When to Push Ads: Optimal Mobile Ad Campaign Strategy under Markov Customer Dynamics

Guokai Li[1] Pin Gao[2] Zizhuo Wang[2]

[1] Smith School of Business, Queen's University, Kingston, Ontario, K7L 3N6, Canada
[2] School of Data Science, School of Management and Economics, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China
guokai.li@queensu.ca, gaopin@cuhk.edu.cn, wangzizhuo@cuhk.edu.cn

Problem definition: We investigate a seller's optimal advertising strategy targeting customers who interact with the seller over time. We model each customer's engagement by a continuous-time (hidden) Markov chain with two states, active and inactive. While in the active state, customers make purchases according to a Poisson process; in contrast, customers in the inactive state make no purchases. The seller can use advertisements to activate customers and the activation probability may be affected by customers' fatigue to promotion. The objective of the seller is to maximize the long-term average expected profit by designing an optimal ad campaign strategy based on customers' purchase histories. **Methodology/results:** For a single customer without promotion budget constraints, we demonstrate the optimality of a triple-threshold policy based on both the elapsed time since the last purchase or ad campaign and the number of repeated ad campaigns. When confronted with a budget constraint on the overall ad cost, we suggest first solving a relaxed problem and then implementing a separate triple-threshold policy for each customer with a specified budget. We establish the asymptotic optimality of this policy and show that the order of loss is $O\left(\sqrt{\log T/T}\right)$ for the general problem and $O\left(\sqrt{1/T}\right)$ for a special case. **Managerial implications:** We find that the seller tends to push ads earlier to customers with low ad costs, high purchase rewards, high activation probabilities, high purchase rates and low recapture rates (i.e., low transition rates from an inactive state to an active state). Surprisingly, we also find that the seller should push ads earlier to customers with intermediate churn rates (i.e., intermediate transition rates from an active state to an inactive state) compared to those with small or large churn rates. Overall, our work reveals important insights of the optimal ad campaign problem and provides a useful strategy in practice.

Key words: mobile ad campaign, Markov customer dynamics, revenue management

1. Introduction

Based on a report by Statista (2025), mobile devices accounted for an impressive 63% of all website traffic in the second quarter of 2025. This significant user base has prompted sellers to invest substantial effort and resources in leveraging mobile advertisements to boost customer engagement. In recent years, various industries (e.g., retailing, financial services, and entertainment) have embraced mobile ad campaigns such as SMS/MMS messages, push notifications, and in-app messages. Meanwhile, the global mobile engagement market size was valued at $12.5 billion in 2024,

with an estimated growth rate of 38% per year, expected to reach a staggering $130.3 billion by 2031 (VMR 2024). As a consequence, the operation of mobile ad campaigns has become a critical issue, garnering substantial attention from businesses. A key challenge in mobile ad operations is to determine the target and the timing of the ad campaigns. Specifically, the ad campaign strategy should specify when and to whom the seller should push ads to maximize the seller's profit.

This paper studies the optimal ad push policy for a seller to maximize the average expected profit under Markov customer dynamics. For practical concerns, we consider the case where the seller cannot observe customers' engagement states but can observe their purchase histories. Specifically, each customer's (hidden) engagement state evolves according to a continuous-time Markov chain with two states, active and inactive. Only when a customer is active will he make purchases (according to a Poisson process). The seller can employ personalized ad campaigns (at a cost) to activate customers. For example, a seller may utilize the "Custom Audience" feature on Meta Ads to target customers based on their emails or phone numbers (see Meta 2024). Upon receiving ad campaigns, customers in the inactive state will instantaneously transition to the active state with some probability depending on the last event (whether it is a purchase or an ad), while customers already in the active state remain unaffected. It is worth noting that pushing ads does not directly generate rewards but affects customers' states, subsequently influencing future rewards. Moreover, due to the existence of ad costs, a crucial trade-off arises in designing the ad campaign strategy: Frequent ad campaigns may lead to excessive ads being delivered to already active customers, resulting in unnecessary ad costs; infrequent ad campaigns may induce delayed activation of inactive customers, resulting in loss of potential revenue. To make things more complicated, as mentioned, the seller encounters challenges related to the incomplete observation of customers' states. These factors collectively contribute to the complexity of designing an optimal ad campaign strategy, posing a significant challenge in practical implementation.

In this paper, to shed light on the aforementioned problem, we study the optimal mobile ad campaign policy for a seller operating within a framework of Markov customer dynamics. In particular, we are interested in answering the following questions:

1. *What is the structure of the optimal ad push policy in the absence of budget constraints?*
2. *How do customer characteristics (e.g., purchase rates and transition rates between active and inactive states) affect the optimal policy?*
3. *How does the optimal policy change when there are budget constraints and a heterogeneous group of customers? Is there an efficient heuristic policy with a good performance guarantee?*

To answer the first question, we begin by considering one representative customer, as decisions for different customers are independent in the absence of budget constraints. In this scenario, we establish the optimality of a triple-threshold policy based on the *silence time*, (i.e., the time elapsed

after the last purchase or ad campaign), and the number of *repeated ads* (i.e., the number of ad campaigns since the last purchase). Specifically, the seller will not push ads until the silence time reaches a certain threshold depending on the number of repeated ads. Then, we investigate the comparative statics of the optimal thresholds to shed more light on this problem. To determine the optimal thresholds, the seller balances the trade-off between two countervailing effects: *timeliness*, which measures how timely the seller delivers ad campaigns to inactive customers, and *effectiveness*, which measures the expected benefit of sending an ad, calculated by the probability of a customer being inactive (when sent an ad) multiplied by the activation probability and the increased profit of accurate activation, minus the ad cost. Through extensive numerical experiments, we find that the seller is inclined to push ads earlier to customers with high rewards, low ad costs, high activation probabilities, high purchase rates, and low recapture rates (transition rate from inactive to active state). Furthermore, we find, somewhat surprisingly, that the seller should push ads earlier to customers with intermediate churn rates (transition rate from active to inactive states) compared to those with low or high churn rates. We also provide theoretical verification for these comparative statics in a special case (referred to as the "efficient-activation case") where ad campaigns can activate inactive customers with certainty (i.e., the activation probabilities equal one).

Having understood the optimal policy for a representative customer without budget constraints, we then consider the problem when the seller faces a group of heterogeneous customers with budget constraints over a finite time horizon. In such a case, the seller's challenge is prioritizing customers of different types and making the best use of limited resources over a finite horizon T. To address this problem, we propose an easy-to-implement heuristic policy based on the optimal solution to a relaxed problem. Then, we prove that the asymptotic optimality gap of the proposed policy is $O\left(\sqrt{\log T/T}\right)$ for the general problem and $O\left(\sqrt{1/T}\right)$ for the efficient-activation case. Moreover, to test the robustness of our results and enrich the insights, we consider the case when customers may react strategically to the seller's policy. We show that our main results still qualitatively hold and obtain some additional insights in this setting.

To summarize, to the best of our knowledge, our work presents the first comprehensive prescriptive analysis for the mobile ad campaign strategy when customer behaviors are governed by a hidden Markov model. Our results provide crucial insights and practical guidelines for sellers to design efficient mobile ad campaign strategies. By addressing these challenges, we contribute to the advancement of mobile advertising in the digital marketing era, empowering businesses to optimize their mobile ad campaigns.

The remainder of this paper is organized as follows. In the rest of this section, we review literature related to our work. In Section 2, we formally introduce the model setup. In Sections 3 and 4, we analyze the optimal ad push policy without budget constraints. In Section 5, we study the problem

under budget constraints and propose an easy-to-implement and asymptotically optimal algorithm. In Section 6, we incorporate strategic customer behaviors into the base model. We conclude the paper in Section 7. All proofs and additional materials are relegated to the Appendix.

1.1. Literature Review

Our work is broadly connected to several streams of literature. In the following, we provide a review for each stream respectively.

Customer relationship management. The behavioral modeling in our paper is closely related to customer relationship management, in which the seller analyzes and manages her interactions with customers. In this area, some researchers focus on deriving behavioral models to estimate customers' marketing values. For example, the pioneering work by Schmittlein et al. (1987) establishes a behavioral model where an active customer makes purchases according to a Poisson process and becomes inactive (i.e., churns) after an exponentially distributed lifetime. Since the seller cannot observe each customer's exact state, the authors provide methods to estimate each customer's active probability and future transactions based on the purchase history. Then, Fader et al. (2005), Fader et al. (2010) and Bemmaor and Glady (2012) propose some variants of this model and report the effectiveness of such kind of models on multiple datasets. However, the above studies ignore the possibility that an inactive customer can become active again (see Jain and Singh 2002). To solve this problem, Netzer et al. (2008) adopts a hidden Markov model to depict customers' behaviors. Specifically, each customer's (hidden) engagement state evolves according to a Markov chain and customers make purchases with state-dependent probabilities. Given the purchase history, they provide a method to compute the conditional distribution over different states and the expected number of future transactions. Such a model is also reported to fit well on several datasets, e.g., Schwartz et al. (2014), Huang et al. (2019) and Zhang et al. (2019). In this paper, we also adopt a hidden Markov chain behavioral model but analyze the problem from the operational perspective, which is different from the above literature that emphasizes estimation.

Then, we review papers in this field focusing on operational problems. First, there are works considering customers who learn from historical interactions. For example, Aflaki and Popescu (2014) study the optimal dynamic quality control when each customer infers the service quality based on historical experiences, and find that a fixed-quality policy is optimal in the long run. Following a similar setup, when service quality cannot be controlled, DeCroix et al. (2021) show that dynamic personalized pricing can fully combat the revenue loss due to service variability. Then, given that each customer learns his service utility after his first purchase, Bernstein et al. (2022) investigate the optimal (open-loop) intertemporal quality investment policy to maximize the aggregate demand. The above literature implicitly assumes that customers are always active along the

time horizon and hence ignores the churn behavior, whereas in our paper customers' engagement states are time-variant and the churn behavior is studied. We also note some papers considering customers' churn behavior. For example, Bastani et al. (2022) propose a dynamic recommendation algorithm for the case where a customer arrives with unknown preferences and may churn after receiving irrelevant recommendations. For customers who may churn due to unsatisfied services, Kanoria et al. (2024) investigate the optimal dynamic service mode control to maximize the revenue before the customer churns. These works assume that customers' states are fully observable, whereas our work studies partially observable engagement states.

Moreover, we would like to highlight a closely related work, Sheng et al. (2022). In Sheng et al. (2022), the authors investigate level-based games where each customer's engagement state can be represented by his (fully observable) level, and analyze when the game publisher should provide the incented option (which allows players to get benefits by watching ads). They prove the optimality of a threshold policy, which provides the incented option if and only if the level is below a threshold. Then, we explain the salient features of our paper. First, we establish the optimality of a triple-threshold policy under a partially observable Markov decision process (POMDP), which is significantly harder than the fully observable case. As Sheng et al. (2022) noted, "in the situation where ... a POMDP model would be required ... the challenge of establishing the existence of threshold policies is likely to be prohibitive." Second, the managerial insights are different. The purpose of the threshold policy in Sheng et al. (2022) is to utilize the incented option to reduce low-engaged customers' churn behaviors and prevent it from cannibalizing in-app purchases from high-engaged customers. In contrast, due to partial information, the purpose of the threshold policy in our paper is to balance the effectiveness and timeliness of ad campaigns.

Pulse marketing. The structure of our optimal policy shares some similarities with pulse marketing, in which advertising effort concentrates on discrete points. There is much literature that investigates the advantages of pulse marketing. For example, given an S-shaped sales response to advertising effort, Mahajan and Muller (1986) show that pulse marketing can be optimal. Then, Hahn and Hyun (1991) and Aravindakshan and Naik (2015) consider additional practical factors (e.g., advertisement costs and memory effects) and identify more conditions when pulse marketing is optimal. In our paper, we prove the optimality of a triple-threshold policy, under which the advertising effort is also exerted at discrete time points, which shares some similarities with pulse marketing. However, the research on pulse marketing mainly focuses on the aggregate market share and takes the same action to a large group of (non-atomic) customers, which is a common setting in the area of promotion design (see Neslin et al. 2009). Since the uncertainty in customers' behaviors cancels out in a large market, the proposed open-loop policies in these works can often

attain optimality. In contrast, motivated by the mobile ad practice, we study a personalized marketing scenario where uncertainty in each (atomic) customer's behaviors plays a crucial role, and the consideration of partial information further complicates our model. In this case, an open-loop policy cannot attain optimality, and we derive the optimal closed-loop policy in this paper.

Network revenue management. The budget-constrained case in our paper is broadly related to network revenue management (NRM). In NRM problems, a seller manages limited inventory of multiple resources over a finite time horizon, and customers with varying resource requests and bid prices sequentially arrive at the system. Once a customer arrives, the seller makes an irrevocable decision whether to accept the request, in order to maximize the total profit under the inventory constraints. For this problem, researchers typically propose heuristic policies and then provide regret bounds (see Gallego et al. 2019). For example, Talluri and Van Ryzin (1998), Cooper (2002) and Reiman and Wang (2008) propose algorithms with $O(\sqrt{T})$ regret bounds, implying their asymptotic optimality. Then, incorporating resolving techniques, researchers propose algorithms with constant regret bounds, e.g., Jasin and Kumar (2012), Bumpensanti and Wang (2020), Vera et al. (2021) and Li et al. (2024). The above literature focuses on independent demand, i.e., the arrivals in different periods are independent. In contrast, our model captures the intertemporal correlation by a hidden Markovian engagement state. Recently, there have been a few papers considering the intertemporal correlation. For example, Bai et al. (2022) consider the case where customers churn simultaneously after a random lifetime, and provide an asymptotically optimal policy. Jiang (2023) studies the NRM problem with customer types evolving according to a Markov chain, and derives a constant-approximation algorithm. Lan et al. (2026) investigate the problem when customers' preferences evolve according to a Markov chain if their requests are rejected by the seller, and provide an asymptotically optimal policy. Compared with the above literature, the consideration of partial information in our paper introduces an additional layer of complexity to our problem.

Restless Bandit. The budget-constrained case in our paper is also related to the restless bandits. In this problem, there are multiple arms, each of which evolves regardless of whether it is pulled, and a principal can pull a limited number of arms in each period to maximize his total profit (see Whittle 1988). There are also some asymptotically optimal policies in this area. For example, Brown and Smith (2020) and Brown and Zhang (2022) propose heuristic policies for the finite-horizon problem based on Lagrangian relaxation. These policies are asymptotically optimal when the number of arms tends to infinity and the number of arms the principal can pull increases proportionally. Our paper is different from this line of literature in two aspects. First, the restless bandit problem considers a budget constraint in each period while we consider an overall budget constraint over the whole time horizon. Second, the asymptotic results are different. Due to the

difficulty of the per-period constraint, there has not been any algorithm that is asymptotically optimal in the sense that the time horizon tends to infinity (see Brown and Zhang 2023 for a review).

Finally, our work is also related to the literature on recommendation systems and we refer readers to Portugal et al. (2018) and Da'u and Salim (2020) for comprehensive reviews on this literature. Typically, researchers in this area assume that each recommendation instantly generates a reward (e.g., a click-through) while its impact on the customer's hidden state is less considered. In contrast, our work uses a hidden Markov model to depict the state of the customers and thus provides an alternative approach to characterize the impact of a recommendation.

2. Base Model

We consider a seller (she) selling a single type of product to a pool of customers (he/them) over an infinite time horizon. Each customer's (hidden) engagement state evolves according to a continuous-time Markov chain with two states, active (A) or inactive (I). A customer in the active state will purchase from the seller according to a Poisson process, whereas a customer in the inactive state will make no purchases. Moreover, the seller can push ad campaigns (e.g., SMS/MMS messages, push notifications, and in-app messages) to activate customers at a cost. Observing only individual purchase histories, the seller aims to maximize the long-term average expected profit. In the following, we specify customers' behaviors and the seller's optimization problem.

The Customer. We first model customers' purchase behaviors without ad campaigns. As Brodie et al. (2011) state, engagement is a dynamic and iterative psychological process, which is typically not observable. In this paper, similar to Netzer et al. (2008), we adopt a hidden Markov behavioral model with a Poisson purchase process. Such kind of behavioral models have been reported to fit well on multiple datasets, e.g., Schwartz et al. (2014), Huang et al. (2019) and Zhang et al. (2019). Specifically, we assume that each customer has a hidden engagement state $s_t \in \{A, I\}$ at time t. The state dynamics is depicted as a continuous-time Markov chain, where a churn event (with rate $\gamma \in \mathbb{R}_+$) represents an active customer becoming inactive and a recapture event (with rate $\theta \in \mathbb{R}_+$) represents the opposite.

When a customer is in the active state, he will purchase from the seller according to a Poisson process with rate $\lambda \in \mathbb{R}_+$; otherwise, the customer will not make any purchase. The parameters are assumed to be known, as they can be estimated from historical behavioral data (see, e.g., Zhang et al. 2019).[1] To facilitate understanding customers' behaviors, we provide an illustration in Figure 1.

[1] In practice, there may be insufficient data to estimate these parameters in the beginning, in which case a learning algorithm may need to be considered. Since this paper focuses on the optimal ad push policy under the Markov dynamics, we leave the analysis of the learning problem as a future research direction. We note that there are several learning-while-doing algorithms with sub-linear regrets for average-reward MDP problems with constraints, e.g., Chen et al. (2022a) and Ghosh et al. (2023). Those algorithms may be adopted in the learning problem.

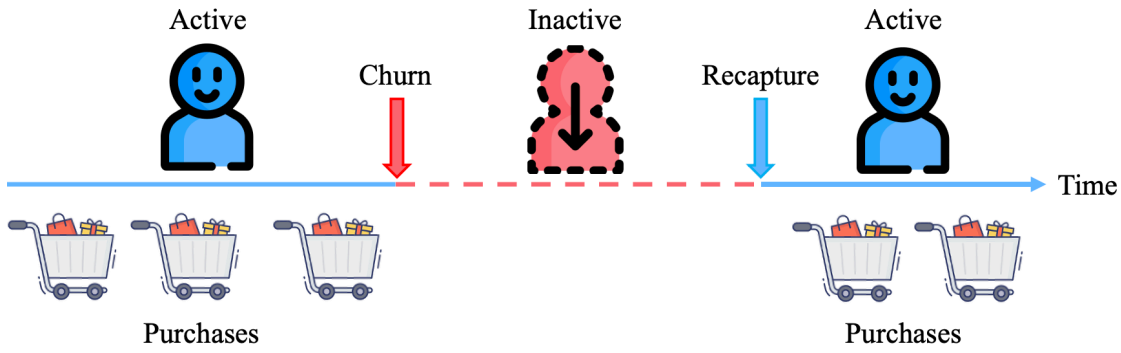


Figure 1 **Illustration of customers' behaviors. Only purchases can be observed by the seller.**

The Seller. According to field experiments (see, e.g., Goldfarb and Tucker 2011, Andrews et al. 2016, Rafieian and Yoganarasimhan 2021), ad campaigns can significantly boost customers' engagement. Hence, we assume that when receiving an ad campaign, inactive customers become active instantaneously with some *activation probability*, while active customers remain active.[2] Moreover, we assume that the activation probability depends on the last event (whether a purchase or an ad). Specifically, the activation probability is p_p (p_a, resp.) if the last event is a purchase (an ad, resp.). We set $p_p \geq p_a$ to capture the marketing fatigue phenomenon, i.e., repeated ad campaigns may lead to a lower click-through rate (e.g., StackAdapt 2023). To facilitate understanding, we illustrate the activation probabilities in Figure 2.

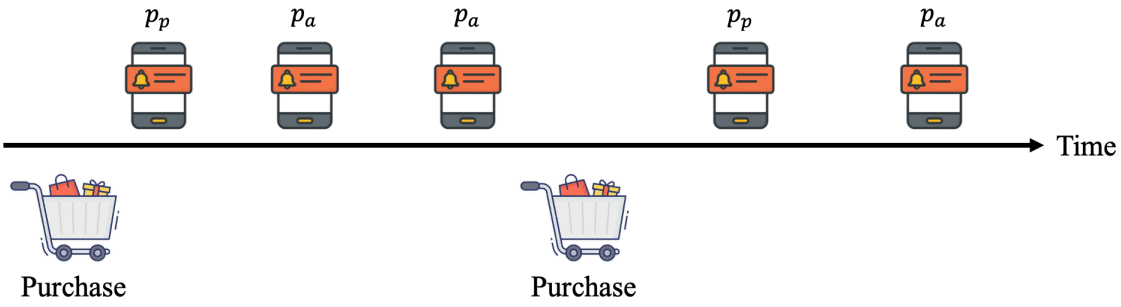


Figure 2 **Illustration of activation probabilities.**

Without loss of generality, the production cost is normalized to zero. The revenue of each purchase is denoted as $r \in \mathbb{R}_+$, and each ad campaign will incur an immediate cost $c \in \mathbb{R}_+$. In this work, we focus on the scenario where the seller determines a personalized ad push policy. In many practical settings, the seller is able to conduct personalized ad campaigns based on customers' emails or phone numbers (see, e.g., Meta 2024). Without coupled constraints (e.g., a budget constraint on

[2] Our main results still hold if we incorporate the phenomenon that given an ad campaign, an active customer may become inactive with some probability (see Appendix B for more details).

the total ad cost), decisions on different customers are independent, and hence we can focus on a single representative customer for analysis. As mentioned, a customer's engagement state evolves along the time horizon but cannot be perfectly observed. In practice, the seller needs to infer the customer's state and make decisions based on observed purchase and ad campaign records. In the following, we formally formulate this problem as a partially observable Markov decision process (POMDP).

From the seller's perspective, the engagement state $s_t \in \{A, I\}$ is insufficient to capture the system dynamics due to the last-event-dependent activation probability. Thus, we define $x_t \in \mathbb{X} := \{A_p, I_p, A_a, I_a, P\}$ as the hidden state at time t. Specifically, both A_p and I_p indicate that the last event was a purchase, implying that the activation probability is p_p, but they differ in the current engagement state s_t, which is active or inactive. Likewise, A_a and I_a represent the case where the last event is an ad push, implying a lower activation probability p_a. The last state P represents that a purchase happens, which is similar to A_p except for an additional reward r. Note that the engagement state is P only at every purchase time point, and transits to the state A_p instantly. Let $a_t \in \mathbb{A} := \{0, 1\}$ denote whether an ad campaign is conducted at time t. The transition kernel $\mathcal{P}(x_{t+\Delta t} \mid x_t, a_t)$ is specified in Appendix A.2. Let the observation $y_t \in \mathbb{Y} := \{0, 1\}$ denote whether the seller observes a purchase at time t. As mentioned, a customer only makes purchases when his engagement state is P. Thus, the observation kernel $\mathcal{Q}(y_t \mid x_t)$ is specified as $\mathcal{Q}\left(y_t = 1 \mid x_t = P\right) = 1$ and $\mathcal{Q}\left(y_t = 0 \mid x_t \neq P\right) = 1$.

Then, we specify admissible policies and the objective. Assume that each customer enters the seller's ad campaign system after one of his purchases, i.e., the initial state x_0 is known as A_p. Then, the observed history at time t is $h_t := \{x_0\} \cup \{(y_\ell, a_\ell)\}_{\ell \in [0,t)} \cup \{y_t\}$. An *admissible policy* μ for the problem is defined as a sequence $\mu = \{\mu_t\}_{t \geq 0}$ of stochastic kernels $\mu_t : h_t \to [0, 1]$ mapping from the observed history to the probability of pushing ads. The reward function is specified as $R(x_t, a_t) = \mathbb{1}\{x_t = P\} \cdot r - a_t \cdot c$. Let $N_p^\mu(t)$ and $N_a^\mu(t)$ denote the counting processes of purchases and ad campaigns under a policy μ, respectively. The seller will determine an admissible ad push policy μ to maximize its long-term average expected profit:

$$\sup_{\mu \in \Pi} \left\{ \liminf_{T \to \infty} \frac{1}{T} \mathbb{E}\left[r \cdot N_p^\mu(T) - c \cdot N_a^\mu(T) \right] \right\}, \tag{1}$$

where Π denotes the set of all admissible policies.

3. Ad Push Policy

In this section, we investigate the optimal ad push policy for a representative customer. As mentioned, we consider the practical scenario where the seller can only observe the customer's purchase and ad campaign records, resulting in a POMDP formulation. Although deriving the optimal policy

for a POMDP in general is hard, we show that there exists a succinct and insightful policy that is optimal for this problem.

A standard approach for solving POMDPs is to convert it into an equivalent belief-MDP (see, e.g., Feinberg et al. 2016), in which the state is the conditional distribution over the state space $\mathbb{X}$ given the observed history h_t. Moreover, at any time t, given the records of ad campaigns and purchases, there are only three possible cases: $x_t \in \{A_p, I_p\}$, $x_t \in \{A_a, I_a\}$ and $x_t \in \{P\}$. Thus, it is sufficient to define the state as $\boldsymbol{b}_t \in \mathbb{B} := \left\{\boldsymbol{b} \in [0,1]^3 \,|\, \|\boldsymbol{b}\|_0 \le 1, b_3 \in \{0,1\}\right\}$, with $b_{t,1}$, $b_{t,2}$ and $b_{t,3}$ denoting the probabilities that the state is in A_p, A_a and P, respectively. Since the problem starts with $x_0 = A_p$, we have $\boldsymbol{b}_0 = [1,0,0]$. The transition kernel $\mathcal{P}_b(\boldsymbol{b}_{t+\Delta t} \,|\, \boldsymbol{b}_t, a_t)$ and the reward function $R_b(\boldsymbol{b}, a)$ are specified in Appendix A.2. In this case, an admissible (belief) policy μ^B is defined as a sequence $\mu^B = \{\mu_t^B\}_{t\ge 0}$ of stochastic kernels $\mu_t^B : \{\boldsymbol{b}_\ell\}_{\ell\in[0,t]} \to [0,1]$.

Among all admissible policies, there exist a class of special policies, *stationary policies*, that do not change over time and map the current belief vector into a deterministic action, i.e., $\mu_t^B\left(\{\boldsymbol{b}_\ell\}_{\ell\in[0,t]}\right) = \mu_S^B(\boldsymbol{b}_t) \in \{0,1\}$ for $t \ge 0$. Note that we follow the definition of stationary policies in Feinberg et al. (2012) and Feinberg et al. (2016), and such policies are sometimes also referred to as *nonrandomized/deterministic stationary policies* (e.g., Bertsekas and Shreve 1996, Guo and Rieder 2006). Then, similar to Bai et al. (2022), we prove the existence of a stationary policy that is optimal to the belief-MDP by checking assumptions in Theorem 4 in Feinberg et al. (2012) (see Appendix A.1 for the proof sketch and Appendix A.2 for the detailed proof.).

THEOREM 1. *There exists a stationary policy $\mu_S^B : \mathbb{B} \to \{0,1\}$ that is optimal to the belief-MDP of problem (1).*

According to Theorem 1, we can focus on the class of stationary policies. In the following, we show that the optimal stationary policy can be represented by a succinct policy, such that the search space can be further narrowed down. We first define a class of succinct policies and then show the mapping from an optimal stationary policy to such policies. We call a purchase or an ad campaign an *event*, and define the *silence time* η_t at time t as the amount of time elapsed since the last event, i.e., $\eta_t = t - \sup\{\ell \,|\, y_\ell = 1 \text{ or } a_\ell = 1\}$. Moreover, we define the number n_t of repeated ads to track the number of ad campaigns since the last purchase, i.e., $n_t = \left|\{\ell \le t \,|\, a_\ell = 1,\ \ell \ge \sup\{\tilde{\ell} \le t \,|\, y_{\tilde{\ell}} = 1\}\}\right|$. Then, we introduce a class of policies called triple-threshold policies.

DEFINITION 1. A policy $\varphi^{\boldsymbol{\omega}}$ is called a *triple-threshold policy* if there exist three thresholds $\boldsymbol{\omega} = (\omega_0, \omega_1, \omega_2)$ such that $\varphi(\eta, n) = \mathbb{1}(\eta \ge \omega_0 \text{ and } n = 0) + \mathbb{1}(\eta \ge \omega_1 \text{ and } n = 1) + \mathbb{1}(\eta \ge \omega_2 \text{ and } n \ge 2)$. That is, if there is no ad campaign since the last purchase, then the seller pushes ads once the silence time reaches ω_0. If there is one ad campaign (at least two ad campaigns, resp.) since the last purchase, then the seller pushes ads once the silence time reaches ω_1 (ω_2, resp.).

According to Definition 1, under a triple-threshold policy, the threshold depends on the number of ad campaigns since the last purchase. Then, we establish that there exists such a policy that is optimal to (1).

THEOREM 2. *There exists a triple-threshold policy $\varphi^{\boldsymbol{\omega}}$ that is optimal to* (1).

To prove Theorem 2, we first analyze the structural properties of the optimal stationary policy μ_S^B mentioned in Theorem 1. Then, we map the policy μ_S^B into a triple-threshold policy $\varphi^{\boldsymbol{\omega}}$. To facilitate understanding, we provide an illustration in Figure 3.

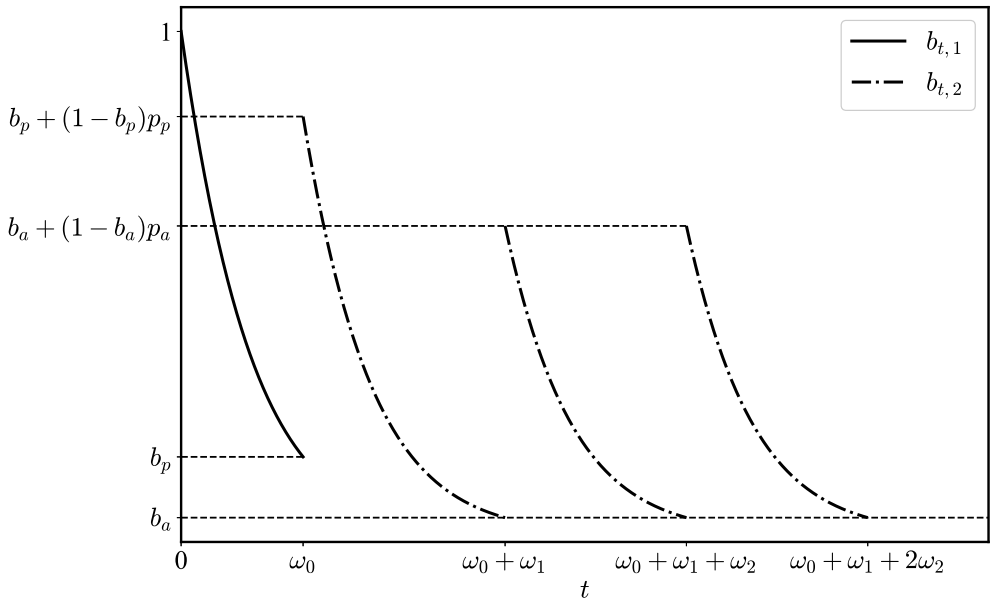


Figure 3 **Illustration of the belief-evolving process when no purchase happens.**

Define two belief thresholds $b_p := \sup\{b \in [0,1] \mid \mu_S^B([b,0,0]) = 1\}$ and $b_a = \sup\{b \in [0,1] \mid \mu_S^B([0,b,0]) = 1\}$. The mapping from μ_S^B to $\varphi^{\boldsymbol{\omega}}$ is as follows (some details are relegated to Appendix A.3): As mentioned, each customer enters the system with the belief state $[1,0,0]$. As illustrated in Figure 3, given the policy μ_S^B and no purchases, the belief $b_{t,1}$ of state A_p gradually decreases and no ads are pushed until the belief state becomes $[b_p,0,0]$, which implies a threshold ω_0. Then, the seller pushes an ad and the belief state becomes $[0, b_p + (1-b_p)p_p, 0]$, where $b_p + (1-b_p)p_p$ is the updated belief of A_a after an ad campaign. Then the belief $b_{t,2}$ of state A_a gradually decreases and no ads are pushed until the belief state $\boldsymbol{b}_t$ becomes $[0,b_a,0]$, which implies a threshold ω_1. After that, the state becomes $[0, b_a + (1-b_a)p_a, 0]$ and similarly induces a threshold ω_2, which is the same for the remaining process. Note that ω_2 can be different from ω_1 because $b_a + (1-b_a)p_a$ and $b_p + (1-b_p)p_p$ may take different values. Once there is a purchase, the state

becomes $[0,0,1]$ and then jumps to the belief state $[1,0,0]$ after the reward is collected. Starting from the new state $[1,0,0]$, the discussion is the same as the above.

According to Theorem 2, the problem is reduced to computing the optimal thresholds $\boldsymbol{\omega}$ embedded in the triple-threshold policy. Then we define each interval between two adjacent purchases as a "cycle". Given a threshold policy $\varphi^{\boldsymbol{\omega}}$, we use $\pi_k(\boldsymbol{\omega})$ and $\kappa_k(\boldsymbol{\omega})$ to denote the profit and the length of the k-th cycle, respectively. These random variables $\pi_k(\boldsymbol{\omega})$'s and $\kappa_k(\boldsymbol{\omega})$'s are independent and identically distributed within their corresponding groups, resulting in a renewal reward process. Let $\Psi(\boldsymbol{\omega})$ denote the long-term average expected profit under the threshold policy $\varphi^{\boldsymbol{\omega}}$. According to Theorem 3.16 in Ross (2013), we have that:

$$\Psi(\boldsymbol{\omega}) = \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}\left[r\, N_p^{\varphi^{\boldsymbol{\omega}}}(T) - c\, N_a^{\varphi^{\boldsymbol{\omega}}}(T)\right] = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{k=1}^{N(T)} \pi_k(\boldsymbol{\omega})\right] = \frac{\mathbb{E}[\pi(\boldsymbol{\omega})]}{\mathbb{E}[\kappa(\boldsymbol{\omega})]}, \tag{2}$$

where $\pi(\boldsymbol{\omega})$ and $\kappa(\boldsymbol{\omega})$ share the same distributions with $\pi_k(\boldsymbol{\omega})$'s and $\kappa_k(\boldsymbol{\omega})$'s, respectively, and $N(T)$ is the counting process of cycles. Based on the number of ad campaigns within a cycle, we first categorize cycles into the following three types.

