Churning While Experimenting: Maximizing User Engagement in Recommendation Platforms

(Authors' names blinded for peer review)

Online media platforms rely on long-term user engagement to generate revenue. The primary operational lever they control is content recommendation, which determines what content should be suggested to each user. However, due to the constantly evolving supply of content and the heterogeneity in user behavior, achieving optimal recommendations is challenging. It necessitates a delicate balance between experimentation (to gauge the effectiveness of new content) and exploitation (recommending high-quality existing content).

Motivated by a real-world dataset, we present a novel framework for the platform recommendation problem with the aim of maximizing long-term user engagement. Our framework comprises a *behavioral* model and a *prescriptive* model. The behavioral component is used to model the period-over-period heterogeneous behavior of users and is grounded in the existing RFM-style models to capture customer lifetime value. The prescriptive component is used to model the dynamic optimization problem faced by the platform, with the experimentation policy being the key lever. Our main contribution is to analytically characterize the structure of the optimal experimentation policy. Doing so allows us to extract important managerial insights as we establish the myopic policy of targeting users based on their churn sensitivity (as advocated in Ascarza (2018)) is long-term optimal under business-relevant scenarios.

Key words: recommendation; supply-side experimentation; demand-side churning; first impression effect; customer lifetime value; marketing-operations interface

1. Introduction

Online media platforms have grown to dominate the music industry over the last decade, with digital music accounting for more than half of all revenues (Watson 2018), and with digital media's market share continuing to grow every year between 2015-2022 (IFPI 2022). As these media platforms grow, they are faced with increasingly complex operational challenges. At a high-level, online media platforms such as Spotify, YouTube Music, and NetEase need to coordinate interactions between content creators who supply new media (e.g., music, videos, etc.), and users who engage with the media, and from whom the platform generates revenue either directly by paid subscriptions, or indirectly by display advertising. In this rich ecosystem, we focus on a particular aspect of coordination between content creators and platform users we term the *platform recommendation problem*, in which the platform must decide how to recommend content to users.

The platform recommendation problem is complicated by a mix of factors related to the nature of the content (supply), user behavior (demand), and the platform's information about each. In terms of supply, the specific catalog of media the platform can recommend is constantly updating as creators upload new content. There is also significant heterogeneity in the quality of such content. For content that has been widely displayed on the platform, it may be possible to estimate its click-through rate (CTR) and optimize recommendations accordingly. However, old content quickly becomes stale, and with new uploaded content, there is little information to inform the platform's recommendation decision. Further, gathering information on new content requires careful experimentation, as showing low-quality content to a user may decrease their future engagement with the platform.

Adding to the complexity, user engagement can itself be heterogeneous depending on the "state" of the user. For instance, as we demonstrate through our analysis of a real-world impressions dataset (cf. Section 2), new users are more likely to *churn* as a result of a single "poor" interaction compared to regular users. This observation is often referred to as the *first impression effect* (Lindgaard et al. 2006), where the future engagement of new users with the platform disproportionately depends on the outcomes of their first few interactions.

Taken together, the platform recommendation problem, then, is how to maximize long-term total engagement[1] while appropriately experimenting with new creatives, considering the heterogeneous churn behavior of users. To address this problem, in this work, we present a novel data-inspired framework for the platform recommendation problem with the aim of maximizing long-term user engagement. Our framework comprises a *behavioral* model and a *prescriptive* model. The behavioral component is used to model the period-over-period heterogeneous behavior of users and is grounded in the existing RFM-style (recency, frequency, and monetary) models to capture customer lifetime

[1] Similar to Netzer et al. (2008) and Fader et al. (2010), we focus on the engagement as opposed to the dollar values.

value. The prescriptive component is used to model the dynamic optimization problem faced by the platform with the experimentation policy being the key lever. Under our model, a candidate for an optimal policy is the one advocated by Ascarza (2018): experiment in increasing order of users' churn sensitivity to experimentation (we label this policy "churn minimization"). However, as discussed in Lemmens and Gupta (2020), churn minimization is sub-optimal in general as it is blind to the long-term value of the users. Remarkably, our key analytical result establishes that churn minimization is long-term optimal under simple conditions on the parameters governing the behavior of users.

1.1. Our Contributions

We study the inherent tension between experimenting with new content and recommending high-quality established content, in the presence of heterogeneous churning behavior. A summary of our key contributions is as follows.

First, we explore the NetEase Cloud Music dataset (Zhang et al. 2020), the second largest music media platform in China with more than 200+ million active daily users (Dredge 2023). In this dataset, we find the content (supply) to be of heterogeneous quality, and also find observational evidence of a significant first impression effect on the demand side. Specifically, we find that new users who do not engage with (click) the recommended creative are around 5 percentage points less likely to return to the platform than new users who do engage with the recommendations. On the other hand, we find this number to be significantly lower for "regular" users (cf. Section 2.2). Further, we find evidence that in spite of this heterogeneous user behavior, the platform might be showing new creatives to users regardless of how long they have used the platform, a policy we refer to as *blind randomization* (cf. Section 2.3).

Second, based on our data analysis, we propose a novel framework for platform recommendation with the goal of maximizing long-term user engagement (cf. Section 3). Our framework captures two key features in this marketplace: (1) demand-side heterogeneous churning (via an RFM-style behavioral model of users behavior) and (2) supply-side experimentation to learn the CTR of newly created content (via a dynamic optimization model that builds upon the behavioral model). To the best of our knowledge, we are the first to analytically integrate an RFM-style model with a prescriptive model to characterize the structure of the optimal experimentation policy.

Third, we study algorithms for the platform recommendation problem in our model, placing special emphasis on two simple experimentation policies, blind randomization (BR) and churn minimization (CM). BR is a natural baseline policy that ignores user state and *blindly* experiments. We show that BR policies can exhibit arbitrarily poor performance (cf. Section 4). In response, we analyze CM as an alternative policy. While many approaches feel intuitively "reasonable", including experimenting on the oldest users, users with the lowest overall churn probability, users with the lowest future value to the platform, and so on, we focus on one simple and easily implementable heuristic, CM, which experiments on users who are *least sensitive* to experimentation (Ascarza 2018). We show that under interpretable restrictions on our parameters, CM is the long-term optimal experimentation policy (cf. Theorems 1 and 2 and Proposition 2). We supplement our analytical results with numerics performed over a wide range of parameters which confirm the efficacy of CM beyond the regimes we study analytically (cf. Section 6).

Overall, we demonstrate the importance of carefully conducting experimentation for content recommendation. We show that policies that do not account for user state can significantly underperform due to mishandling the first impression effect. Instead, we provide a rigorous analysis characterizing the regimes when the simple targeting rule of churn minimization is long-term optimal. This managerial insight is surprising as on the surface, CM is only myopically optimal (cf. Section 5). We note that though our writing focuses on recommendation platforms, our modeling framework is general and can potentially be applied to other two-sided markets such as labor, transportation, and lodging. We discuss this further in Section 7.

1.2. Literature Review

Our work intersects with several streams of literature in marketing, machine learning, operations management, and behavioral psychology. Here, we overview some of these streams and explain how our work contributes to and/or differs from each.

