for both the budget constraints and the click-through requirements, one needs to design weight functions handling them jointly. In a case where one ad has little remaining budget, but also falls behind the schedule of its click-through requirements, it is unclear how we should design the weight functions to adjust the weight of this ad. Even worse, the click-through requirements are imposed at the ad by subset of customer types level, which may not be compatible with the ad level budget constraints. Therefore, even without incorporating the fairness metric, it is highly nontrivial to extend the IB algorithm that embeds the click-through requirements. So we remove these modeling features in the comparison between DWO and IB policies. We consider the same numerical setup as Section 6 with the click-through requirements and the fairness metric removed, and the identical per-click value r_i of each ad i across all customer types, sampled from a uniform distribution on the interval $[10, 50]$.

CP	LF	EIB	LIB	DWO
0.1	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.44% (0.18%)	99.44% (0.25%)	99.20% (0.48%)
	0.75	91.79% (0.21%)	90.66% (0.28%)	99.00% (0.63%)
	0.5	89.19% (0.31%)	86.72% (0.32%)	98.96% (0.49%)
1	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.43% (0.31%)	99.45% (0.24%)	99.40% (0.39%)
	0.75	91.75% (0.17%)	90.60% (0.24%)	98.81% (0.63%)
	0.5	89.23% (0.29%)	86.86% (0.24%)	98.67% (0.64%)
10	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.54% (0.21%)	99.52% (0.27%)	99.37% (0.39%)
	0.75	91.64% (0.17%)	90.53% (0.25%)	99.19% (0.41%)
	0.5	89.20% (0.23%)	86.76% (0.24%)	98.75% (0.51%)
100	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.59% (0.25%)	99.56% (0.20%)	99.35% (0.40%)
	0.75	91.68% (0.19%)	90.53% (0.21%)	99.29% (0.29%)
	0.5	89.20% (0.18%)	86.71% (0.23%)	98.93% (0.62%)

Table 4 Numerical Results (Standard Error Relative to the Theoretical Upper Bound in Parentheses)

We report the results on the comparison between our DWO policy and the EIB and LIB algorithms in Table 4, with the ratio between the standard error of the total advertising revenue for each policy examined to the theoretical upper-bound of advertising revenue included in the parenthesis. The most important takeaway from our experiments is that the DWO policy outperforms the EIB and LIB algorithms when LF is low, especially when $LF < 1$. In this case, the budget constraints are not binding, so discounting the ad value when the budget is low is not helpful. On the other hand, when the loading factor is high, the budgets are more likely to be exhausted, so the discount functions of the EIB and LIB algorithms can help smoothly allocate the ad budgets, thus giving rise to higher efficiency performance than the DWO algorithm. To conclude this section, we remark that, because of the difficulty to incorporate the nonlinear fairness metric into the IB policies, this family of algorithms are not amenable to address the algorithmic fairness issue, which can be well handled by our DWO policy.