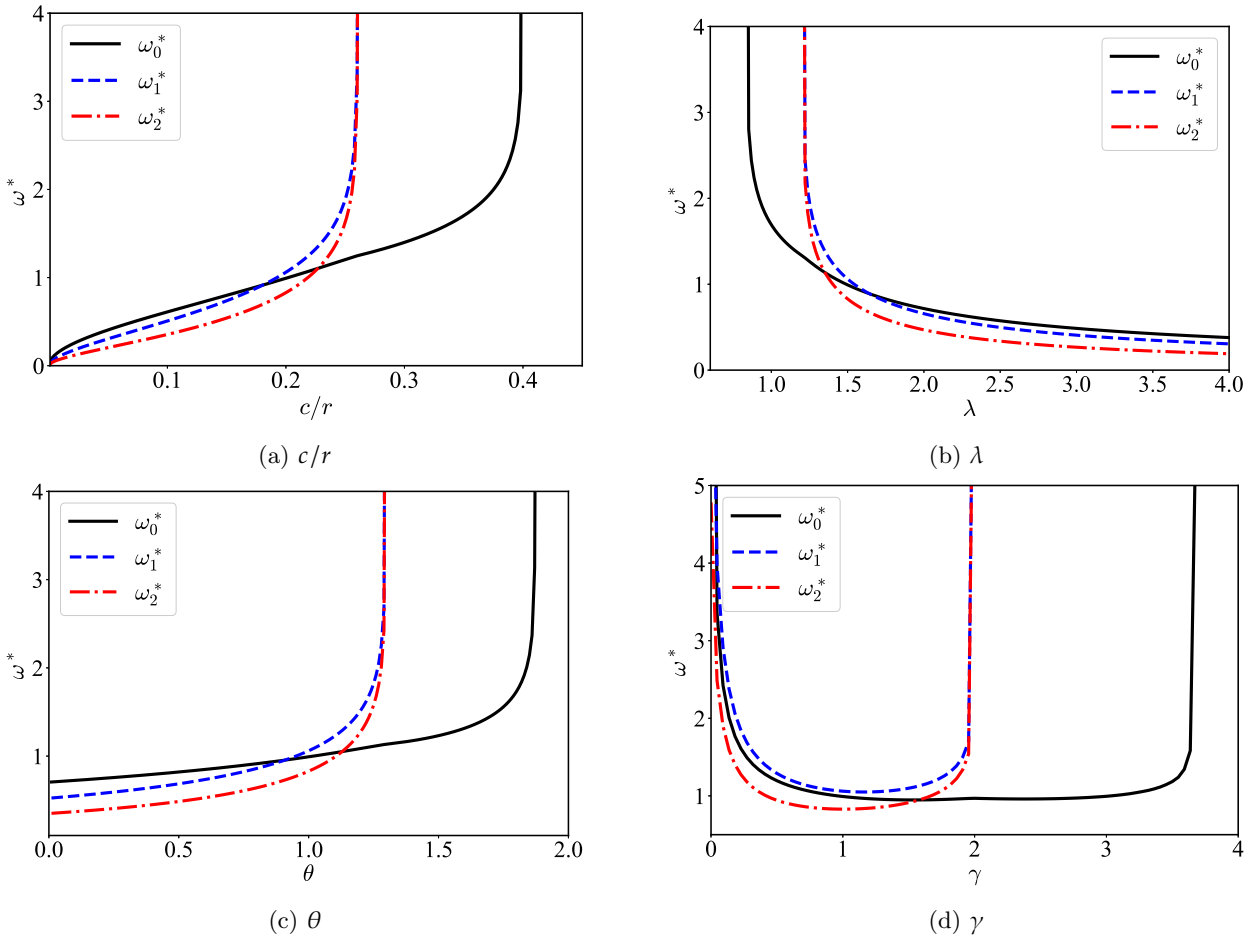
1. *No-ad cycle*: No ad campaign is pushed. In this case, we have $\pi(\boldsymbol{\omega}) = r$ and $\kappa(\boldsymbol{\omega}) < \omega_0$.
2. *Single-ad cycle*: One ad campaign is pushed. In this case, we have $\pi(\boldsymbol{\omega}) = r - c$ and $\kappa(\boldsymbol{\omega}) \in [\omega_0, \omega_0 + \omega_1)$.
3. *Multiple-ad cycle*: At least two ad campaigns are pushed. In this case, if m ads are pushed, then we have $\pi(\boldsymbol{\omega}) = r - mc$ and $\kappa(\boldsymbol{\omega}) \in [\omega_0 + \omega_1 + (m-2)\omega_2, \omega_0 + \omega_1 + (m-1)\omega_2)$.

According to the mapping in Theorem 2 and Figure 3, we can focus on the case where the thresholds $\boldsymbol{\omega}$ can be expressed as functions of b_p and b_a, such that the computation is more tractable. Given the thresholds $\boldsymbol{\omega}$, we can derive the expected cycle profit and cycle length, $\mathbb{E}[\pi(\boldsymbol{\omega})]$ and $\mathbb{E}[\kappa(\boldsymbol{\omega})]$. Then, we can compute the optimal long-term average expected profit $\Psi(\boldsymbol{\omega}) = \mathbb{E}[\pi(\boldsymbol{\omega})]/\mathbb{E}[\kappa(\boldsymbol{\omega})]$ and expand it as an expression of b_p and b_a. Lastly, we solve the optimal belief thresholds (b_p^*, b_a^*) and derive the corresponding optimal thresholds $\boldsymbol{\omega}^*$. For detailed computation, see Appendix A.5.

4. Comparative Analysis of Optimal Thresholds

In this section, we investigate the comparative statics of the optimal thresholds $\boldsymbol{\omega}^*$. To start with, we first numerically illustrate the impacts of several parameters, including the cost-reward ratio c/r, the purchase rate λ, the recapture rate θ and the churn rate γ (these observations will soon be theoretically proved in some special cases). The results are shown in Figure 4.

In the following, we summarize several observations in Figure 4 and then present valuable insights into these results. Before proceeding, we first introduce general insights for the optimal thresholds



(a) c/r (b) λ

(c) θ (d) γ

Figure 4 $\boldsymbol{\omega}^*$ **as functions of different parameters (**$\lambda = 1.5$, $\gamma = 1$, $\theta = 1$, $r = 1$, $c = 0.2$, $p_p = 0.8$ **and** $p_a = 0.5$**).**

$\boldsymbol{\omega}^*$. Given that the expected average profit $\Psi(\boldsymbol{\omega})$ is obtained by dividing the expected cycle profit $\mathbb{E}[\pi(\boldsymbol{\omega})]$ by the expected cycle length $\mathbb{E}[\kappa(\boldsymbol{\omega})]$, we can identify the key factors that significantly influence the threshold policy in the following manner:

1. **Effectiveness**: To maximize the profit contribution of each ad campaign, the seller shall postpone the ad campaigns (i.e., enlarge $\boldsymbol{\omega}$) in order to target inactive customers more accurately. In other words, effectiveness can induce larger expected cycle profit $\mathbb{E}[\pi(\boldsymbol{\omega})]$, thereby increasing the profit $\Psi(\boldsymbol{\omega})$. To be precise, the effectiveness can be conceptually represented as follows:

$$\text{Effectiveness} = \text{Inactive Probability} \times \text{Activation Probability} \times \text{Activation Profit} - \text{Ad Cost}, \quad (3)$$

where the *inactive probability* means the likelihood of the targeted customer being inactive when sent an ad (i.e., $1 - \sum_{i=1}^{3} b_{t,i}$), which increases in the silence time. The *activation probability* can be p_p or p_a due to customers' fatigue to promotion. The *activation profit* measures the profit difference between customers in the active and the inactive states. As the thresholds $\boldsymbol{\omega}$ increase, the effectiveness is enhanced due to the increased inactive probability. Moreover,

due to customers' fatigue to promotion (i.e., $p_p \geq p_a$), the effectiveness corresponding to ω_0^* is greater than that corresponding to ω_1^* and ω_2^*.

2. **Timeliness**: To sustain high customer engagement, the seller has the incentive to advance the ad campaigns (i.e., reduce $\boldsymbol{\omega}$) to activate inactive customers earlier. In other words, timeliness can induce short expected cycle length $\mathbb{E}[\kappa(\boldsymbol{\omega})]$, thereby increasing the average profit $\Psi(\boldsymbol{\omega})$. Apparently, as the threshold $\boldsymbol{\omega}$ increases, the timeliness is diminished.

The optimal threshold $\boldsymbol{\omega}^*$ achieves a balance between the effectiveness and the timeliness, with effectiveness being more significantly impacted by the model parameters compared to timeliness. Consequently, as the effectiveness of ad campaigns is diminished (enhanced, resp.) at any given $\boldsymbol{\omega}$, the optimal threshold $\boldsymbol{\omega}^*$ should increase (decrease, resp.) in order to maintain this balance. Then, we are ready to discuss several observations from Figure 4.

Observation I (Figure 4a): impact of the cost-reward ratio c/r. All thresholds ω_i^*'s increase in the cost-reward ratio c/r and tend to infinity when c/r is sufficiently large. Moreover, as c/r increases, ω_1^* and ω_2^* simultaneously tend to infinity first, and then ω_0^* tends to infinity.

As the cost c increases or the reward r decreases, the effectiveness in (3) is diminished at any given $\boldsymbol{\omega}$, resulting in increased thresholds $\boldsymbol{\omega}^*$. When the cost-reward ratio is larger than some cut-off value, the effectiveness becomes non-positive and hence the corresponding optimal thresholds tend to infinity. Since $p_a \leq p_p$, the effectiveness corresponding to ω_0^* is greater than that corresponding to ω_1^* and ω_2^*, such that ω_1^* and ω_2^* tend to infinity before ω_0^*. This suggests that the seller should push ad campaigns earlier to customers with low ad costs, as it is more effective and cost-efficient in reaching the desired target audience. In recent years, it has been increasingly harder to reach customers due to the rise of ad blockers, resulting in an increase in the ad cost (see The New York Times 2019). As a result, marketing managers may need to evaluate the ad cost and strategically slow down or even abandon advertisements to high-cost customers. By doing so, they can potentially save costs while still activating their target audience effectively.

Observation II (Figure 4b): impact of the purchase rate λ. All thresholds ω_i^*'s decrease in the purchase rate λ and tend to infinity when λ is sufficiently small. Moreover, as λ decreases, ω_1^* and ω_2^* simultaneously tend to infinity first, and then ω_0^* tends to infinity.

As the purchase rate λ increases, both the activation profit and the inactive probability increase because an active customer makes more purchases and the same silence time implies a higher inactive probability. In this case, the effectiveness of ad campaigns (as measured by (3)) is enhanced at any given $\boldsymbol{\omega}$, leading to decreased thresholds $\boldsymbol{\omega}^*$. That is, the seller should push ad campaigns earlier to customers with high purchase rates. Similar to Observation I, as λ decreases, ω_1^* and ω_2^*

tend to infinity before ω_0^* because $p_p \geq p_a$. As the survey UNCTAD (2020) shows, the COVID-19 pandemic accelerates the shift toward a more digital world, resulting in customers' increased purchase frequency of online shopping. In response to such changes, online sellers shall push ad campaigns earlier in this post-pandemic era.

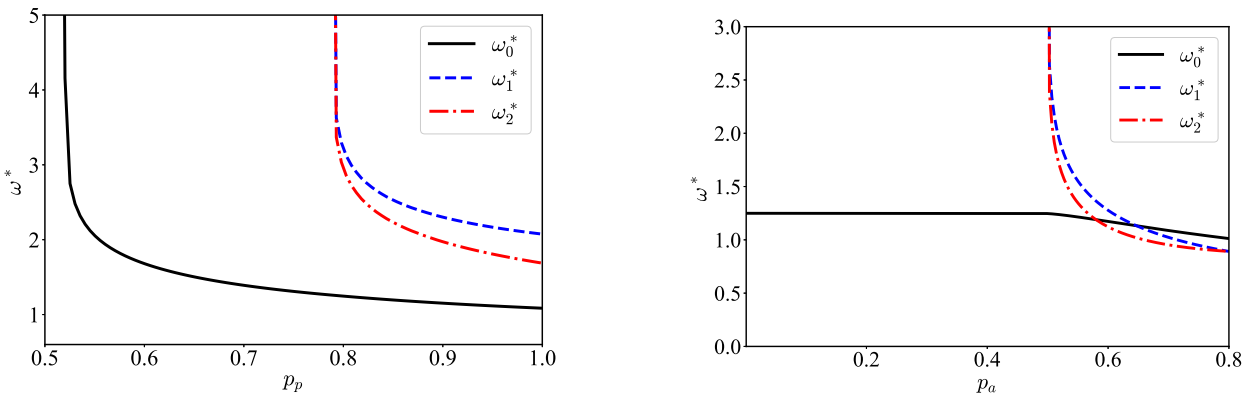
Observation III (Figure 4c): impact of the recapture rate θ. All thresholds ω_i^*'s increase in the recapture rate θ and tend to infinity when θ is sufficiently large. Moreover, as θ increases, ω_1^* and ω_2^* simultaneously tend to infinity first, and then ω_0^* tends to infinity.

As the recapture rate θ increases, inactive customers spontaneously become active after a shorter time, resulting in lower inactive probability and lower activation profit at any given $\boldsymbol{\omega}$. In this case, the effectiveness is diminished and hence the optimal thresholds increase in θ. Therefore, the seller should push ads later to customers with high recapture rates. Moreover, when θ is sufficiently large, the effectiveness is non-positive and hence it is not worthwhile to push ads.

Observation IV (Figure 4d): impact of the churn rate γ. All thresholds ω_i^*'s first decrease and then increase in the churn rate γ and tend to infinity when γ is either sufficiently large or small. Moreover, as γ increases, ω_1^* and ω_2^* simultaneously tend to infinity first, and then ω_0^* tends to infinity.

As the churn rate γ increases, active customers transition to the inactive state after a shorter time, resulting in higher inactive probability but lower activation profit at any given $\boldsymbol{\omega}$. When γ is small, the customer will stay in the active state with a high probability, resulting in a minuscule inactive probability. When γ is large, an active customer will transition to the inactive state soon, resulting in a minuscule activation profit. In both cases, the effectiveness is minuscule such that the optimal threshold is large and even infinity. In contrast, when γ is intermediate, both the inactive probability and the activation profit are away from zero, resulting in a relatively significant effectiveness (and hence a small optimal threshold). This implies that sellers should mainly focus on customers with intermediate churn rates. Numerous marketing surveys (see, e.g., KPMG 2020 and Deloitte 2021) have revealed substantial impacts of COVID-19 on customer behaviors, which further induce noticeable shifts in churn rates and recapture rates. For instance, TechSee (2022) reveals a noteworthy decline in product-dependent brand loyalty and an increase in churn events when compared to the pre-COVID era. In light of such changes, sellers are suggested to re-evaluate customers' behavioral parameters and adapt their policies accordingly.

Then, we investigate the impacts of the activation probabilities (p_p and p_a). We illustrate the numerical results in Figure 5 when c is relatively large, and then discuss some observations. Moreover, if c is extremely small, there could be some additional observations due to the subtle correlations between the three thresholds and we provide more discussions in Appendix EC.2.1.



(a) p_p (b) p_a

Figure 5 ω^* **as functions of activation probabilities** p_p **and** p_a **(**$\lambda = 1.5$, $\gamma = 1$, $\theta = 1$, $r = 1$, $c = 0.26$, $p_p = 0.8$ **and** $p_a = 0.5$**).**

Observation V (Figure 5a): impact of the activation probability p_p for non-fatigued customers. All thresholds ω_i^*'s decrease in p_p and tend to infinity when p_p is sufficiently small. Moreover, as p_p decreases, ω_1^* and ω_2^* simultaneously tend to infinity first, and then ω_0^* tends to infinity.

As p_p increases, the enhanced effectiveness of ad campaigns for non-fatigued customers leads to a reduction in the optimal threshold ω_0^* for non-fatigued customers. Meanwhile, it also enlarges the difference between active customers (without fatigue) and inactive customers, i.e., the activation profit of ad campaigns for fatigued customers. In other words, the effectiveness of ad campaigns for fatigued customers is enhanced by the spillover effect from ad campaigns for non-fatigued customers. Thus, ω_1^* and ω_2^* also decrease in p_p. Since the spillover effect is limited, as p_p decreases, ω_1^* and ω_2^* tend to infinity before ω_0^*.

Observation VI (Figure 5b): impact of the activation probability p_a for fatigued customers. All thresholds ω_i^* decrease in p_a, and ω_1^* and ω_2^* tend to infinity when p_a is sufficiently small.

Similar to the discussion for Observation V, as p_a increases, the effectiveness of ad campaigns for fatigued (non-fatigued, resp.) customers is enhanced due to the increased activation probability (spillover effect, resp.). When p_a is sufficiently small, the effectiveness of ad campaigns for fatigued customers is non-positive and hence ω_1^* and ω_2^* tend to infinity.

Next, we verify the observations in some special cases. In particular, we establish structural properties of the optimal thresholds $\boldsymbol{\omega}^*$ in the following proposition.

PROPOSITION 1. *If* $p_p = p_a$, *then* $\omega_0^* \geq \omega_1^* = \omega_2^*$. *If* $p_p = p_a = 1$, *then* $\omega_0^* = \omega_1^* = \omega_2^*$.

According to Proposition 1, the triple-threshold policy degenerates to a single-threshold policy when $p_p = p_a = 1$, which is referred to as the "*efficient-activation case*". In this case, the formulas are significantly simplified so that analytical results can be derived. The computation details are relegated to Appendix EC.3.1 and the objective function is as follows:

$$\Psi^E(\omega) = \frac{\lambda\theta\left(2\varsigma r - (c+r)e^{-\frac{1}{2}\omega(\varsigma+\gamma+\lambda+\theta)}\left((\gamma-\lambda+\theta)e^{\omega\varsigma} + \varsigma e^{\omega\varsigma} + \varsigma - \gamma + \lambda - \theta\right)\right)}{e^{-\frac{1}{2}\omega(\gamma+\lambda+\theta)}\left(-2\left(\gamma^2+\gamma(\lambda+2\theta)+\theta(\theta-\lambda)\right)\sinh\left(\frac{\omega\varsigma}{2}\right) - 2\varsigma\left((\gamma+\theta)\cosh\left(\frac{\omega\varsigma}{2}\right)\right)\right) + 2\varsigma(\gamma+\theta)}, \tag{4}$$

where $\varsigma = \sqrt{(\lambda+\gamma+\theta)^2 - 4\lambda\theta}$ is defined to simplify the formula, and $\sinh(\cdot)$ and $\cosh(\cdot)$ are the hyperbolic sine and cosine functions, respectively. Let $\omega^* = \arg\max_{\omega\geq 0}\Psi^E(\omega)$ denote the optimal threshold. In this case, we are able to theoretically prove Observations I-IV in the following proposition.

PROPOSITION 2. *For the efficient-activation case, the analytical properties of the optimal threshold ω^* are as follows (the detailed formulas of α, $\tilde{\lambda}$, $\tilde{\theta}$, $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ are provided in Appendix EC.1):*

- *When $c/r \in (0, \alpha)$, ω^* increases in c and decreases in r; otherwise, ω^* tends to infinity.*
- *When $\lambda \in (\tilde{\lambda}, \infty)$, ω^* decreases in λ; otherwise, ω^* tends to infinity.*
- *When $\theta \in (0, \tilde{\theta})$, ω^* increases in θ; otherwise, ω^* tends to infinity.*
- *When $\gamma \in (\tilde{\gamma}_1, \tilde{\gamma}_2)$ and $\theta = 0$, ω^* first decreases and then increases in γ; otherwise, ω^* tends to infinity.*

According to Proposition 2, Observations I-IV are theoretically verified for the efficient-activation case. Note that even in this special case, it is nontrivial to derive the theoretical results due to the complicated expression of $\Psi^E(\omega)$ in (4). Moreover, if we consider the efficient-activation case with $\theta = 0$, i.e., no recapture events, then there exist closed-form expressions for the optimal threshold and the corresponding profit (see Appendix EC.3.5), which can be useful for practical deployment.

5. Ad Push Policy with Budget Constraints

In Section 3, we analyzed the seller's optimal ad push policy for a representative customer without constraints. In practice, the seller usually sets a budget constraint for each cluster of customers (e.g., customers in the same region) over a finite horizon (e.g., a month or a quarter). In this case, decisions for a cluster of customers are coupled and we focus on a representative cluster, which is a large pool of heterogeneous customers, each with a different set of parameters. Specifically, suppose there are M customers and customer j is characterized by the parameter set $(c_j, r_j, \lambda_j, \gamma_j, \theta_j, p_{p,j}, p_{a,j})$.

In the following, we slightly abuse notations. Let $\boldsymbol{x}_t \in \mathbb{X}^M$ denote the state vector of all customers, $\boldsymbol{y}_t \in \mathbb{Y}^M$ denote the purchase vector, and $\boldsymbol{a}_t \in \mathbb{A}^M$ denote the action vector. Similar to the base model, we assume that the initial state $\boldsymbol{x}_0$ satisfies $x_{0,j} = A_p$ for all j. Then, the observed history

at time t is $h_t = \{\boldsymbol{x}_0\} \cup \{(\boldsymbol{y}_\ell, \boldsymbol{a}_\ell)\}_{\ell\in[0,t)} \cup \{\boldsymbol{y}_t\}$. An admissible policy μ is defined as a sequence $\mu = \{\mu_t\}_{t\geq 0}$ where each $\mu_t : h_t \to [0,1]^M$ is a mapping from the observed history to a probability vector of pushing ads. Given any policy μ, we use $N^\mu_{p,j}$ and $N^\mu_{a,j}$ to denote the counting processes of purchases and ad campaigns of customer j. As mentioned, the seller has a budget constraint for the total ad cost $\sum_{j=1}^{M} c_j \cdot N^\mu_{a,j}(T) \leq B \cdot T$, where B measures the average budget per unit of time. Thus, the seller chooses an ad push policy to maximize its average expected profit as follows:

$$\begin{aligned} V(T) = \max_{\mu\in\Pi_T} \quad & \frac{1}{T}\mathbb{E}\left[\sum_{j=1}^{M}\left(r_j N^\mu_{p,j}(T) - c_j N^\mu_{a,j}(T)\right)\right] \\ \text{s.t.} \quad & \sum_{j=1}^{M} c_j N^\mu_{a,j}(T) \leq B \cdot T \qquad (a.s.), \end{aligned} \tag{5}$$

where Π_T is the set of all admissible policies for the problem over the time horizon T.

For a continuous-time partially observable MDP with constraints, efficiently deriving the exact optimal policy, or even a near-optimal policy, is in general intractable. In the following, we aim to provide a heuristic policy that achieves a certain performance guarantee. Specifically, we first analyze a relaxed problem and then design an asymptotically optimal policy based on the optimal solution to the relaxed problem.

First, we define a relaxed problem $\bar{V}(T)$ with the almost-sure budget constraint replaced by an expectation constraint as follows:

$$\begin{aligned} \bar{V}(T) = \max_{\mu\in\Pi_T} \quad & \frac{1}{T}\mathbb{E}\left[\sum_{j=1}^{M}\left(r_j N^\mu_{p,j}(T) - c_j N^\mu_{a,j}(T)\right)\right] \\ \text{s.t.} \quad & \mathbb{E}\left[\sum_{j=1}^{M} c_j N^\mu_{a,j}(T)\right] \leq B \cdot T. \end{aligned} \tag{6}$$

The following lemma describes the relationship between $V(T)$ and $\bar{V}(T)$.

Lemma 1. *The optimal objective of (6) is an upper bound of that of (5), i.e., $\bar{V}(T) \geq V(T)$.*

According to Lemma 1, the relaxed problem can serve as an upper bound of the initial problem. However, problem (6) is still intractable. In the following, we modify the time horizon in the relaxed problem to infinity and the resulting problem is referred to as the "infinite-horizon relaxed problem". Then, we demonstrate that its optimal solution exhibits some favorable structural properties, allowing for an efficient optimization algorithm. Building on this finding, we introduce a simple yet asymptotically optimal heuristic policy for the initial problem (5). Our discussion begins by examining the infinite-horizon relaxed problem.

5.1. Infinite-Horizon Relaxed Problem

In this section, we study the infinite-horizon version of the relaxed problem $\bar{V}(T)$. Specifically, the infinite-horizon relaxed problem is as follows:

$$\begin{aligned} \bar{V}_\infty = \max_{\mu\in\Pi_\infty} \quad & \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{j=1}^{M}\left(r_j N^{\mu}_{p,j}(T) - c_j N^{\mu}_{a,j}(T)\right)\right] \\ \text{s.t.} \quad & \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{j=1}^{M} c_j \cdot N^{\mu}_{a,j}(T)\right] \le B, \end{aligned} \tag{7}$$

where Π_∞ is the set of all admissible policies for the infinite-horizon problem.

Similar to the base model, we first define a policy that applies a triple-threshold policy to each individual customer. We call such a policy a "decentralized triple-threshold policy." We first have the following result.

Proposition 3. *For (7), there exists an optimal policy that is a decentralized triple-threshold policy.*

According to Proposition 3, the seller only needs to determine the optimal thresholds $\boldsymbol{\omega}_j$ for each customer j. In this case, the infinite-horizon problem (7) can be rewritten as the following format (see Theorem 3.16 in Ross 2013):

$$\begin{aligned} \bar{V}_\infty = \max_{\boldsymbol{\omega}_1,\boldsymbol{\omega}_2,\ldots,\boldsymbol{\omega}_M\ge 0} \quad & \sum_{j=1}^{M} \frac{\mathbb{E}[\pi_j(\boldsymbol{\omega}_j)]}{\mathbb{E}[\kappa_j(\boldsymbol{\omega}_j)]} \\ \text{s.t.} \quad & \sum_{j=1}^{M} \frac{\mathbb{E}[\psi_j(\boldsymbol{\omega}_j)]}{\mathbb{E}[\kappa_j(\boldsymbol{\omega}_j)]} \le B, \end{aligned} \tag{8}$$

where $\psi_j(\boldsymbol{\omega}_j) = r_j - \pi_j(\boldsymbol{\omega}_j)$ is the random ad cost of each cycle of customer j. Let $\boldsymbol{\omega}^u_j$ denote the unconstrained optimal threshold derived in the base model and $q_j(\boldsymbol{\omega}) = \mathbb{E}[\psi_j(\boldsymbol{\omega}_j)]/\mathbb{E}[\kappa_j(\boldsymbol{\omega}_j)]$ denote the budget consumption of customer j under the triple-threshold policy $\varphi^{\boldsymbol{\omega}_j}$, respectively. Then, the property of the optimal solution to (8) is summarized in the following lemma.

Lemma 2 **(Optimal Threshold Policy under Budget Constraints)**.

1. *When* $B \ge \sum_{j=1}^{M} q_j(\boldsymbol{\omega}^u_j)$, *the optimal solution to (8) is* $\boldsymbol{\omega}^*_j = \boldsymbol{\omega}^u_j$ *for any* j.
2. *When* $B < \sum_{j=1}^{M} q_j(\boldsymbol{\omega}^u_j)$, *the optimal solution to (8) satisfies that* $\sum_{j=1}^{M} q_j(\boldsymbol{\omega}^*_j) = B$ *and there exists a constant* $\ell > 0$ *such that the optimal threshold for customer* j *can be solved from (2) by plugging in the parameter set* $((1+\ell)c_j, r_j, \lambda_j, \gamma_j, \theta_j, p_{p,j}, p_{a,j})$.

According to Lemma 2, when the budget is sufficient, the optimal policy is to implement the unconstrained optimal policy for each customer. Otherwise, there is a penalty factor ℓ capturing the scarcity of the budget, and the optimal threshold for each customer should be solved from a cost-penalized problem. In this case, we can find the optimal policy by a 1-D search of the penalty factor ℓ.

In the following subsection, we propose a heuristic policy for (5) based on the optimal policy of (8), and then prove its asymptotic optimality in the following subsection.

5.2. Budget Allocation with Thresholds (BAT) Policy

In this subsection, based on the optimal solution to (8), we construct a heuristic policy for the finite-horizon problem (5) under the almost-sure constraint, named the Budget Allocation with Thresholds (BAT) policy. The details of the policy are given in Algorithm 1.

Algorithm 1 Budget allocation with thresholds (BAT) policy

Solving (8) to obtain the optimal thresholds and the corresponding budget consumption $(\boldsymbol{\omega}_j^*, q_j^* = q_j(\boldsymbol{\omega}_j^*))$ for each customer j;

Allocate $Q_j = q_j^* T$ budget to each customer j;

Apply the triple-threshold policy $\varphi^{\boldsymbol{\omega}_j^*}$ to customer j until the allocated budget Q_j is exhausted.

With the BAT policy, the seller is only required to solve the optimization problem (8) at the beginning. Afterward, the seller can assign a budget and separately implement a triple-threshold policy for each customer until the corresponding budget runs out. Compared to solving the exact optimal solution to $\bar{V}(T)$, such a simple policy does not suffer from the curse of dimensionality and is easy to implement in practice. In addition to this advantage, in the following proposition, we provide an upper bound for the optimality gap (which may also be referred to as "regret") of the BAT policy (denoted by $\mathcal{A}$).

THEOREM 3. *Let $V^{\mathcal{A}}(T)$ denote the average expected profit of the BAT policy over the time horizon T. Then, we have:*

$$V(T) - V^{\mathcal{A}}(T) \le \bar{V}(T) - V^{\mathcal{A}}(T) = O\left(\sqrt{\frac{\log T}{T}}\right).$$

Moreover, for the efficient-activation case (where $p_p = p_a = 1$), we have

$$V(T) - V^{\mathcal{A}}(T) \le \bar{V}(T) - V^{\mathcal{A}}(T) = O\left(\sqrt{\frac{1}{T}}\right).$$

According to Theorem 3, the BAT policy is asymptotically optimal as the time horizon tends to infinity. The detailed proof of Theorem 3 is relegated to Appendix A.9. Here, we provide a sketch of the proof. First, the inequality is direct from Lemma 1, and hence we can focus on the gap $\bar{V}(T) - V^{\mathcal{A}}(T)$. Then, we decompose the optimality gap into two components:

$$V(T) - V^{\mathcal{A}}(T) \le \bar{V}(T) - V^{\mathcal{A}}(T) = \underbrace{\bar{V}(T) - \bar{V}_\infty}_{\Delta_1} + \underbrace{\bar{V}_\infty - V^{\mathcal{A}}(T)}_{\Delta_2}. \tag{9}$$

Before proceeding, we would like to highlight several technical features that are different from traditional NRM literature. First, the BAT policy is directly associated with the infinite-horizon relaxed problem $\bar{V}_\infty$ rather than the finite-horizon upper bound $\bar{V}(T)$. Thus, we need to provide an upper bound for Δ_1, a term that is absent in traditional NRM literature. Second, due to the incorporation of intertemporal correlation and partial information, we analyze Δ_2 by cycles defined in the base model. These features entail special treatments in our analysis.

In (9), the first term Δ_1 measures the difference between the optimal values of the finite-horizon and the infinite-horizon problems. We can first use Lagrangian relaxation to simplify Δ_1 such that the budget constraints can be ignored. In this case, we use μ_T^* to denote the optimal policy for $\bar{V}(T)$, and construct a feasible infinite-horizon policy for $\bar{V}_\infty$. For each of exposition, we illustrate the constructed policy in Figure 6.

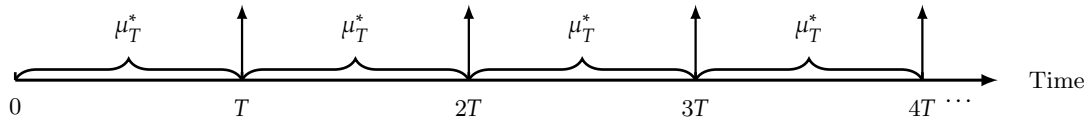


Figure 6 **Illustration of constructed infinite-horizon policy.**

Specifically, as Figure 6 shows, we consider a feasible infinite-horizon policy μ_T^∞ as follows: For each $k \in \mathbb{Z}_+$, the policy μ_T^∞ implements the policy μ_T^* during the interval $((k-1)T, kT)$ and utilizes an "efficient-activation" action at time kT. In this case, the profit difference between μ_T^* and μ_T^∞ only happens at the time points $\{kT\}_{k=0,1,\dots}$. For each customer j, we can prove that each efficient-activation action's contribution to the total profit is upper bounded by a constant $\tilde{C}_j$, which is independent of T. Then, we can deduce the following inequality:

$$\Delta_1 = \bar{V}(T) - \bar{V}_\infty \le \bar{V}(T) - \bar{V}_\infty^{\mu_T^\infty} = \sum_{j=1}^{M} \frac{\tilde{C}_j}{T} = O(1/T).$$

The second term Δ_2 measures the average profit difference between the infinite-horizon relaxed problem and the BAT policy. In this case, we analyze the problem by cycles, whose cycle profits and cycle lengths are i.i.d. random variables. To provide a lower bound for $V^{\mathcal{A}}(T)$, we consider

$n_j = \lfloor T/\mathbb{E}[\kappa(\boldsymbol{\omega}_j^*)] - \epsilon_j \rfloor$ cycles of customer j, where $\epsilon_j = O(\sqrt{T\log T})$. Then, we can use the Chernoff bound to prove that the following conditions (we call an event a good event if it satisfies these conditions) are satisfied with probability no less than $1-2/T$:

1. The system does not terminate before the end of the n_j-th cycle.
2. The budget is not exhausted at the end of the n_j-th cycle.
3. The total reward during the first n_j cycles is $n_j r_j$.
4. The total ad cost is no greater than $q_j^* T$ due to allocated budget.

Under the good event, during the time horizon T, the total reward is no less than $n_j r_j$ and the total ad cost is no greater than $q_j^* T$. Then, the regret incurred under the good event is no greater than

$$\left(1-\frac{2}{T}\right)\left(\frac{\mathbb{E}\left[\pi_j(\boldsymbol{\omega}_j^*)\right]}{\mathbb{E}\left[\kappa_j(\boldsymbol{\omega}_j^*)\right]}-\frac{1}{T}(n_j r_j - q_j^* T)\right) \le \frac{r_j}{\mathbb{E}\left[\kappa_j(\boldsymbol{\omega}_j^*)\right]} - \frac{1}{T}\left(\frac{T}{\mathbb{E}\left[\kappa(\boldsymbol{\omega}_j^*)\right]} - \epsilon_j - 1\right) r_j = O\left(\sqrt{\frac{\log T}{T}}\right),$$

where the inequality is because $1-2/T \le 1$, $q_j^* = q_j(\boldsymbol{\omega}_j^*) = \mathbb{E}\left[r_j - \pi_j(\boldsymbol{\omega}_j^*)\right]/\mathbb{E}\left[\kappa_j(\boldsymbol{\omega}_j)\right]$ and $n_j \ge T/\mathbb{E}\left[\kappa(\boldsymbol{\omega}_j^*)\right] - \epsilon_j - 1$. As for the bad event (which means an event in which some of conditions 1-4 do not hold), the regret incurred under the good event is no greater than $\frac{2}{T}\mathbb{E}\left[\pi_j(\boldsymbol{\omega}_j^*)\right]/\mathbb{E}\left[\kappa_j(\boldsymbol{\omega}_j^*)\right] = O\left(\frac{1}{T}\right)$. Thus, we have

$$\Delta_2 \le O\left(\sqrt{\frac{\log T}{T}}\right) + O\left(\frac{1}{T}\right) = O\left(\sqrt{\frac{\log T}{T}}\right).$$

Finally, combining the above two terms, we derive the order for (9) as

$$\bar{V}(T) - V^{\mathcal{A}}(T) \le \Delta_1 + \Delta_2 \le O(\frac{1}{T}) + O\left(\sqrt{\frac{\log T}{T}}\right) = O\left(\sqrt{\frac{\log T}{T}}\right),$$

which finishes the proof for the general case. In the following, we provide a better upper bound for the regret under the efficient-activation case where $p_p = p_a = 1$. Under the efficient-activation case, an ad campaign can activate customers with certainty and hence the intertemporal correlation can be stopped once there is an ad campaign. Such a weaker correlation leads to a better upper bound.

In this case, the upper bound for Δ_1 is still $O(1/T)$. As for Δ_2, we further decompose it as follows:

$$\Delta_2 = \bar{V}_\infty - V^{\mathcal{A}}(T) = \left(\bar{V}_\infty - R^{\mathcal{A}_B}(T)\right) + \left(R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T)\right),$$

where $\mathcal{A}_B$ is the policy that implements the decentralized threshold policy in $\mathcal{A}$ over the time horizon T without the budget constraint, and $R^{\mathcal{A}_B}(T)$ is the average expected profit under such a policy. First, similar to the analysis for Δ_1, we can construct an infinite-horizon policy based on $\mathcal{A}_B$ and then show that $\bar{V}_\infty - R^{\mathcal{A}_B}(T) \le O(1/T)$.

Second, the profit difference $R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T)$ is due to the additional ad campaigns delivered by the policy $\mathcal{A}_B$. Since actions on different customers are independent, we can focus on a representative customer and temporarily omit the subscript j. In this case, since an ad campaign can activate customers with certainty, we define each interval between two adjacent ad campaigns as an "ad cycle". Let ν_k denote the random cycle length of the k-th ad cycle. Similar to Section 4, ν_k's are i.i.d. variables. Since each ad cycle corresponds to an ad campaign, the remaining budget is less than c after time $T_A = \sum_{k=1}^{\lfloor qT/c \rfloor} \nu_k$, and there are at most $\frac{(T-T_A)^+}{\omega^*}$ additional ad campaigns. As we shown in the analysis for Δ_1, the profit contribution of an ad campaign is upper bounded by a constant $\tilde{C}$. Then, by Proposition 1 in Gallego (1992), we can provide an upper bound for the average profit difference as

$$\frac{1}{T}\mathbb{E}\left[\frac{(T-T_A)^+\tilde{C}}{\omega^*}\right] = \frac{\tilde{C}}{T\omega^*} \times \mathbb{E}[(T-T_A)^+] \le \frac{\tilde{C}}{T\omega^*} \times \frac{\sqrt{Var(T_A)} + 2\mathbb{E}[v_k]}{2} = O\left(\frac{1}{\sqrt{T}}\right),$$

where $Var(T_A)$ is the variance of T_A and its order is $O(T)$. Then, we can derive that $R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T) \le O(1/\sqrt{T})$.