Customer Lifetime Value (CLV) and Churn Management. There has been a sustained interest in the marketing community to understand CLV. Seminal works such as Fader et al. (2005), Fader and Hardie (2007), and Fader et al. (2010) have proposed dynamic RFM-style models for users behavior to capture CLV. Our stochastic model of users behavior is grounded in this literature, but captures certain elements that are novel. For instance, we allow the user behavior (retention and churn parameters) to vary as a function of the user's state. Furthermore, in addition to the behavior being Markovian, we do not assume any distribution over the parameters (contrast with Fader et al. (2010) assuming a Beta distribution in their Assumptions 4 and 5). More importantly, for us, the user behavior model serves as an input to the downstream optimization problem, which is in stark contrast to Fader et al. (2005), Fader and Hardie (2007), and Fader et al. (2010) as they use such a model to primarily predict CLV and there is little (if any) prescriptive component. Our optimization problem focuses on the lever of experimentation / targeting with the goal of maximizing long-term user engagement. This inextricably requires one to manage churn, a topic of central interest in the literature. Ascarza (2018) has established that merely directing retention efforts towards customers deemed mostly likely to churn is insufficient to achieve optimal revenue. Platforms must also account for how sensitive customers are to these retention efforts. This in fact forms the basis of the experimentation policy ("churn minimization") we analyze rigorously. However, as Lemmens and Gupta (2020) rightly point out, such a policy is sub-optimal in general as it myopically focuses on minimizing churn but does not necessarily account for the long-term value of the users. Perhaps surprisingly, our main result establishes that churn minimization is in fact the long-term optimal policy under a simple condition on the parameters governing users behavior. As such, we provide a meaningful bridge between the views in Ascarza (2018) and Lemmens and Gupta (2020), under a well-grounded RFM-style model of user behavior.

First Impressions and Churning Behavior. A key motivation for our model is the first impression effect, where new users are more likely to churn than regular users as the result of one poor interaction. There is a growing literature that models churning and emphasizes the importance of the first platform interaction with a user (Liu et al. 2016, Padilla and Ascarza 2017, Agnew et al. 2018a). There is also a rich behavioral literature that documents the importance of first impressions for humans (Rabin and Schrag 1999, Lindgaard et al. 2006, Agnew et al. 2018b). Our work is one of the first to directly leverage this theory to derive and analyze policies for content recommendation.

Learning from Experimentation. There is a recent surge in learning-from-experimentation based approaches for pricing and recommendation problems on online platforms (Bastani et al. 2022, Misra et al. 2019). Such approaches leverage general learning theory (Sutton and Barto 2018) and tailor it to a specific problem of interest, which allows one to extract application-specific insights. Our work follows a similar track in that we focus on the problem of experimenting with new/uncertain content in the presence of heterogeneous churning. Though there exist a few learning-based approaches for content recommendation (Kveton et al. 2015, Schwartz et al. 2017, McInerney et al. 2018, Xu et al. 2022), most of them are "bandit" inspired, whereas our focus is on coordinating how to experiment while managing user churn (as opposed to the studying the learning itself).

Recommendation Systems. More generally, our work contributes to the broader literature on recommendation systems for online platforms. Given the ubiquity of recommendation problems (Alexander 2020), it would be impossible to describe all of the work done in this space. Instead, we highlight several reference works in this domain (Wei et al. 2007, Melville and Sindhwani 2010), including the well-known literature on the "Netflix prize" (Bennett et al. 2007). The high-level idea in these works is to use historical data to predict the behavior of the next user and then optimize the recommendation accordingly. A key limitation of much work in this area is that they are myopic in nature, i.e., their objective is to maximize "immediate reward" (e.g., probability user clicks on the recommendation), and they do not necessarily capture the "long-term value". In this work, we also study a type of myopic policy, but one which we explicitly connect to the long-run value of a recommendation by optimizing for the steady-state of the platform.

The rest of this paper is organized as follows. In Section 2, we perform an exploratory analysis of the NetEase Cloud Music dataset and highlight some of the salient features. In Section 3, we build on our data analysis and propose our modeling framework, which captures supply-side experimentation and demand-side heterogeneous churning. In Section 4, we highlight that "blind" policies can be highly sub-optimal for the objective of maximizing user engagement. In Section 5, we analyze a simple policy, churn minimization (CM), and characterize when it is optimal. In Section 6, we numerically study CM for a broad class of models, and demonstrate its performance is excellent and robust. Finally, in Section 7, we summarize our contributions and highlight avenues for future work. Various details (including most proofs) are deferred to the appendices. We code in `MATLAB` to numerically simulate/optimize various instances and provide the code at `BLINDED`.

2. Exploratory Data Analysis

We now introduce and analyze impressions data from the platform NetEase Cloud Music (Zhang et al. 2020). First, in Section 2.1, we explore the supply side of NetEase, the digital media produced by content creators (referred to as cards). Next, in Section 2.2, we study the demand side, the NetEase users who interact with the media. Finally, in Section 2.3, we discuss the platform, the operational levers under its control, and its objectives. Our key findings here are that the quality of supply-side (cards) is heterogeneous and so is the churn behavior of users (depending on their visit number and the recent interaction with the platform). These findings inform the development of the model in Section 3.

The NetEase Cloud Music Dataset. NetEase Cloud Music is the second largest online music platform in China, and with more than 200+ million daily active users, one of the largest in the world. In 2020, NetEase released a uniquely rich dataset consisting of more than 57 million impressions from more than 2 million users, collected in the month of November 2019 (Zhang et al. 2020). Each impression contains information about the user and the content they were recommended, which we use to construct detailed browsing histories. From these histories, we can see which content led to a conversion (click, like, etc.) and which did not, allowing us to study the value of each recommendation. By studying these histories and the associated recommendation outcomes, we get a rare glimpse into the inner workings of a massive online media platform and can directly observe how users behave and vary in their behaviors.

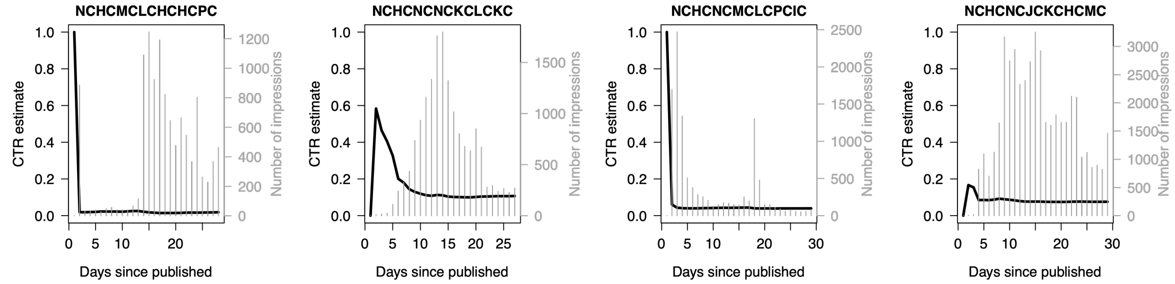
Our analysis of this dataset connects with a rich literature of work on extracting modeling elements from exploratory analysis of online media platform data such as at *Hulu* (Schwartz et al. 2011), *Reddit* (Glenski and Weninger 2017), and *Spotify* (Zhang et al. 2013). Indeed, given the complexity of these media environments, it is natural to look for guidance by seeing how users behave in data. However, relative to other publicly available datasets, the NetEase dataset is unrivaled in terms of both size and fine-grained detail, which as we will see, allows us to study user trends at a scale which has hitherto only been available behind proprietary walls of data access.