To summarize, we can deduce that

$$\Delta_2 = \left(\bar{V}_\infty - R^{\mathcal{A}_B}(T)\right) + \left(R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T)\right) = O(1/T) + O(1/\sqrt{T}) = O(1/\sqrt{T}),$$

which implies that $V(T) - V^{\mathcal{A}}(T) \le \Delta_1 + \Delta_2 = O(1/\sqrt{T})$ for the efficient-activation case.

6. Extension: Strategic Customers

In the base model, customers are assumed to be myopic. That is, they make purchases according to an exogenous Poisson process when they are in the active state. This assumption is reasonable if the ad campaign does not provide additional benefits to the customer. However, in practice, some ad campaigns may come with a benefit for the customer (e.g., Starbucks may send a push with a free upsize coupon and customers can have extra gains when purchasing with the coupon). In such a case, customers may have an incentive to determine their strategies in response to the seller's policy if they find it beneficial to do so. In this part, we incorporate such customer strategic behaviors and analyze the optimal policy. Specifically, an active customer raises purchase impulses according to a Poisson process with rate λ. When a purchase impulse comes up, the customer can choose to purchase immediately, which gives him an immediate surplus $u - r$ (where r is the price disutility). However, a strategic customer may choose to hold the impulse and wait for a certain amount of time in order to receive an ad campaign with benefits.[3] We assume the customer has a

[3] The strategic customer is also assumed to hold one unit of demand during the waiting period. In reality, even after strategic waiting, customers typically only purchase one unit of the commodity (e.g., a cup of coffee).

time discount factor δ and there is an additional utility u_a if he purchases with the given benefit (we also assume there is no additional cost for the seller since for many industries such as coffee and video games, the variable cost of providing service is negligible compared to the revenue).[4] Then, if he waits for t time units and receives an ad campaign with benefits, then the utility is $e^{-\delta t}(u+u_a-r)$. Therefore, a customer determines whether to wait for the ad campaign that arrives after t time units according to the following problem:

$$\max\{\underbrace{u-r,}_{\text{Do Not Wait}}\ \underbrace{e^{-\delta t}(u+u_a-r)}_{\text{Wait}}\}. \tag{10}$$

When a tie happens in (10), we assume that the customer does not wait. From (10), we can solve for the maximal time the customer will wait as $\xi := \frac{1}{\delta}\ln\left(1+\frac{u_a}{u-r}\right)$. Given a triple-threshold policy $\varphi^{\boldsymbol{\omega}}$, the strategic behaviors of customers with different impulse generation times are illustrated in Figure 7.

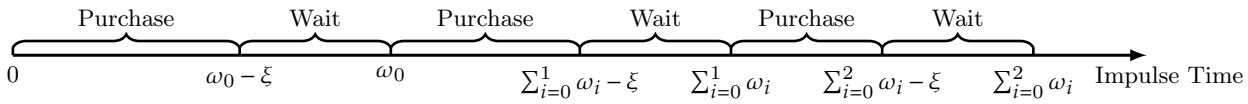


Figure 7 **Illustration of strategic customers' behaviors with different impulse generation times under triple-threshold policy $\varphi^{\boldsymbol{\omega}}$ with $\boldsymbol{b}_0=[1,0,0]$.**

According to Figure 7, as an impulse is generated and the current threshold is ω_i, if the silence time η is greater than $\omega_i-\xi$, then the customer will wait until an ad is pushed rather than making a purchase instantly. Then, we compute the long-term average expected profit under triple-threshold policies. Similar to the base model, we still define each interval between two adjacent purchases as a "cycle", and use $\pi^S(\boldsymbol{\omega})$ and $\kappa^S(\boldsymbol{\omega})$ to denote the random cycle profit and cycle length under a triple-threshold policy $\varphi^{\boldsymbol{\omega}}$. Then, according to Figure 7, the cycles can be categorized into the following three types.

1. *No-ad cycle with direct purchase* (applicable when $\omega_0 \geq \xi$): No ad campaign has been pushed since the last purchase, and the purchase impulse is generated at least ξ time units before the next ad campaign time. In this case, $\pi^S(\boldsymbol{\omega})=r$ and $\kappa^S(\boldsymbol{\omega})\in[0,\omega_0-\xi]$.

[4] We assume that there is no time discount for the seller. That is, the seller aims to maximize the long-term revenue. Such discrepancy in discount factor is reasonable and has been adopted in prior literature. In particular, the customer's discount is due to the delay of satisfaction, which is usually significant over time. In contrast, the seller's discount is mainly due to the discount in money's value, which is usually much less significant. Such treatments of discount factors (different discount factors between seller and customer) can be found in many previous works, e.g., Amin et al. (2013), Mohri and Munoz (2015) and Chen et al. (2022b).

2. *Single-ad cycle with waiting*: No ad campaign has been pushed since the last purchase. After the purchase impulse is generated, the customer waits until the next ad campaign is pushed. In this case, $\pi^S(\boldsymbol{\omega}) = r - c$ and $\kappa^S(\boldsymbol{\omega}) = \omega_0$.
3. *Single-ad cycle with direct purchase* (applicable when $\omega_1 \geq \xi$): One ad campaign has been pushed since the last purchase, and the customer makes a purchase once an impulse is generated. In this case, $\pi^S(\boldsymbol{\omega}) = r - c$ and $\kappa^S(\boldsymbol{\omega}) \in [\omega_0, \omega_0 + \omega_1 - \xi]$.
4. *Double-ad cycle with waiting*: One ad campaign has been pushed since the last purchase. After the purchase impulse is generated, the customer waits until the next ad campaign is pushed. In this case, $\pi^S(\boldsymbol{\omega}) = r - 2c$ and $\kappa^S(\boldsymbol{\omega}) = \omega_0 + \omega_1$.
5. *Multi-ad cycle with direct purchase* (applicable when $\omega_2 \geq \xi$): At least two ad campaigns have been pushed since the last purchase, and the customer makes a purchase once an impulse is generated. In this case, if $m \geq 2$ ads have been pushed since the last purchase, then $\pi^S(\boldsymbol{\omega}) = r - mc$ and $\kappa^S(\boldsymbol{\omega}) \in [\omega_0 + \omega_1 + (m-2)\omega_2,\ \omega_0 + \omega_1 + (m-1)\omega_2 - \xi]$.
6. *Multi-ad cycle with waiting*: At least two ad campaigns have been pushed since the last purchase. After the purchase impulse is generated, the customer waits until the next ad campaign is pushed. In this case, if $m \geq 2$ ads have been pushed since the last purchase, then $\pi^S(\boldsymbol{\omega}) = r - (m+1)c$ and $\kappa^S(\boldsymbol{\omega}) = \omega_0 + \omega_1 + (m-1)\omega_2$.

Then, the objective function can be computed as $\Psi^S(\boldsymbol{\omega}) := \mathbb{E}[\pi^S(\boldsymbol{\omega})]/\mathbb{E}[\kappa^S(\boldsymbol{\omega})]$, and then the optimal thresholds $\boldsymbol{\omega}^{S,*}$ can be solved. The detailed computation is relegated to Appendix EC.3.2.

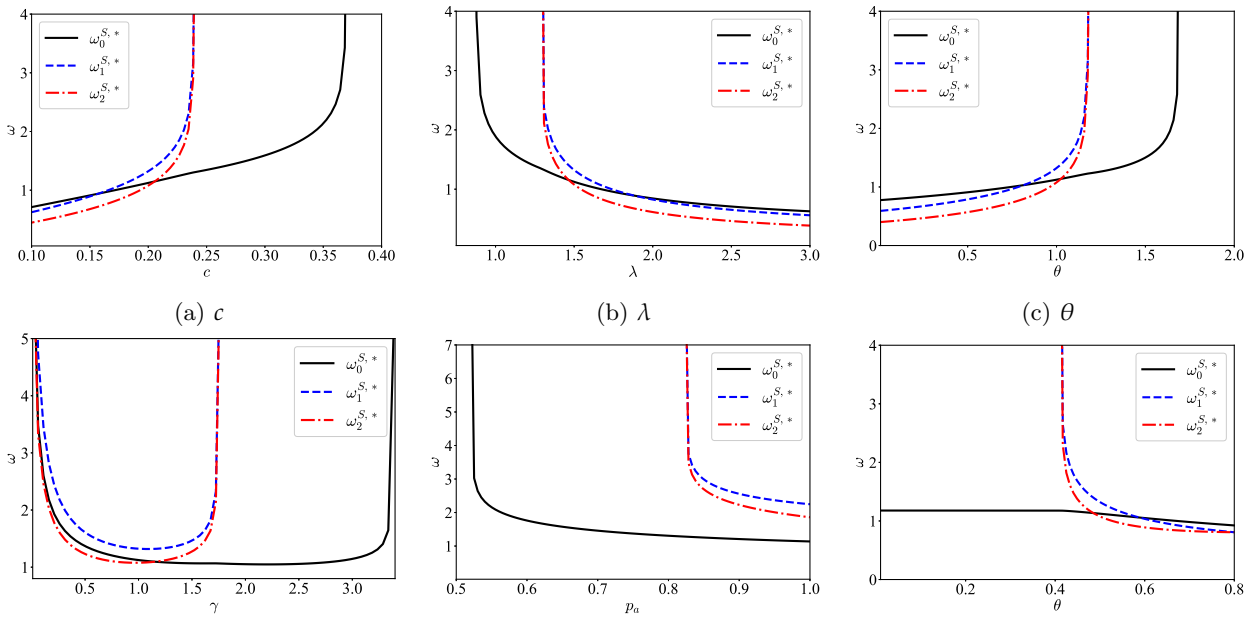
6.1. Unconstrained Case

In this subsection, similar to Section 4, we start with the unconstrained case and investigate the comparative statics of the optimal thresholds $\boldsymbol{\omega}^{S,*}$. First, in Figure 8, we illustrate the impacts of the parameters investigated in Section 4 on the optimal thresholds $\boldsymbol{\omega}^{S,*}$ through numerical experiments. According to Figure 8, Observations I-VI still hold when customers are strategic.[5]

Then, in Figure 9, we illustrate the impacts of the parameters related to the strategic behavior (i.e., δ, u, u_a, and r) through numerical experiments. Then, we discuss some additional observations from Figure 9.

Observation VII (Figures 9a-9c): impact of the discount factor δ and the utilities u and u_a. All thresholds $\omega_i^{S,*}$'s decrease in δ and u, and increase in u_a. To explain, as δ and u increases or as u_a decreases, the maximal waiting time ξ is shortened. This mitigates strategic customer behavior, thereby improving the activation profit. Thus, the effectiveness is enhanced and the optimal thresholds $\omega_i^{S,*}$'s decrease.

[5] Note that Observation I holds when r is fixed.



(a) c (b) λ (c) θ

(d) γ (e) p_p $(c = 0.24)$ (f) p_a

Figure 8 $\omega^{S,*}$ **as functions of different parameters (**$\lambda = 1.5$, $\gamma = 1$, $\theta = 1$, $r = 1$, $c = 0.2$, $p_p = 0.8$, $p_a = 0.5$, $u = 2$, $u_a = 1$ **and** $\delta = 5$**).**

Observation VIII (Figure 9d): impact of the price r. All thresholds $\omega_i^{S,*}$'s first decrease and then increase in r. To explain, as r increases, both the reward r of each purchase and the strategic waiting time ξ increase, resulting in countervailing effects on the activation profit. Recall that $\xi = \frac{1}{\delta}\ln\left(1 + \frac{u_a}{u-r}\right)$. When r is small, the increment in purchase reward is more significant, and the optimal thresholds decrease due to the enhanced effectiveness. When r is large, the impact of the enlarged waiting time is more significant, and hence the optimal thresholds increase.

6.2. Constrained Case

In this subsection, similar to Section 5, we consider M customers and each customer j is characterized by the parameter sets $(c_j, r_j, \lambda_j, \gamma_j, \theta_j, \xi_j, p_{p,j}, p_{a,j})$. Given a time horizon T, the total ad cost budget is $B \cdot T$. Similar to (8), if we focus on triple-threshold policies, the optimal triple-threshold policy for the infinite-time case can be solved from the following problem:

$$\begin{aligned} \bar{V}_\infty^S = \max_{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_M \geq 0} \quad & \sum_{j=1}^{M} \frac{\mathbb{E}[\pi_j^S(\boldsymbol{\omega}_j)]}{\mathbb{E}[\kappa_j^S(\boldsymbol{\omega}_j)]} \\ \text{s.t.} \quad & \sum_{j=1}^{M} \frac{\mathbb{E}[\psi_j^S(\boldsymbol{\omega}_j)]}{\mathbb{E}[\kappa_j^S(\boldsymbol{\omega}_j)]} \leq B, \end{aligned} \tag{11}$$

where $\psi_j^S(\boldsymbol{\omega}_j) = r_j - \pi_j^S(\boldsymbol{\omega})$. Define $q_j^S(\boldsymbol{\omega}) := \mathbb{E}[\psi_j^S(\boldsymbol{\omega}_j)]/\mathbb{E}[\kappa_j^S(\boldsymbol{\omega}_j)]$. Based on the optimal solution to (11), we construct a modified BAT policy in Algorithm 2, which is denoted by $\mathcal{A}_S$.

Algorithm 2 Budget allocation with thresholds policy for strategic customers (BAT-S policy)

Solving (11) to obtain the optimal thresholds and the corresponding budget consumption $(\boldsymbol{\omega}_j^*, q_j^* = q_j^S(\boldsymbol{\omega}_j^*))$ for each customer j;

Allocate $Q_j = q_j^* T$ budget to each customer j;

Apply the triple-threshold policy $\varphi^{\boldsymbol{\omega}_j^*}$ to customer j until the allocated budget Q_j is exhausted.

Let $V^{S,\mathcal{A}_S}(T)$ denote the average expected profit of the BAT-S policy over the time horizon T. In the following proposition, we provide an upper bound for the gap between $\bar{V}_\infty^S$ and $V^{S,\mathcal{A}_S}(T)$.

PROPOSITION 4. *For the general case, we have* $\bar{V}_\infty^S - V^{S,\mathcal{A}_S}(T) = O\left(\sqrt{\log T/T}\right)$. *For the efficient-activation case (where* $p_p = p_a = 1$*), we have* $\bar{V}_\infty^S - V^{S,\mathcal{A}_S}(T) = O\left(\sqrt{1/T}\right)$.

According to Proposition 4, as the time horizon T increases, the average profit under the BAT-S policy converges to that under the optimal triple-threshold policy for the infinite-horizon case.

6.3. Benefit of Randomness

In this subsection, we demonstrate that different from the base model, the triple-threshold policy may not be optimal, i.e., Theorem 2 does not hold in this case. Specifically, in the following, we use an example to show that a randomized policy can achieve a higher profit than that under the optimal triple-threshold policy.

EXAMPLE 1. Consider the case when $\lambda = 3$, $\gamma = 1$, $\theta = 0.5$, $p_p = p_a = 1$, $r = 1$, $c = 0.2$, $u = 2$, $u_a = 0.6$ and $\delta = 0.5$. In this case, given a deterministic threshold, the maximal waiting time is $\xi = 0.94$. Since $p_p = p_a = 1$, the optimal triple-threshold policy degenerates to a single-threshold policy. The optimal deterministic threshold is $\omega^* = 1.74$ and the corresponding long-term average expected profit is $\Psi^S(\omega^*) = 1.30$.

Now we consider a two-point distributed random threshold Ω with $\omega_1^r = 0.93$, $\omega_2^r = 1.87$, $p_1 = 0.33$ and $p_2 = 0.67$. By some calculations, we can verify that the average expected profit of the corresponding policy is $\Psi_r^S(\Omega) = 1.36$, which is about 5% greater than that of the optimal deterministic policy. □

Here, we provide insights into why the randomized policy above can improve the profit. Given the silence time η, each customer determines his action according to the following formula.

$$V(\eta) = \begin{cases} \max\{\underbrace{u-r}_{\text{Do not wait}}, \underbrace{p_1 e^{-\delta(\omega_1^r-\eta)}(u+u_a-r) + p_2 e^{-\delta(\omega_1^r-\eta)} V(\omega_1^r)}_{\text{Wait}}\}, & \eta \in [0, \omega_1^r) \\ \max\{\underbrace{u-r}_{\text{Do not wait}}, \underbrace{e^{-\delta(\omega_2^r-\eta)}(u+u_a-r)}_{\text{Wait}}\}, & \eta \in [\omega_1^r, \omega_2^r). \end{cases} \tag{12}$$

To facilitate understanding, we illustrate customers' strategic behaviors in Figure 10 when $\omega_2^r - \omega_1^r \geq \xi$. As Figure 10 shows, if the purchase impulse comes when $\eta \geq \omega_1^r$ and no ad campaign is

provided, then the customer can infer that an ad campaign will be pushed when $\eta = \omega_2$. Therefore, the decision problem with $\eta \geq \omega_1^r$ in (12) is similar to (10) and hence his maximal waiting time is also ξ (see the interval $[\omega_1^r, \omega_2^r]$ in Figure 10). In Example 1, since $\omega_2^r - \omega_1^r \geq \xi$, an impulse generated when $\eta = \omega_1^r$ will instantaneously induce a purchase and hence $V(\omega_1^r) = u - r$. Then we discuss purchase impulses that come when $\eta < \omega_1^r$. In this case, if the customer waits until $\eta = \omega_1^r$ but the realized threshold is ω_2^r, then the customer will purchase without a benefit and get the utility $u - r < u + u_a - r$. Then we have η^r satisfying

$$u - r = e^{-\delta(\omega_1^r - \eta^r)} \left(u - r + p_1 u_a\right).$$

Since the term in the brackets is smaller than that in (10) (i.e., $u - r + p_1 u_a < u - r + u_a$), the maximal waiting time $\omega_1^r - \eta^r$ before the first threshold is smaller than ξ. In other words, compared with the optimal deterministic threshold policy, under a well-designed randomized policy, the maximal waiting time may decrease due to the diminished value of waiting. Therefore, the randomized policy can combat customers' strategic waiting and hence increase profit.

7. Conclusions and Discussions

In this paper, we examine the optimal ad campaign push policy in the presence of Markov customer dynamics. Each customer's engagement state evolves according to a continuous-time Markov chain and a customer will only make purchases when he is in the active state. The seller cannot observe customers' engagement states but can observe their purchase records. In order to maximize the long-term average expected profit, the seller may push ad campaigns at a cost to activate customers, thus encouraging future purchases. Initially, we focus on the unconstrained scenario and demonstrate that the optimal policy is a triple-threshold policy based on the number of repeated ads and the silence time, namely the amount of time elapsed since the last purchase or ad campaign. Then, we investigate the comparative statics of the optimal thresholds. Through extensive numerical experiments, we find that the optimal thresholds increase in the ad cost, the recapture rate, and the activation probabilities, but decrease in the purchase reward and the purchase rate. Moreover, as the churn rate increases, the optimal thresholds first decrease and then increase. The results are further verified by theoretical results derived under the efficient-activation case where the activation probabilities are set to one.

Subsequently, we study the case with a budget constraint on the overall ad cost over a finite horizon. For such a problem, we provide an easy-to-implement algorithm, the BAT policy, which solves a relaxed problem at the beginning and then implements a triple-threshold policy for each individual customer. Additionally, we prove the optimality gap of the BAT policy to be $O\left(\sqrt{\log T/T}\right)$ for the general problem and $O\left(1/\sqrt{T}\right)$ for the efficient-activation case, implying its asymptotic

optimality as the time horizon T tends to infinity. Moreover, to test the robustness of our results and to enrich the insights, we further incorporate strategic customer behaviors into the base model. Encouragingly, our main results qualitatively hold in this case and some additional insights are derived.

Lastly, we discuss some future research directions. First, the seller may access more data as the information technology advances or fewer data as the privacy regulation tightens. In these cases, the model may be modified to exploit the new information structure. For example, suppose the seller observes customers' history of visiting, which happens according to a Poisson process and brings with a purchase with a certain probability $\hat{p}$. In this case, replacing the reward r with $\hat{p}r$, we can directly use the results in this paper. The second interesting direction is to consider the seller's ability to cultivate customers. In this work, we assume the parameter set $(\lambda, \gamma, \theta, r, c, p_p, p_a)$ is exogenously given and will not change over time. Nevertheless, in practice, customers can be cultivated and the parameter set may vary according to the seller's policy. We believe that our model in this paper has the potential to be extended to incorporate these features.

References

Aflaki S, Popescu I (2014) Managing retention in service relationships. *Management Science* 60(2):415–433.

Amin K, Rostamizadeh A, Syed U (2013) Learning prices for repeated auctions with strategic buyers. *Proceedings of the 27th Annual Conference on Neural Information Processing Systems*, 1169–1177.

Andrews M, Luo X, Fang Z, Ghose A (2016) Mobile ad effectiveness: Hyper-contextual targeting with crowdedness. *Marketing Science* 35(2):218–233.

Aravindakshan A, Naik PA (2015) Understanding the memory effects in pulsing advertising. *Operations Research* 63(1):35–47.

Åström KJ (1969) Optimal control of Markov processes with incomplete state information II: The convexity of the loss function. *Journal of Mathematical Analysis and Applications* 26(2):403–406.

Bai X, Chen X, Stolyar AL (2022) Average cost optimality in partially observable lost-sales inventory systems. *Operations Research* 71(6):2390–2396.

Bastani H, Harsha P, Perakis G, Singhvi D (2022) Learning personalized product recommendations with customer disengagement. *Manufacturing & Service Operations Management* 24(4):2010–2028.

Bemmaor AC, Glady N (2012) Modeling purchasing behavior with sudden "death": A flexible customer lifetime model. *Management Science* 58(5):1012–1021.

Bernstein F, Chakraborty S, Swinney R (2022) Intertemporal content variation with customer learning. *Manufacturing & Service Operations Management* 24(3):1664–1680.

Bertsekas D, Shreve SE (1996) *Stochastic Optimal Control: The Discrete-Time Case*, volume 5 (Athena Scientific).

Brodie RJ, Hollebeek LD, Jurić B, Ilić A (2011) Customer engagement: Conceptual domain, fundamental propositions, and implications for research. *Journal of Service Research* 14(3):252–271.

Brown DB, Smith JE (2020) Index policies and performance bounds for dynamic selection problems. *Management Science* 66(7):3029–3050.

Brown DB, Zhang J (2022) Dynamic programs with shared resources and signals: Dynamic fluid policies and asymptotic optimality. *Operations Research* 70(5):3015–3033.

Brown DB, Zhang J (2023) Fluid policies, reoptimization, and performance guarantees in dynamic resource allocation. *Operations Research* 73(2):1029–1045.

Bumpensanti P, Wang H (2020) A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science* 66(7):2993–3009.

Chen L, Jain R, Luo H (2022a) Learning infinite-horizon average-reward Markov decision process with constraints. *Proceedings of the 39th International Conference on Machine Learning*, 3246–3270.

Chen X, Gao J, Ge D, Wang Z (2022b) Bayesian dynamic learning and pricing with strategic customers. *Production and Operations Management* 31(8):3125–3142.

Cooper WL (2002) Asymptotic behavior of an allocation policy for revenue management. *Operations Research* 50(4):720–727.

Da'u A, Salim N (2020) Recommendation system based on deep learning methods: A systematic review and new directions. *Artificial Intelligence Review* 53(4):2709–2748.

DeCroix G, Long X, Tong J (2021) How service quality variability hurts revenue when customers learn: Implications for dynamic personalized pricing. *Operations Research* 69(3):683–708.

Deloitte (2021) COVID-19 broke the orthodoxies of customer loyalty and retention. URL https://www2.deloitte.com/us/en/pages/consulting/articles/the-orthodoxies-of-loyalty.html.

Ely JC (2017) Beeps. *American Economic Review* 107(1):31–53.

Fader PS, Hardie BG, Lee KL (2005) "Counting your customers" the easy way: An alternative to the Pareto/NBD model. *Marketing Science* 24(2):275–284.

Fader PS, Hardie BG, Shang J (2010) Customer-base analysis in a discrete-time noncontractual setting. *Marketing Science* 29(6):1086–1108.

Feinberg EA, Kasyanov PO, Zadoianchuk NV (2012) Average cost Markov decision processes with weakly continuous transition probabilities. *Mathematics of Operations Research* 37(4):591–607.

Feinberg EA, Kasyanov PO, Zadoianchuk NV (2013) Berge's theorem for noncompact image sets. *Journal of Mathematical Analysis and Applications* 397(1):255–259.

Feinberg EA, Kasyanov PO, Zgurovsky MZ (2016) Partially observable total-cost Markov decision processes with weakly continuous transition probabilities. *Mathematics of Operations Research* 41(2):656–681.

Gallego G (1992) A minmax distribution free procedure for the (Q, R) inventory model. *Operations Research Letters* 11(1):55–60.

Gallego G, Topaloglu H, et al. (2019) *Revenue Management and Pricing Analytics* (Springer).

Ghosh A, Zhou X, Shroff N (2023) Achieving sub-linear regret in infinite horizon average reward constrained MDP with linear function approximation. *Proceedings of the 17th International Conference on Learning Representations*.

Goldfarb A, Tucker C (2011) Online display advertising: Targeting and obtrusiveness. *Marketing Science* 30(3):389–404.

Guo X, Rieder U (2006) Average optimality for continuous-time Markov decision processes in Polish spaces. *The Annals of Applied Probability* 16(2):730–756.

Hahn M, Hyun JS (1991) Advertising cost interactions and the optimality of pulsing. *Management Science* 37(2):157–169.

Huang Y, Jasin S, Manchanda P (2019) "Level up": Leveraging skill and engagement to maximize player game-play in online video games. *Information Systems Research* 30(3):927–947.

Jain D, Singh SS (2002) Customer lifetime value research in marketing: A review and future directions. *Journal of Interactive Marketing* 16(2):34–46.

Jasin S, Kumar S (2012) A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research* 37(2):313–345.

Jiang J (2023) Constant approximation for network revenue management with Markovian-correlated customer arrivals. *arXiv preprint arXiv:2305.05829* .

Kanoria Y, Lobel I, Lu J (2024) Managing customer churn via service mode control. *Mathematics of Operations Research* 49(2):1192–1222.

KPMG (2020) Customer experience in the new reality. URL https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2020/07/customer-experience-in-the-new-reality.pdf.

Lan H, Gallego G, Wang Z, Ye Y (2026) LP-based control for network revenue management under markovian demands. *Proceedings of the 20th Conference on Web and Internet Economics*, 476–493.

Li G, Wang Z, Zhang J (2024) Infrequent resolving algorithm for online linear programming. *arXiv preprint arXiv:2408.00465* .

Mahajan V, Muller E (1986) Advertising pulsing policies for generating awareness for new products. *Marketing Science* 5(2):89–106.

Meta (2024) Customer list formatting guidelines for custom audiences. URL https://www.facebook.com/business/help/2082575038703844.

Mohri M, Munoz A (2015) Revenue optimization against strategic buyers. *Proceedings of the 29th Annual Conference on Neural Information Processing Systems*, 2521–2529.

Neslin SA, Van Heerde HJ, et al. (2009) Promotion dynamics. *Foundations and Trends® in Marketing* 3(4):177–268.

Netzer O, Lattin JM, Srinivasan V (2008) A hidden Markov model of customer relationship dynamics. *Marketing Science* 27(2):185–204.

Norris JR (1998) *Markov Chains*. Number 2 (Cambridge University Press).

Papaspiliopoulos O (2020) *High-Dimensional Probability: An Introduction with Applications in Data Science* (Cambridge University Press).

Portugal I, Alencar P, Cowan D (2018) The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems with Applications* 97:205–227.

Rafieian O, Yoganarasimhan H (2021) Targeting and privacy in mobile advertising. *Marketing Science* 40(2):193–218.

Reiman MI, Wang Q (2008) An asymptotically optimal policy for a quantity-based network revenue management problem. *Mathematics of Operations Research* 33(2):257–282.

Ross SM (2013) *Applied Probability Models with Optimization Applications* (Courier Corporation).

Schmittlein DC, Morrison DG, Colombo R (1987) Counting your customers: Who are they and what will they do next? *Management Science* 33(1):1–24.

Schwartz EM, Bradlow ET, Fader PS (2014) Model selection using database characteristics: Developing a classification tree for longitudinal incidence data. *Marketing Science* 33(2):188–205.

Sheng L, Ryan CT, Nagarajan M, Cheng Y, Tong C (2022) Incentivized actions in freemium games. *Manufacturing & Service Operations Management* 24(1):275–284.

StackAdapt (2023) 5 tips for avoiding ad fatigue. URL https://www.stackadapt.com/resources/blog/ad-fatigue.

Statista (2025) Share of global mobile website traffic. URL https://www.statista.com/statistics/277125/share-of-website-traffic-coming-from-mobile-devices/.

Talluri K, Van Ryzin G (1998) An analysis of bid-price controls for network revenue management. *Management Science* 44(11):1577–1593.

TechSee (2022) New data shows post pandemic impact of customer loyalty resulting in increased likelihood to churn from poor service. URL https://www.prnewswire.com/news-releases/new-data-shows-post-pandemic-impact-of-customer-loyalty-resulting-in-increased-likelihood-to-churn-from-poor-service-301606607.html.

The New York Times (2019) The advertising industry has a problem: People hate ads. URL https://www.nytimes.com/2019/10/28/business/media/advertising-industry-research.html.

UNCTAD (2020) COVID-19 has changed online shopping forever, survey shows. URL https://unctad.org/news/covid-19-has-changed-online-shopping-forever-survey-shows.

Vera A, Banerjee S, Gurvich I (2021) Online allocation and pricing: Constant regret via bellman inequalities. *Operations Research* 69(3):821–840.

VMR (2024) Global mobile engagement market size. URL https://www.verifiedmarketresearch.com/product/mobile-engagement-market/.

Whittle P (1988) Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability* 25(A):287–298.

Zhang Y, Li B, Luo X, Wang X (2019) Personalized mobile targeting with user engagement stages: Combining a structural hidden Markov model and field experiment. *Information Systems Research* 30(3):787–804.

Appendix for "When to Push Ads: Optimal Mobile Ad Campaign Strategy under Markov Customer Dynamics"

Guokai Li, Pin Gao, Zizhuo Wang

Appendix A: Omitted Proofs in Main Text

In this section, we provide proofs omitted from the main text.

A.1. Proof Sketch of Theorem 1

In the following, we provide the proof sketch of Theorem 1. In the proof, we mainly study the discretized problem and the property of the continuous-time problem is proved by taking the discretization granularity to zero. First, following the definitions in Sections 2 and 3, a POMDP can be represented by a tuple $(\mathbb{X},\mathbb{Y},\mathbb{A},\mathcal{P},\mathcal{Q},R)$. According to Section 3 in Feinberg et al. (2016), we can transform a POMDP into an equivalent belief-MDP, which is a fully observable MDP whose states are posterior state distributions, i.e., $\tilde{\boldsymbol{b}}_t \in \tilde{\mathbb{B}} = \{\boldsymbol{b} \in [0,1]^5 \mid \sum_{i=1}^5 b_i = 1\}$. Moreover, based on the patterns of the system dynamics, we can further simplify the state as $\boldsymbol{b}_t \in \mathbb{B} = \{\boldsymbol{b} \in [0,1]^3 \mid \|\boldsymbol{b}\|_0 \le 1, b_3 \in \{0,1\}\}$. The resulting belief-MDP is represented by a tuple $(\mathbb{B},\mathbb{A},\mathcal{P}_b,R_b)$, and can be treated as a fully observable MDP with Borel state and action sets. Thus, we can prove the optimality of stationary policies under the belief-MDP by checking assumptions in Feinberg et al. (2012) that establish the optimality of stationary policies under long-term average objective.

Before proceeding, we first define some functions. First, given a discount factor $\beta \in (0,1)$, we define the discounted revenue and the average revenue as follows:

$$V_\beta(\boldsymbol{b}) := \sup_{\mu\in\Pi^B} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \cdot R_b(\boldsymbol{b}_t^\mu, a_t^\mu) \mid \boldsymbol{b}_0 = \boldsymbol{b}\right]$$

$$w(\boldsymbol{b}) := \sup_{\mu\in\Pi^B} \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=0}^{T-1} R_b(\boldsymbol{b}_t^\mu, a_t^\mu) \mid \boldsymbol{b}_0 = \boldsymbol{b}\right],$$

where Π^B denotes the set of all admissible policies for the belief-MDP, and $\boldsymbol{b}_t^\mu$ and a_t^μ are the random state and action at period t under the policy μ. Then, we define $m_\beta := \sup_{\boldsymbol{b}\in\mathbb{B}} V_\beta(\boldsymbol{b})$, $u_\beta(\boldsymbol{b}) := m_\beta - V_\beta(\boldsymbol{b})$, and $C(\boldsymbol{b},a) := -R_b(\boldsymbol{b},a)$. According to Theorem 4 in Feinberg et al. (2012), in order to prove the optimality of stationary policies under the average revenue objective, we only need to check Assumptions W^* and B in Feinberg et al. (2012). In the following, we first present an equivalent version of Assumption W^*.

ASSUMPTION 1 **(Equivalent Version of Assumption W^* in Feinberg et al. 2012).**

(i) *$C(\boldsymbol{b},a)$ is K-inf-compact and bounded below on $\mathbb{B}\times\mathbb{A}$. That is, for any finite $c_0 \in \mathbb{R}$ and any compact subset $K \subseteq \mathbb{B}$, the level set $\{(\boldsymbol{b},a) \in K\times\mathbb{A} \mid C(\boldsymbol{b},a) \le c_0\}$ is compact.*

(ii) *The state transition kernel $\mathcal{P}_b(\cdot \mid \boldsymbol{b},a)$ is weakly continuous in $(\boldsymbol{b},a) \in \mathbb{B}\times\mathbb{A}$. That is, for any sequence $\{(\boldsymbol{b}_k,a_k) \in \mathbb{B}\times\mathbb{A}\}_{k=0,1,2\ldots}$ converging to $(\boldsymbol{b},a)$ and any bounded continuous function $g : \mathbb{B} \to \mathbb{R}$, we have*

$$\int_{\mathbb{B}} g(\boldsymbol{b}') \cdot \mathcal{P}_b(d\boldsymbol{b}' \mid \boldsymbol{b}_k, a_k) \to \int_{\mathbb{B}} g(\boldsymbol{b}') \cdot \mathcal{P}_b(d\boldsymbol{b}' \mid \boldsymbol{b}, a) \qquad \text{as } k \to \infty.$$

According to Lemma 2.5 in Feinberg et al. (2013), Assumption 1(i) is equivalent to Assumptions W^*(i)-(ii). Assumption 1(ii) is the same as Assumptions W^*(iii). Thus, Assumption 1 is equivalent to Assumptions W^*.

Second, we present an equivalent version of Assumption B in Feinberg et al. (2012).

ASSUMPTION 2 **(Equivalent Version of Assumption B in Feinberg et al. 2012).**

(i) $w^* := \sup_{\boldsymbol{b}\in\mathbb{B}} w(\boldsymbol{b}) > -\infty$.

(ii) $\sup_{\beta\in[0,1)} u_\beta(\boldsymbol{b}) < \infty$ *for all* $\boldsymbol{b}\in\mathbb{B}$.

Assumption 2(i) is the same as Assumption B(i). According to Lemma 5 in Feinberg et al. (2012), if $C(\cdot,\cdot)$ is lower bounded and Assumption B(i) holds, then Assumption 2(ii) is equivalent to Assumption B(ii). Thus, given that Assumption 1(i) holds, Assumption 2 is equivalent to Assumption B.

To summarize, if Assumptions 1 and 2 hold, then the optimality of stationary policies is established for the belief-MDP, which finishes the proof of Theorem 1. □

A.2. Proof of Theorem 1

As mentioned, the problem is formulated as a POMDP with the hidden state $x_t \in \mathbb{X} = \{A_p, I_p, A_a, I_a, P\}$, the observation $y_t \in \mathbb{Y} = \{0,1\}$ and the action $a_t \in \mathbb{A} = \{0,1\}$. In the following, similar to Ely (2017), we approximate the continuous-time problem by discretizing the time horizon into intervals of sufficiently small length $h < \min\left\{\frac{1}{2(\lambda+\gamma)}, \frac{p_a}{4(\lambda+\gamma+\theta)}\right\}$. Given the state x_t and the action a_t at period t, the transition probabilities $\mathcal{P}(x_{t+1} \mid x_t, a_t)$ (with positive values) are as follows (the $o(h)$ terms are omitted):

$$\mathcal{P}\left(P \mid A_p, 0\right) = \lambda h \qquad \mathcal{P}\left(I_p \mid A_p, 0\right) = \gamma h$$
$$\mathcal{P}\left(A_p \mid A_p, 0\right) = 1-(\lambda+\gamma)h \qquad \mathcal{P}\left(A_p \mid I_p, 0\right) = \theta h$$
$$\mathcal{P}\left(I_p \mid I_p, 0\right) = 1-\theta h \qquad \mathcal{P}\left(P \mid A_a, 0\right) = \lambda h$$
$$\mathcal{P}\left(I_a \mid A_a, 0\right) = \gamma h \qquad \mathcal{P}\left(A_a \mid A_a, 0\right) = 1-(\lambda+\gamma)h$$
$$\mathcal{P}\left(A_a \mid I_a, 0\right) = \theta h \qquad \mathcal{P}\left(I_a \mid I_a, 0\right) = 1-\theta h$$
$$\mathcal{P}\left(P \mid P, 0\right) = \lambda h \qquad \mathcal{P}\left(I_p \mid P, 0\right) = \gamma h$$
$$\mathcal{P}\left(A_p \mid P, 0\right) = 1-(\lambda+\gamma)h$$
$$\mathcal{P}\left(P \mid A_p, 1\right) = \lambda h \qquad \mathcal{P}\left(I_a \mid A_p, 1\right) = \gamma h$$
$$\mathcal{P}\left(A_a \mid A_p, 1\right) = 1-(\lambda+\gamma)h \qquad \mathcal{P}\left(P \mid I_p, 1\right) = p_p\lambda h$$
$$\mathcal{P}\left(A_a \mid I_p, 1\right) = p_p\left(1-(\lambda+\gamma)h\right)+(1-p_p)\theta h \qquad \mathcal{P}\left(I_a \mid I_p, 1\right) = p_p\gamma h+(1-p_p)(1-\theta h)$$
$$\mathcal{P}\left(P \mid A_a, 1\right) = \lambda h \qquad \mathcal{P}\left(I_a \mid A_a, 1\right) = \gamma h$$
$$\mathcal{P}\left(A_a \mid A_a, 1\right) = 1-(\lambda+\gamma)h \qquad \mathcal{P}\left(P \mid I_a, 1\right) = p_a\lambda h$$
$$\mathcal{P}\left(A_a \mid I_a, 1\right) = p_a\left(1-(\lambda+\gamma)h\right)+(1-p_a)\theta h \qquad \mathcal{P}\left(I_a \mid I_a, 1\right) = p_a\gamma h+(1-p_a)(1-\theta h)$$
$$\mathcal{P}(P \mid P, 1) = \lambda h \qquad \mathcal{P}(I_a \mid P, 1) = \gamma h$$
$$\mathcal{P}(A_a \mid P, 1) = 1-(\lambda+\gamma)h.$$

The reward function is specified as $R(x_t, a_t) = \mathbb{1}\{x_t = P\}\cdot r - a_t \cdot c$. Given the state x_t, the observation generating probabilities are specified by the observation kernel $Q(y_t \mid x_t) = y_t \cdot \mathbb{1}\{x_t = P\} + (1-y_t)\cdot \mathbb{1}\{x_t \neq P\}$. Given the initial state $x_0 = A_p$, the observed history at period t is $h_t := (x_0, y_0, a_0, y_1, a_1, \ldots, y_t)$. An

admissible policy can be represented as $\mu = \{\mu_t\}_{t=0,1,\dots}$ where $\mu_t : h_t \to [0,1]$. The objective is to minimize the long-term average expected profit over the set Π of all admissible policies:

$$\sup_{\mu\in\Pi}\left\{\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{t=0}^{T-1}R(x_t^\mu, a_t^\mu)\,|\,x_0 = A_p\right]\right\},$$

where x_t^μ (a_t^μ, resp.) is the random state (action, resp.) at period t under the policy μ. As mentioned in Appendix A.1, the problem can be transformed into an equivalent belief-MDP with the state $\tilde{\boldsymbol{b}}_t$ as the conditional distribution over $\mathbb{X}$, and it can be further simplified into a compact state representation $\boldsymbol{b}_t$ as follows:

$$\boldsymbol{b}_t = \left[\mathbb{P}\left(x_t = A_p \,|\, h_t\right),\ \mathbb{P}\left(x_t = A_a \,|\, h_t\right),\ \mathbb{P}\left(x_t = P \,|\, h_t\right)\right] \in \mathbb{B},$$

where $\mathbb{B} = \{\boldsymbol{b}\in[0,1]^3 \,|\, \|\boldsymbol{b}\|_0 \le 1, b_3 \in \{0,1\}\}$. In the following, we use this compact belief representation $\boldsymbol{b}$ as the state. Given the state $\boldsymbol{b}$ and the action a of the last period and the observation y of the current period, the belief state of the current period is specified by a function $\mathcal{T}(\boldsymbol{b}, a, y)$. In particular, we have

$$\begin{aligned}
\mathcal{T}([b_1,0,0],0,0) &= \left[\frac{b_1(1-\lambda h-\gamma h)+(1-b_1)\theta h}{b_1(1-\lambda h)+1-b_1}, 0, 0\right]\\
\mathcal{T}([0,b_2,0],0,0) &= \left[0, \frac{b_2(1-\lambda h-\gamma h)+(1-b_2)\theta h}{b_2(1-\lambda h)+1-b_2}, 0\right]\\
\mathcal{T}([b_1,0,0],1,0) &= \left[0, \frac{\left(b_1+(1-b_1)p_p\right)(1-\lambda h-\gamma h)+(1-b_1)(1-p_p)\theta h}{\left(b_1+(1-b_1)p_p\right)(1-\lambda h)+(1-b_1)(1-p_p)}, 0\right],\\
\mathcal{T}([0,b_2,0],1,0) &= \left[0, \frac{\left(b_2+(1-b_2)p_a\right)(1-\lambda h-\gamma h)+(1-b_2)(1-p_a)\theta h}{\left(b_2+(1-b_2)p_a\right)(1-\lambda h)+(1-b_2)(1-p_a)}, 0\right],\\
\mathcal{T}([0,0,1],0,0) &= \left[\frac{1-\lambda h-\gamma h}{1-\lambda h}, 0, 0\right]\\
\mathcal{T}([0,0,1],1,0) &= \left[0, \frac{1-\lambda h-\gamma h}{1-\lambda h}, 0\right]\\
\mathcal{T}(\boldsymbol{b},a,1) &= [0,0,1] \qquad\qquad \forall \boldsymbol{b}, a.
\end{aligned}$$

Since $h < \frac{1}{\lambda+\gamma+\theta}$, we can deduce that $\mathcal{T}(\boldsymbol{b}, a, y)$ is nondecreasing in $\boldsymbol{b}$.

Then, we can specify the transition kernel $\mathcal{P}_b(\boldsymbol{b}_{t+1} \,|\, \boldsymbol{b}_t, a_t)$ as follows:

$$\begin{aligned}
\mathcal{P}_b(\mathcal{T}(\boldsymbol{b},0,1)\,|\,\boldsymbol{b},0) &= (b_1+b_2+b_3)\lambda h\\
\mathcal{P}_b(\mathcal{T}(\boldsymbol{b},0,0)\,|\,\boldsymbol{b},0) &= 1-(b_1+b_2+b_3)\lambda h\\
\mathcal{P}_b(\mathcal{T}(\boldsymbol{b},1,1)\,|\,\boldsymbol{b},1) &= \left(b_1+(1-b_1)p_p+b_2+(1-b_2)p_a+b_3\right)\lambda h\\
\mathcal{P}_b(\mathcal{T}(\boldsymbol{b},1,0)\,|\,\boldsymbol{b},1) &= 1-\left(b_1+(1-b_1)p_p+b_2+(1-b_2)p_a+b_3\right)\lambda h.
\end{aligned}$$

The reward function is $R_b(\boldsymbol{b}, a) = b_3 \cdot r - a \cdot c$. Let $h_t^B = \{\boldsymbol{b}_\ell\}_{\ell=0}^t$ denote the observed history up to period t. In this case, an admissible policy μ^B is defined as a sequence $\mu^B = \{\mu_t^B\}_{t\ge0}$ of stochastic kernels $\mu_t^B : h_t^B \to [0,1]$, and the objective becomes

$$V(\boldsymbol{b}) = \max_{\mu^B\in\Pi^B} V^{\mu_B}(\boldsymbol{b}) = \sup_{\mu^B\in\Pi^B}\left\{\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{t=0}^{T-1}\left(b_{t,3}^{\mu^B}r - a_t^{\mu^B}c\right)\,|\,\boldsymbol{b}_0 = \boldsymbol{b}\right]\right\},$$

where $\boldsymbol{b}_t^{\mu^B}$ is the random belief state at period t under the policy μ^B. Among all admissible policies, there exists a class of *stationary policies* with $\mu_S^B : \mathbb{B} \to \{0,1\}$ and $\mu_t^B(h_t^B) = \mu_S^B(\boldsymbol{b}_t)$. As mentioned in Appendix A.1, we will prove the optimality of stationary policies by checking Assumptions 1 and 2.

We start with Assumption 1. First, we check Assumption 1(i). Since $C(\boldsymbol{b}, a) = -b_3 \cdot r + a \cdot c \in [-r, c]$, we can derive that $C(\boldsymbol{b}, a)$ is bounded. Given any compact subset $K \subseteq \mathbb{B}$ and any finite constant c_0, we can deduce that $\{(\boldsymbol{b}, a) \in K \times \mathbb{A} \mid C_b(\boldsymbol{b}, a) \le c_0\} = \{(\boldsymbol{b}, a) \in \mathbb{B} \times \mathbb{A} \mid -b_3 \cdot r + a \cdot c \le c_0\} \cap \{(\boldsymbol{b}, a) \in K \times \mathbb{A}\}$ is compact. Thus, Assumption 1(i) is satisfied. Then, we check Assumption 1(ii). Since the sequence $\{(\boldsymbol{b}_k, a_k)\}_{k=0,1,2,\ldots}$ converges to $(\boldsymbol{b}, a)$ and $\mathbb{A}$ only has two items, we can fix $a_k = a$ when proving the convergence. Then, given any continuous and bounded function $g(\cdot)$, we have

$$\begin{aligned}
\int_{\mathbb{B}} g(\boldsymbol{b}') \mathcal{P}_b(d\boldsymbol{b}' \mid \boldsymbol{b}_k, a_k) &= \left(\mathbf{1}^T \boldsymbol{b}_k + a(1 - b_{k,1})p_p + a(1 - b_{k,2})p_a\right) \lambda h g\left([0,0,1]\right) \\
&\quad + \left(1 - \left(\mathbf{1}^T \boldsymbol{b}_k + a(1 - b_{k,1})p_p + a(1 - b_{k,2})p_a\right) \lambda h\right) g\left(\mathcal{T}(\boldsymbol{b}_k, a, 0)\right) \\
&\to \left(\mathbf{1}^T \boldsymbol{b} + a(1 - b_1)p_p + a(1 - b_2)p_a\right) \lambda h g\left([0,0,1]\right) \\
&\quad + \left(1 - \left(\mathbf{1}^T \boldsymbol{b} + a(1 - b_1)p_p + a(1 - b_2)p_a\right) \lambda h\right) g\left(\mathcal{T}(\boldsymbol{b}, a, 0)\right) \\
&= \int_{\mathbb{B}} g(\boldsymbol{b}') \mathcal{P}_b(d\boldsymbol{b}' \mid \boldsymbol{b}, a),
\end{aligned}$$

where the convergence is because $g(\cdot)$ and $\mathcal{T}(\cdot, a, y)$ are bounded continuous functions.

Second, for Assumption 2(i), we consider a stationary policy $\bar{\mu}^B$ with $\bar{\mu}^B(\boldsymbol{b}) = 0$ for any $\boldsymbol{b}$ (i.e., never use ad campaigns), under which there is zero ad cost and hence $w(\boldsymbol{b}) \ge 0$ for any $\boldsymbol{b} \in \mathbb{B}$. Therefore, we have $w^* \ge w([0,0,1]) \ge 0 > -\infty$, implying that Assumption 2(i) is satisfied.

Lastly, we need to check Assumption2(ii). We prove the following results for any $\beta \in [0,1)$. According to Theorem 2 in Feinberg et al. (2012), when Assumption 1 is satisfied, there exists a stationary policy that is optimal to the discounted problem $v_\beta(\boldsymbol{b})$. Then, we can write the Bellman equation as follows:

$$\begin{aligned}
V_\beta(\boldsymbol{b}) = \max \Big\{ & b_3 r + \beta(b_1 + b_2 + b_3)\lambda h \cdot V_\beta\left(\mathcal{T}(\boldsymbol{b}, 0, 1)\right) + \beta\left(1 - (b_1 + b_2 + b_3)\lambda h\right) \cdot V_\beta\left(\mathcal{T}(\boldsymbol{b}, 0, 0)\right), \\
& b_3 r - c + \beta\left(b_1 + (1 - b_1)p_p + b_2 + (1 - b_2)p_a + b_3\right) \lambda h \cdot V_\beta\left(\mathcal{T}(\boldsymbol{b}, 1, 1)\right) \\
& + \beta\left(1 - \left(b_1 + (1 - b_1)p_p + b_2 + (1 - b_2)p_a + b_3\right) \lambda h\right) \cdot V_\beta\left(\mathcal{T}(\boldsymbol{b}, 1, 0)\right) \Big\}.
\end{aligned}$$

If a tie happens, then we choose to make an ad push. We will prove that $u_\beta(\tilde{\boldsymbol{b}}) = \sup_{\boldsymbol{b}' \in \mathbb{B}} V_\beta(\boldsymbol{b}') - V_\beta(\boldsymbol{b}) < \infty$ for all $\boldsymbol{b} \in \mathbb{B}$ in two steps.

Step 1: Structural Properties of $V_\beta(\boldsymbol{b})$.

Define $\boldsymbol{b}^* := [0,0,1]$. Based on the value iteration, we will use the induction argument to prove the following statement: Both $V_\beta([b,0,0])$ and $V_\beta([0,b,0])$ increase in b, $V_\beta([b,0,0]) \ge V_\beta([0,b,0])$, and $V_\beta(\boldsymbol{b}) \le V_\beta(\boldsymbol{b}^*)$ for any $\boldsymbol{b} \in \mathbb{B}$. Then, we start the value iteration with $V_\beta^0(\boldsymbol{b}) = 0$ for any $\boldsymbol{b} \in \mathbb{B}$ such that the statement holds at the beginning. Suppose the statement holds in the k-th iteration. Then, for the $(k+1)$-iteration, we can prove the statement by the Bellman equation as follows ($b_1 \le b_2$):

$$\begin{aligned}
V_\beta^{k+1}([b_1,0,0]) = \max \Big\{ & \beta b_1 \lambda h V_\beta^k([0,0,1]) + \beta(1 - b_1\lambda h) V_\beta^k\left(\mathcal{T}([b_1,0,0],0,0)\right), \\
& -c + \beta(b_1 + (1 - b_1)p_p)\lambda h V_\beta^k([0,0,1]) + \beta(1 - (b_1 + (1 - b_1)p_p)\lambda h) V_\beta^k\left(\mathcal{T}([b_1,0,0],1,0)\right) \Big\}
\end{aligned}$$

$$\le \max\Big\{\beta b_2\lambda h V_\beta^k([0,0,1]) + \beta(1-b_2\lambda h)V_\beta^k\left(\mathcal{T}([b_2,0,0],0,0)\right),$$
$$-c + \beta(b_2+(1-b_2)p_p)\lambda h V_\beta^k([0,0,1]) + \beta(1-(b_2+(1-b_2)p_p)\lambda h)V_\beta^k\left(\mathcal{T}([b_2,0,0],1,0)\right)\Big\}$$
$$= V_\beta^{k+1}([b_2,0,0]).$$

Similarly, we have $V_\beta^{k+1}([0,b_1,0]) \le V_\beta^{k+1}([0,b_2,0])$ for any $b_1 \le b_2$ and $V_\beta^{k+1}([b,0,0]) \ge V_\beta^{k+1}([0,b,0])$ for any b. Furthermore, we have $V_\beta^{k+1}(\boldsymbol{b}^*) = r + V_\beta^{k+1}([1,0,0]) > V_\beta^{k+1}(\boldsymbol{b})$ for all $\boldsymbol{b} \in \mathbb{B}\setminus\boldsymbol{b}^*$. Due to the convergence of the value iteration algorithm (see Proposition 9.17 in Bertsekas and Shreve 1996), we have $V_\beta(\boldsymbol{b}) = V_\beta^\infty(\boldsymbol{b})$ for all $\boldsymbol{b}\in\mathbb{B}$ and hence the statement is proved.

Step 2: $V_\beta(\boldsymbol{b}^*) - V_\beta(\boldsymbol{b}) < \infty$ for all $\boldsymbol{b}\in\mathbb{B}$.

We fix $\boldsymbol{b}$ in the following. We first present some properties of the transition function $\mathcal{T}(\cdot,\cdot,\cdot)$. For any $b\in[0,1]$, we have

$$\mathcal{T}([b,0,0],1,0)_2 = \frac{\left(b+(1-b)p_p\right)(1-\lambda h-\gamma h)+(1-b)(1-p_p)\theta h}{\left(b+(1-b)p_p\right)(1-\lambda h)+(1-b)(1-p_p)} \ge \left(b+(1-b)p_p\right)(1-\lambda h-\gamma h) \ge \frac{p_p}{2}$$
$$\mathcal{T}([0,b,0],1,0)_2 = \frac{\left(b+(1-b)p_a\right)(1-\lambda h-\gamma h)+(1-b)(1-p_a)\theta h}{\left(b+(1-b)p_a\right)(1-\lambda h)+(1-b)(1-p_a)} \ge \left(b+(1-b)p_a\right)(1-\lambda h-\gamma h) \ge \frac{p_a}{2}$$
$$b - \mathcal{T}([0,b,0],0,0)_2 = b - \frac{b(1-\lambda h-\gamma h)+(1-b)\theta h}{b(1-\lambda h)+1-b} \le b - (b(1-\lambda h-\gamma h)+(1-b)\theta h) \le (\lambda+\gamma+\theta)\,h,$$

where the subscript '2' indicates the second entry. Therefore, there exist two positive constants $c_1 := \frac{p_a}{2}$ and $c_2 := \lambda+\gamma+\theta$ such that for any $b\in[0,1]$, we have $\mathcal{T}([b,0,0],1,0)_2 \ge c_1$, $\mathcal{T}([0,b,0],1,0)_2 \ge c_1$ and $b - \mathcal{T}([0,b,0],0,0)_2 \le c_2 h$. Then, we consider two systems A and B starting from $\boldsymbol{b}^*$ and $\boldsymbol{b}$, respectively. System A uses the optimal policy. In contrast, system B uses a suboptimal policy: We push ads at periods $\left\{k\left\lfloor\frac{c_1}{2c_2h}\right\rfloor \mid k\in\mathbb{Z}_+\right\}$ ($\left\lfloor\frac{c_1}{2c_2h}\right\rfloor > 2$ due to the choice of h) when $t<\tau$ and use the optimal policy when $t\ge\tau$, where τ is the first purchase period of system B. Let $\boldsymbol{b}_t^A$ and $\boldsymbol{b}_t^B$ denote the belief state of system A and B, respectively, at period t. Due to the design of the policy, for each time period $t\in[1,\tau)$, we have $b_{t,2}^B \ge \frac{c_1}{2}$. Since the probability that a purchase happens at period $t+1$ is $b_{t,2}^B\lambda h \ge \frac{c_1\lambda h}{2}$, the expectation of the stopping time $\mathbb{E}[\tau]$ is no greater than $\frac{2}{c_1\lambda h}+1 \le \frac{4}{c_1\lambda h}$. Then, we split the time horizon into $[0,\tau)$ and $[\tau,\infty)$ and derive the following results:

$$\begin{aligned}
V_\beta(\boldsymbol{b}^*) - V_\beta(\boldsymbol{b}) &\le \mathbb{E}\left[\sum_{t=0}^{\tau-1}\beta^t\left(R(x_t^A,a_t^A) - R(x_t^B,a_t^B)\right) + \beta^\tau\left(V_\beta(\boldsymbol{b}_\tau^A) - V_\beta(\boldsymbol{b}^*)\right)\right]\\
&\le \mathbb{E}\left[\sum_{t=0}^{\tau-1}\beta^t\left(\lambda h r + a_t^B\cdot c\right) + \beta^\tau\left(V_\beta(\boldsymbol{b}_\tau^A) - V_\beta(\boldsymbol{b}^*)\right)\right]\\
&\le \mathbb{E}[\tau]\lambda h r + \left(\frac{\mathbb{E}[\tau]}{\lfloor c_1/(2c_2h)\rfloor}+1\right)c\\
&\le \frac{4r}{c_1} + \left(\frac{16c_2}{c_1^2\lambda}+1\right)c < \infty,
\end{aligned}$$

where the first inequality is because system B uses a suboptimal policy, and the belief of system B is $\boldsymbol{b}^*$ at the purchase period τ. The second inequality is because per period expected revenue of system A is no greater than $\lambda h r$. The third inequality is because $\beta\le 1$ and $V_\beta(\boldsymbol{b}^*) \ge V_\beta(\boldsymbol{b})$ for any $\boldsymbol{b}\in\mathbb{B}$. The last inequality is because $\mathbb{E}[\tau] \le \frac{4}{c_1\lambda h}$ and $\lfloor c_1/(2c_2h)\rfloor \ge c_1/(4c_2h)$. Therefore, Assumption 2(ii) is satisfied.

To summarize, both Assumptions 1 and 2 hold. According to Feinberg et al. (2012), there exists a stationary policy that is optimal for the discretized problem and $V(\boldsymbol{b}) = \lim_{\beta\to 1}(1-\beta)V_\beta(\boldsymbol{b})$. As h tends to 0, we can deduce that there exists a stationary policy that is optimal to the belief-MDP. □

A.3. Proof of Theorem 2

In this proof, we still use the discretized POMDP in Appendix A.2 to analyze the structural properties of the optimal stationary policy μ_S^B. We fix h to be a sufficiently small value. Let $\mu_S^{B,\beta}(\cdot)$ denote the optimal stationary policy for the discounted problem under the discount factor β. According to Theorem 5 in Feinberg et al. (2012), there exists a sequence of $\{\beta_n\}_{n=1}^{\infty}$ with $\lim_{n\to\infty}\beta_n = 1$ such that $\mu_S^{B,\beta_n}(\cdot)$ converges to $\mu_S^B(\cdot)$ as n tends to infinity. Therefore, we can focus on the structural property of the optimal stationary discounted policy as β tends to 1. Define $b_p := \sup\{b\in[0,1] \mid \mu_S^B([b,0,0]) = 1\}$, $b_a := \sup\{b\in[0,1] \mid \mu_S^B([0,b,0]) = 1\}$ and $f(b) := \frac{b(1-\lambda h-\gamma h)+(1-b)\theta h}{b(1-\lambda h)+1-b}$.

Fact I: $\mu_S^B([1,0,0]) = 0$ and hence $b_p < 1$.

Given a discount factor β, the Bellman equation for the discounted problem is

$$V_\beta([1,0,0]) = \max\Big\{\beta\lambda h V_\beta([0,0,1]) + \beta(1-\lambda h)V_\beta(\mathcal{T}([1,0,0],0,0)),\\ -c + \beta\lambda h V_\beta([0,0,1]) + \beta(1-\lambda h)V_\beta(\mathcal{T}([1,0,0],1,0))\Big\}.$$

We can deduce that the revenue gain of pushing ads is $-c+\beta(1-\lambda h)\left(V_\beta([0,f(1),0]) - V_\beta([f(1),0,0])\right) < 0$. Therefore, as β tends to 1, we have $\mu_S^B([1,0,0]) = 0$ and hence $b_p < 1$.

Fact II: There exists a constant b_a such that $\mu_S^B([0,b,0]) = 0$ when $b > b_a$ and $\mu_S^B([0,b,0]) = 1$ when $b \le b_a$.

Similarly, given a discount factor β, the Bellman equation for the discounted problem is

$$V_\beta([0,b,0]) = \max\Big\{\beta b\lambda h V_\beta([0,0,1]) + \beta(1-b\lambda h)V_\beta(\mathcal{T}([0,b,0],0,0)),\\ -c + \beta\left(b+(1-b)p_a\right)\lambda h V_\beta([0,0,1]) + \beta\left(1-\left(b+(1-b)p_a\right)\lambda h\right)V_\beta(\mathcal{T}([0,b,0],1,0))\Big\}.$$

Consider $b_2 \le b_1$ and $\mu_S^{B,\beta}([0,b_1,0]) = 1$. In this case, we can derive the following inequalities:

$$\begin{aligned}0 \le & -c + \beta\left(b_1+(1-b_1)p_a\right)\lambda h V_\beta([0,0,1]) + \beta\left(1-\left(b_1+(1-b_1)p_a\right)\lambda h\right)V_\beta(\mathcal{T}([0,b_1,0],1,0))\\ & - \left(\beta b_1\lambda h V_\beta([0,0,1]) + \beta(1-b_1\lambda h)V_\beta(\mathcal{T}([0,b_1,0],0,0))\right)\\ = & -c + \beta(1-b_1)p_a\lambda h\left(V_\beta([0,0,1]) - V_\beta(\mathcal{T}([0,b_1,0],0,0))\right)\\ & + \beta\left(1-\left(b_1+(1-b_1)p_a\right)\lambda h\right)\left(V_\beta(\mathcal{T}([0,b_1,0],1,0)) - V_\beta(\mathcal{T}([0,b_1,0],0,0))\right).\end{aligned}$$

According to Åström (1969), given the existence of an optimal stationary policy, $V_\beta(\boldsymbol{b})$ is concave in $\boldsymbol{b}$. Therefore, we can derive the following inequalities:

$$\begin{aligned}V_\beta(\mathcal{T}([0,b_1,0],1,0)) - V_\beta(\mathcal{T}([0,b_1,0],0,0)) =& V_\beta\left(\left[0, f(b_1+(1-b_1)p_a), 0\right]\right) - V_\beta\left(\left[0, f(b_1), 0\right]\right)\\ \le& V_\beta\left(\left[0, f(b_1+(1-b_1)p_a) + f(b_2) - f(b_1), 0\right]\right) - V_\beta\left(\left[0, f(b_2), 0\right]\right)\\ \le& V_\beta\left(\left[0, f(b_2+(1-b_2)p_a), 0\right]\right) - V_\beta\left(\left[0, f(b_2), 0\right]\right)\end{aligned}$$

$$= V_\beta\left(\mathcal{T}([0,b_2,0],1,0)\right) - V_\beta\left(\mathcal{T}([0,b_2,0],0,0)\right),$$

where the first inequality is because $f(b_2)-f(b_1)\le 0$ when h is sufficiently small and $V_\beta([0,b,0])$ is concave in b, and the second inequality is because $f(b_1+(1-b_1)p_a)+f(b_2)-f(b_1)-f(b_2+(1-b_2)p_a)=(b_2-b_1)p_a - O(h)\le 0$ when h is sufficiently small. Then, we have

$$\begin{aligned}
&-c+\beta(1-b_2)p_a\lambda h\left(V_\beta([0,0,1])-V_\beta\left(\mathcal{T}([0,b_2,0],0,0)\right)\right)\\
&+\beta\left(1-\left(b_2+(1-b_2)p_a\right)\lambda h\right)\left(V_\beta\left(\mathcal{T}([0,b_2,0],1,0)\right)-V_\beta\left(\mathcal{T}([0,b_2,0],0,0)\right)\right)\\
\ge\ & c+\beta(1-b_1)p_a\lambda h\left(V_\beta([0,0,1])-V_\beta\left(\mathcal{T}([0,b_1,0],0,0)\right)\right)\\
&+\beta\left(1-\left(b_1+(1-b_1)p_a\right)\lambda h\right)\left(V_\beta\left(\mathcal{T}([0,b_1,0],1,0)\right)-V_\beta\left(\mathcal{T}([0,b_1,0],0,0)\right)\right)\\
\ge\ & 0,
\end{aligned}$$

which implies that $\mu_S^{B,\beta}([0,b_2,0]) = 1$. Thus, we can deduce that there exists $b_a^\beta \in [0,1]$ such that $\mu_S^{B,\beta}([0,b,0])=0$ when $b>b_a^\beta$ and $\mu_S^{B,\beta}([0,b,0])=1$ when $b\le b_a^\beta$. As β tends to 1, we finish the proof of Fact II.

Fact III: $b_a < b_p+(1-b_p)p_p$**.**

Define $b_p^\beta := \sup\left\{b\in[0,1]\,|\,\mu_S^{B,\beta}([b,0,0])=1\right\}$, $b_a^\beta := \sup\left\{b\in[0,1]\,|\,\mu_S^{B,\beta}([0,b,0])=1\right\}$, $\bar{b}_a^\beta := b_p^\beta+(1-b_p^\beta)p_p$, and $\tilde{b}_a^\beta := \frac{p_p-p_a+b_p^\beta(1-p_p)}{1-p_a}$. First, given the state $[0,1,0]$, the revenue gain of pushing ads is $-c<0$. Thus, there exist constants $c_4\in(0,1)$ and $c_5>0$ such that $b_a^\beta\le c_4$ and $\frac{\partial}{\partial b}V_\beta([0,b,0])\,|_{b=b_a}\ge c_5$ for any β. For ease of exposition, we temporarily ignore the superscript β.

Due to the definition of b_p, we have

$$\begin{aligned}
&\beta b_p\lambda h V_\beta([0,0,1])+\beta(1-b_p\lambda h)V_\beta\left(\mathcal{T}([b_p,0,0],0,0)\right)\\
=&-c+\beta(b_p+(1-b_p)p_p)\lambda h V_\beta([0,0,1])+\beta(1-(b_p+(1-b_p)p_p)\lambda h)V_\beta(\mathcal{T}([b_p,0,0],1,0)).
\end{aligned}\tag{A.1}$$

Suppose $b_a\ge\bar{b}_a$. Since $\tilde{b}_a<\bar{b}_a\le b_a$, according to Fact II, we have $\mu_S^B([0,\tilde{b}_a,0])=1$. Since $\mu_S^B([b_p,0,0])=1$ and $\tilde{b}_a+(1-\tilde{b}_a)p_a = b_p+(1-b_p)p_p$, we have

$$V_\beta([0,\tilde{b}_a,0]) = -c+\beta(\tilde{b}_a+(1-\tilde{b}_a)p_a)\lambda h V_\beta([0,0,1])+\beta(1-(\tilde{b}_a+(1-\tilde{b}_a)p_a)\lambda h)V_\beta(\mathcal{T}([0,\tilde{b}_a,0],1,0)) = V_\beta([b_p,0,0]).$$

Similar to the proof for Fact II, we have

$$\begin{aligned}
&V_\beta\left(\mathcal{T}([0,b_a,0],1,0)\right)-V_\beta\left(\mathcal{T}([0,b_a,0],0,0)\right)\\
\le& V_\beta\left(\left[0,f(b_a+(1-b_a)p_a)+f(\tilde{b}_a)-f(b_a),0\right]\right)-V_\beta\left(\left[0,f(\tilde{b}_a),0\right]\right)\\
=& V_\beta\left(\mathcal{T}([0,\tilde{b}_a,0],1,0)\right)-V_\beta\left(\mathcal{T}([0,\tilde{b}_a,0],0,0)\right)\\
&+V_\beta\left(\left[0,f(b_a+(1-b_a)p_a)+f(\tilde{b}_a)-f(b_a),0\right]\right)-V_\beta\left(\left[0,\tilde{b}_a+(1-\tilde{b}_a)p_a,0\right]\right)\\
=& V_\beta\left(\mathcal{T}([0,\tilde{b}_a,0],1,0)\right)-V_\beta\left(\mathcal{T}([0,\tilde{b}_a,0],0,0)\right)-(b_a-\tilde{b}_a)c_5+O(h).
\end{aligned}\tag{A.2}$$

Then, given the belief state $[0,b_a,0]$, the revenue gain of pushing ads is

$$-c+\beta(1-b_a)p_a\lambda h\left(V_\beta([0,0,1])-V_\beta(\mathcal{T}([0,b_a,0],0,0))\right)$$

$$
\begin{aligned}
&+\beta(1-(b_a+(1-b_a)p_a)\lambda h)\left(V_\beta(\mathcal{T}([0,b_a,0],1,0))-V_\beta(\mathcal{T}([0,b_a,0],0,0))\right)\\
=&\beta b_p\lambda h V_\beta([0,0,1])+\beta(1-b_p\lambda h)V_\beta\left(\mathcal{T}([b_p,0,0],0,0)\right)-\beta(b_p+(1-b_p)p_p)\lambda h V_\beta([0,0,1])\\
&-\beta(1-(b_p+(1-b_p)p_p)\lambda h)V_\beta(\mathcal{T}([b_p,0,0],1,0))+\beta(1-b_a)p_a\lambda h\left(V_\beta([0,0,1])-V_\beta(\mathcal{T}([0,b_a,0],0,0))\right)\\
&+\beta(1-(b_a+(1-b_a)p_a)\lambda h)\left(V_\beta(\mathcal{T}([0,b_a,0],1,0))-V_\beta(\mathcal{T}([0,b_a,0],0,0))\right)\\
=&\beta V_\beta\left(\mathcal{T}([b_p,0,0],0,0)\right)-\beta V_\beta\left(\mathcal{T}([b_p,0,0],1,0)\right)+\beta\left(V_\beta(\mathcal{T}([0,b_a,0],1,0))-V_\beta(\mathcal{T}([0,b_a,0],0,0))\right)+O(h)\\
\le&\beta V_\beta\left([0,\tilde{b}_a,0]\right)-\beta V_\beta\left(\mathcal{T}([0,\tilde{b}_a,0],1,0)\right)+\beta\left(V_\beta(\mathcal{T}([0,b_a,0],1,0))-V_\beta(\mathcal{T}([0,b_a,0],0,0))\right)+O(h)\\
\le&-\beta(\bar{b}_a-\tilde{b}_a)c_5+O(h)<0,
\end{aligned}
$$

which raises a contradiction with $\mu_S([0,b_a,0])=1$. Here, the first equality is due to (A.1); the second equality is because $V_\beta(\cdot)$ is bounded; the first inequality is because $V_\beta\left(\mathcal{T}([b_p,0,0],0,0)\right)\le V_\beta\left([b_p,0,0]\right)=V_\beta\left([0,\tilde{b}_a,0]\right)$ and $\mathcal{T}([0,\tilde{b}_a,0],1,0)=\mathcal{T}([b_p,0,0],1,0)$; the second inequality is due to (A.2); the last inequality is because h is sufficiently small. Therefore, as β tends to 1 and h tends to 0, we have $b_a\le b_p+(1-b_p)p_p$.

Lastly, we provide a rigorous mapping from the optimal stationary policy to a triple-threshold policy (see Figure 3 for an illustration).

1. First, the customer enters the system with the state $\boldsymbol{b}_0=[1,0,0]$. According to Fact I, we have $b_p<1$ and hence the optimal policy will not push ads at time 0. Given no purchases, the belief $b_{t,1}$ gradually decreases and no ads are pushed until the belief state becomes $[b_p,0,0]$, which implies a positive threshold ω_0. Note that the optimal actions for states $[b,0,0]$ with $b<b_p$ do not affect this mapping because the system never attains these states.

2. At time ω_0, the optimal policy makes an ad push and the state becomes $[0,b_p+(1-b_p)p_p,0]$. According to Facts II and III, we can derive that $\sup\{b\in[0,b_p+(1-b_p)p_p]\mid\mu_S^B([0,b,0])=1\}=b_a$ and $b_a<b_p+(1-b_p)p_p$. Thus, the optimal policy does not make another ad push at time ω_0. Then, given no purchases, the belief $b_{t,2}$ decreases and no ads are pushed until the state becomes $[0,b_a,0]$, which implies a positive threshold ω_1.