2.1. Content Creators

At NetEase, the supply in the market corresponds to the various *cards* (combinations of song and video) generated by content creators. Each card can have various intrinsic dimensions, such as the song/video it contains and the underlying artist. All such dimensions ultimately affect the "clickability" of the card, which can be summarized by its *click-through rate* (CTR), i.e., the probability it will be clicked if shown to a random user. For our modeling purposes, we primarily focus on the CTR of each card as the measure of its quality. Of course, not every card is created equal, and there is heterogeneity among the CTRs of various cards. Furthermore, the CTR of a card is a parameter that the platform must estimate by *experimenting*, i.e., by showing the card to a sufficiently large number of users and observing whether they click or not. In Fig. 1, we plot the evolution of a running sample estimate of the CTR for four randomly chosen cards.

Fig. 1 demonstrates the heterogeneity among the cards' CTRs, which stabilize after a few thousand impressions (e.g., the left-most card's CTR stabilizes to a value close to 0, whereas the right-most card's CTR tends to a value around 0.15). Naturally, learning this underlying "true" CTR for each card is valuable to the platform, as it can use such information to recommend a better portfolio of cards to the users, resulting in higher engagement. Indeed, we find the true

Figure 1 Evolution of the CTR estimates.



Note. Depicted are CTR estimates for four randomly chosen cards (ID is given at the top of each plot) from the NetEase dataset over time. The x-axis denotes the days since the card was published. The left side of the y-axis and the thick black line within each plot correspond to the empirical CTR estimate, i.e., the number of times the card was clicked divided by the number of times the card was shown over the period (accounting for all the impressions till this day). The right side of the y-axis and the gray vertical bars within each plot denote the number of times the card was shown on a given day.

underlying distribution of the CTRs of the cards can be excellently modeled via a Beta distribution (see Section A.1)[2], suggesting there is significant value in searching for high-value cards out on the long tail of the distribution. For our purposes, it suffices that the heterogeneity is significant enough to motivate a platform to experiment.

2.2. Platform Users

In a recommendation platform, the demand side of the market corresponds to the platform's millions of users. At a high-level, one way to think about the user base is to split it into two categories: (1) regular users and (2) new users. Regular users are the ones who have already spent considerable time on the platform and are less likely to churn than new users. To substantiate this view of users, in Fig. 2a, we plot the churn behavior of a subset of new users as they spend time on the platform. Out of 16083 new users, only 8002, 4437, 2644, 1645, 1089, and 746 returned to the platform for a second, third, fourth, fifth, sixth, and seventh visit, respectively. This suggests a churn rate of around 50% ($1-8002/16083$) after the first visit, monotonically decreasing to around 30% ($1-746/1089$) after the sixth visit (black dots in Fig. 2b), meaning the churn likelihood

[2] We note that modeling responses rates via a Beta distribution is common in the RFM literature (see for instance Assumption 4 of Fader and Hardie (2010)). Our data analysis provides real-world support for this assumption.

decreases as a user becomes more regular. In fact, as shown by the red dotted line in Fig. 2b, the following exponential decay explains the data very well:

$$\text{Churn rate after visit } k = 0.4957 \times \exp\left(-0.0938(k-1)\right) \text{ for } k = 1, 2, \ldots. \tag{1}$$

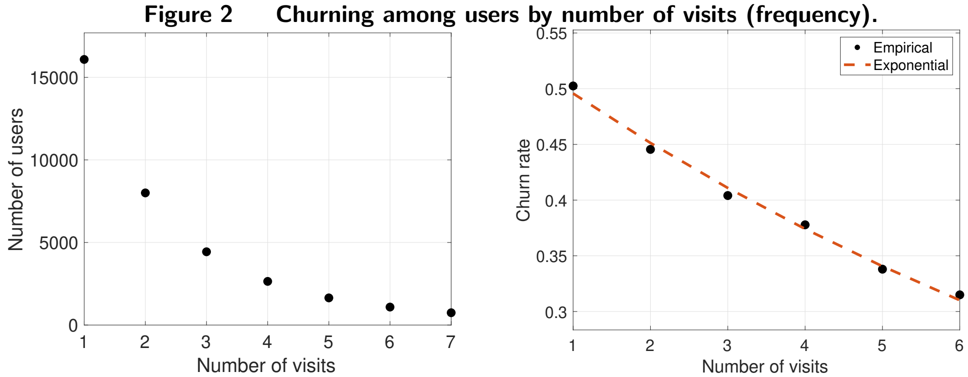
The two parameters (0.4957 and 0.0938) are calibrated to the data by minimizing the total squared error. Thus, like supply-side CTR, demand-side churn behavior is heterogeneous, and depends on the current "state" of each user (e.g., number of visits so far, i.e., the F (frequency) in RFM).

Moreover, we find that whether or a not a user clicks has a significant impact on their churn rate, which relates to the R (recency) in RFM. On an additional sample of 8846 users (users whose first visits were on days 11, 12, and 13 of November), we find the churn proportion (after the first visit) to be 47.27% if a user clicks during the visit, and 52.39% if they do not click. This suggests an increase of around 5 percentage points ("delta") if a user does not click in their first visit. Furthermore, in Section A.2, we repeat our analysis and find the churn rate in "regular" users to be much lower (around 5% given a click), with an increase of only 1 to 2 percentage points in the absence of a click, lending evidence to the existence of a first impression effect. Hence, in addition to the "base" churn rate, the "delta" also depends on the user state, which we will capture in our user behavior model. Such state-dependent user behavior is consistent with models discussed in Gupta and Lehmann (2005) and Fader and Hardie (2007). (Note that these works discuss "retention rate" to be increasing, which is equivalent to the churn rate to be decreasing as retention rate $=1-$ churn rate.)

2.3. Platform Recommendations

A recommendation platform is responsible for matching supply (cards) with demand (users) by recommending cards to the users when they use the app. As stated in Zhang et al. (2020), a goal of such a platform is to maximize the *long-term* user activity (e.g., number of clicks over a horizon), which naturally relates to the M (monetary) dimension of RFM. Given the heterogeneity in both supply (Section 2.1) and demand (Section 2.2), performing this coordination optimally is

Figure 2 Churning among users by number of visits (frequency).