3. At time $\omega_0+\omega_1$, the optimal policy makes an ad push and the state becomes $[0,b_a+(1-b_a)p_a,0]$. According to Facts II and III, we have $\sup\{b\in[0,b_a+(1-b_a)p_a]\mid\mu_S^B([0,b,0])\}=b_a$. Then, given no purchases, the belief $b_{t,2}$ decreases and no ads are pushed until the state becomes $[0,b_a,0]$, which implies a positive threshold ω_2. Note that ω_1 and ω_2 can be different if $b_p+(1-b_p)p_p\ne b_a+(1-b_a)p_a$.

4. At time $\omega_0+\omega_1+\omega_2$, the optimal policy makes an ad push and the state becomes $[0,b_a+(1-b_a)p_a,0]$. Similarly, given no purchases, the belief $b_{t,2}$ decreases and no ads are pushed until the state becomes $[0,b_a,0]$. In this case, the resulting threshold is the same as ω_2 because the starting belief is the same as that for ω_2. After that, the dynamics are similar and hence we have the same threshold ω_2 for the following process until a purchase happens.

5. If a purchase happens, we can collect a reward r. Then, the belief state becomes $[1,0,0]$ which is the same as the initial state, and the discussion is the same as the above. □

A.4. Proof of Proposition 1

If $p_p = p_a$, then we can deduce that the states $[b,0,0]$ and $[0,b,0]$ are equivalent, and hence $b_p^* = b_a^*$. According to the mapping in Appendix A.3, the threshold ω_0^* is induced by the belief evolvement from $[1,0,0]$ to $[0,b_p^*,0]$, the threshold ω_1^* is induced by the belief evolvement from $[0,b_p^*+(1-b_p^*)p_p,0]$ to $[0,b_a^*,0]$, and the threshold ω_2^* is induced by the belief evolvement from $[0,b_a^*+(1-b_a^*)p_a,0]$ to $[0,b_a^*,0]$. Since $b_p^*+(1-b_p^*)p_p \le 1$, $b_a^* = b_p^*$, and $[b,0,0]$ is equivalent to $[0,b,0]$ for any $b \in [0,1]$, we can deduce that $\omega_0^* \ge \omega_1^*$. Since $b_p^*+(1-b_p^*)p_p = b_a^*+(1-b_a^*)p_a$, we have $\omega_1^* = \omega_2^*$. Therefore, we have $\omega_0^* \ge \omega_1^* = \omega_2^*$. Furthermore, if $p_p = p_a = 1$, then we can deduce that $b_p^*+(1-b_p^*)p_p = 1$ and hence $\omega_0^* = \omega_1^* = \omega_2^*$. □

A.5. Computation Details of Equation (2)

To simplify the formulas, we define $\varsigma = \sqrt{(\lambda+\gamma+\theta)^2 - 4\lambda\theta}$, and $\sinh(\cdot)$ and $\cosh(\cdot)$ are the hyperbolic sine and cosine functions, respectively.

Before we start the computation, we first compute the cumulative distribution function $\rho_A(\cdot)$ of the first purchase time of an initially active customer. We consider an auxiliary continuous-time Markov chain $\{z_t, t \ge 0\}$ with three states: Active (A), Inactive (I) and Purchase (P). The transition graph is illustrated in Figure 11.

The corresponding transition rate matrix of Figure 11 is as follows:

$$G = \begin{bmatrix} -(\lambda+\gamma) & \gamma & \lambda \\ \theta & -\theta & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -\frac{\lambda+\gamma-\theta+\varsigma}{2\theta} & -\frac{\lambda+\gamma-\theta-\varsigma}{2\theta} & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{2\lambda\theta}{\varsigma-(\lambda+\gamma+\theta)} & 0 & 0 \\ 0 & -\frac{2\lambda\theta}{\varsigma+(\lambda+\gamma+\theta)} & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} -\frac{\lambda+\gamma-\theta+\varsigma}{2\theta} & -\frac{\lambda+\gamma-\theta-\varsigma}{2\theta} & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^{-1}.$$

We define τ_A as a stopping time $\tau_A = \min\{t \in \mathbb{R}_+ \mid z_t = P; z_0 = A\}$, which is the first purchase time of an initially active customer. Similarly, we can define the stopping time for an initially inactive customer (a customer with active probability b, resp.) as τ_I (τ, resp.). We can easily couple the auxiliary system with the initial system before time τ_A or τ_I. Then, we consider the policy that never pushes ads and compute some useful functions:

1. Starting from an active state, the c.d.f of the first purchase time is $\rho_A(t) = \mathbb{P}(N_p(t) \ge 1 \mid s_0 = A)$. Since P is the only absorbing state in the auxiliary system, we have $\mathbb{P}(\tau_A \le t) = \mathbb{P}(z_t = P \mid z_0 = A)$ and hence the cumulative density function of τ_A can be computed as

$$\begin{aligned}\rho_A(t) &= \mathbb{P}(\tau_A \le t \mid z_0 = A) = \mathbb{P}(z_t = P \mid z_0 = A) \\ &= \mathrm{Expm}(tG)_{1,3} \cdot \mathbb{1}(t \ge 0) = \left(1 - e^{-\frac{1}{2}(\gamma+\lambda+\theta)t}\left(\cosh\left(\frac{\varsigma t}{2}\right) + \frac{\gamma+\theta-\lambda}{\varsigma}\sinh\left(\frac{\varsigma t}{2}\right)\right)\right) \cdot \mathbb{1}(t \ge 0),\end{aligned}$$

where $\mathrm{Expm}(tG)_{i,j}$ is the (i,j) item of the matrix exponential of tG (see Norris 1998). Similarly, we have a similar c.d.f. for the first purchase time of an initially inactive customer, $\rho_I(t) = \mathbb{P}(N_p(t) \ge 1 \mid s_0 = I)$.

$$\begin{aligned}\rho_I(t) &= \mathbb{P}(\tau_I \le t \mid s_0 = I) = \mathbb{P}(z_t = P \mid z_0 = I) = \mathrm{Expm}(tG)_{2,3} \cdot \mathbb{1}(t \ge 0) \\ &= \left(1 - e^{-\frac{1}{2}(\gamma+\lambda+\theta)t}\left(\cosh\left(\frac{\varsigma t}{2}\right) + \frac{\lambda+\gamma+\theta}{\varsigma}\sinh\left(\frac{\varsigma t}{2}\right)\right)\right) \cdot \mathbb{1}(t \ge 0).\end{aligned}$$

Similarly, starting from an active probability b, the c.d.f. of the first purchase time is

$$\rho_B(b,t) = \mathbb{P}(z_0 = A)\mathbb{P}(z_t = P \mid z_0 = A) + \mathbb{P}(z_0 = I)\mathbb{P}(z_t = P \mid z_0 = I) = b\rho_A(t) + (1-b)\rho_I(t).$$

2. Furthermore, starting from an active probability b, given that $\tau \le t$, the conditional expectation of the first purchase time τ is

$$\Xi(b,t) = \mathbb{E}[\tau \mid \tau \le t, b_0 = b] = \frac{1}{\rho_B(b,t)} \int_0^t x d\rho_B(b,x).$$

3. Given no ads, the probability that an initially active customer makes no purchase and is active after t time units is

$$\chi_A(t) = \mathbb{P}(s_t = A, N_p(t) = 0 \mid s_0 = A) = \mathbb{P}(z_t = A \mid z_0 = A) = \mathrm{Expm}(tG)_{1,1} = \frac{\cosh\left(\frac{\varsigma t}{2}\right) - \frac{\gamma+\lambda-\theta}{\varsigma}\sinh\left(\frac{\varsigma t}{2}\right)}{e^{\frac{\lambda+\gamma+\theta}{2}t}}.$$

Similarly, the corresponding probability of an initially inactive customer is

$$\chi_I(t) = \mathbb{P}(s_t = A, N_p(t) = 0 \mid s_0 = I) = \mathbb{P}(x_t = A \mid x_0 = I) = \mathrm{Expm}(tG)_{2,1} = \frac{2\theta}{\varsigma} e^{-\frac{\lambda+\gamma+\theta}{2}t} \sinh\left(\frac{\varsigma t}{2}\right).$$

4. Given no ads and no purchases, the probability that a customer (starting with the active probability b) is active after t time units is

$$\begin{aligned}
\Gamma(b,t) &= \mathbb{P}(s_t = A \mid b_0 = b, N_p(t) = 0) = \frac{\mathbb{P}(s_t = A, b_0 = b, N_p(t) = 0)}{\mathbb{P}(b_0 = b, N_p(t) = 0)} \\
&= \frac{\mathbb{P}(s_t = A, N_p(t) = 0 \mid s_0 = A)\mathbb{P}(s_0 = A \mid b_0 = b) + \mathbb{P}(s_t = A, N_p(t) = 0 \mid s_0 = I)\mathbb{P}(s_0 = I \mid b_0 = b)}{\mathbb{P}(N_p(t) = 0 \mid s_0 = A)\mathbb{P}(s_0 = A \mid b_0 = b) + \mathbb{P}(N_p(t) = 0 \mid s_0 = I)\mathbb{P}(s_0 = I \mid b_0 = b)} \\
&= \frac{b \cdot \chi_A(t) + (1-b) \cdot \chi_I(t)}{b(1-\rho_A(t)) + (1-b)(1-\rho_I(t))}.
\end{aligned}$$

Then, we can compute the expected cycle profit and cycle length.

$$\begin{aligned}
\mathbb{E}[\pi(\boldsymbol{\omega})] &= \rho_A(\omega_0) \cdot r + (1-\rho_A(\omega_0)) \cdot \rho_B\left(b_p + (1-b_p)p_p, \omega_1\right) \cdot (r-c) \\
&\quad + \sum_{i=2}^{\infty} (1-\rho_A(\omega_0)) \cdot \left(1 - \rho_B\left(b_p + (1-b_p)p_p, \omega_1\right)\right) \cdot \left(1 - \rho_B\left(b_a + (1-b_a)p_a, \omega_2\right)\right)^{i-2} \cdot \rho_B\left(b_a + (1-b_a)p_a, \omega_2\right) \cdot (r - ic) \\
&= r - (1-\rho_A(\omega_0))c - \frac{(1-\rho_A(\omega_0))\left(1 - (b_p + (1-b_p)p_p)\rho_A(\omega_1) - (1-b_p)(1-p_p)\rho_I(\omega_1)\right)}{(b_a + (1-b_a)p_a)\rho_A(\omega_2) + (1-b_a)(1-p_a)\rho_I(\omega_2)} c \\
\mathbb{E}[\kappa(\boldsymbol{\omega})] &= \rho_A(\omega_0) \cdot \mathbb{E}[\tau \mid \tau < \omega_0] + (1-\rho_A(\omega_0)) \cdot \rho_B\left(b_p + (1-b_p)p_p, \omega_1\right) \cdot \mathbb{E}[\tau \mid \tau \in [\omega_0, \omega_0 + \omega_1)] \\
&\quad + \sum_{i=2}^{\infty} (1-\rho_A(\omega_0)) \cdot \left(1 - \rho_B\left(b_p + (1-b_p)p_p, \omega_1\right)\right) \cdot \left(1 - \rho_B\left(b_a + (1-b_a)p_a, \omega_2\right)\right)^{i-2} \\
&\qquad \times \rho_B\left(b_a + (1-b_a)p_a, \omega_2\right) \cdot \mathbb{E}[\tau \mid \tau \in [\omega_0 + \omega_1 + (i-2)\omega_2, \omega_0 + \omega_1 + (i-1)\omega_2)] \\
&= \rho_A(\omega_0) \cdot \Xi(1, \omega_0) + (1-\rho_A(\omega_0)) \cdot \rho_B\left(b_p + (1-b_p)p_p, \omega_1\right) \cdot \left(\omega_0 + \Xi\left(b_p + (1-b_p)p_p, \omega_1\right)\right) \\
&\quad + \sum_{i=2}^{\infty} (1-\rho_A(\omega_0)) \cdot \left(1 - \rho_B\left(b_p + (1-b_p)p_p, \omega_1\right)\right) \cdot \left(1 - \rho_B\left(b_a + (1-b_a)p_a, \omega_2\right)\right)^{i-2} \\
&\qquad \times \rho_B\left(b_a + (1-b_a)p_a, \omega_2\right) \cdot \left(\omega_0 + \omega_1 + (i-2)\omega_2 + \Xi\left(b_a + (1-b_a)p_a, \omega_2\right)\right) \\
&= \rho_A(\omega_0) \cdot \Xi(1, \omega_0) + (1-\rho_A(\omega_0))\omega_0 + (1-\rho_A(\omega_0))\left((b_p + (1-b_p)p_p)\rho_A(\omega_1) + (1-b_p)(1-p_p)\rho_I(\omega_1)\right)\Xi(b_p + (1-b_p)p_p, \omega_1) \\
&\quad + (1-\rho_A(\omega_0))\left(1 - (b_p + (1-b_p)p_p)\rho_A(\omega_1) - (1-b_p)(1-p_p)\rho_I(\omega_1)\right)\left(\omega_1 + \Xi(b_a + (1-b_a)p_a, \omega_2)\right) \\
&\quad + (1-\rho_A(\omega_0))\left(1 - (b_p + (1-b_p)p_p)\rho_A(\omega_1) - (1-b_p)(1-p_p)\rho_I(\omega_1)\right) \cdot \frac{1 - (b_a + (1-b_a)p_a)\rho_A(\omega_2) - (1-b_a)(1-p_a)\rho_I(\omega_2)}{(b_a + (1-b_a)p_a)\rho_A(\omega_2) + (1-b_a)(1-p_a)\rho_I(\omega_2)} \omega_2.
\end{aligned}$$

Then, the long-term average expected profit is $\Psi(\boldsymbol{\omega}) = \mathbb{E}[\pi(\boldsymbol{\omega})] / \mathbb{E}[\kappa(\boldsymbol{\omega})]$. Moreover, we have the relationship between $\boldsymbol{\omega}$ and (b_p, b_a) as follows:

$$\Gamma(1, \omega_0) = b_p \qquad \Gamma(b_p + (1-b_p)p_p, \omega_1) = b_a \qquad \Gamma(b_a + (1-b_a)p_a, \omega_2) = b_a,$$

which implies that

$$\omega_0 = \frac{1}{\varsigma}\log\left(\frac{(b_p+1)\gamma+(1-b_p)(\lambda-\theta+\varsigma)}{(b_p+1)\gamma-(1-b_p)(\theta-\lambda+\varsigma)}\right)$$
$$\omega_1 = \frac{1}{\varsigma}\log\left(\frac{b_a(\lambda+\gamma+\theta-\varsigma)+(b_p+(1-b_p)p_p)(\lambda-2b_a\lambda+\gamma+\theta+\varsigma)-2\theta}{b_a(\lambda+\gamma+\theta+\varsigma)+(b_p+(1-b_p)p_p)(\lambda-2b_a\lambda+\gamma+\theta-\varsigma)-2\theta}\right)$$
$$\omega_2 = \frac{1}{\varsigma}\log\left(\frac{b_a(\lambda+\gamma+\theta-\varsigma)+(b_a+(1-b_a)p_a)(\lambda-2b_a\lambda+\gamma+\theta+\varsigma)-2\theta}{b_a(\lambda+\gamma+\theta+\varsigma)+(b_a+(1-b_a)p_a)(\lambda-2b_a\lambda+\gamma+\theta-\varsigma)-2\theta}\right).$$

Then, we can plug the above expressions into $\Psi(\boldsymbol{\omega})$, solve the problem to get the optimal belief thresholds (b_p^*, b_a^*), and then derive the corresponding optimal thresholds $\boldsymbol{\omega}^*$. □

A.6. Proof of Lemma 1

Any policy μ feasible to (5) is also a feasible policy to (6). Therefore, we have $\bar{V}(T) \geq V(T)$. □

A.7. Proof of Proposition 3

For any Lagrangian multiplier $\ell \geq 0$, we have an upper bound of $\bar{V}_\infty$.

$$\bar{V}_\infty \leq V_\infty^\ell = \max_{\mu\in\Pi_\infty}\left\{\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{j=1}^{M}\left(r_j N_{p,j}^\mu(T) - c_j N_{a,j}^\mu(T)\right)\right] + \ell\left(B - \limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{j=1}^{M} c_j N_{a,j}^\mu(T)\right]\right)\right\}$$
$$= \max_{\mu\in\Pi_\infty}\sum_{j=1}^{M}\left\{\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\left(r_j N_{p,j}^\mu(T) - (1+\ell)c_j N_{a,j}^\mu(T)\right)\right]\right\} + \ell B.$$

The problem V_∞^ℓ can be decomposed into solving the optimal policy for each customer j with the parameter set $\left((1+\ell)c_j, r_j, \lambda_j, \gamma_j, \theta_j\right)$. According to Theorem 2, the optimal policy for each customer is a triple-threshold policy. Given the multiplier ℓ, let $\boldsymbol{\omega}_j(\ell)$ denote the optimal thresholds of customer j, and $\mu(\ell)$ denote the decentralized triple-threshold policy that implements the triple-threshold policy $\varphi^{\boldsymbol{\omega}_j(\ell)}$ to each customer j. In addition, the average expected cost under such a decentralized triple-threshold policy is

$$C(\ell) = \limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{j=1}^{M} c_j N_{a,j}^{\mu(\ell)}(T)\right] = \sum_{j=1}^{M}\frac{\mathbb{E}\left[\pi_j\left(\boldsymbol{\omega}_j(\ell)\right)\right]}{\mathbb{E}\left[\kappa_j\left(\boldsymbol{\omega}_j(\ell)\right)\right]}.$$

Let $R_\infty^{\mu(\ell)}$ denote the long-term average expected profit under the policy $\mu(\ell)$). Then, we consider two cases and prove the optimality of the decentralized triple-threshold policy.

1. $C(0) > B$. Since $\lim_{\ell\to\infty} C(\ell) = 0$, there exists a finite and positive constant ℓ^* such that $C(\ell^*) = B$. Then, the policy $\mu(\ell^*)$ is feasible and we can derive the following equality.

$$R_\infty^{\mu(\ell^*)} \leq \bar{V}_\infty \leq V_\infty^{\ell^*} = \max_{\mu\in\Pi_\infty}\left\{\liminf_{T\to\infty}\frac{1}{T}\mathbb{E}\left[\sum_{j=1}^{M}\left(r\cdot N_{p,j}^\mu(T) - (1+\ell^*)c\cdot N_{a,j}^\mu(T)\right)\right]\right\} + \ell^* B$$
$$= R_\infty^{\mu(\ell^*)} - \ell^* C(\ell^*) + \ell^* B = R_\infty^{\mu(\ell^*)},$$

   which implies that $R_\infty^{\mu(\ell^*)} = \bar{V}_\infty$, and hence the decentralized triple-threshold policy $\mu(\ell^*)$ is optimal. Moreover, we also refer to ℓ^* as the optimal Lagrangian multiplier.

2. $C(0) \leq B$. In this case, we can deduce that $\mu(0)$ is an admissible policy satisfying the constraint in (7). Then, similarly, we can deduce that $R_\infty^{\mu(\ell^*)} = \bar{V}_\infty$ with $\ell^* = 0$, and hence the decentralized triple-threshold policy $\mu(0)$ is optimal.

Therefore, there always exists an optimal policy to (7) that is a decentralized triple-threshold policy. □

A.8. Proof of Lemma 2

Similar to the proof in Appendix A.7, we also consider two cases.

1. When $\sum_{j=1}^{M} q_j(\boldsymbol{\omega}_j^u) \le B$, we have $C(0) \le B$ and hence $\boldsymbol{\omega}_j^u = \boldsymbol{\omega}_j(0)$ are the optimal thresholds for customer j.
2. When $\sum_{j=1}^{M} q_j(\boldsymbol{\omega}_j^u) > B$, we have $C(0) > B$ and hence there exists a constant ℓ^* such that $C(\ell) = B$ and $\boldsymbol{\omega}_j(\ell^*)$ are the optimal thresholds for customer j. □

A.9. Proof of Theorem 3

In this proof, we use superscripts (e.g., $\mathcal{A}$) to denote the policy used. First, according to Lemma 1, we have $V(T) \le \bar{V}(T)$, and hence the inequality holds. Then, we can decompose the optimality gap into two components as follows:

$$V(T) - V^{\mathcal{A}}(T) \le \bar{V}(T) - V^{\mathcal{A}}(T) = \underbrace{\bar{V}(T) - \bar{V}_\infty}_{\Delta_1} + \underbrace{\bar{V}_\infty - V^{\mathcal{A}}(T)}_{\Delta_2}.$$

Step 1. Bound $\bar{V}(T) - \bar{V}_\infty$.

Similar to the Lagrangian relaxation in Appendix A.7, we have

$$\begin{aligned}
\bar{V}(T) - \bar{V}_\infty &= \bar{V}(T) - \max_{\mu_\infty \in \Pi_\infty} \left\{ \liminf_{t\to\infty} \frac{1}{t} \mathbb{E}\left[\sum_{j=1}^{M} \left(r_j \cdot N_{p,j}^{\mu_\infty}(t) - c_j \cdot N_{a,j}^{\mu_\infty}(t) \right) \right] + \ell^* \left(B - \liminf_{t\to\infty} \frac{1}{t} \mathbb{E}\left[\sum_{j=1}^{M} c_j N_{a,j}^{\mu_\infty}(t) \right] \right) \right\} \\
&\le \max_{\mu_T \in \Pi_T} \left\{ \frac{1}{T} \mathbb{E}\left[\sum_{j=1}^{M} \left(r_j \cdot N_{p,j}^{\mu_T}(T) - c_j \cdot N_{a,j}^{\mu_T}(T) \right) \right] + \ell^* \left(B - \frac{1}{T} \mathbb{E}\left[\sum_{j=1}^{M} c_j N_{a,j}^{\mu_T}(T) \right] \right) \right\} \\
&\quad - \max_{\mu_\infty \in \Pi_\infty} \left\{ \liminf_{t\to\infty} \frac{1}{t} \mathbb{E}\left[\sum_{j=1}^{M} \left(r_j \cdot N_{p,j}^{\mu_\infty}(t) - c_j \cdot N_{a,j}^{\mu_\infty}(t) \right) \right] + \ell^* \left(B - \liminf_{t\to\infty} \frac{1}{t} \mathbb{E}\left[\sum_{j=1}^{M} c_j N_{a,j}^{\mu_\infty}(t) \right] \right) \right\} \\
&= \max_{\mu_T \in \Pi_T} \left\{ \frac{1}{T} \mathbb{E}\left[\sum_{j=1}^{M} \left(r_j \cdot N_{p,j}^{\mu_T}(T) - (1+\ell^*) c_j \cdot N_{a,j}^{\mu_T}(T) \right) \right] \right\} \\
&\quad - \max_{\mu_\infty \in \Pi_\infty} \left\{ \liminf_{t\to\infty} \frac{1}{t} \mathbb{E}\left[\sum_{j=1}^{M} \left(r_j \cdot N_{p,j}^{\mu_\infty}(t) - (1+\ell^*) c_j \cdot N_{a,j}^{\mu_\infty}(t) \right) \right] \right\}, \qquad \text{(A.3)}
\end{aligned}$$

where $\ell^* \ge 0$ denotes the optimal Lagrangian multiplier for the problem $\bar{V}_\infty$ in Appendix A.7, and is a finite constant. Since the resource constraint is relaxed in (A.3), we can focus on each customer independently. Consider a representative customer with the parameter set $\left((1+\ell^*)c_j, r_j, \lambda_j, \gamma_j, \theta_j, p_{p,j}, p_{a,j}\right)$. To simplify the notation, we temporarily omit the subscript 'j'.

We recall some results in Appendix A.2. For the discretized problem with a sufficiently small h and any discount factor $\beta \in (0,1)$, we can bound $V_\beta(\boldsymbol{b}_1) - V_\beta(\boldsymbol{b}_2) \le V_\beta([0,0,1]) - V_\beta(\boldsymbol{b}_2) \le \frac{4r}{c_1} + \left(\frac{16c_2}{c_1^2 \lambda} + 1 \right) c$ for any $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \mathbb{B}$. As $h \to 0$ and $\beta \to 1$, we can deduce that the total profit difference between two systems with different initial belief states $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ is upper bounded by a constant $\tilde{C}$, which is independent of T. Since an ad campaign just changes the belief state from $\tilde{\boldsymbol{b}}$ to $\tilde{\boldsymbol{b}}'$, the additional profit brought by an ad campaign is upper bounded by $\tilde{C}$. Similarly, if we consider an additional "efficient-activation" action that can change the

belief state to $[1,0,0]$ with certainty, its contribution to the total profit is also upper bounded by $\tilde{C}$. Then, consider an alternative POMDP, which is the same as the base model except for an additional "efficient-activation" action with a cost $2\tilde{C}$. Specifically, the action set becomes $\mathbb{A}' = \{0,1,2\}$ with action 2 representing the efficient-activation action, and the reward function becomes $R(x,a) = \mathbb{1}\{x = P\} \cdot r - \mathbb{1}\{a = 1\} \cdot c - \mathbb{1}\{a = 2\} \cdot 2\tilde{C}$. Since the cost of the efficient-activation action is greater than its maximal contribution to total profit, the policy for the initial POMDP is still optimal for the auxiliary POMDP, and hence their objective values are the same. The above statements can be similarly generalized to the finite-horizon case.

Then, we focus on the auxiliary POMDP. Let μ_T^* denote the optimal policy for the finite horizon problem. Then, to bridge the finite-horizon problem with the infinite-horizon problem, we construct a feasible policy μ_T^∞ for the infinite-horizon problem (see Figure 6 for an illustration): For each $k \in \mathbb{Z}_+$, the policy μ_T^∞ implements the policy μ_T^* during the interval $((k-1)T, kT)$ and takes the efficient-activation action at time kT. The long-term average expected profit of such a policy is

$$\liminf_{t\to\infty} \frac{1}{t}\mathbb{E}\left[r \cdot N_p^{\mu_T^\infty}(t) - (1+\ell^*)c \cdot N_a^{\mu_T^\infty}(t)\right] = \frac{1}{T}\mathbb{E}\left[r \cdot N_p^{\mu_T}(T) - (1+\ell^*)c \cdot N_a^{\mu_T}(T)\right] - \frac{2\tilde{C}}{T}.$$

Subsequently, we can derive the following bound

$$\begin{aligned}
&\max_{\mu_T \in \Pi_T}\left\{\frac{1}{T}\mathbb{E}\left[r \cdot N_p^{\mu_T}(T) - (1+\ell^*)c \cdot N_a^{\mu_T}(T)\right]\right\} - \max_{\mu_\infty \in \Pi_\infty}\left\{\liminf_{t\to\infty} \frac{1}{t}\mathbb{E}\left[r \cdot N_p^{\mu_\infty}(t) - (1+\ell^*)c \cdot N_a^{\mu_\infty}(t)\right]\right\} \\
=&\frac{1}{T}\mathbb{E}\left[r \cdot N_p^{\mu_T}(T) - (1+\ell^*)c \cdot N_a^{\mu_T}(T)\right] - \max_{\mu_\infty \in \Pi_\infty}\left\{\liminf_{t\to\infty} \frac{1}{t}\mathbb{E}\left[r \cdot N_p^{\mu_\infty}(t) - (1+\ell^*)c \cdot N_a^{\mu_\infty}(t)\right]\right\} \\
\le&\frac{1}{T}\mathbb{E}\left[r \cdot N_p^{\mu_T}(T) - (1+\ell^*)c \cdot N_a^{\mu_T}(T)\right] - \liminf_{t\to\infty} \frac{1}{t}\mathbb{E}\left[r \cdot N_p^{\mu_T^\infty}(t) - (1+\ell^*)c \cdot N_a^{\mu_T^\infty}(t)\right] \\
=&\frac{2\tilde{C}}{T} = O(\frac{1}{T}).
\end{aligned}$$

Similar to the above argument, for each customer j, we can derive an upper bound $\tilde{C}_j$ for the total revenue contribution of any action. Then, we derive the following result

$$\Delta_1 = \bar{V}(T) - \bar{V}_\infty \le \sum_{j=1}^{M} \frac{2\tilde{C}_j}{T} = O(1/T).$$

Step 2. Bound $\bar{V}_\infty - V^{\mathcal{A}}(T)$.

According to Proposition 3, there exists an optimal policy for (7) that is a decentralized triple-threshold policy. Let $\boldsymbol{\omega}_j^*$ denote the optimal thresholds of customer j. Thus, we have $\bar{V}_\infty = \sum_{j=1}^M \mathbb{E}\left[\pi_j(\boldsymbol{\omega}_j^*)\right] / \mathbb{E}\left[\kappa_j(\boldsymbol{\omega}_j^*)\right]$.

In the following, we study the profit under the BAT policy, i.e., $V^{\mathcal{A}}(T)$. First, we focus on customer j. If $\pi_j(\boldsymbol{\omega}_j^*) = 0$, then the policy never implements ad campaigns, resulting in zero optimality gap. Thus, we focus on case when $\pi_j(\boldsymbol{\omega}_j^*) > 0$. For ease of exposition, we temporarily omit the subscript 'j' and the superscript '$*$'.

We consider a sequence of cycles with their random cycle profits X_k's and random cycle length Y_k's. According to the computation in Appendix A.5, we define $p_1 = \rho(\omega_0)$, $p_2 = (b_p + (1-b_p)p_p)\rho_A(\omega_1) + (1-b_p)(1-p_p)\rho_I(\omega_1)$ and $p_3 = (b_a + (1-b_a)p_a)\rho_A(\omega_2) + (1-b_a)(1-p_a)\rho_I(\omega_2)$. First, we provide concentration bounds for these random variables (see Appendix EC.4 for details). We can prove that X_k (Y_k, resp.) is a

sub-exponential variable (see Papaspiliopoulos 2020) with a constant K_X (K_Y, resp.). Then, we focus on the first n cycles, where

$$n := \left\lfloor \frac{T}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} - \epsilon \right\rfloor, \qquad \epsilon = \max\left\{ \sqrt{\frac{4K_Y^2 T \log T}{(\mathbb{E}[\kappa(\boldsymbol{\omega})])^3}}, \sqrt{\frac{16K_X^2 T \log T}{(\mathbb{E}[\kappa(\boldsymbol{\omega})])\,(r - \mathbb{E}[\pi(\boldsymbol{\omega})])^2}} \right\}.$$

Using the Chernoff bound, we can prove that the probability that either one of the following conditions happens is no greater than $1/T$:

1. The total cycle length of the first n cycles is greater than T.
2. The total profit of the first n cycles is less than $nr - qT$.

Then, we define the good event as the case when the above two events do not happen, and the probability of the good event is no less than $1 - \frac{2}{T}$. Under the good event, we have

1. The system does not terminate before the end of the n-th cycle, i.e., $\sum_{k=1}^{n} Y_k \le T$.
2. The budget is not exhausted at the end of the n-th cycle, i.e., $nr - \sum_{k=1}^{n} Y_k \le qT$.
3. The total reward during the first n cycles is nr.

Furthermore, the total ad cost is no greater than qT due to the allocated budget. Therefore, for the single-customer case, we can provide an upper bound for $\bar{V}_\infty - V^{\mathcal{A}}(T)$ as follows:

$$\begin{aligned}
\bar{V}_\infty - V^{\mathcal{A}}(T) &\le \frac{2}{T}\frac{\mathbb{E}[\pi(\boldsymbol{\omega})]}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} + \left(1 - \frac{2}{T}\right)\left(\frac{\mathbb{E}[\pi(\boldsymbol{\omega})]}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} - \frac{nr - qT}{T}\right) \\
&\le \frac{2}{T}\frac{\mathbb{E}[\pi(\boldsymbol{\omega})]}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} + \left(\frac{\mathbb{E}[\pi(\boldsymbol{\omega})]}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} - \frac{r}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} + \frac{(\epsilon+1)r}{T} + \frac{r - \mathbb{E}[\pi(\boldsymbol{\omega})]}{\mathbb{E}[\kappa(\boldsymbol{\omega})]}\right) \\
&= \frac{2}{T}\frac{\mathbb{E}[\pi(\boldsymbol{\omega})]}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} + \frac{(\epsilon+1)r}{T} = O(1/T) + O\left(\sqrt{\frac{\log T}{T}}\right) = O\left(\sqrt{\frac{\log T}{T}}\right),
\end{aligned}$$

where the first inequality is because the probability of the good event is no less than $1 - \frac{2}{T}$, the second inequality is because $n \ge \frac{T}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} - \epsilon - 1$ and $q = \frac{\mathbb{E}[r - \pi(\boldsymbol{\omega})]}{\mathbb{E}[\kappa(\boldsymbol{\omega})]}$. For the multiple-customer case, we just take the summation of all M customers' optimality gaps, resulting in an $O(\log T/\sqrt{T})$ total optimality gap.

To summarize, for the general case, we have

$$V(T) - V^{\mathcal{A}}(T) \le \Delta_1 + \Delta_2 \le O(1/T) + O\left(\sqrt{\frac{\log T}{T}}\right) = O\left(\sqrt{\frac{\log T}{T}}\right).$$

Moreover, we can provide a better bound for the efficient-activation case, where an ad campaign can activate customers with certainty, i.e., $p_p = p_a = 1$. In this case, the triple-threshold policy degenerates to a single-threshold policy and we use ω_j^* to denote the optimal threshold for customer j. The analysis of Δ_1 is the same as Step 1 and we can provide an upper bound $O(1/T)$, while the analysis of Δ_2 is different. Specifically, we further decompose Δ_2 as follows:

$$\Delta_2 = \bar{V}_\infty - V^{\mathcal{A}}(T) = \left(\bar{V}_\infty - R^{\mathcal{A}_B}(T)\right) + \left(R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T)\right),$$

where $\mathcal{A}_B$ is the policy that implements the decentralized threshold policy in $\mathcal{A}$ ignoring the budget constraint, and $R^{\mathcal{A}_B}(T)$ is the average expected profit under such a policy. In the following, we provide upper bounds for the above two terms.

First, we study the first term, $\bar{V}_\infty - R^{\mathcal{A}_B}(T)$. Since both policies treat customers separately, we can focus on a representative customer and omit the subscript 'j'. Note that each ad campaign can certainly activate inactive customers in this case. Similar to Step 1, to bridge the finite-horizon problem and the infinite-horizon problem, we consider an infinite-horizon policy $\mathcal{A}_B^\infty$: For each $k \in \mathbb{Z}_+$, the policy $\mathcal{A}_B^\infty$ implements the policy $\mathcal{A}_B$ during the interval $((k-1)T, kT)$ and delivers an ad campaign at time kT. The long-term average profit of such a policy is

$$\liminf_{t\to\infty} \frac{1}{t}\mathbb{E}\left[r \cdot N_p^{\mathcal{A}_B^\infty}(t) + c \cdot N_a^{\mathcal{A}_B^\infty}(t)\right] = \frac{1}{T}\mathbb{E}\left[r \cdot N_p^{\mathcal{A}_B}(t) + c \cdot N_a^{\mathcal{A}_B}(t)\right] - \frac{c}{T}. \tag{A.4}$$

Then, we compare the constructed policy $\mathcal{A}_B^\infty$ and the optimal threshold policy for $\bar{V}_\infty$, referred to as μ_∞^*. Given the time horizon t, we consider two systems. The policies implemented in systems A and B are μ_∞^* and $\mathcal{A}_B^\infty$, respectively. As mentioned in Appendix EC.3.1, we can redefine each interval between two adjacent events (purchases or ads) as a cycle and these cycles have i.i.d. cycle profits and cycle lengths. We couple the two systems as follows:

1. Initialize system A with an array $\mathcal{I}$. After each purchase or ad campaign, we sample the first purchase time $\tau_k \sim \rho_A(\cdot)$ and add τ_k to $\mathcal{I}$. If $\tau_k \le \omega^*$, then a purchase happens; if $\tau_k > \omega^*$, an ad campaign is delivered.

2. Then, we run system B. After each purchase or ad campaign, we pop the purchase time with the lowest index in $\mathcal{I}$. If the set $\mathcal{I}$ is empty, then we sample a purchase time according to the distribution $\rho_A(\cdot)$. Similarly, if an ad campaign is delivered before the purchase time, then an ad cost is incurred; otherwise, a purchase happens.