(a) Churn progression (b) Churn rate

Note. We consider the churn behavior of users whose first visit was between days 7 and 10 (inclusive) (16083 such users). We do not include users with a later first visit since the dataset only contains information for 30 days and hence, including a user with a first visit on day 25 for example might not give them enough time to re-visit the platform multiple times. Subplot (a) shows the number of users (y-axis) who visited the platform at least x times. The black dots in subplot (b) display the corresponding churn rate after each visit. For example, out of the 8002 users who visited the platform twice, only 4437 returned for a third visit, implying a churn rate of $1 - 4437/8002 \approx 45\%$ after visit 2. Clearly, as the user spends more time on the platform, the churn rate decreases. Such a decay is modeled by the exponential fit to the data (red dotted line), given in Eq. (1).

challenging. In terms of supply, estimating the CTR of newly created cards is vital, and thus the platform must experiment with new creatives. With regards to the demand, controlling the churn rate across users is critical. All else being equal, a higher churn rate results in a lower number of clicks in the future, and thus, the platform needs to control the churn if it wishes to maximize clicks in the long term. Otherwise, if the platform optimizes myopically, it might end up maximizing the clicks in the short term and lose out on the long-term benefits due to high user churn.

To learn the CTR of any new card, the platform must experiment with the users by showing them the new card and observing their click behavior. Hence, the main algorithmic paradigm we focus on in this work is the *experimentation policy* used by the platform. Given the heterogeneous user behavior discussed in Section 2.2, it is natural to think of the policy as a function of the user state. In particular, given a new card and a corresponding *experimentation budget* (i.e., the number of times to show the card in order to learn its CTR), the platform needs to decide on how to optimally split this budget between the users in various states (e.g., regular and new users).

One simple experimentation policy is *blind randomization*, where the allocation of the experimentation budget is done independent of the users' state. As an illustration, consider the following example. Suppose there are 100 regular users, 100 new users, and 1 new card. Given an exogenous experimentation budget of 50 impressions, the blind experimentation policy allocates 25 impressions to regular users and 25 impressions to new users. In other words, 25 regular and 25 new users are shown the new card, whereas the remaining 150 users are shown some other card (an old card for which the platform has already learned the CTR). Intuitively, given the state-specific churn behavior discussed above, such blind randomization seems sub-optimal. However, it is worth highlighting that *any* experimentation policy that does not account for the state of the user will result in blind randomization automatically. In fact, by analyzing the NetEase dataset, we find evidence that blind randomization might be prevalent at NetEase (or perhaps even *more* experimentation on new users than on regular users, see Section A.3). Irrespective of the specific policy at NetEase, understanding the implications of blind randomization is important, especially in the presence of heterogeneous churning behavior. In this work, we not only do so but also focus on designing optimal experimentation policies, with the goal of maximizing the long-term expected number of clicks.

3. Model and Preliminaries

In this section, we introduce our modeling framework, inspired by the insights from our data analysis in Section 2. Our overall framework consists of a *behavioral* model for users behavior and a *prescriptive* model to characterize the optimal experimentation policy. In Section 3.1, we propose a stochastic model of user behavior and discuss how it relates to the existing RFM-style CLV models (Fader et al. 2005, 2010, Fader and Hardie 2007). We supplement our general model with two special cases that shed light on how it can capture a spectrum of user behaviors, including the one we discussed in our data analysis in Section 2 (Fig. 2). In Section 3.2, we formally describe the platform's optimization problem. In Section 3.3, we present a tractable upper bound to the optimization problem, which serves as a performance benchmark for our subsequent analysis.

3.1. A Stochastic Model for User Behavior

Motivated by our data analysis in Section 2.2, we model user behavior on the platform via a multi-period state-based Markov model. In a given period, each (non-churned) user is in some observable state $s \in \mathbb{S}$, where $\mathbb{S} := \{1, \ldots, m\}$ has m states. For example, our state can capture the R (recency) and F (frequency) dimensions of RFM-style models.[3] In addition, consistent with the RFM models of Fader et al. (2005, 2010), we assume users go through two stages in their lifetime: they are active for some period and then become permanently inactive (see for example Assumption 1 of Fader et al. (2010)). We model a user being "permanently inactive" via an absorbing state, which we call the *quit* state q, denoting the user has churned. We define the state space augmented with the quit state by $\mathbb{S}^+ := \mathbb{S} \cup \{q\}$. (Our analysis works for both "contractual" (churning behavior observed) and "non-contractual" (unobserved) settings. This is because our optimization problem will focus on the users in *aggregation* as opposed to each *individual* user. This will become clear when we define the optimization problem in Section 3.2.)

In each period, a user can either receive experimentation (e.g., be shown new card) or not (e.g., be shown old card). This maps to a "transaction opportunity" in the language of Fader et al. (2010).[4] Whether a user receives experimentation is determined by the platform's experimentation policy (defined in Section 3.2). We denote the platform's underlying action space by $\mathbb{A} := \{1, 2\}$, where action $a = 1$ corresponds to a user being shown an *old* card (no experimentation), and $a = 2$ corresponds to a user been show an experimental, *new* card. If a user is in state $s \in \mathbb{S}$ and receives action $a \in \mathbb{A}$, then the user transitions to state $s' \in \mathbb{S}^+$ with fixed, state and action dependent probability, $p_{sas'}$. As in the RFM models of Fader et al. (2005, 2010), these parameters are independent across users (conditioned on the state s); see for example Assumption 6 of Fader et al. (2010). We call the probability of transition to the quit state, p_{saq}, the *churn probability* for

[3] This will become clear when we discuss concrete examples (binary and funnel state spaces) of state spaces below.

[4] As in Fader et al. (2010), we work with a discrete-time system. An alternative would be to work in continuous time (Fader et al. 2005), which might be more apt for our application. However, this simplification enables a tractable closed-form analysis of the downstream optimization problem (platform recommendation) we study.

state s under action a. Such a treatment of churning allows for some customers to become inactive much sooner than others, as advocated by Fader et al. (2005, 2010) (see for example Assumption 6 of Fader et al. (2005)). Since the quit state q is absorbing, we have $p_{qaq} = 1$ for all $a \in \mathbb{A}$. Furthermore, we will call $r_s := 1 - p_{s1q}$ the *return probability* from $s \in \mathbb{S}$ under no experimentation ($a = 1$). It is worth mentioning that the return probability in our model is state-dependent, i.e., r_s depends on s. This is a generalization of the existing models such as Fader et al. (2010) where they assume the "dropout probability" to be state-independent (see Assumption 3 in Fader et al. (2010)). Our view here is consistent with the models discussed in Gupta and Lehmann (2005) and Fader and Hardie (2007). Further, observe that Fig. 2 provides observational evidence for it.

For each state $s \in \mathbb{S}$, we define $d_s := p_{s2q} - p_{s1q}$ as the difference in the churn probabilities between the two actions and refer to it as the *churn delta*. We assume that the experimental action ($a = 2$) is more likely to cause a user to churn than the tried-and-tested action ($a = 1$); thus, $d_s \geq 0$ for all $s \in \mathbb{S}$. This assumption is intuitive; if the platform has experimented enough with a card to estimate its CTR and finds that its value is low, it is natural to imagine the platform would discard it; hence, only high-value old cards are used. Mapping this to the RFM models (Fader et al. 2005, 2010), the delta here pertains to the R (recency effect) as it allows the user behavior to vary as a function of the most recent interaction. It is also possible to allow the user behavior to depend on the *second* most recent interaction (and so on) via an appropriate definition of the state and we will discuss this later by incorporating "momentum" effects in our state.