By this coupling argument, system B replicates most cycles in system A except for cycles corresponding to ad campaigns at time points $\{kT, k = 1, 2, \ldots, \lfloor t/T \rfloor\}$ and the newly-generated cycles because $\mathcal{I}$ becomes empty. First, an early ad campaign may transform a purchase cycle into an ad campaign cycle, resulting in a $(r+c)$ difference and a reduction in the cycle length by at most ω^*. Second, since the remaining time when $\mathcal{I}$ is found empty is at most $\lfloor t/T \rfloor \cdot \omega^*$, the total profit of the newly-generated cycles is no less than $-\lfloor t/T \rfloor \cdot c$. In this case, we can bound the profit difference as

$$\frac{1}{t}\mathbb{E}\left[r \cdot N_p^{\mu_\infty^*}(t) + c \cdot N_a^{\mu_\infty^*}(t)\right] - \frac{1}{t}\mathbb{E}\left[r \cdot N_p^{\mathcal{A}_B^\infty}(t) + c \cdot N_a^{\mathcal{A}_B^\infty}(t)\right] \le \frac{1}{t}\left(\left\lfloor \frac{t}{T} \right\rfloor \cdot (r+c) + \left\lfloor \frac{t}{T} \right\rfloor \cdot c\right) \le \frac{r+2c}{T}. \tag{A.5}$$

Therefore, combining (A.4) and (A.5), we can deduce that

$$\bar{V}_\infty - R^{\mathcal{A}_B}(T) \le \sum_j \left(\frac{c_j}{T} + \frac{r_j + 2c_j}{T}\right) = O\left(\frac{1}{T}\right).$$

Then, we study the second term, $R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T)$. Since both policies treat customers separately, we can focus on a representative customer and omit the subscript 'j'. These two policies are the same until the allocated budget is exhausted (or the remaining budget is less than c). We define each interval between two adjacent ad campaigns as an "ad cycle". Let ν_k denote the random length of the k-th ad cycle. Similar to the cycles for the efficient-activation case, ad cycles also have i.i.d. cycle profits and cycle lengths, and we can derive that

$$\frac{c}{\mathbb{E}[\nu]} = \frac{\mathbb{E}[\psi(\omega^*)]}{\mathbb{E}[\kappa(\omega^*)]} = q,$$

where ν shares the same distribution with ν_k's. Then, we can define $T_A = \sum_{k=1}^{\lfloor qT/c \rfloor} \nu_k$ and derive the time $\tau := \min\{T, T_A\}$ when the allocated budget is exhausted. Moreover, we can deduce that

$$\mathbb{E}[T_A] = \left\lfloor \frac{qT}{c} \right\rfloor \mathbb{E}[\nu] \le T, \qquad Var(T_A) = \left\lfloor \frac{qT}{c} \right\rfloor Var(\nu) = O(T).$$

Since the system dynamics can be easily coupled before time T_A, we can bound the term $R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T)$ for the representative customer by the profit difference during the period $[T_A, T]$:

$$\begin{aligned} R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T) \le \frac{1}{T}\mathbb{E}\left[\frac{\max\{T - T_A, 0\}}{\omega^*} \times \tilde{C}\right] &\le \frac{\sqrt{(T - \mathbb{E}[T_A])^2 + Var(T_A)} + (T - \mathbb{E}[T_A])}{2T\omega^*} \times \tilde{C} \\ &\le \frac{2c/q + \sqrt{Var(T_A)}}{2T\omega^*} \times \tilde{C} \\ &= O(1/\sqrt{T}), \end{aligned}$$

where the first inequality is because an ad campaign can only be delivered after ω^* times units and an ad campaign brings at most $\tilde{C}$ profit, the second inequality is due to Proposition 1 in Gallego (1992), and the last inequality is because $\sqrt{a^2 + b^2} \le |a| + |b|$ and $T - \mathbb{E}[T_A] \le \mathbb{E}[\nu] = c/q$. For the multi-customer case, we just need to take the summation of the optimality gaps of M customers, and hence we have $R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T) \le O(1/\sqrt{T})$.

Therefore, we can derive that

$$\Delta_2 = \bar{V}_\infty - V^{\mathcal{A}}(T) = \left(\bar{V}_\infty - R^{\mathcal{A}_B}(T)\right) + \left(R^{\mathcal{A}_B}(T) - V^{\mathcal{A}}(T)\right) \le O\left(\frac{1}{T}\right) + O\left(\frac{1}{\sqrt{T}}\right) = O\left(\frac{1}{\sqrt{T}}\right),$$

which implies that $V(T) - V^{\mathcal{A}}(T) \le O\left(1/\sqrt{T}\right)$ for the efficient-activation case. □

A.10. Proof of Proposition 4

The proof is similar to Step 2 in Appendix A.9, except that the distributions of cycle profits and cycle lengths are different. Similar to Appendix EC.4, we can still prove the sub-exponential properties, then we can use Chernoff bounds for the sub-exponential variables to derive the high-probability event. Then, we use the same methodology to derive the $O(\sqrt{\log T/T})$ regret bound for the general and the $O(\sqrt{1/T})$ regret bound for the efficient-activation case. □

Appendix B: Extension: Negative Effect of Ad Campaigns

In the base model, we assume that customers in the active state are not affected by ad campaigns. However, as mentioned in Footnote 2, given an ad campaign, an active customer may become inactive with some deactivation probability. In this section, we extend the base model to incorporate this negative effect. In particular, the deactivation probability is q_p (q_a, resp.) if the last event is a purchase (an ad, resp.). We set $q_p \le q_a$ to capture the negative effect of too many ad campaigns. Moreover, we assume that $p_q + q_q < 1$ and $p_a + q_a < 1$ such that both $b(1 - q_p) + (1 - b)p_p$ and $b(1 - q_a) + (1 - b)p_a$ increase in b.

In this case, we only need to replace the activated beliefs $b + (1 - b)p_p$ and $b + (1 - b)p_a$ with $b(1 - q_p) + (1 - b)p_p$ and $b(1 - q_a) + (1 - b)p_a$, respectively. Specifically, we update the functions $\mathcal{T}(\boldsymbol{b}, a, y)$ as follows:

$$\mathcal{T}([b_1, 0, 0], 0, 0) = \left[\frac{b_1(1 - \lambda h - \gamma h) + (1 - b_1)\theta h}{b_1(1 - \lambda h) + 1 - b_1}, 0, 0\right], \quad \mathcal{T}([0, b_2, 0], 0, 0) = \left[0, \frac{b_2(1 - \lambda h - \gamma h) + (1 - b_2)\theta h}{b_2(1 - \lambda h) + 1 - b_2}, 0\right]$$

$$\mathcal{T}([b_1,0,0],1,0)=\left[0,\frac{\left(b_1(1-q_q)+(1-b_1)p_p\right)(1-\lambda h-\gamma h)+\left((1-b_1)(1-p_p)+b_1q_q\right)\theta h}{\left(b_1(1-q_p)+(1-b_1)p_p\right)(1-\lambda h)+((1-b_1)(1-p_p)+b_1q_q)},0\right],$$

$$\mathcal{T}([0,b_2,0],1,0)=\left[0,\frac{\left(b_2(1-q_a)+(1-b_2)p_a\right)(1-\lambda h-\gamma h)+\left((1-b_2)(1-p_a)+b_2q_a\right)\theta h}{\left(b_2(1-q_a)+(1-b_2)p_a\right)(1-\lambda h)+\left((1-b_2)(1-p_a)+b_2q_a\right)},0\right],$$

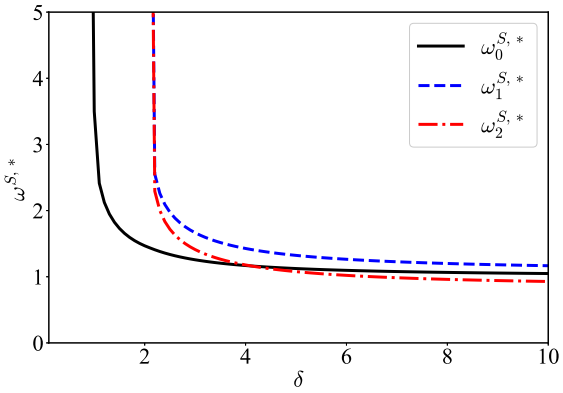
$$\mathcal{T}([0,0,1],0,0)=\left[\frac{1-\lambda h-\gamma h}{1-\lambda h},0,0\right],\quad \mathcal{T}([0,0,1],1,0)=\left[0,\frac{1-\lambda h-\gamma h}{1-\lambda h},0\right]$$

$$\mathcal{T}(\boldsymbol{b},a,1)=[0,0,1]\qquad \forall \boldsymbol{b},a.$$

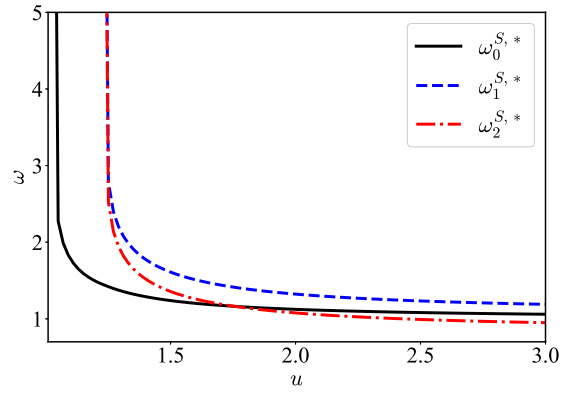
The Bellman equation for the discounted problem is updated as follows:

$$\begin{aligned}V_\beta(\boldsymbol{b})=\max\Big\{&b_3r+\beta(b_1+b_2+b_3)\lambda h\cdot V_\beta\left(\mathcal{T}(\boldsymbol{b},0,1)\right)+\beta\left(1-(b_1+b_2+b_3)\lambda h\right)\cdot V_\beta\left(\mathcal{T}(\boldsymbol{b},0,0)\right),\\ &b_3r-c+\beta\left(b_1(1-q_p)+(1-b_1)p_p+b_2(1-q_a)+(1-b_2)p_a+b_3\right)\lambda h\cdot V_\beta\left(\mathcal{T}(\boldsymbol{b},1,1)\right)\\ &\quad+\beta\left(1-\left(b_1(1-q_p)+(1-b_1)p_p+b_2(1-q_a)+(1-b_2)p_a+b_3\right)\lambda h\right)\cdot V_\beta\left(\mathcal{T}(\boldsymbol{b},1,0)\right)\Big\}.\end{aligned}$$

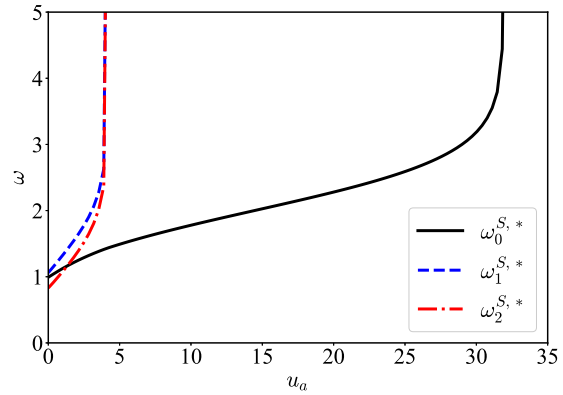
Then, given that $p_p+q_p<1$ and $p_a+q_a<1$, we can prove the optimality of triple-threshold policies using proofs similar to Theorems 1 and 2. □



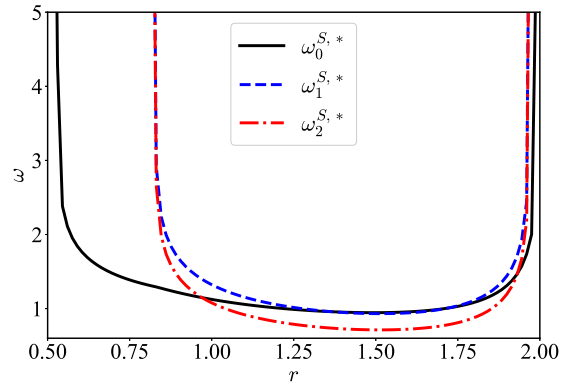
(a) δ



(b) u



(c) u_a



(d) r

Figure 9 $\omega^{S,*}$ **as functions of different parameters (**$\lambda = 1.5$, $\gamma = 1$, $\theta = 1$, $r = 1$, $c = 0.2$, $p_p = 0.8$, $p_a = 0.5$, $u = 2$, $u_a = 1$ **and** $\delta = 5$**).**

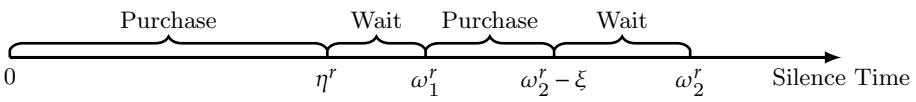


Figure 10 **Illustration of strategic customers' behaviors when $\omega_2 - \omega_1 \geq \xi$.**

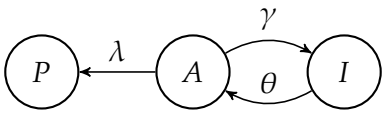


Figure 11 **State transition illustration of auxiliary Markov chain**

E-Companion of "When to Push Ads: Optimal Mobile Ad Campaign Strategy under Markov Customer Dynamics"

EC.1 Proof of Proposition 2

For ease of exposition, we omit the superscript 'E' in the following. We also use $\diamond(g) = h$ to denote that g and h have the same sign. We first prove that there exists a finite optimal threshold if and only if $\frac{c}{r} < \alpha := \frac{\lambda-\gamma-\theta+\varsigma}{2(\gamma+\theta)}$.

To analyze the property of $\Psi(\omega)$, we compute the gradient $\frac{\partial\Psi(\omega)}{\partial\omega}$. By eliminating some positive items, we get the following formula:

$$\begin{aligned} l_1(\omega) &= \diamond\left[\frac{\partial\Psi(\omega)}{\partial\omega}\right] \\ &= e^{\frac{\lambda+\gamma+\theta}{2}\omega}\left(\varsigma\left(c(\gamma+\theta)+\gamma r\right)\cosh\left(\frac{\omega\varsigma}{2}\right) - \left(c(\gamma+\theta)(\gamma+\lambda-\theta)+\gamma r(\gamma+\lambda+\theta)\right)\sinh\left(\frac{\omega\varsigma}{2}\right)\right) - \gamma(c+r)\varsigma. \end{aligned}$$

Furthermore, we compute the gradient of $l_1(\omega)$ as follows:

$$l_2(\omega) = \diamond\left[\frac{\partial l_1(\omega)}{\partial\omega}\right] = \frac{c}{r} - \frac{2\gamma\lambda}{(\gamma+\theta)\left((\gamma-\lambda+\theta)+\varsigma\coth\left(\frac{\omega\varsigma}{2}\right)\right)}, \tag{EC.1}$$

which decreases in ω. Having these derivatives, we study different cases.

Case 1: $\frac{c}{r} < \alpha$.

In this case, since $l_2(0) = c/r > 0$ and $\lim_{\omega\to\infty} l_2(\omega) = \frac{c}{r} - \frac{\lambda-\gamma-\theta+\varsigma}{2(\gamma+\theta)} < 0$, there exists $\tilde{\omega}$ such that $l_2(\tilde{\omega}) = 0$ and hence $l_1(\omega)$ first decreases and then increases in ω. After that, we further analyze the function $l_1(\omega)$. When ω tends to infinity, we have $\lim_{\omega\to\infty}\left(l_1(\omega)\times e^{\frac{1}{2}(\varsigma-\lambda-\gamma-\theta)\omega}\right) = (\gamma+\theta)(\varsigma-\lambda-\gamma+\theta)c - \gamma(\lambda+\gamma+\theta-\varsigma)r < 0$. Then, we have $\lim_{\omega\to\infty} l_1(\omega) < 0$ and $l_1(0) = c\theta\varsigma > 0$. There exists a point ω^* such that $\frac{\partial\Psi(\omega)}{\partial\omega} > 0$ when $\omega < \omega^*$ and $\frac{\partial\Psi(\omega)}{\partial\omega} < 0$ when $\omega > \omega^*$. Therefore, $\Psi(\omega)$ first increases and then decreases in ω and there exists a unique maximizer ω^*.

Case 2: $\frac{c}{r} \geq \alpha$.

In this case, since $l_2(\omega)$ decreases in ω, we have $l_2(\omega) \geq \lim_{\omega\to\infty} l_2(\omega) = \frac{c}{r} - \frac{\lambda-\gamma-\theta+\varsigma}{2(\gamma+\theta)} \geq 0$. Then, $l_1(\omega)$ always increases in ω and hence $l_1(\omega) \geq l_1(0) = c\theta\varsigma > 0$. Therefore, $\frac{\partial\Psi(\omega)}{\partial\omega} \geq 0$ and hence $\Psi(\omega)$ always increases in ω.

In the following, we derive some equivalent conditions to $c/r < \alpha$, which helps further understand the role each parameter plays in determining whether there exists an optimal threshold.

1. $\frac{c}{r} \in (0, \alpha)$.
2. $\lambda \in (\tilde{\lambda}, \infty)$, where $\tilde{\lambda} = \frac{c(c+r)(\gamma+\theta)^2}{r(\gamma c+\gamma r+\theta c)}$.
3. $\theta \in (0, \tilde{\theta})$, where $\tilde{\theta} = \frac{\lambda r+\sqrt{(4\gamma+\lambda+4\gamma\frac{r}{c})\lambda\cdot r}-2\gamma c-2\gamma r}{2(c+r)}$.

4. $\gamma \in (\tilde{\gamma}_1, \tilde{\gamma}_2)$, where $(\tilde{\gamma}_1, \tilde{\gamma}_2) = \begin{cases} (\frac{\lambda r - 2\theta c - \sqrt{(\lambda - 4\theta\frac{c}{c+r})\lambda}\cdot r}{2c}, \frac{\lambda r - 2\theta c + \sqrt{(\lambda - 4\theta\frac{c}{c+r})\lambda}\cdot r}{2c}), & \frac{\lambda-\theta}{\theta} \le \frac{c}{r} < \frac{\lambda}{4\theta-\lambda}, \\ (0, \frac{\lambda r - 2\theta c + \sqrt{(\lambda - 4\theta\frac{c}{c+r})\lambda}\cdot r}{2c}), & \frac{c}{r} \le \frac{\lambda-\theta}{\theta}, \\ (0,0), & \text{otherwise.} \end{cases}$

Then, we analyze the comparative statics for different parameters when either of the above conditions hold. Since $l_1(\omega^*) = 0$, we derive the following equation.

$$\frac{c}{r} = \frac{\gamma\left(e^{\frac{1}{2}\omega^*(\gamma+\lambda+\theta)}\left(\varsigma\cosh\left(\frac{1}{2}\omega^*\varsigma\right) - (\gamma+\lambda+\theta)\sinh\left(\frac{1}{2}\omega^*\varsigma\right)\right) - \varsigma\right)}{(\gamma+\theta)e^{\frac{1}{2}\omega^*(\gamma+\lambda+\theta)}\left((\gamma+\lambda-\theta)\sinh\left(\frac{1}{2}\omega^*\varsigma\right) - \varsigma\cosh\left(\frac{1}{2}\omega^*\varsigma\right)\right) + \gamma\varsigma} := h(\omega^*). \tag{EC.2}$$

In the following, we derive the comparative statics with the assistance of the Mathematica software. The general approach is to derive the gradient of the optimal threshold ω^* by implicit differentiation and prove that the sign of the gradient for any positive ω. For example, for the parameter λ, we first derive the gradient $\frac{\partial\omega^*}{\partial\lambda}$ as $f(\omega^*)$, and then show that $f(\omega) \le 0$ for all $\omega \ge 0$. Therefore, we prove that ω^* decreases in λ.

Part I: Cost-Reward Ratio c/r.

First, we consider the cost-reward ratio c/r. We compute the gradient of the function $h(\omega)$.

$$\begin{aligned} \diamond\left[\frac{\partial h(\omega)}{\partial\omega}\right] &= -\left(\gamma^2+\gamma(\lambda+2\theta)+\theta(\theta-\lambda)\right)\sinh\left(\frac{1}{2}\omega\varsigma\right) + (\gamma+\theta)\varsigma e^{\frac{1}{2}\omega(\gamma+\lambda+\theta)} - (\gamma+\theta)\varsigma\cosh\left(\frac{1}{2}\omega\varsigma\right) \\ &\ge -\left(\gamma^2+\gamma(\lambda+2\theta)+\theta(\theta-\lambda)\right)\sinh\left(\frac{1}{2}\omega\varsigma\right) + (\gamma+\theta)\varsigma e^{\frac{1}{2}\omega\varsigma} - (\gamma+\theta)\varsigma\cosh\left(\frac{1}{2}\omega\varsigma\right) \\ &= \left((\gamma+\theta)\varsigma - \gamma^2 - \gamma(\lambda+2\theta) - \theta(\theta-\lambda)\right)\sinh\left(\frac{1}{2}\omega\varsigma\right) \ge 0. \end{aligned}$$

The first inequality holds because $\varsigma \le (\lambda+\gamma+\theta)$, and the second inequality holds because $\sinh(x) \ge 0$, $\forall x \ge 0$ and $\varsigma \ge \frac{\gamma^2+\gamma(\lambda+2\theta)+\theta(\theta-\lambda)}{\gamma+\theta}$. Therefore, $h(\omega)$ increases in ω. Since $c/r = h(\omega^*)$, ω^* increases in c and decreases in r.

Part II: Purchase Rate λ.

Since $h(\omega^*) = \frac{c}{r}$, we derive the gradient $\frac{\partial\omega^*}{\partial\lambda}$ by taking implicit differentiation.

$$\begin{aligned} g_1(\omega^*) &= \diamond\left[\frac{\partial\omega^*}{\partial\lambda}\right] = \diamond\left[-\frac{\partial h(\omega^*)}{\partial\lambda} \Big/ \frac{\partial h(\omega^*)}{\partial\omega^*}\right] \\ &= \gamma\omega^*(\gamma+\lambda+\theta)\varsigma e^{-\frac{1}{2}\omega^*(\gamma+\lambda+\theta)}\cosh\left(\frac{\omega^*\varsigma}{2}\right) + (\gamma+\theta)(\gamma+\lambda-\theta)(\sinh(\omega^*\varsigma) - \omega^*\varsigma) \\ &\quad + \gamma e^{-\frac{1}{2}\omega^*(\gamma+\lambda+\theta)}\left(\omega^*\left(\gamma^2+2\gamma(\lambda+\theta)+(\lambda-\theta)^2\right) - 2(\gamma+\lambda+\theta)\right)\sinh\left(\frac{\omega^*\varsigma}{2}\right) - (\gamma+\theta)\varsigma(\cosh(\omega^*\varsigma) - 1). \end{aligned}$$

Then, we will show that $g_1(\omega) \le 0$ for any $\omega \ge 0$. We compute the gradient of $g_1(\omega)$ as follows:

$$\begin{aligned} \diamond\left[\frac{\partial g_1(\omega)}{\partial\omega}\right] &= 2\sinh\left(\frac{1}{2}\omega\varsigma\right)\left(\gamma\left((\gamma+\lambda)^2+2\gamma\theta+\theta^2\right)e^{-\frac{1}{2}\omega(\gamma+\lambda+\theta)} + (\gamma+\theta)\left((\gamma+\lambda-\theta)\varsigma\sinh\left(\frac{1}{2}\omega\varsigma\right)\right.\right. \\ &\quad \left.\left. - \left((\gamma+\lambda+\theta)^2 - 4\lambda\theta\right)\cosh\left(\frac{1}{2}\omega\varsigma\right)\right)\right) - 2\gamma\lambda\theta\omega\varsigma e^{-\frac{1}{2}\omega(\gamma+\lambda+\theta)}\cosh\left(\frac{1}{2}\omega\varsigma\right) \\ &\le 2\sinh\left(\frac{1}{2}\omega\varsigma\right)\left(\gamma\left((\gamma+\lambda)^2+2\gamma\theta+\theta^2\right)e^{-\frac{1}{2}\omega(\gamma+\lambda+\theta)} + (\gamma+\theta)\left((\gamma+\lambda-\theta)\varsigma\sinh\left(\frac{1}{2}\omega\varsigma\right) - \left((\gamma+\lambda+\theta)^2 - 4\lambda\theta\right)\cosh\left(\frac{1}{2}\omega\varsigma\right)\right)\right) \\ &\le 2\sinh\left(\frac{1}{2}\omega\varsigma\right)\left(\gamma\left((\gamma+\lambda)^2+2\gamma\theta+\theta^2\right)e^{-\frac{1}{2}\omega\varsigma} + (\gamma+\theta)\left((\gamma+\lambda-\theta)\varsigma\sinh\left(\frac{1}{2}\omega\varsigma\right) - \left((\gamma+\lambda+\theta)^2 - 4\lambda\theta\right)\cosh\left(\frac{1}{2}\omega\varsigma\right)\right)\right) \end{aligned}$$

$$
\begin{aligned}
&= 2\sinh\left(\frac{1}{2}\omega\varsigma\right)\left(-\theta\left(\gamma^2+(\lambda-\theta)^2+2\gamma\theta\right)\cosh\left(\frac{1}{2}\omega\varsigma\right)+\left((\gamma+\theta)(\gamma+\lambda-\theta)\varsigma-\gamma\left((\gamma+\lambda)^2+2\gamma\theta+\theta^2\right)\right)\sinh\left(\frac{1}{2}\omega\varsigma\right)\right)\\
&\le -2\sinh\left(\frac{1}{2}\omega\varsigma\right)^2(\gamma+\theta)\sqrt{(\lambda+\gamma+\theta)^2-4\lambda\theta}\left(\sqrt{(\lambda+\gamma+\theta)^2-4\lambda\theta}-(\lambda+\gamma-\theta)\right)\\
&\le 0.
\end{aligned}
$$

The first inequality holds because $\cosh(x)\ge 0$, the second inequality holds because $\varsigma\le(\lambda+\gamma+\theta)$, the third inequality holds because $\sinh(x)\le\cosh(x)$, and the last inequality holds because $\varsigma\ge(\lambda+\gamma-\theta)$. Therefore, we have $g_1(\omega)\le g_1(0)=0$ and hence ω^* decreases in λ.

Part III: Recapture Rate θ.

Since $h(\omega^*)=\frac{c}{r}$, we derive the gradient $\frac{\partial\omega^*}{\partial\theta}$ by taking implicit differentiation.

$$
\begin{aligned}
g_2(\omega^*) &= \diamond\left[\frac{\partial\omega^*}{\partial\theta}\right]=\diamond\left[-\frac{\partial h(\omega^*)}{\partial\theta}\Big/\frac{\partial h(\omega^*)}{\partial\omega^*}\right]\\
&=\Big(-\gamma^2(\lambda+3\theta)(\theta\omega^*+1)+\varsigma\left(\theta\omega^*\left(\gamma^2+2\gamma\theta+(\lambda-\theta)^2\right)+(\gamma+\lambda+\theta)^2-4\lambda\theta\right)-\left(\gamma^3(\theta\omega^*+1)\right)+\gamma\left(\lambda^2+\theta\omega^*(\lambda-\theta)(\lambda+3\theta)-2\lambda\theta-3\theta^2\right)\\
&\quad+\lambda\left(-\theta(\gamma+3\lambda)+(\gamma+\lambda-2\theta)\varsigma+(\gamma+\lambda)^2+2\theta^2\right)e^{\frac{1}{2}\omega^*(-\varsigma+\gamma+\lambda+\theta)}\left(e^{\omega^*\varsigma}-1\right)+(\lambda-\theta)^3(\theta\omega^*+1)\Big)\\
&\quad-e^{\omega^*\varsigma}\Big(-\gamma^2(\lambda+3\theta)(\theta\omega^*+1)-\varsigma\left(\theta\omega^*\left(\gamma^2+2\gamma\theta+(\lambda-\theta)^2\right)+(\gamma+\lambda+\theta)^2-4\lambda\theta\right)-\left(\gamma^3(\theta\omega^*+1)\right)\\
&\quad+\gamma\left(\lambda^2+\theta\omega^*(\lambda-\theta)(\lambda+3\theta)-2\lambda\theta-3\theta^2\right)+\lambda\left(\theta(\gamma+3\lambda)+(\gamma+\lambda-2\theta)\varsigma-(\gamma+\lambda)^2-2\theta^2\right)e^{\frac{1}{2}\omega^*(-\varsigma+\gamma+\lambda+\theta)}\left(e^{\omega^*\varsigma}-1\right)\\
&\quad+(\lambda-\theta)^3(\theta\omega^*+1)\Big)-2\varsigma e^{\frac{1}{2}\omega^*(\gamma+\lambda+\theta)}e^{\frac{1}{2}\omega^*\varsigma}\left(\theta\omega^*(\gamma+\theta)(\gamma-\lambda+\theta)+(\gamma+\lambda+\theta)^2-4\lambda\theta\right).
\end{aligned}
$$

Next, we show $g_2(\omega)\ge 0$ for any $\omega\ge 0$. To show this, we use the Mathematica software to obtain the k-th derivative of $g_2(\omega)$ for $k=1,2,\ldots,6$ and verify that $g_2^{(k)}(0)\ge 0$ for $k=1,2,\ldots,6$ and $g_2^{(6)}(\omega)\ge 0$ for all $\omega\ge 0$. Then, we have $g_2(\omega)\ge g_2(0)=0$ for all $\omega\ge 0$. Therefore, we have ω^* increases in θ. For ease of exposition, we leave the detailed calculation to Appendix EC.3.3.

Part IV: Churn Rate γ.

For the churn rate γ, we find that ω^* first decreases and then decreases for the general model. However, due to the complicated formula, we only provide proof for a special case when $\theta=0$. In this case, the function $h(\omega)$ becomes

$$
h(\omega)=\frac{\lambda\gamma\left(e^{(\lambda+\gamma)\omega}+-(\lambda+\gamma)\omega-1\right)}{\gamma^2e^{(\lambda+\gamma)\omega}+\gamma^2\lambda\omega+\gamma\lambda^2\omega+2\lambda\gamma+\lambda^2}.
$$

By implicit differentiation, we have

$$
\frac{\partial\omega^*}{\partial\gamma}=-\frac{\partial h(\omega^*)}{\partial\gamma}\Big/\frac{\partial h(\omega^*)}{\partial\omega^*}=-\frac{\gamma^2\left(-e^{2\omega(\gamma+\lambda)}\right)+e^{w(\gamma+\lambda)}\left(\gamma\left(\gamma\omega^2(\gamma+\lambda)^2+w(\gamma+\lambda)^2+\gamma\right)+\lambda^2\right)+\lambda\left(-w(\gamma+\lambda)^2-\lambda\right)}{\gamma(\gamma+\lambda)^2\left(e^{w(\gamma+\lambda)}(\gamma\omega(\gamma+\lambda)+\lambda)-\lambda\right)},
$$

and

$$
\frac{\partial^2\omega^*}{\partial\gamma^2}=\frac{\partial}{\partial\gamma}\left(\frac{\partial\omega^*}{\partial\gamma}\right)+\frac{\partial}{\partial\omega^*}\left(\frac{\partial\omega^*}{\partial\gamma}\right)\frac{\partial\omega^*}{\partial\gamma}.
$$

Then, we define $g_3(\omega)$ as follows:

$$
g_3(\omega)=\diamond\left[\frac{\partial^2\omega^*}{\partial\gamma^2}\right]
$$

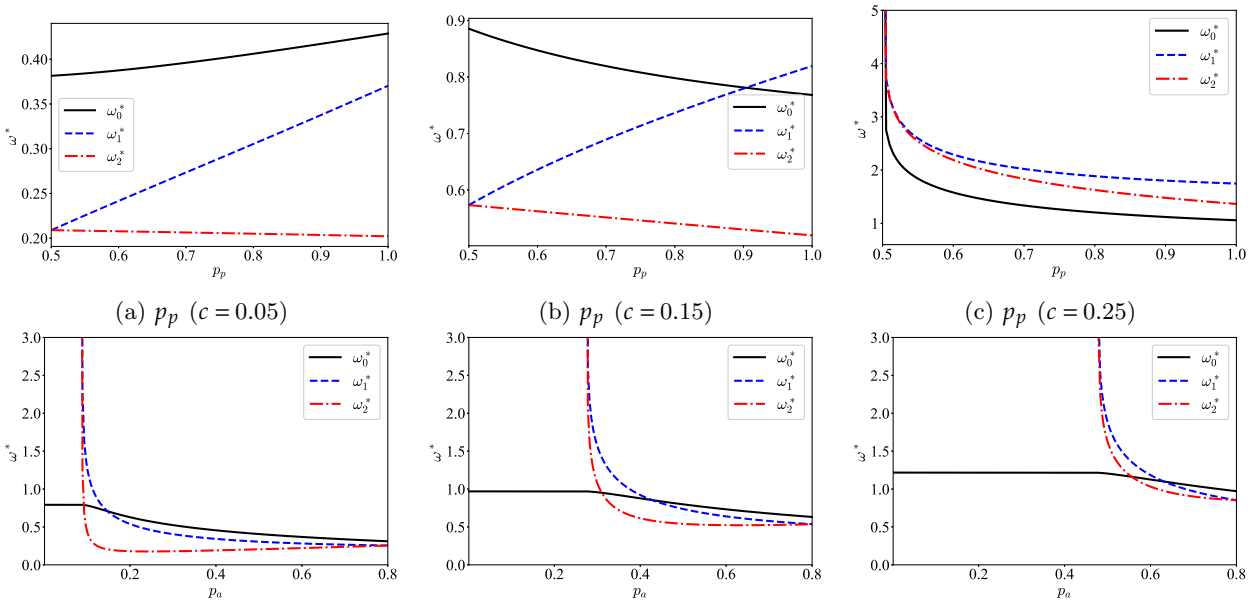
$$
\begin{aligned}
&= -2\lambda e^{2\omega(\gamma+\lambda)}\Big(\gamma^2\omega^3(\gamma+\lambda)^3\left(3\gamma^2+2\gamma\lambda+\lambda^2\right)+\gamma\omega(\gamma+\lambda)\left(4\gamma^3+13\gamma\lambda^2+4\lambda^3\right)+2\gamma^2\lambda^2+2\gamma^3\lambda+\gamma^4+5\gamma\lambda^3+\gamma\omega^2(\gamma+\lambda)^4(3\gamma+\lambda)+2\lambda^4\Big)\\
&\quad+\lambda^2 e^{\omega(\gamma+\lambda)}\Big(\omega^2(\gamma+\lambda)^2\left(10\gamma^2\lambda+6\gamma^3+\gamma\lambda^2-\lambda^3\right)+\omega(\gamma+\lambda)\left(18\gamma^2\lambda+2\gamma^3+9\gamma\lambda^2+2\lambda^3\right)+2\gamma^2\lambda+2\gamma^3-\gamma\lambda^2\omega^3(\gamma+\lambda)^3+11\gamma\lambda^2+5\lambda^3\Big)\\
&\quad+e^{3\omega(\gamma+\lambda)}\Big(\gamma^2\omega^2(\gamma+\lambda)^2\left(5\gamma^2\lambda+3\gamma^3+10\gamma\lambda^2+2\lambda^3\right)-\gamma\omega(\gamma+\lambda)\left(2\gamma^2\lambda^2-14\gamma^3\lambda+\gamma^4-10\gamma\lambda^3-3\lambda^4\right)+2\gamma^3\lambda^2+2\gamma^2\lambda^3+2\gamma^5\omega^4(\gamma+\lambda)^4\\
&\quad+3\gamma^4\omega^3(\gamma+\lambda)^3(\gamma+2\lambda)+5\gamma^4\lambda-\gamma^5+3\gamma\lambda^4+\lambda^5\Big)-2\gamma^4 e^{4\omega(\gamma+\lambda)}\left(2\gamma\omega^2(\gamma+\lambda)^2+3\lambda\omega(\gamma+\lambda)-\gamma+2\lambda\right)\\
&\quad+\gamma^4 e^{5\omega(\gamma+\lambda)}(\gamma\omega(\gamma+\lambda)-\gamma+\lambda)-2\lambda^3\omega(\gamma+\lambda)^3-2\lambda^4(2\gamma+\lambda).
\end{aligned}
$$

Next, we show $g_3(\omega)\geq 0$ for any $\omega\geq 0$. To show this, we use the Mathematica software to obtain the k-th derivative of $g_3(\omega)$ for $k=1,2,\ldots,10$ and verify that $g_3^{(k)}(0)\geq 0$ for $k=1,2,\ldots,10$ and $g_3^{(10)}(\omega)\geq 0$ for all $\omega\geq 0$. Then, we can derive that $g_3(\omega)\geq 0$ for all $\omega\geq 0$. Subsequently, we have $\frac{\partial^2\omega^*}{\partial\gamma^2}\geq 0$ and hence ω^* is convex in γ. As mentioned, the optimal threshold ω^* is infinite when γ is either sufficiently low or high. Therefore, the optimal solution ω^* first decreases and then increases in γ. For ease of exposition, we leave the detailed calculation to Appendix EC.3.4. □

EC.2 Additional Numerical Experiments

EC.2.1. Additional Numerical Experiments for Section 4

In the following, we provide additional numerical experiments for the activation probabilities p_p and p_a. In Figure EC.1, we illustrate the impacts of activation probabilities p_p and p_a in three representative cases with different cost-reward ratios.



(a) p_p $(c=0.05)$ (b) p_p $(c=0.15)$ (c) p_p $(c=0.25)$

(d) p_a $(c=0.05)$ (e) p_a $(c=0.15)$ (f) p_a $(c=0.25)$

Figure EC.1 ω^* **as functions of activation probabilities** p_p **and** p_a **(**$\lambda=1.5$, $\gamma=1$, $\theta=1$, $r=1$, $c\in\{0.05,0.15.0.25\}$, $p_p=0.8$ **and** $p_a=0.5$**).**

Then, we explain the additional insights behind the results in Figure EC.1. In Figures EC.1a and EC.1b, as the activation probability p_p increases, in addition to Observation V, we also observe some cases where ω_0^* or ω_1^* increase. First, in addition to activating inactive customer, an ad campaign also induces customers' fatigue to promotion, which diminishes the effectiveness. As p_p increases, the difference between non-fatigue customers and fatigue customers becomes more significant, resulting in a weaker effectiveness. If this negative effect dominates, then the threshold ω_0^* increases in p_p (see Figure EC.1a).

Second, according to the mapping in Appendix A.3, since the threshold ω_1^* is induced by the belief evolvement from $[0, b_p^* + (1-b_p^*)p_p, 0]$ to $[0, b_a^*, 0]$. as p_p increases, ω_1^* may increase if the starting belief $b_p^* + (1-b_p^*)p_p$ significantly increases (see Figures EC.1a and EC.1b).

In Figures EC.1a and EC.1b, as the activation probability p_a increases, in addition to Observation V, we also observe some cases where ω_2^* slightly increases. Similar to ω_1^*, since the threshold ω_2^* is induced by the belief evolvement from $[0, b_a^* + (1-b_a^*)p_a, 0]$ to $[0, b_a^*, 0]$, as p_a increases, ω_2^* may increase if the starting belief $b_a^* + (1-b_a^*)p_a$ significantly increases. □

EC.2.2. Additional Numerical Experiments for Section 6.1

In Figure EC.2, we illustrate the impact of the impact factor δ when the ad cost c is extremely small.

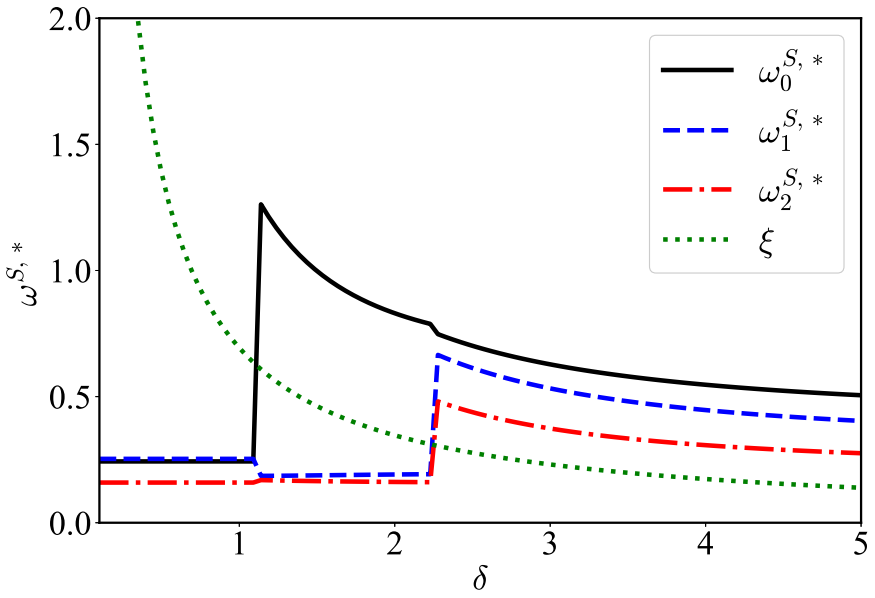


Figure EC.2 $\omega^{S,*}$ **as function of** δ **(**$\lambda = 1.5$, $\gamma = 1$, $\theta = 1$, $r = 1$, $c = 0.05$, $p_p = 0.8$, $p_a = 0.5$, $u = 2$, $u_a = 1$**).**

According to Figure EC.2, if both the discount factor δ and the ad cost c are small, then the optimal thresholds may be smaller than ξ. The reason is that the maximal waiting time ξ is extremely large such that choosing thresholds greater than ξ will induce a low profit. □

EC.3 Omitted Computation

EC.3.1. Computation Details of Efficient-Activation Case

In the following, we slightly abuse the notation and define each interval between two adjacent events as a "cycle". Similar to the base model, given a single-threshold policy φ^ω, we use $\pi_k^E(\omega)$ and $\kappa_k^E(\omega)$ to denote the random profit and length of the k-th cycle, respectively. We can deduce that $\pi_k^E(\omega)$'s and $\kappa_k^E(\omega)$'s are i.i.d. with their corresponding groups. Then, we can compute the long-term average expected profit $\Psi^E(\omega)$ under the single-threshold policy φ^ω as follows:

$$\Psi^E(\omega) = \frac{\mathbb{E}\left[\pi^E(\omega)\right]}{\mathbb{E}\left[\kappa^E(\omega)\right]},$$

where $\pi^E(\omega)$ and $\kappa^E(\omega)$ share the same distributions with $\pi_k^E(\omega)$'s and $\kappa_k^E(\omega)$'s, respectively. The cycles can be categorized into two types:

1. *Purchase Cycle*: The cycle ends with a purchase. In this case, we have $\pi^E(\omega) = r$ and $\kappa^E(\omega) < \omega$.
2. *Ad Campaign Cycle*: The cycle ends with an ad campaign. In this case, we have $\pi^E(\omega) = -c$ and $\kappa^E(\omega) = \omega$.

Suppose the probabilities of the purchase and the ad campaign cycle are $\rho_A(\omega)$ and $1-\rho_A(\omega)$, respectively. Then, we can compute the objective as follows:

$$\begin{aligned}\Psi^E(\omega) &= \frac{\rho_A(\omega) r - \left(1-\rho_A(\omega)\right) c}{\rho_A(\omega)\Xi(1,\omega) + \left(1-\rho_A(\omega)\right)\omega} \\ &= \frac{\lambda\theta\left(2\varsigma r - (c+r)e^{-\frac{1}{2}\omega(\varsigma+\gamma+\lambda+\theta)}\left((\gamma-\lambda+\theta)e^{\omega\varsigma} + \varsigma e^{\omega\varsigma} + \varsigma - \gamma + \lambda - \theta\right)\right)}{e^{-\frac{1}{2}\omega(\gamma+\lambda+\theta)}\left(-2\left(\gamma^2+\gamma(\lambda+2\theta)+\theta(\theta-\lambda)\right)\sinh\left(\frac{\omega\varsigma}{2}\right) - 2\varsigma\left((\gamma+\theta)\cosh\left(\frac{\omega\varsigma}{2}\right)\right)\right) + 2\varsigma(\gamma+\theta)}.\end{aligned}$$

where the notations are the same as those in Appendix A.5. □

EC.3.2. Computation Details of Section 6

Similar to Appendix A.5, we compute the expected cycle profit and cycle length as follows:

$$\begin{aligned}\mathbb{E}[\pi^S(\boldsymbol{\omega})] =& \rho_A(\omega_0-\xi)\cdot r + \left(\rho_A(\omega_0)-\rho_A(\omega_0-\xi)\right)\cdot(r-c) + \left(1-\rho_A(\omega_0)\right)\cdot\rho_B\left(b_p+(1-b_p)p_p,\omega_1-\xi\right)\cdot(r-c) \\ &+ \left(1-\rho_A(\omega_0)\right)\cdot\left(\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right) - \rho_B\left(b_p+(1-b_p)p_p,\omega_1-\xi\right)\right)\cdot(r-2c) \\ &+ \sum_{i=2}^{\infty}\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right)\right)\cdot\left(1-\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)^{i-2}\cdot\rho_B\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)\cdot(r-ic) \\ &+ \sum_{i=2}^{\infty}\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right)\right)\cdot\left(1-\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)^{i-2}\cdot\left(\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)-\rho_B\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)\right)\cdot(r-(i+1)c), \\ =& r - \left(1-\rho_A(\omega_0-\xi)\right)\cdot c - \left(1-\rho_A(\omega_0)\right)\cdot\left(\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right)-\rho_B\left(b_p+(1-b_p)p_p,\omega_1-\xi\right)\right)\cdot c \\ &- \frac{\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B(b_p+(1-b_p)p_p,\omega_1)\right)\cdot\rho_B\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)}{\left(\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)^2}\cdot c \\ &- \frac{\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B(b_p+(1-b_p)p_p,\omega_1)\right)\cdot\left(\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)-\rho_B\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)\right)\left(1+\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)}{\left(\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)^2}\cdot c, \\ \mathbb{E}[\kappa^S(\boldsymbol{\omega})] =& \rho_A(\omega_0-\xi)\cdot\Xi(1,\omega_0-\xi) + \left(\rho_A(\omega_0)-\rho_A(\omega_0-\xi)\right)\cdot\omega_0 + \left(1-\rho_A(\omega_0)\right)\cdot\rho_B\left(b_p+(1-b_p)p_p,\omega_1-\xi\right)\cdot\left(\omega_0+\Xi\left(b_p+(1-b_p)p_p,\omega_1-\xi\right)\right)\end{aligned}$$

$$
\begin{aligned}
&+\left(1-\rho_A(\omega_0)\right)\cdot\left(\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right)-\rho_B\left(b_p+(1-b_p)p_p,\omega_1-\xi\right)\right)\cdot(\omega_0+\omega_1)\\
&+\sum_{i=2}^{\infty}\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right)\right)\cdot\left(1-\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)^{i-2}\cdot\rho_B\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)\cdot\left(\omega_0+\omega_1+(i-2)\omega_2+\Xi\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)\right)\\
&+\sum_{i=2}^{\infty}\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right)\right)\cdot\left(1-\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)^{i-2}\cdot\left(\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)-\rho_B\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)\right)\cdot(\omega_0+\omega_1+(i-1)\omega_2)\\
=&\rho_A(\omega_0-\xi)\cdot\Xi(1,\omega_0-\xi)+\left(1-\rho_A(\omega_0-\xi)\right)\cdot\omega_0+\left(1-\rho_A(\omega_0)\right)\cdot\rho_B\left(b_p+(1-b_p)p_p,\omega_1-\xi\right)\cdot\Xi\left(b_p+(1-b_p)p_p,\omega_1-\xi\right)\\
&+\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B(b_p+(1-b_p)p_p,\omega_1-\xi)\right)\cdot\omega_1+\frac{\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right)\right)\cdot\rho_B\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)}{\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)}\cdot\Xi\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)\\
&+\frac{\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right)\right)\cdot\rho_B\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)\cdot\left(1-\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)}{\left(\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)^2}\cdot\omega_2\\
&+\frac{\left(1-\rho_A(\omega_0)\right)\cdot\left(1-\rho_B\left(b_p+(1-b_p)p_p,\omega_1\right)\right)\cdot\left(\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)-\rho_B\left(b_a+(1-b_a)p_a,\omega_2-\xi\right)\right)}{\left(\rho_B\left(b_a+(1-b_a)p_a,\omega_2\right)\right)^2}\cdot\omega_2.
\end{aligned}
$$

The objective function can be computed as $\Psi^S(\boldsymbol{\omega})=\mathbb{E}[\pi^S(\boldsymbol{\omega})]/\mathbb{E}[\kappa^S(\boldsymbol{\omega})]$. Then, we plug the expressions of ω_i's into $\Psi^S(\boldsymbol{\omega})$, and solve the optimal belief threshold $(b_p^{S,*},b_a^{S,*})$ and the corresponding optimal thresholds $\boldsymbol{\omega}^{S,*}$. □

EC.3.3. Detailed Calculation of $g_2(\omega)$ in Appendix EC.1

In the following, we show that $g_2(\omega)\geq 0$ for any $\omega\geq 0$.

$$
\begin{aligned}
\diamond\left[\frac{\partial g_2(\omega)}{\partial\omega}\right]=&-\Bigg(e^{\frac{1}{2}(2\varsigma)\omega}\Big((\theta\omega+1)\gamma^4+(\lambda(\theta\omega+3)+(\theta\omega+1)(4\theta+\varsigma))\gamma^3+\left((3-\theta\omega)\lambda^2+\theta(4-\theta\omega)\lambda+2\varsigma\lambda+3\theta^2(2\theta\omega+\varsigma\omega+2)+5\theta\varsigma\right)\gamma^2\\
&+\Big((1-\theta\omega)\lambda^3+(\theta\omega(2\theta-\varsigma)+\varsigma)\lambda^2+\theta(\varsigma-\theta(5\theta\omega+1))\lambda+\theta^2(\theta(4\theta\omega+3\varsigma\omega+4)+7\varsigma)\Big)\gamma\\
&+\theta^2\left((3\theta\omega-\varsigma\omega+1)\lambda^2-(\theta(3\theta\omega+2)+\varsigma)\lambda-\lambda^3\omega+\theta(\omega(\theta+\varsigma)\theta+\theta+3\varsigma)\right)\Big)\\
&-e^{\frac{3}{2}\varsigma\omega-\frac{1}{2}(\gamma+\lambda+\theta)\omega}\Big((\theta\omega+1)\gamma^4+\left(2\lambda(\theta\omega+2)+\theta(4\theta\omega+\varsigma\omega+5)+\varsigma\right)\gamma^3+\Big(2(\theta\omega+3)\lambda^2+(\omega\varsigma\theta+5\theta+\varsigma)\lambda\\
&+\theta(3\theta(2\theta\omega+\varsigma\omega+3)+4\varsigma)\Big)\gamma^2+\Big(2(\theta\omega+2)\lambda^3-(\omega\varsigma\theta+5\theta+\varsigma)\lambda^2-2\theta(\theta(3\theta\omega+\varsigma\omega+3)-\varsigma)\lambda+\theta^2\Big(\theta(4\theta\omega+3\varsigma\omega+7)\\
&+5\varsigma\Big)\Big)\gamma-(\lambda-\theta)^2(-\lambda+\theta+\varsigma)(\theta\omega\lambda+\lambda-\theta(\theta\omega+2))\Big)+\lambda\theta(\gamma(\lambda+2\theta)+(\lambda-2\theta)(\lambda-\theta+\varsigma))+e^{\frac{1}{2}(4\varsigma)\omega}\lambda\Big(\gamma^3-(\varsigma-3\lambda)\gamma^2\\
&+\left(3\lambda^2-5\theta\lambda-2\varsigma\lambda-5\theta^2+\theta\varsigma\right)\gamma-(\lambda-2\theta)^2(-\lambda+\theta+\varsigma)\Big)+e^{\frac{1}{2}\varsigma\omega-\frac{1}{2}(\gamma+\lambda+\theta)\omega}\theta\Big(\gamma^3+(\lambda+3\theta-\varsigma)\gamma^2\\
&-\left(\lambda^2+2\theta\lambda-3\theta^2+2\theta\varsigma\right)\gamma-(\lambda-\theta)^2(\lambda-\theta+\varsigma)\Big)\Bigg)\\
:=&\,\Theta_1\\
\diamond\left[\frac{\partial\Theta_1}{\partial\omega}\right]=&-\Bigg(-e^{\varsigma\omega}\Big((\theta\omega+1)\gamma^5+(\lambda(3\theta\omega+2)+\theta(5\theta\omega+\varsigma\omega+7)+\varsigma)\gamma^4+\Big((\theta\omega-2)\lambda^2+\theta(3\theta\omega+2\varsigma\omega+10)\lambda+5\varsigma\lambda+2\theta^2(5\theta\omega+2\varsigma\omega+9)\\
&+6\theta\varsigma\Big)\gamma^3+\Big((\theta(-4\theta\omega+3\varsigma\omega+11)+9\varsigma)\lambda^2-\theta\left(9\omega\theta^2+\omega\varsigma\theta+\theta-5\varsigma\right)\lambda-\lambda^3(5\theta\omega+8)+2\theta^2\left(\theta(5\theta\omega+3\varsigma\omega+11)+6\varsigma\right)\Big)\gamma^2\\
&+\Big((\theta(7\theta\omega+4\varsigma\omega+20)+7\varsigma)\lambda^3+\theta(\theta(9\theta\omega-2)-11\varsigma)\lambda^2-\theta^2(\theta(15\theta\omega+8\varsigma\omega+24)+12\varsigma)\lambda-\lambda^4(6\theta\omega+7)\\
&+\theta^3(\theta(5\theta\omega+4\varsigma\omega+13)+10\varsigma)\Big)\gamma+(\lambda-\theta)^2\left(2(\theta\omega+1)\lambda^2-3\theta(\theta\omega+2)\lambda+\theta^2(\theta\omega+3)\right)(-\lambda+\theta+\varsigma)\Big)\\
&+e^{\frac{1}{2}(\gamma+\lambda+\theta+\varsigma)\omega}\Big((\theta\omega+1)\gamma^5+(2\lambda(\theta\omega+2)+\theta(5\theta\omega+\varsigma\omega+8)+\varsigma)\gamma^4+\Big(6\lambda^2+\theta(4\theta\omega+\varsigma\omega+14)\lambda+3\varsigma\lambda\\
&+2\theta^2(5\theta\omega+2\varsigma\omega+11)+5\theta\varsigma\Big)\gamma^3+\Big((4-2\theta\omega)\lambda^3+(-\varsigma\omega\theta+4\theta+3\varsigma)\lambda^2+\theta(-\varsigma\omega\theta+6\theta+4\varsigma)\lambda\\
&+\theta^2(2\theta(5\theta\omega+3\varsigma\omega+14)+9\varsigma)\Big)\gamma^2+\Big((1-\theta\omega)\lambda^4+(\varsigma-\theta(\omega\varsigma+2))\lambda^3+\theta(2\omega\varsigma\theta+6\theta-\varsigma)\lambda^2\\
&-\theta^2(\theta(4\theta\omega+5\varsigma\omega+14)+\varsigma)\lambda+\theta^3\Big(\theta(5\theta\omega+4\varsigma\omega+17)+7\varsigma\Big)\Big)\gamma-(\lambda-\theta)\theta^2\Big(\omega\lambda^3+(-\theta\omega+\varsigma\omega+2)\lambda^2
\end{aligned}
$$

$$
\begin{aligned}
&- \theta\left(\theta\omega + 2\varsigma\omega + 6\right)\lambda + \theta\left(\theta\left(\omega\left(\theta+\varsigma\right)+4\right)+2\varsigma\right)\Big)\Big) - \theta\Big(\gamma^4 + \left(2\lambda + 4\theta - \varsigma\right)\gamma^3 + \left(\lambda^2 + \theta\lambda - \varsigma\lambda + 6\theta^2 - 3\theta\varsigma\right)\gamma^2 \\
&+ \theta\left(4\theta^2 - 4\lambda\theta - 3\varsigma\theta + \lambda\varsigma\right)\gamma - \left(\lambda-\theta\right)^2\theta\left(\lambda-\theta+\varsigma\right)\Big) - 2e^{\frac{1}{2}\left(\gamma+\lambda+\theta+3\varsigma\right)\omega}\lambda\Big(\gamma^4 + \left(4\lambda+\theta-\varsigma\right)\gamma^3 \\
&+ \left(6\lambda^2 - 4\theta\lambda - 3\varsigma\lambda + 3\theta^2\right)\gamma^2 + \left(4\lambda^3 - \left(11\theta + 3\varsigma\right)\lambda^2 + \theta\left(4\theta+5\varsigma\right)\lambda + \theta^2\left(7\theta+5\varsigma\right)\right)\gamma - \left(\lambda-2\theta\right)^2\left(\lambda-\theta\right)\left(-\lambda+\theta+\varsigma\right)\Big)\Big) \\
&:= \Theta_2 \\
\diamond\left[\frac{\partial\Theta_2}{\partial\omega}\right] &= -\Big(-\Big(\left(\left(\lambda+\gamma+\theta\right)^2 - 4\lambda\theta\right)\Big(\left(\theta\omega+1\right)\gamma^4 + \left(\lambda\left(2\theta\omega+5\right) + \theta\left(4\theta\omega+\varsigma\omega+7\right)+\varsigma\right)\gamma^3 + \Big(3\left(\theta\omega+3\right)\lambda^2 + \theta\left(-\theta\omega+\varsigma\omega+6\right)\lambda \\
&+ 3\theta\left(\theta\left(2\theta\omega+\varsigma\omega+5\right)+2\varsigma\right)\Big)\gamma^2 + \Big(\left(4\theta\omega+7\right)\lambda^3 - \left(2\omega\varsigma\theta + 13\theta + 3\varsigma\right)\lambda^2 - \theta\Big(\theta\left(8\theta\omega+3\varsigma\omega+15\right) - 2\varsigma\Big)\lambda + \theta^2\left(\theta\left(4\theta\omega+3\varsigma\omega+13\right)+9\varsigma\right)\Big)\gamma \\
&- \left(\lambda-\theta\right)\left(-\lambda+\theta+\varsigma\right)\Big(2\left(\theta\omega+1\right)\lambda^2 - \theta\left(3\theta\omega+8\right)\lambda + \theta^2\left(\theta\omega+4\right)\Big)\Big)\Big) + e^{\frac{1}{2}\left(\gamma+\lambda+\theta-\varsigma\right)\omega}\Big(\left(\theta\omega+1\right)\gamma^6 + \left(\lambda\left(3\theta\omega+5\right) + \theta\left(6\theta\omega+\varsigma\omega+9\right)+\varsigma\right)\gamma^5 \\
&+ \Big(2\left(\theta\omega+5\right)\lambda^2 + \theta\left(9\theta\omega+2\varsigma\omega+24\right)\lambda + 4\varsigma\lambda + 5\theta^2\left(3\theta\omega+\varsigma\omega+6\right) + 8\theta\varsigma\Big)\gamma^4 + 2\Big(\left(5-\theta\omega\right)\lambda^3 + \left(\theta\left(\theta\omega+9\right)+3\varsigma\right)\lambda^2 + \theta\Big(\theta\left(3\theta\omega+2\varsigma\omega+17\right) \\
&+ 7\varsigma\Big)\lambda + \theta^2\left(5\theta\left(2\theta\omega+\varsigma\omega+5\right)+11\varsigma\right)\Big)\gamma^3 + \Big(\left(5-3\theta\omega\right)\lambda^4 - 2\varsigma\left(\theta\omega-2\right)\lambda^3 + 2\theta\left(\theta\left(\theta\omega+4\right)+2\varsigma\right)\lambda^2 + 2\theta^2\left(\theta\left(4-3\theta\omega\right)+4\varsigma\right)\lambda \\
&+ \theta^3\left(5\theta\left(3\theta\omega+2\varsigma\omega+9\right)+28\varsigma\right)\Big)\gamma^2 + \Big(\left(1-\theta\omega\right)\lambda^5 + \left(\varsigma - \theta\left(\omega\varsigma+3\right)\right)\lambda^4 - 2\theta\left(\theta\left(\theta\omega-1\right)+\varsigma\right)\lambda^3 + 2\theta^2\left(3\omega\theta^2 + \theta + 2\varsigma\right)\lambda^2 \\
&- \theta^3\left(\theta\left(9\theta\omega+4\varsigma\omega+15\right)+10\varsigma\right)\lambda + \theta^4\left(\theta\left(6\theta\omega+5\varsigma\omega+21\right)+17\varsigma\right)\Big)\gamma - \left(\lambda-\theta\right)\theta^2\Big(\omega\lambda^4 + \left(-2\theta\omega+\varsigma\omega+2\right)\lambda^3 + \left(\theta\left(2\theta\omega-\varsigma\omega-2\right)+2\varsigma\right)\lambda^2 \\
&- \theta\Big(\theta\left(2\theta\omega+\varsigma\omega+4\right)+4\varsigma\Big)\lambda + \theta^2\left(\theta\omega+4\right)\left(\theta+\varsigma\right)\Big)\Big) - 2e^{\frac{1}{2}\left(\gamma+\lambda+\theta+\varsigma\right)\omega}\lambda\Big(-\gamma^5 + \left(-5\lambda-2\theta+\varsigma\right)\gamma^4 + \Big(-10\lambda^2 + 2\left(\theta+2\varsigma\right)\lambda \\
&+ \theta\left(8\theta+\varsigma\right)\Big)\gamma^3 + \Big(-10\lambda^3 + 6\left(3\theta+\varsigma\right)\lambda^2 - 5\theta\left(\varsigma-3\theta\right)\lambda + \theta^2\left(26\theta+7\varsigma\right)\Big)\gamma^2 + \Big(-5\lambda^4 + \left(22\theta+4\varsigma\right)\lambda^3 - \theta\left(18\theta+13\varsigma\right)\lambda^2 \\
&+ \theta^2\left(7\varsigma-20\theta\right)\lambda + 5\theta^3\left(5\theta+3\varsigma\right)\Big)\gamma + \left(\lambda-\theta\right)\left(\lambda-2\theta\right)^3\left(-\lambda+\theta+\varsigma\right)\Big)\Big) \\
&:= \Theta_3 \\
\diamond\left[\frac{\partial\Theta_3}{\partial\omega}\right] &= -\Big(e^{\frac{1}{2}\left(\gamma+\lambda+\theta-\varsigma\right)\omega}\lambda\theta\Big(\left(\theta\omega+1\right)\gamma^5 + \left(\left(2\lambda+5\theta\right)\left(\theta\omega+2\right) + \theta\omega\varsigma + \varsigma\right)\gamma^4 + \Big(6\lambda^2 + \left(4\theta\left(\theta\omega+4\right) + \theta\omega\varsigma + 3\varsigma\right)\lambda + \theta\Big(2\theta\left(5\theta\omega+2\varsigma\omega+15\right)+5\varsigma\Big)\Big)\gamma^3 \\
&+ \Big(\left(4-2\theta\omega\right)\lambda^3 + \left(-\varsigma\omega\theta + 2\theta + 3\varsigma\right)\lambda^2 + \theta\left(-\varsigma\omega\theta + 4\theta + 4\varsigma\right)\lambda + \theta^2\left(2\theta\left(5\theta\omega+3\varsigma\omega+20\right)+9\varsigma\right)\Big)\gamma^2 \\
&+ \Big(\left(1-\theta\omega\right)\lambda^4 + \left(\varsigma - \theta\left(\omega\varsigma+4\right)\right)\lambda^3 + \theta\left(2\theta\left(\omega\varsigma+5\right)-\varsigma\right)\lambda^2 - \theta^2\left(\theta\left(4\theta\omega+5\varsigma\omega+24\right)+\varsigma\right)\lambda + \theta^3\left(\theta\left(5\theta\omega+4\varsigma\omega+25\right)+7\varsigma\right)\Big)\gamma \\
&- \left(\lambda-\theta\right)\theta^2\Big(\omega\lambda^3 + \left(-\theta\omega+\varsigma\omega+4\right)\lambda^2 - \theta\left(\theta\omega+2\varsigma\omega+10\right)\lambda + \theta\left(\theta\left(\omega\left(\theta+\varsigma\right)+6\right)+2\varsigma\right)\Big)\Big) \\
&+ e^{\frac{1}{2}\left(\gamma+\lambda+\theta-\varsigma\right)\omega}\theta\left(\gamma+\theta\right)\left(\gamma-\lambda+\theta\right)\Big(\left(\lambda+\gamma+\theta\right)^2 - 4\lambda\theta\Big)\Big(\gamma^2 + \left(2\lambda+2\theta+\varsigma\right)\gamma + \lambda^2 + \lambda\varsigma + \theta\left(\theta+\varsigma\right)\Big) \\
&- \Big(\left(\lambda+\gamma+\theta\right)^2 - 4\lambda\theta\Big)\Big(\theta\gamma^4 + \theta\left(2\lambda+4\theta+\varsigma\right)\gamma^3 + \theta\Big(3\lambda^2 + \left(\varsigma-\theta\right)\lambda + 3\theta\left(2\theta+\varsigma\right)\Big)\gamma^2 + \theta\Big(4\lambda^3 - 2\varsigma\lambda^2 - \theta\left(8\theta+3\varsigma\right)\lambda \\
&+ \theta^2\left(4\theta+3\varsigma\right)\Big)\gamma - \left(\lambda-\theta\right)^2\left(2\lambda-\theta\right)\theta\left(-\lambda+\theta+\varsigma\right)\Big) - e^{\frac{1}{2}\left(\gamma+\lambda+\theta+\varsigma\right)\omega}\lambda\left(\gamma+\lambda+\theta+\varsigma\right)\Big(-\gamma^5 + \left(-5\lambda-2\theta+\varsigma\right)\gamma^4 \\
&+ \Big(-10\lambda^2 + 2\left(\theta+2\varsigma\right)\lambda + \theta\left(8\theta+\varsigma\right)\Big)\gamma^3 + \Big(-10\lambda^3 + 6\left(3\theta+\varsigma\right)\lambda^2 - 5\theta\left(\varsigma-3\theta\right)\lambda + \theta^2\left(26\theta+7\varsigma\right)\Big)\gamma^2 + \Big(-5\lambda^4 + \left(22\theta+4\varsigma\right)\lambda^3 \\
&- \theta\left(18\theta+13\varsigma\right)\lambda^2 + \theta^2\left(7\varsigma-20\theta\right)\lambda + 5\theta^3\left(5\theta+3\varsigma\right)\Big)\gamma + \left(\lambda-\theta\right)\left(\lambda-2\theta\right)^3\left(-\lambda+\theta+\varsigma\right)\Big)\Big) \\
&:= \Theta_4 \\
\diamond\left[\frac{\partial\Theta_4}{\partial\omega}\right] &= -\Big(\lambda\left(\gamma+\theta\right)\Big(\gamma^3 + \left(\lambda+3\theta\right)\gamma^2 - \Big(\lambda^2 + 2\theta\lambda - 3\theta^2\Big)\gamma - \left(\lambda-\theta\right)^3\Big)\left(\gamma+\lambda+\theta+\varsigma\right)\theta^2 + \frac{1}{2}\lambda\Big(\left(\theta\omega+1\right)\gamma^5 + \left(\left(2\lambda+5\theta\right)\left(\theta\omega+2\right) + \theta\omega\varsigma + \varsigma\right)\gamma^4 \\
&+ \Big(6\lambda^2 + \left(4\theta\left(\theta\omega+4\right) + \theta\omega\varsigma + 3\varsigma\right)\lambda + \theta\Big(2\theta\left(5\theta\omega+2\varsigma\omega+15\right)+5\varsigma\Big)\Big)\gamma^3 + \Big(\left(4-2\theta\omega\right)\lambda^3 + \left(\theta\left(2-\varsigma\omega\right)+3\varsigma\right)\lambda^2 + \theta\left(\theta\left(4-\varsigma\omega\right)+4\varsigma\right)\lambda \\
&+ \theta^2\left(2\theta\left(5\theta\omega+3\varsigma\omega+20\right)+9\varsigma\right)\Big)\gamma^2 + \Big(\left(1-\theta\omega\right)\lambda^4 + \left(\varsigma - \theta\left(\omega\varsigma+4\right)\right)\lambda^3 + \theta\left(2\theta\left(\omega\varsigma+5\right)-\varsigma\right)\lambda^2 - \theta^2\left(\theta\left(4\theta\omega+5\varsigma\omega+24\right)+\varsigma\right)\lambda \\
&+ \theta^3\left(\theta\left(5\theta\omega+4\varsigma\omega+25\right)+7\varsigma\right)\Big)\gamma - \left(\lambda-\theta\right)\theta^2\Big(\omega\lambda^3 + \left(-\theta\omega+\varsigma\omega+4\right)\lambda^2 - \theta\left(\theta\omega+2\varsigma\omega+10\right)\lambda + \theta\left(\theta\left(\omega\left(\theta+\varsigma\right)+6\right)+2\varsigma\right)\Big)\Big)\Big(\gamma+\lambda+\theta
\end{aligned}
$$

$$
\begin{aligned}
&-\varsigma\Big)\theta+\frac{1}{2}(\gamma+\theta)(\gamma-\lambda+\theta)\Big((\lambda+\gamma+\theta)^2-4\lambda\theta\Big)(\gamma+\lambda+\theta-\varsigma)\Big(\gamma^2+(2\lambda+2\theta+\varsigma)\gamma+\lambda^2+\lambda\varsigma+\theta(\theta+\varsigma)\Big)\theta\\
&-\frac{1}{2}e^{\varsigma\omega}\lambda(\gamma+\lambda+\theta+\varsigma)^2\Big(-\gamma^5+(-5\lambda-2\theta+\varsigma)\gamma^4+\Big(-10\lambda^2+2(\theta+2\varsigma)\lambda+\theta(8\theta+\varsigma)\Big)\gamma^3+\Big(-10\lambda^3+6(3\theta+\varsigma)\lambda^2-5\theta(\varsigma-3\theta)\lambda\\
&+\theta^2(26\theta+7\varsigma)\Big)\gamma^2+\Big(-5\lambda^4+(22\theta+4\varsigma)\lambda^3-\theta(18\theta+13\varsigma)\lambda^2+\theta^2(7\varsigma-20\theta)\lambda+5\theta^3(5\theta+3\varsigma)\Big)\gamma+(\lambda-\theta)(\lambda-2\theta)^3(-\lambda+\theta+\varsigma)\Big)\Big)\\
&:=\Theta_5\\
\diamond\Big[\frac{\partial\Theta_5}{\partial\omega}\Big]&=-\Big(2\lambda^2\theta^3(\gamma+\theta)\Big(\gamma^2(\lambda+3\theta)+\gamma^3-\gamma\Big(\lambda^2+2\lambda\theta-3\theta^2\Big)-(\lambda-\theta)^3\Big)-\frac{1}{2}\lambda\varsigma(\varsigma+\gamma+\lambda+\theta)^2\Big(\gamma^2\Big(6\lambda^2(\varsigma+3\theta)+\theta^2(7\varsigma+26\theta)-5\lambda\theta(\varsigma-3\theta)-10\lambda^3\Big)\\
&+\gamma\Big(\lambda^3(4\varsigma+22\theta)-\lambda^2\theta(13\varsigma+18\theta)+\lambda\theta^2(7\varsigma-20\theta)+5\theta^3(3\varsigma+5\theta)-5\lambda^4\Big)+\gamma^3\Big(2\lambda(2\varsigma+\theta)+\theta(\varsigma+8\theta)-10\lambda^2\Big)+\gamma^4(\varsigma-5\lambda-2\theta)\\
&+(\lambda-\theta)(\lambda-2\theta)^3(\varsigma-\lambda+\theta)-\gamma^5\Big)e^{\omega\varsigma}\Big)\\
&:=\Theta_6\\
\diamond\Big[\frac{\partial\Theta_6}{\partial\omega}\Big]&=-\Big(\Big(\gamma^2+2\gamma(\lambda+\theta)+(\lambda-\theta)^2\Big)\Big(3\gamma^2\lambda+\gamma^3+3\gamma(\lambda+\theta)(\lambda-3\theta)+(\lambda-2\theta)^3\Big)\Big)\\
&-\Big(\theta^2\Big(7\gamma^2+7\gamma\lambda+18\lambda^2\Big)+5\theta^3(3\gamma-4\lambda)+\theta(\gamma-7\lambda)(\gamma+\lambda)^2+(\gamma+\lambda)^4+8\theta^4\Big)\varsigma\\
&\geq-\frac{2\gamma\theta^2\left(2\theta\left(12\gamma^2+4\gamma\lambda+\lambda^2\right)+\theta^2(24\gamma-11\lambda)+(\gamma+\lambda)^2(8\gamma+3\lambda)+8\theta^3\right)}{\gamma+\lambda}\\
&\geq 0;\\
\Theta_6&\geq\Theta_6\,|_{\omega=0}=2\lambda\theta^2((\lambda+\gamma+\theta)^2-4\lambda\theta)^{\frac{3}{2}}\Big(\Big(8\gamma^2+4\gamma(\lambda+4\theta)+\lambda^2-5\lambda\theta+8\theta^2\Big)\varsigma+12\gamma^2(\lambda+2\theta)+8\gamma^3+3\gamma\Big(\lambda^2+8\theta^2\Big)-(\lambda-2\theta)^3\Big)\\
&\geq 2\lambda\theta^2((\lambda+\gamma+\theta)^2-4\lambda\theta)^{\frac{3}{2}}\Big(\frac{\theta^2\left(48\gamma^2+7\gamma\lambda+\lambda^2\right)+\gamma\theta\left(48\gamma^2+47\gamma\lambda+9\lambda^2\right)+8\gamma(\gamma+\lambda)^2(2\gamma+\lambda)+16\gamma\theta^3}{\gamma+\lambda}\Big)\\
&\geq 0;\\
\Theta_5&\geq\Theta_5\,|_{\omega=0}=2\lambda\theta^2\left((\lambda+\gamma+\theta)^2-4\lambda\theta\right)^{\frac{3}{2}}\Big((6\gamma-\lambda+6\theta)\varsigma+8\gamma^2+4\gamma(\lambda+4\theta)+\lambda^2-5\lambda\theta+8\theta^2\Big)\\
&\geq 2\lambda\theta^2\left((\lambda+\gamma+\theta)^2-4\lambda\theta\right)^{\frac{3}{2}}\Big(8\gamma^2+4\gamma(\lambda+4\theta)+\lambda^2-5\lambda\theta+8\theta^2-\lambda(\lambda+\gamma+\theta)+6(\lambda+\theta)\varsigma\Big)\\
&\geq 0;\\
\Theta_4&\geq\Theta_4\,|_{\omega=0}=2\lambda\theta^2\left((\lambda+\gamma+\theta)^2-4\lambda\theta\right)^{\frac{3}{2}}\Big(3\varsigma+5\gamma+5\theta-2\lambda\Big)\geq 0;\\
\Theta_3&\geq\Theta_3\,|_{\omega=0}=4\lambda\theta^2\Big(\gamma^2+2\gamma(\lambda+\theta)+(\lambda-\theta)^2\Big)^{3/2}\geq 0\\
\Theta_2&\geq\Theta_2\,|_{\omega=0}=0\\
\Theta_1&\geq\Theta_1\,|_{\omega=0}=0.
\end{aligned}
$$