The state transition process (subject to some experimentation policy) repeats in the next period, and so on (infinite-horizon as in Fader et al. (2005, 2010) and Fader and Hardie (2007)). We suppose the system starts with no users ("period 0"), and λ_s users arrive to state $s \in \mathbb{S}$ at the beginning of each period.[5] To ensure the market size (i.e., total number of users) converges to a steady state, we assume every user eventually churns under any experimentation policy, i.e., $p_{saq} > 0$ (strict) for all $(s, a) \in \mathbb{S} \times \mathbb{A}$. We call this the "leakage" assumption.

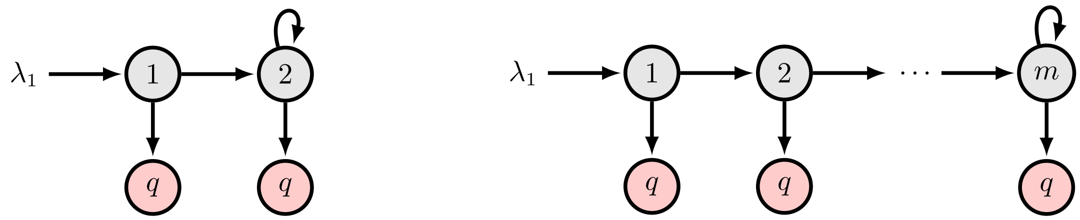
[5] We can allow for a *random* number of arrivals such that the *expected* arrivals equal $[\lambda_s]_s$. None of our results will change. This is because we will analyze a "fluid" system (as it will become clear in Section 3.2).

As alluded to above, our model is general enough to capture an array of user behaviors. We give special attention to two state space configurations: (1) the *binary* state space and (2) the *funnel* state space. The funnel state space is not only grounded in the RFM-style models but also captures the dynamics uncovered in Section 2 (recall Eq. (1)), whereas the binary state space serves as the simplest non-trivial setting to convey intuition and can be seen as a "warm-up" for the funnel setting. Furthermore, it is worth emphasizing that though we give a structure to the state space, we do not assume any additional parametric form over the transition probabilities $[p_{sas'}]_{(s,a,s')}$ but work with a generic form for now. Later on, we show how our general results allow us to extract clean insights when we give a parametric form to the transition probabilities as in Eq. (1).

Binary State Space. In Section 2.2, we described evidence for a first-impression effect where new users to the platform are more sensitive to experimental content and thus more likely to churn as the result of one poor interaction. We can capture these dynamics in our model by describing users as being in one of two states: (1) new and (2) regular. The "new" state corresponds to the user's first visit to the platform, after which they are classified as a "regular" user in their subsequent visits. When there are only two states, we will refer to the state space as *binary*. For the binary state space, we will assume users only arrive into the first state (i.e., $\lambda_2 = 0$), and there is a self-loop only in state 2 (i.e., $p_{1a1} = 0$ for all $a \in \mathbb{A}$), reflecting the mechanics of a recommendation platform. The binary state space under these assumptions is the simplest non-trivial model, and is appropriate to describe users who are highly sensitive to the quality of their recommendation during the first visit but for whose sensitivity diminishes after visit #1, which is consistent with our findings from the NetEase data. In Fig. 3 (left panel), we visualize how our binary model captures such user behavior.

Funnel State Space. A more granular way to represent user behavior is using their number of visits, i.e., F (frequency) of RFM, as the state. For instance, instead of doing a hard classification in terms of a "new" or a "regular" user, we can model the user state via an m-step spectrum. This generalizes the binary state model. Specifically, we will call a state space *funnel* if there is a

Figure 3 **Binary (left panel) and funnel (right panel) state spaces for user behavior.**



Note. To map the binary space (left) to our Section 2 data analysis, state 1 represents a "new" user (visit #1) and state 2 represents a "regular" user (visit #2 or higher). Arrows represent possible transitions. For the funnel (right), state $s \in \{1, \ldots, m\}$ can represent the number of visits of the user, i.e., the F (frequency) dimension of RFM, with state m denoting the visit number after which user behavior does not change. Note for both models we make the natural assumption that arrivals occur only into the first state, and self-loops only occur in the terminal state. Though not visible in the figures, both state spaces capture the R (recency) dimension of RFM via the "churn delta" discussed above, and the M (monetary) dimension of RFM via the "immediate reward" and "platform's objective" discussed in Section 3.2 below.

labeling of the states $1, 2, \ldots, m$ such that for every state $s < m$, users can only transition to state $s+1$ or churn, and for the terminal state m, users can only transition back into m or churn. As for the binary state space, for the funnel, we will assume users only arrive into the first state (i.e., $\lambda_s = 0$ for $s \neq 1$), and there is a self-loop only in the terminal state (i.e., $p_{sas} = 0$ for all $a \in \mathbb{A}$ and $s \neq m$). In this way, the funnel model under these assumptions captures instances where each state represents the visit number of the user (frequency), as shown in Fig. 3 (right panel), and where the m^{th} state captures our notion of a "regular" user. Naturally, the funnel captures the user model presented in Eq. (1) and Fig. 2 if we parameterize the transition probabilities accordingly. In fact, in Section 5.2, after we establish our general results for the funnel, we will interpret them in the light of such a parameterization.

Our funnel state space is data-inspired (Eq. (1) and Fig. 2) and captures the the R (recency) and F (frequency) dimensions of RFM via the churn delta and state definition (number of visits), respectively. Furthermore, we will discuss how our model captures the M (monetary) dimension in Section 3.2 below.[6] Nonetheless, we reckon that the funnel simplifies certain aspects of the

[6] Our decoupling of M from R and F is equivalent to the decoupling in Fader et al. (2005, 2010) (more in Section 3.2).

existing RFM models. For instance, we capture the R (recency) dimension via the *most* recent interaction and it is possible that the *second* most recent interaction affects user behavior as well ("momentum"). Such simplifications facilitate a tractable closed-form analysis of the downstream optimization problem (platform recommendation) and allow us to extract managerial insights into the structure of the optimal experimentation policy.[7] Irrespective, we will relax such simplifications by allowing for momentum effects for instance in our numerical study (Section 6). We next define the optimization problem faced by a platform aiming to maximize the long-term value.

3.2. The Platform Recommendation Problem

At the start of each period $t \in \{1, 2, \ldots\}$ (infinite-horizon), after the $\boldsymbol{\lambda} := [\lambda_s]_{s=1}^m$ arrivals, the *market state* is denoted by $\boldsymbol{\Lambda}^{(t)} := [\Lambda_s^{(t)}]_{s=1}^m$, where $\Lambda_s^{(t)}$ equals the aggregate number of users in state s at the beginning of period t (including the period t arrivals). We initialize the market state at zero, i.e., $\boldsymbol{\Lambda}^{(0)} = (0, \ldots, 0)$. The key operational lever we focus on is the *experimentation policy* employed by the platform, i.e., which of the $\boldsymbol{\Lambda}^{(t)}$ users to experiment on. In particular, as discussed in Section 2, in order to maintain a healthy balance between the supply and demand sides of the market, the platform needs to experiment with its user base. For example, in content recommendation, doing so enables the platform to learn the quality (e.g., click probability) of the new content (supply side), enabling it to separate out the high-quality content from the low-quality content. The goal of the platform is to maximize the long-term value, which maps to the M (monetary) dimension of RFM. We elaborate on the optimization problem by discussing the various modules next.