We mainly use the facts that $\varsigma\geq\lambda+\gamma+\frac{\gamma-\lambda}{\lambda+\gamma}\theta$ and $\theta^2\left(7\gamma^2+7\gamma\lambda+18\lambda^2\right)+5\theta^3(3\gamma-4\lambda)+\theta(\gamma-7\lambda)(\gamma+\lambda)^2+(\gamma+\lambda)^4+8\theta^4\geq 0$ and $2\theta\left(12\gamma^2+4\gamma\lambda+\lambda^2\right)+\theta^2(24\gamma-11\lambda)+(\gamma+\lambda)^2(8\gamma+3\lambda)+8\theta^3\geq 0$ when $\lambda,\gamma,\theta\geq 0$. Therefore, we have $g_2(\omega)\geq g_2(0)=0$ for any $\omega\geq 0$. □

EC.3.4. Detailed Calculation of $g_3(\omega)$ in Appendix EC.1

In the following, we show that $g_3(\omega)\geq 0$ for any $\omega\geq 0$.

$$
\begin{aligned}
\diamond\Big[\frac{\partial g_3(\omega)}{\partial\omega}\Big]&=-2\lambda e^{2\omega(\gamma+\lambda)}\Big(2\gamma^2\omega^3(\gamma+\lambda)^3\Big(3\gamma^2+2\gamma\lambda+\lambda^2\Big)+\gamma\omega^2(\gamma+\lambda)^2\Big(20\gamma^2\lambda+15\gamma^3+13\gamma\lambda^2+2\lambda^3\Big)+2\gamma\omega(\gamma+\lambda)\Big(7\gamma^2\lambda+7\gamma^3+18\gamma\lambda^2+5\lambda^3\Big)\\
&+17\gamma^2\lambda^2+4\gamma^3\lambda+6\gamma^4+14\gamma\lambda^3+4\lambda^4\Big)+\lambda^2e^{\omega(\gamma+\lambda)}\Big(\omega^2(\gamma+\lambda)^2\Big(10\gamma^2\lambda+6\gamma^3-2\gamma\lambda^2-\lambda^3\Big)+\gamma\omega(\gamma+\lambda)\Big(14\gamma^2+38\gamma\lambda+11\lambda^2\Big)+20\gamma^2\lambda\\
&+4\gamma^3-\gamma\lambda^2\omega^3(\gamma+\lambda)^3+20\gamma\lambda^2+7\lambda^3\Big)+e^{3\omega(\gamma+\lambda)}\Big(3\gamma^2\omega^2(\gamma+\lambda)^2\Big(11\gamma^2\lambda+6\gamma^3+10\gamma\lambda^2+2\lambda^3\Big)\\
&+\gamma\omega(\gamma+\lambda)\Big(14\gamma^2\lambda^2+52\gamma^3\lambda+3\gamma^4+34\gamma\lambda^3+9\lambda^4\Big)+4\gamma^3\lambda^2+16\gamma^2\lambda^3+6\gamma^5\omega^4(\gamma+\lambda)^4+\gamma^4\omega^3(\gamma+\lambda)^3(17\gamma+18\lambda)+29\gamma^4\lambda-4\gamma^5\\
&+12\gamma\lambda^4+3\lambda^5\Big)-2\gamma^4e^{4\omega(\gamma+\lambda)}\Big(8\gamma\omega^2(\gamma+\lambda)^2+4\omega(\gamma+\lambda)(\gamma+3\lambda)-4\gamma+11\lambda\Big)+\gamma^4e^{5\omega(\gamma+\lambda)}(5\gamma\omega(\gamma+\lambda)-4\gamma+5\lambda)-2\lambda^3(\gamma+\lambda)^2\\
&:=\Pi_1\\
\diamond\Big[\frac{\partial\Pi_1}{\partial\omega}\Big]&=-4\lambda e^{\omega(\gamma+\lambda)}\Big(2\gamma^2\omega^3(\gamma+\lambda)^3\Big(3\gamma^2+2\gamma\lambda+\lambda^2\Big)+2\gamma\omega^2(\gamma+\lambda)^2\Big(13\gamma^2\lambda+12\gamma^3+8\gamma\lambda^2+\lambda^3\Big)+\gamma\omega(\gamma+\lambda)\Big(34\gamma^2\lambda+29\gamma^3+49\gamma\lambda^2+12\lambda^3\Big)
\end{aligned}
$$

$$+35\gamma^2\lambda^2+11\gamma^3\lambda+13\gamma^4+19\gamma\lambda^3+4\lambda^4\Big)+\lambda^2\Big(\omega^2(\gamma+\lambda)^2\left(10\gamma^2\lambda+6\gamma^3-5\gamma\lambda^2-\lambda^3\right)+\omega(\gamma+\lambda)\left(58\gamma^2\lambda+26\gamma^3+7\gamma\lambda^2-2\lambda^3\right)+58\gamma^2\lambda$$
$$+18\gamma^3-\gamma\lambda^2\omega^3(\gamma+\lambda)^3+31\gamma\lambda^2+7\lambda^3\Big)+e^{2\omega(\gamma+\lambda)}\Big(3\gamma^2\omega^2(\gamma+\lambda)^2\left(51\gamma^2\lambda+35\gamma^3+30\gamma\lambda^2+6\lambda^3\right)$$
$$+3\gamma\omega(\gamma+\lambda)\left(34\gamma^2\lambda^2+74\gamma^3\lambda+15\gamma^4+38\gamma\lambda^3+9\lambda^4\right)+26\gamma^3\lambda^2+82\gamma^2\lambda^3+18\gamma^5\omega^4(\gamma+\lambda)^4+3\gamma^4\omega^3(\gamma+\lambda)^3(25\gamma+18\lambda)+139\gamma^4\lambda$$
$$-9\gamma^5+45\gamma\lambda^4+9\lambda^5\Big)-8\gamma^4e^{3\omega(\gamma+\lambda)}\left(8\gamma\omega^2(\gamma+\lambda)^2+4\omega(\gamma+\lambda)(2\gamma+3\lambda)-3\gamma+14\lambda\right)+5\gamma^4e^{4\omega(\gamma+\lambda)}(5\gamma\omega(\gamma+\lambda)-3\gamma+5\lambda)$$
$$:=\Pi_2$$

$$\diamond\left[\frac{\partial\Pi_2}{\partial\omega}\right]=-4\lambda e^{\omega(\gamma+\lambda)}\Big(2\gamma^2\omega^3(\gamma+\lambda)^3\left(3\gamma^2+2\gamma\lambda+\lambda^2\right)+2\gamma\omega^2(\gamma+\lambda)^2\left(19\gamma^2\lambda+21\gamma^3+11\gamma\lambda^2+\lambda^3\right)+\gamma\omega(\gamma+\lambda)\left(86\gamma^2\lambda+77\gamma^3+81\gamma\lambda^2+16\lambda^3\right)$$
$$+84\gamma^2\lambda^2+45\gamma^3\lambda+42\gamma^4+31\gamma\lambda^3+4\lambda^4\Big)+e^{2\omega(\gamma+\lambda)}\Big(3\gamma^2\omega^2(\gamma+\lambda)^2\left(156\gamma^2\lambda+145\gamma^3+60\gamma\lambda^2+12\lambda^3\right)$$
$$+6\gamma\omega(\gamma+\lambda)\left(64\gamma^2\lambda^2+125\gamma^3\lambda+50\gamma^4+44\gamma\lambda^3+9\lambda^4\right)+154\gamma^3\lambda^2+278\gamma^2\lambda^3+36\gamma^5\omega^4(\gamma+\lambda)^4+6\gamma^4\omega^3(\gamma+\lambda)^3(37\gamma+18\lambda)+500\gamma^4\lambda$$
$$+27\gamma^5+117\gamma\lambda^4+18\lambda^5\Big)+\lambda^2\left(\gamma^3(\lambda\omega(32-3\lambda\omega)+26)+2\gamma^2\lambda(\lambda\omega(5-3\lambda\omega)+29)+12\gamma^4\omega+\gamma\lambda^2(7-3\lambda\omega(\lambda\omega+4))-2\lambda^3(\lambda\omega+1)\right)$$
$$-8\gamma^4e^{3\omega(\gamma+\lambda)}\left(24\gamma\omega^2(\gamma+\lambda)^2+4\omega(\gamma+\lambda)(10\gamma+9\lambda)-\gamma+54\lambda\right)+5\gamma^4e^{4\omega(\gamma+\lambda)}(20\gamma\omega(\gamma+\lambda)-7\gamma+20\lambda)$$
$$:=\Pi_3$$

$$\diamond\left[\frac{\partial\Pi_3}{\partial\omega}\right]=-2\lambda e^{\omega(\gamma+\lambda)}\Big(2\gamma^2\omega^3(\gamma+\lambda)^3\left(3\gamma^2+2\gamma\lambda+\lambda^2\right)+2\gamma\omega^2(\gamma+\lambda)^2\left(25\gamma^2\lambda+30\gamma^3+14\gamma\lambda^2+\lambda^3\right)+\gamma\omega(\gamma+\lambda)\left(162\gamma^2\lambda+161\gamma^3+125\gamma\lambda^2+20\lambda^3\right)$$
$$+165\gamma^2\lambda^2+131\gamma^3\lambda+119\gamma^4+47\gamma\lambda^3+4\lambda^4\Big)+e^{2\omega(\gamma+\lambda)}\Big(6\gamma^2\omega^2(\gamma+\lambda)^2\left(105\gamma^2\lambda+128\gamma^3+30\gamma\lambda^2+6\lambda^3\right)$$
$$+3\gamma\omega(\gamma+\lambda)\left(188\gamma^2\lambda^2+406\gamma^3\lambda+245\gamma^4+100\gamma\lambda^3+18\lambda^4\right)+346\gamma^3\lambda^2+410\gamma^2\lambda^3+36\gamma^5\omega^4(\gamma+\lambda)^4+6\gamma^4\omega^3(\gamma+\lambda)^3(49\gamma+18\lambda)$$
$$+875\gamma^4\lambda+177\gamma^5+144\gamma\lambda^4+18\lambda^5\Big)-\lambda^2\left(\gamma^2\lambda(3\lambda\omega-10)-6\gamma^3+\gamma\lambda^2(3\lambda\omega+5)+\lambda^3\right)$$
$$-4\gamma^4e^{3\omega(\gamma+\lambda)}\left(72\gamma\omega^2(\gamma+\lambda)^2+12\omega(\gamma+\lambda)(14\gamma+9\lambda)+37\gamma+198\lambda\right)+20\gamma^4e^{4\omega(\gamma+\lambda)}(10\gamma\omega(\gamma+\lambda)-\gamma+10\lambda)$$
$$:=\Pi_4$$

$$\diamond\left[\frac{\partial\Pi_4}{\partial\omega}\right]=-2\lambda e^{\omega(\gamma+\lambda)}\Big(2\gamma^2\omega^3(\gamma+\lambda)^3\left(3\gamma^2+2\gamma\lambda+\lambda^2\right)+2\gamma\omega^2(\gamma+\lambda)^2\left(31\gamma^2\lambda+39\gamma^3+17\gamma\lambda^2+\lambda^3\right)+\gamma\omega(\gamma+\lambda)\left(262\gamma^2\lambda+281\gamma^3+181\gamma\lambda^2+24\lambda^3\right)$$
$$+290\gamma^2\lambda^2+293\gamma^3\lambda+280\gamma^4+67\gamma\lambda^3+4\lambda^4\Big)+e^{2\omega(\gamma+\lambda)}\Big(6\gamma^2\omega^2(\gamma+\lambda)^2\left(264\gamma^2\lambda+403\gamma^3+60\gamma\lambda^2+12\lambda^3\right)$$
$$+6\gamma\omega(\gamma+\lambda)\left(248\gamma^2\lambda^2+616\gamma^3\lambda+501\gamma^4+112\gamma\lambda^3+18\lambda^4\right)+1256\gamma^3\lambda^2+1120\gamma^2\lambda^3+72\gamma^5\omega^4(\gamma+\lambda)^4+12\gamma^4\omega^3(\gamma+\lambda)^3(61\gamma+18\lambda)$$
$$+2968\gamma^4\lambda+1089\gamma^5+342\gamma\lambda^4+36\lambda^5\Big)-36\gamma^4e^{3\omega(\gamma+\lambda)}\left(24\gamma\omega^2(\gamma+\lambda)^2+36\omega(\gamma+\lambda)(2\gamma+\lambda)+31\gamma+78\lambda\right)$$
$$+40\gamma^4e^{4\omega(\gamma+\lambda)}(20\gamma\omega(\gamma+\lambda)+3\gamma+20\lambda)-3\gamma\lambda^4$$
$$:=\Pi_5$$

$$\diamond\left[\frac{\partial\Pi_5}{\partial\omega}\right]=\lambda\Big(-2\gamma^2\omega^3(\gamma+\lambda)^3\left(3\gamma^2+2\gamma\lambda+\lambda^2\right)-2\gamma\omega^2(\gamma+\lambda)^2\left(37\gamma^2\lambda+48\gamma^3+20\gamma\lambda^2+\lambda^3\right)-\gamma\omega(\gamma+\lambda)\left(386\gamma^2\lambda+437\gamma^3+249\gamma\lambda^2+28\lambda^3\right)$$
$$-471\gamma^2\lambda^2-555\gamma^3\lambda-561\gamma^4-91\gamma\lambda^3-4\lambda^4\Big)+4e^{\omega(\gamma+\lambda)}\Big(3\gamma^2\omega^2(\gamma+\lambda)^2\left(159\gamma^2\lambda+293\gamma^3+30\gamma\lambda^2+6\lambda^3\right)$$
$$+3\gamma\omega(\gamma+\lambda)\left(154\gamma^2\lambda^2+440\gamma^3\lambda+452\gamma^4+62\gamma\lambda^3+9\lambda^4\right)+500\gamma^3\lambda^2+364\gamma^2\lambda^3+18\gamma^5\omega^4(\gamma+\lambda)^4+3\gamma^4\omega^3(\gamma+\lambda)^3(73\gamma+18\lambda)$$
$$+1204\gamma^4\lambda+648\gamma^5+99\gamma\lambda^4+9\lambda^5\Big)-54\gamma^4e^{2\omega(\gamma+\lambda)}\left(24\gamma\omega^2(\gamma+\lambda)^2+4\omega(\gamma+\lambda)(22\gamma+9\lambda)+55\gamma+90\lambda\right)$$
$$+320\gamma^4e^{3\omega(\gamma+\lambda)}(5\gamma\omega(\gamma+\lambda)+2\gamma+5\lambda)$$
$$:=\Pi_6$$

$$\diamond\left[\frac{\partial\Pi_6}{\partial\omega}\right]=\gamma\lambda\Big(-6\gamma\omega^2(\gamma+\lambda)^2\left(3\gamma^2+2\gamma\lambda+\lambda^2\right)-4\omega(\gamma+\lambda)\left(37\gamma^2\lambda+48\gamma^3+20\gamma\lambda^2+\lambda^3\right)-386\gamma^2\lambda-437\gamma^3-249\gamma\lambda^2-28\lambda^3\Big)$$
$$+4e^{\omega(\gamma+\lambda)}\Big(3\gamma^2\omega^2(\gamma+\lambda)^2\left(213\gamma^2\lambda+512\gamma^3+30\gamma\lambda^2+6\lambda^3\right)+3\gamma\omega(\gamma+\lambda)\left(214\gamma^2\lambda^2+758\gamma^3\lambda+1038\gamma^4+74\gamma\lambda^3+9\lambda^4\right)+962\gamma^3\lambda^2$$
$$+550\gamma^2\lambda^3+18\gamma^5\omega^4(\gamma+\lambda)^4+3\gamma^4\omega^3(\gamma+\lambda)^3(97\gamma+18\lambda)+2524\gamma^4\lambda+2004\gamma^5+126\gamma\lambda^4+9\lambda^5\Big)$$
$$-108\gamma^4e^{2\omega(\gamma+\lambda)}\left(24\gamma\omega^2(\gamma+\lambda)^2+4\omega(\gamma+\lambda)(28\gamma+9\lambda)+99\gamma+108\lambda\right)+320\gamma^4e^{3\omega(\gamma+\lambda)}(15\gamma\omega(\gamma+\lambda)+11\gamma+15\lambda)$$
$$:=\Pi_7$$

$$\diamond\left[\frac{\partial\Pi_7}{\partial\omega}\right]=e^{\omega(\gamma+\lambda)}\Big(3\gamma^2\omega^2(\gamma+\lambda)^2\left(267\gamma^2\lambda+803\gamma^3+30\gamma\lambda^2+6\lambda^3\right)+3\gamma\omega(\gamma+\lambda)\left(274\gamma^2\lambda^2+1184\gamma^3\lambda+2062\gamma^4+86\gamma\lambda^3+9\lambda^4\right)+1604\gamma^3\lambda^2$$
$$+772\gamma^2\lambda^3+18\gamma^5\omega^4(\gamma+\lambda)^4+3\gamma^4\omega^3(\gamma+\lambda)^3(121\gamma+18\lambda)+4798\gamma^4\lambda+5118\gamma^5+153\gamma\lambda^4+9\lambda^5\Big)-\gamma\lambda\Big(3\gamma\omega(\gamma+\lambda)\left(3\gamma^2+2\gamma\lambda+\lambda^2\right)$$
$$+37\gamma^2\lambda+48\gamma^3+20\gamma\lambda^2+\lambda^3\Big)-54\gamma^4e^{2\omega(\gamma+\lambda)}\left(24\gamma\omega^2(\gamma+\lambda)^2+4\omega(\gamma+\lambda)(34\gamma+9\lambda)+155\gamma+126\lambda\right)$$
$$+240\gamma^4e^{3\omega(\gamma+\lambda)}(15\gamma\omega(\gamma+\lambda)+16\gamma+15\lambda)$$
$$:=\Pi_8$$

$$\diamond\left[\frac{\partial\Pi_8}{\partial\omega}\right]=e^{\omega(\gamma+\lambda)}\Big(3\gamma^2\omega^2(\gamma+\lambda)^2\left(321\gamma^2\lambda+1166\gamma^3+30\gamma\lambda^2+6\lambda^3\right)+3\gamma\omega(\gamma+\lambda)\left(334\gamma^2\lambda^2+1718\gamma^3\lambda+3668\gamma^4+98\gamma\lambda^3+9\lambda^4\right)+2426\gamma^3\lambda^2$$
$$+1030\gamma^2\lambda^3+18\gamma^5\omega^4(\gamma+\lambda)^4+3\gamma^4\omega^3(\gamma+\lambda)^3(145\gamma+18\lambda)+8350\gamma^4\lambda+11304\gamma^5+180\gamma\lambda^4+9\lambda^5\Big)-3\gamma^2\lambda\left(3\gamma^2+2\gamma\lambda+\lambda^2\right)$$

$$
\begin{aligned}
&\quad -108\gamma^4 e^{2\omega(\gamma+\lambda)}\left(24\gamma\omega^2(\gamma+\lambda)^2+4\omega(\gamma+\lambda)(40\gamma+9\lambda)+223\gamma+144\lambda\right)+2160\gamma^4 e^{3\omega(\gamma+\lambda)}(5\gamma\omega(\gamma+\lambda)+7\gamma+5\lambda)\\
&:=\Pi_9\\
\diamond\left[\frac{\partial\Pi_9}{\partial\omega}\right]&=2\gamma^3\lambda^2(63\lambda\omega(\lambda\omega+12)+1714)+\gamma^2\lambda^3(3\lambda\omega(6\lambda\omega+119)+1324)+3\gamma^8\omega^3(24\lambda\omega+169)+3\gamma^7\omega^2\left(-1728e^{\omega(\gamma+\lambda)}+3\lambda\omega(12\lambda\omega+175)+1601\right)\\
&\quad +3\gamma^6\omega\left(-576(6\lambda\omega+23)e^{\omega(\gamma+\lambda)}+10800e^{2\omega(\gamma+\lambda)}+\lambda\omega(3\lambda\omega(8\lambda\omega+187)+3577)+6000\right)\\
&\quad +3\gamma^5\left(720(15\lambda\omega+26)e^{2\omega(\gamma+\lambda)}-72(4\lambda\omega(6\lambda\omega+55)+303)e^{\omega(\gamma+\lambda)}+\lambda\omega(\lambda\omega(\lambda\omega(6\lambda\omega+223)+2381)+8360)+7436\right)\\
&\quad +\gamma^4\lambda\left(-3888(2\lambda\omega+9)e^{\omega(\gamma+\lambda)}+32400e^{2\omega(\gamma+\lambda)}+27\lambda\omega(\lambda\omega(2\lambda\omega+49)+306)+13504\right)+18\gamma^9\omega^4+9\gamma\lambda^4(3\lambda\omega+23)+9\lambda^5\\
&:=\Pi_{10}\\
\diamond\left[\frac{\partial\Pi_{10}}{\partial\omega}\right]&=2\gamma\omega(\gamma+\lambda)\left(375\gamma^2\lambda+1601\gamma^3+30\gamma\lambda^2+6\lambda^3\right)+394\gamma^2\lambda^2+24\gamma^4\omega^3(\gamma+\lambda)^3+3\gamma^3\omega^2(\gamma+\lambda)^2(169\gamma+18\lambda)+2360\gamma^3\lambda+6000\gamma^4+110\gamma\lambda^3\\
&\quad +9\lambda^4+72\gamma^3e^{\omega(\gamma+\lambda)}\Big(-4\gamma^2\omega\left(-75e^{\omega(\gamma+\lambda)}+12\lambda\omega+58\right)-24\gamma^3\omega^2+\gamma\left(10(30\lambda\omega+67)e^{\omega(\gamma+\lambda)}-4\lambda\omega(6\lambda\omega+67)-487\right)\\
&\quad -6\lambda\left(-50e^{\omega(\gamma+\lambda)}+6\lambda\omega+33\right)\Big)\\
&\geq 2\gamma\omega(\gamma+\lambda)\left(375\gamma^2\lambda+1601\gamma^3+30\gamma\lambda^2+6\lambda^3\right)+394\gamma^2\lambda^2+24\gamma^4\omega^3(\gamma+\lambda)^3+3\gamma^3\omega^2(\gamma+\lambda)^2(169\gamma+18\lambda)+2360\gamma^3\lambda+6000\gamma^4\\
&\quad +110\gamma\lambda^3+9\lambda^4+216\gamma^3e^{\omega(\gamma+\lambda)}\left(92\gamma\omega^2(\gamma+\lambda)^2+2\omega(\gamma+\lambda)(123\gamma+44\lambda)+61\gamma+34\lambda\right)\geq 0
\end{aligned}
$$

$$
\begin{aligned}
&\Pi_{10}\geq\Pi_{10}|_{\omega=0}=3428\gamma^3\lambda^2+1324\gamma^2\lambda^3+10912\gamma^4\lambda+13020\gamma^5+207\gamma\lambda^4+9\lambda^5\geq 0\\
&\Pi_9\geq\Pi_9|_{\omega=0}=2420\gamma^3\lambda^2+1027\gamma^2\lambda^3+3589\gamma^4\lambda+2340\gamma^5+180\gamma\lambda^4+9\lambda^5\geq 0\\
&\Pi_8\geq\Pi_8|_{\omega=0}=(\gamma+\lambda)\left(609\gamma^2\lambda^2+958\gamma^3\lambda+588\gamma^4+143\gamma\lambda^3+9\lambda^4\right)\geq 0\\
&\Pi_7\geq\Pi_7|_{\omega=0}=(\gamma+\lambda)^2\left(1107\gamma^2\lambda+844\gamma^3+404\gamma\lambda^2+36\lambda^3\right)\geq 0\\
&\Pi_6=\Pi_6|_{\omega=0}=(\gamma+\lambda)^3\left(262\gamma^2+209\gamma\lambda+32\lambda^2\right)\geq 0\\
&\Pi_5\geq\Pi_5|_{\omega=0}=(\gamma+\lambda)^4(93\gamma+28\lambda)\geq 0\\
&\Pi_4\geq\Pi_4|_{\omega=0}=9(\gamma+\lambda)^5\geq 0\\
&\Pi_3\geq\Pi_3|_{\omega=0}=0\\
&\Pi_2\geq\Pi_2|_{\omega=0}=0\\
&\Pi_1\geq\Pi_1|_{\omega=0}=0.
\end{aligned}
$$

In the computation, we mainly use the inequality $e^x\geq x+1$. Therefore, we have $g_3(\omega)\geq g_3(0)=0$ for any $\omega\geq 0$. □

EC.3.5. Closed-Form Expressions for $\theta=0$ Case

In this case, similar to Schmittlein et al. (1987), once a customer becomes inactive, he will not switch back to the active state again (unless an ad is pushed to him). Then, we can compute the objective as

$$\Psi^{NR}(\omega)=\frac{(\lambda+\gamma)\left(e^{(\lambda+\gamma)\omega}(\lambda r-\gamma c)-\lambda(c+r)\right)}{\lambda(e^{(\lambda+\gamma)\omega}-1)+\gamma(\lambda+\gamma)e^{(\lambda+\gamma)\omega}\omega}.$$

Then, similar to Appendix EC.1, we can compute the condition when a finite optimal threshold exists. In addition, the optimal threshold and the corresponding objective can be explicitly solved. As for the comparative statics, we can follow the similar proof in Appendix EC.1. The results are summarized as follows:

1. When $c/r<\lambda/\gamma$, $\Psi^{NR}(\omega)$ first increases and then decreases in ω, and the optimal $\omega^{NR,*}$ is given as

$$\omega^{NR,*}=-\frac{c}{\gamma(c+r)}-\frac{1}{\gamma+\lambda}\left(W_{-1}\left[-\frac{e^{-\frac{2\gamma c+\lambda c+\gamma r}{\gamma(c+r)}}(\lambda r-\gamma c)}{\lambda(c+r)}\right]+1\right),$$

and the corresponding long-term average expected profit is

$$\Psi^{NR}(\omega^{NR,*}) = \frac{(\gamma+\lambda)(c+r)(\gamma c-\lambda r)}{\lambda r-\gamma c-\gamma(c+r)W_{-1}\left[-\frac{e^{-\frac{2\gamma c+\lambda c+\gamma r}{\gamma(c+r)}}(\lambda r-\gamma c)}{\lambda(c+r)}\right]},$$

where $W_{-1}(x)$ is the non-principal branch of the Lambert W function, which is the smaller solution (if exists) of $ye^y = x$ that is smaller than -1. Furthermore, $\omega^{NR,*}$ increases in c, decreases in r and λ, and first decreases and then increases in γ.

2. When $c/r \geq \lambda/\gamma$, $\Psi^{NR}(\omega)$ always increases in ω and hence the optimal policy never pushes ad campaigns.

□

EC.4 Omitted Details in Appendix A.9

For each X_k and any $u > \max\{r, c-r\}$, we have the following inequality

$$\begin{aligned}\mathbb{P}(|X_k| \geq u) = \mathbb{P}(X_k \leq -u) &= \sum_{i=\lceil (r+u)/c\rceil}^{\infty}(1-p_1)(1-p_2)(1-p_3)^{i-2}p_3\\ &= (1-p_1)(1-p_2)(1-p_3)^{\lceil u/c\rceil-2}\\ &\leq (1-p_1)(1-p_2)(1-p_3)^{u/c-2}\\ &= \frac{(1-p_1)(1-p_2)}{(1-p_3)^2}e^{-\frac{1}{c}\log(\frac{1}{1-p_3})u} = O(e^{-K_X u}),\end{aligned}$$

which implies that X_k is a sub-exponential variable with the constant $K_X = \frac{1}{c}\log(\frac{1}{1-p_3})$. Thus, for $s \in [-1/K_X, 1/K_X]$, we have

$$\mathbb{E}\left[e^{(X_k-\mathbb{E}[X_k])s}\right] \leq e^{K_X^2 s^2}.$$

Similarly, for each Y_k and any $u > \omega_0+\omega_1$, we have the following inequality

$$\begin{aligned}\mathbb{P}(|Y_k| \geq u) = \mathbb{P}(Y_k \geq u) &\leq \mathbb{P}\left(Y_k \geq \omega_0+\omega_1+\omega_2\cdot\lfloor (u-\omega_0-\omega_1)/\omega_2\rfloor\right)\\ &= \sum_{i=\lfloor (u-\omega_0-\omega_1)/\omega_2\rfloor}^{\infty}(1-p_1)(1-p_2)(1-p_3)^i p_3 = O(e^{-K_Y u}),\end{aligned}$$

implying that Y_k is a sub-exponential variable with a constant K_Y. Thus, for $s \in [-1/K_Y, 1/K_Y]$, we have

$$\mathbb{E}\left[e^{(Y_k-\mathbb{E}[Y_k])s}\right] \leq e^{K_Y^2 s^2}.$$

Then, we can use the Chernoff bound to derive upper bounds for the probabilities of two low-probability events.

1. The probability that the total cycle length of the first n cycles is greater than T is as follows:

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=1}^{n} Y_k > T\right) &= \mathbb{P}\left(e^{(\sum_{k=1}^{n} Y_k)s} \geq e^{Ts}\right) \\
&\leq \mathbb{E}\left[e^{(\sum_{k=1}^{n} Y_k)s}\right] e^{-Ts} \\
&= e^{-Ts} \times \Pi_{k=1}^{n} \mathbb{E}\left[e^{Y_k s}\right] \\
&\leq \exp\left(nK_Y^2 s^2 + n\mathbb{E}[\kappa(\boldsymbol{\omega})]s - Ts\right) \\
&\leq \exp\left(\frac{T}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} K_Y^2 s^2 + \left(\frac{T}{\mathbb{E}\left[\kappa(\boldsymbol{\omega})\right]} - \epsilon\right)\mathbb{E}[\kappa(\boldsymbol{\omega})]s - Ts\right) \\
&= \exp\left(\frac{T}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} K_Y^2 s^2 - \epsilon\mathbb{E}[\kappa(\boldsymbol{\omega})] \times s\right),
\end{aligned}$$

where $s \in (0, 1/K_Y)$. Define $s_1 := \epsilon \times \frac{(\mathbb{E}[\kappa(\boldsymbol{\omega})])^2}{2TK_Y^2}$. Since $s_1 = O(\sqrt{\log T/T})$, we have $s_1 \in (0, 1/K_Y)$ when T is sufficiently large. Letting $s = s_1$ in the above inequality, we can derive that $\mathbb{P}\left(\sum_{k=1}^{n} Y_k > T\right) \leq \frac{1}{T}$, because $\epsilon \geq \sqrt{\frac{4K_Y^2 T \log T}{(\mathbb{E}[\kappa(\boldsymbol{\omega})])^3}}$.

2. The probability that the total profit of the first n cycles is less than $nr - qT$ is as follows:

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=1}^{n} X_k < nr - qT\right) &= \mathbb{P}\left(e^{(\sum_{k=1}^{n} X_k)s} \geq e^{(nr-qT)s}\right) \\
&\leq \exp\left(nK_X^2 s^2 + n\mathbb{E}[\pi(\boldsymbol{\omega})]s - nrs + qTs\right) \\
&\leq \exp\left(\frac{T}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} K_X^2 s^2 + \left(\frac{T}{\mathbb{E}\left[\kappa(\boldsymbol{\omega})\right]} - \epsilon - 1\right)\mathbb{E}[\pi(\boldsymbol{\omega})]s\right. \\
&\qquad \left. - \left(\frac{T}{\mathbb{E}\left[\kappa(\boldsymbol{\omega})\right]} - \epsilon\right) rs + \frac{r - \mathbb{E}[\pi(\boldsymbol{\omega})]}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} Ts\right) \\
&= \exp\left(\frac{T}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} K_X^2 s^2 - (\epsilon + 1)\mathbb{E}[\pi(\boldsymbol{\omega})]s + \epsilon rs\right) \\
&\leq \exp\left(\frac{T}{\mathbb{E}[\kappa(\boldsymbol{\omega})]} K_X^2 s^2 + (\epsilon r - (\epsilon + 1)\mathbb{E}[\pi(\boldsymbol{\omega}]) s\right) \\
&= \exp\left(-\frac{(\epsilon(r - \mathbb{E}[\pi(\boldsymbol{\omega})]) - \mathbb{E}[\pi(\boldsymbol{\omega})])^2 \mathbb{E}[\kappa(\boldsymbol{\omega})]}{4TK_X^2}\right) \\
&\leq \exp\left(-\frac{\epsilon^2 (r - \mathbb{E}[\pi(\boldsymbol{\omega})])^2 \mathbb{E}[\kappa(\boldsymbol{\omega})]}{16TK_X^2}\right) \\
&\leq \frac{1}{T},
\end{aligned}$$

where $s = -\frac{(\epsilon r - (\epsilon+1)\mathbb{E}[\pi(\boldsymbol{\omega})])\mathbb{E}[\kappa(\boldsymbol{\omega}]}{2TK_X^2} \in (-1/K_X, 0)$ for a sufficiently large T. The last inequality is because $\epsilon \geq \sqrt{\frac{16K_X^2 T \log T}{(\mathbb{E}[\kappa(\boldsymbol{\omega})])(r - \mathbb{E}[\pi(\boldsymbol{\omega})])^2}}$. □