Experimentation Budget. To focus on the nature of the optimal experimentation policy, we abstract away the learning dynamics and instead posit that in each period, the platform must experiment on some fixed fraction of the total market. That is, in period t, it must experiment on

[7] In contrast with existing works such as Fader et al. (2005, 2010) and Fader and Hardie (2007) that employ RFM-style models, our goal is not only to capture the CLV but compute the optimal experimentation policy. This requires some additional considerations when building our model as a rather involved model of user behavior would render the optimization problem analytically intractable.

exactly $B^{(t)} := \frac{\mathbf{1}^\top \boldsymbol{\Lambda}^{(t)}}{\eta}$ users, where $\eta > 1$ is an exogenous parameter. We refer to $B^{(t)}$ as the *experimentation budget* for period t. To illustrate, if, for example, $\eta = 20$, then the platform experiments on $\frac{1}{\eta} = 5\%$ of its user base (this does not imply that the same set of users is experimented upon in each period). Intuitively, such constant fraction experimentation is appropriate when the two sides of the market are growing in accordance with each other, i.e., the number of new supply-side arrivals (e.g., content cards) is proportional to the demand-side market size. For example, such proportional growth may be expected in the case where content is user generated (e.g., NetEase, YouTube, TikTok, etc.), and so growth in the user base implies growth in content.

Experimentation Policy. Given the market state $\boldsymbol{\Lambda}^{(t)}$, we denote the *experimentation policy* by $\Pi(\boldsymbol{\Lambda}^{(t)}) \in \mathbb{R}^m$. The s^{th} element of the vector $\Pi(\boldsymbol{\Lambda}^{(t)})$ equals the number of users in state s that receive experimentation during period t, and we denote it $\pi_s^{(t)}$, with $\boldsymbol{\pi}^{(t)} := [\pi_s^{(t)}]_{s=1}^m$. Naturally, for all periods t, the following constraints must be obeyed:

$$\boldsymbol{\pi}^{(t)} \leq \boldsymbol{\Lambda}^{(t)}, \tag{2a}$$

$$\mathbf{1}^\top \boldsymbol{\pi}^{(t)} = B^{(t)}. \tag{2b}$$

The first set of constraints Eq. (2a) enforces that the platform cannot experiment on more users than it has in each state. The second set of constraints Eq. (2b) ensures that the amount of experimentation adds up to the experimentation budget.

Immediate Reward. As a function of the market state $\boldsymbol{\Lambda}^{(t)}$ and the experimentation policy $\boldsymbol{\pi}^{(t)}$, the platform reaps an expected immediate reward of $R(\boldsymbol{\Lambda}^{(t)}, \boldsymbol{\pi}^{(t)})$, which we denote by $R^{(t)}$. It maps to the M (monetary) dimension of RFM. Given the primitives c_1 and c_2 that represent the expected reward for action 1 (old card) and action 2 (new card), respectively, the expected immediate reward in period t can be expressed as:

$$R^{(t)} = \sum_{s=1}^{m} \left\{ (\Lambda_s^{(t)} - \pi_s^{(t)}) c_1 + \pi_s^{(t)} c_2 \right\}, \tag{3}$$

where the sum is taken over the states. For example, if $[c_a]_a$ represents the click probabilities for tested and experimental content, then $R^{(t)}$ equals the expected number of clicks during period

t, which quantifies user engagement. We note that we model the reward parameters $[c_a]_a$ to be state-independent, which is consistent with existing RFM models (Fader et al. 2005, 2010); see for example Assumption 2 of Fader et al. (2010). This possibly is a simplification as more regular users might spend more time on the platform and we discuss a way to model such a behavior in Section 7. On the upside, as in Eq. (1) ("CLV") of Fader et al. (2005) and the "E(RLV)" equation of Fader et al. (2010), this will allow us to factor out the objective function of maximizing long-term value into two terms (one capturing rewards and the other capturing market size), and focus solely on the market size without loss of optimality (as the term pertaining to rewards would be a constant). This will become clear in Lemma 1 below.

State Transitions. At the end of period t, each existing user undergoes a state transition as dictated by the transition probabilities $[p_{sas'}]_{(s,a,s')}$ defined in Section 3.1. For analytical tractability, we model the transitions as being fluid so that the market state $\mathbf{\Lambda}^{(t+1)} = [\Lambda_s^{(t+1)}]_s$ in period $t+1$ is determined by the following flow balance equation:

$$\Lambda_s^{(t+1)} = \lambda_s + \sum_{s'=1}^{m} \left\{ \pi_{s'}^{(t)} (p_{s'2s} - p_{s'1s}) + \Lambda_{s'}^{(t)} p_{s'1s} \right\} \ \forall s \in \mathbb{S}. \tag{4}$$

The first term in the RHS of Eq. (4) (λ_s) is due to the exogenous arrivals at the start of period $t+1$, and the second term ($\sum_{s'}$) corresponds to the users transitioning into state s at the end of period t. To see this, observe that for each state $s' \in \{1, \ldots, m\}$, $\pi_{s'}^{(t)}$ users receive experimentation ($a=2$) and hence, $\pi_{s'}^{(t)} p_{s'2s}$ of them transition into state s. The remaining $\Lambda_{s'}^{(t)} - \pi_{s'}^{(t)}$ users do not receive experimentation ($a=1$) and hence, $(\Lambda_{s'}^{(t)} - \pi_{s'}^{(t)}) p_{s'1s}$ of them transition into state s. Adding the two gives us $\pi_{s'}^{(t)} (p_{s'2s} - p_{s'1s}) + \Lambda_{s'}^{(t)} p_{s'1s}$.

Our fluid approach to modeling transitions is equivalent to examining the expectations of a stochastic model and is appropriate in markets with a large number of users, such as NetEase, Spotify, and YouTube. We emphasize that this simplification is consistent with the existing work in the space of maximizing long-term user engagement in two-sided markets (see Azevedo and Leshno (2016) and Chen et al. (2021) for example). Furthermore, in our numerics, we obtained nearly identical results when we relaxed this assumption (discussed in Section B.1).

Platform's Objective. Using the experimentation policy as the key lever, the platform's objective is to maximize the long-term value in expectation. We focus on the *steady-state* expected reward of the multi-period system described above, which maps to the M (monetary) dimension of RFM. Given the fluid transitions, it follows that the steady state corresponds to the solution of the flow balance equation Eq. (4). That is, denoting by $\mathbf{\Lambda}$ and $\boldsymbol{\pi}$ the steady-state market and experimentation action, we have:

$$\Lambda_s = \lambda_s + \sum_{s'=1}^{m} \{\pi_{s'}(p_{s'2s} - p_{s'1s}) + \Lambda_{s'} p_{s'1s}\} \ \ \forall s \in \mathbb{S}. \tag{5}$$

Of course, the steady-state market $\mathbf{\Lambda}$ depends on the experimentation policy $\Pi(\cdot)$. Similar to Freund and Hssaine (2021), we restrict our focus to policies for which the steady-state exists (we show the steady-state existence and uniqueness for the policies of interest in Sections 4 and 5). Given $\mathbf{\Lambda}$ and $\boldsymbol{\pi}$, it follows from Eq. (3) that the steady-state expected reward equals $\sum_{s=1}^{m} \{\pi_s c_2 + (\Lambda_s - \pi_s) c_1\}$. Interestingly, maximizing the steady-state expected reward is equivalent to maximizing the steady-state market size $\mathbf{1}^\top \mathbf{\Lambda}$, as we establish next.

Lemma 1 **(Market Size)**. *Maximizing the steady-state expected reward is equivalent to maximizing the steady-state market size* $\mathbf{1}^\top \mathbf{\Lambda}$.

Proof. For a fixed policy $\boldsymbol{\pi}$, observe that the steady-state reward equals steady-state market size times a policy-independent non-negative constant:

$$\begin{aligned}
\sum_{s=1}^{m} \{\pi_s(c_2 - c_1) + \Lambda_s c_1\} &= (c_2 - c_1)\mathbf{1}^\top \boldsymbol{\pi} + c_1 \mathbf{1}^\top \mathbf{\Lambda} \\
&= (c_2 - c_1)\frac{\mathbf{1}^\top \mathbf{\Lambda}}{\eta} + c_1 \mathbf{1}^\top \mathbf{\Lambda} && \text{[follows Eq. (2b)]} \\
&= \underbrace{\left\{\frac{c_2}{\eta} + c_1\left(1 - \frac{1}{\eta}\right)\right\}}_{\geq 0 \text{ (constant)}} \mathbf{1}^\top \mathbf{\Lambda}. && \text{[recall } \eta > 1]
\end{aligned}$$

Thus, maximizing one is equivalent to maximizing the other. $\square$

As mentioned above, this factorization is analogous to that in Fader et al. (2005, 2010)[8] and is enabled by the reward parameters being state-independent. Given this result, moving forward, we will solely focus on the market size without loss of optimality.

The Platform Recommendation Problem. We can now fully articulate the platform recommendation problem. The platform's objective is to maximize the steady-state expected reward, which is equivalent to maximizing the steady-state market size $\mathbf{1}^\top \boldsymbol{\Lambda}$ (by Lemma 1). It does so by controlling the experimentation policy $\Pi(\cdot)$, which must obey the constraints Eq. (2) in each time period. This is a computationally challenging dynamic optimization problem involving an optimization over a rich space of policies subject to non-trivial constraints. Its solution would be the long-term optimal policy. Recognizing that the myopic policy of targeting users based on their churn delta (which we refer to as "churn minimization") discussed in Ascarza (2018) can be sub-optimal, Lemmens and Gupta (2020) advocate for such a long-term perspective. However, it is unclear a priori whether the long-term optimal policy admits any clean structure and whether it can be computed/implemented in our model. On the other hand, the churn minimization policy advocated by Ascarza (2018) is very interpretable and easy to implement. Our main analytical result (Theorem 2) will characterize conditions under which churn minimization is in fact long-term optimal (under a funnel setting). In this direction, next, we develop a tractable upper bound to aid our analysis of experimentation policies.

3.3. A Tractable Upper Bound via Linear Programming

To aid our analysis of experimentation policies, it will be useful to have an upper bound. To that end, consider the following linear program, which we denote by LP:

$$\max_{(\boldsymbol{\pi},\boldsymbol{\Lambda})\geq 0} \mathbf{1}^\top \boldsymbol{\Lambda} \tag{6a}$$

[8] For instance, the "DET" term in Eq. (1) of Fader et al. (2005) maps to our market size term $\mathbf{1}^\top \boldsymbol{\Lambda}$ and the "margin $\times$ revenue/transaction" term of Fader et al. (2005) maps to our $\left\{\frac{c_2}{\eta} + c_1\left(1-\frac{1}{\eta}\right)\right\}$ term. ($\frac{1}{\eta}$ fraction of the users receive experimentation (corresponding reward/transaction equals c_2) and the remaining $1-\frac{1}{\eta}$ do not (corresponding reward/transaction equals c_1).)

$$\text{s.t. } \Lambda_s = \lambda_s + \sum_{s'=1}^{m} \{\pi_{s'}(p_{s'2s} - p_{s'1s}) + \Lambda_{s'} p_{s'1s}\} \qquad \forall s \in \mathbb{S} \tag{6b}$$

$$\mathbf{1}^\top \boldsymbol{\pi} = \frac{\mathbf{1}^\top \boldsymbol{\Lambda}}{\eta} \tag{6c}$$

$$\boldsymbol{\pi} \le \boldsymbol{\Lambda}. \tag{6d}$$

By construction, any optimal solution to LP (denoted by $(\boldsymbol{\pi}^{\mathsf{LP}}, \boldsymbol{\Lambda}^{\mathsf{LP}})$) corresponds to a steady-state (cf. Eq. (6b)) and maximizes the steady-state market size (cf. Eq. (6a)), while obeying the required constraints in steady-state (cf. Eq. (6c) and Eq. (6d)). Therefore, its solution serves as an upper bound to any policy that admits a steady state. We summarize in Lemma 2.

LEMMA 2 **(LP Upper Bound)**. *Consider an arbitrary feasible policy Π for the platform recommendation problem among the set of policies for which steady state exists. Denote by Λ^{Π} the steady-state market under Π. Then, $\mathbf{1}^\top \boldsymbol{\Lambda}^{\Pi} \le \mathbf{1}^\top \boldsymbol{\Lambda}^{\mathsf{LP}}$.*

The linear program LP has $2m$ decision variables and $2m+1$ constraints, lending it computational tractability. In addition to being a tractable performance benchmark, it serves as a tool to analytically characterize the structure of an optimal policy, as we will see in Section 5. It is worth noting that the optimal solution $(\boldsymbol{\pi}^{\mathsf{LP}}, \boldsymbol{\Lambda}^{\mathsf{LP}})$ corresponds to only what happens in steady-state, and LP does not provide any insight on how to reach the steady-state market $\boldsymbol{\Lambda}^{\mathsf{LP}}$. In particular, LP does not output a policy $\Pi(\cdot)$ that maps an arbitrary market state $\boldsymbol{\Lambda}$ to a corresponding experimentation action $\boldsymbol{\pi}$. It only gives a steady-state feasible market size and experimentation action. As such, it provides no prescription on what experimentation to perform before the system reaches the steady state. In fact, given the period-wise constraints Eq. (2), it is unclear a priori if the corresponding steady-state $\boldsymbol{\Lambda}^{\mathsf{LP}}$ can even be achieved. Therefore, we treat LP as an upper bound.

Next, we study simple, easily implementable policies for the platform recommendation problem, giving special emphasis to binary and funnel-style models in order to extract analytical insights. We discuss more complicated state spaces in our numerical simulations in Section 6 and show that the insights extracted from the funnel-style models are robust to such perturbations. Our two managerial goals for the upcoming sections are (1) to understand whether it is okay to ignore user heterogeneity when experimenting and (2) if not, then what is a "good" experimentation policy?

4. The Perils of Blind Randomization (BR)

In this section, we analyze baseline policies that ignore user heterogeneity and experiment "blindly", which we term *blind randomization*. Recall, in Section 2.3 and Section A.3, we described evidence that NetEase experimentation policies in practice are possibly state-blind and thus, an example of blind randomization. We define the policy next.

Definition 1 (Blind Randomization (BR)). Given a market state $\mathbf{\Lambda}$ and experimentation parameter η, *blind randomization* $\Pi^{\mathsf{BR}}(\cdot)$ is defined as follows:

$$\Pi_s^{\mathsf{BR}}(\mathbf{\Lambda}) := \frac{\Lambda_s}{\eta} \ \forall s \in \mathbb{S}.$$

Note that BR corresponds to a platform experimenting on each user with probability $1/\eta$, regardless of the current state they are in. Furthermore, observe that BR is always a feasible policy: Eq. (2a) is satisfied as $\boldsymbol{\pi}^{(t),\mathsf{BR}} = \frac{\mathbf{\Lambda}^{(t)}}{\eta} \le \mathbf{\Lambda}^{(t)}$ for all t (recall $\eta > 1$), and Eq. (2b) holds since $\mathbf{1}^\top \boldsymbol{\pi}^{(t),\mathsf{BR}} = \frac{\mathbf{1}^\top \mathbf{\Lambda}^{(t)}}{\eta} = B^{(t)}$ for all t. We next prove that the steady-state exists uniquely under BR and characterize this steady-state for a general state space $\mathbb{S}$ (all omitted proofs are in Section C).

Lemma 3 **(BR Steady State)**. *For an arbitrary state space $\mathbb{S}$, under BR, the steady-state market $\mathbf{\Lambda}^{\mathsf{BR}}$ exists uniquely and is the solution to the following system of linear equations:*

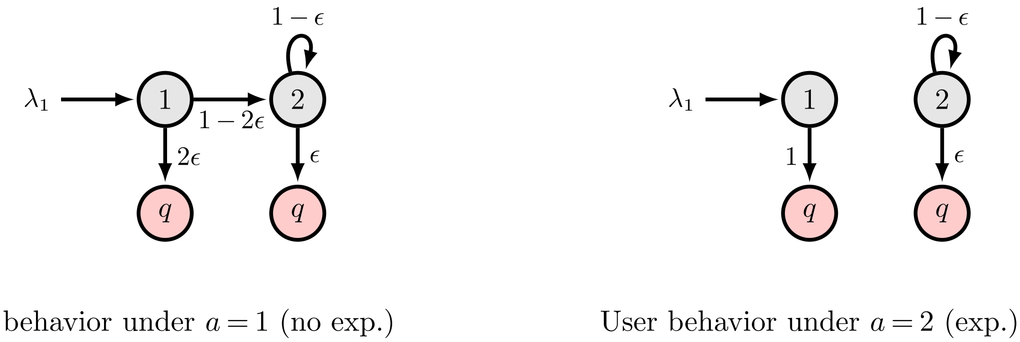
$$\Lambda_s \left(1 - p_{s1s} - \frac{p_{s2s} - p_{s1s}}{\eta}\right) - \sum_{s' \neq s} \Lambda_{s'} \left(p_{s'1s} + \frac{p_{s'2s} - p_{s'1s}}{\eta}\right) = \lambda_s, \ \forall s \in \mathbb{S}.$$

Having characterized the steady-state market under BR, we shift our attention to evaluating the performance of BR. Intuitively, being "blind" to the heterogeneity among the users appears sub-optimal, but it is unclear how necessary state information is for an experimentation policy. We formalize this next, establishing it can result in an arbitrarily poor performance.

Proposition 1 **(BR Arbitrarily Bad)**. *BR can guarantee no constant factor of the optimal steady-state reward.*

Proof Sketch. Consider a family of binary state space instances parameterized by ϵ (Fig. 4). We compare the steady-state market size under BR to that under the "churn minimization" policy (experiment first on users in state 2 (e.g., regular users) and on users in state 1 (e.g., new users) only when we run out of state 2 users to experiment on), and show that this ratio gets arbitrarily close to 0 as ϵ goes to 0. (A formal proof is in Section C.1.) □

Figure 4 Family of binary state space models used in the proof of Proposition 1.



Note. Above we describe a family of user models parameterized by ϵ where there are two states, and for which users arrive only at the first state (i.e., $\lambda_2 = 0$). State transition parameters when the user is shown an old (new) card are written on the arcs in the left (right) panel.

Proposition 1 highlights some fundamental limitations of BR. By not incorporating state information, BR ends up often experimenting on sensitive early state users, inhibiting them from transitioning to later states where they become more tolerant to experimentation. Consequently, the steady-state market size ends up being much smaller than it would be under a policy that considers user heterogeneity. Recall that in Section 2.2, we showed evidence of a first impression effect where new users were relatively more likely to churn as the result of one poor interaction. Viewed in this light, Proposition 1 formalizes that ignoring this first impression effect can be very damaging to a platform's long-term growth.

Moreover, in the proof of Proposition 1, we introduced a simple policy of churn minimization (CM), which prioritizes experimenting on state 2 (low sensitivity "regular") users before doing so on state 1 (high sensitivity "new") users (as advocated in Ascarza (2018)). Intuitively speaking, such prioritization provides short-term gains by preventing immediate churn among new users, but it is unclear if these short-term gains offset the loss that occurred due to decreased retention among the regular users, which have a higher long-term value since they are more likely to return to the platform (analogous to the motivation of Lemmens and Gupta (2020)). To understand this tension, we analyze CM next. Somewhat surprisingly, we establish that CM is in fact the long-term optimal policy under a simple condition on the funnel's primitives.

5. Churn Minimization (CM)

In the previous section, we demonstrated that state-blind experimentation can be arbitrarily sub-optimal. Motivated by Ascarza (2018), we now introduce and analyze a myopic policy, churn minimization (CM), that utilizes state information to perform experimentation in a way that attempts to retain as much of the user base as possible from one period to the next. To do so, CM prioritizes experimenting on users who are minimally sensitive to being shown experimental content. Recall, in Section 3, we introduced the churn delta, $d_s := p_{s2q} - p_{s1q}$, which exactly captures the sensitivity of a user in state s by their difference in churn rate for old vs. experimental content. In the context of NetEase, if we separate the user base into new and regular users, CM corresponds to experimenting first on less sensitive regular users before experimenting on more sensitive new users. We define the policy next.

Definition 2 (Churn minimization (CM)). Given a market state $\boldsymbol{\Lambda}$ and experimentation parameter η, *churn minimization* $\Pi^{\mathsf{CM}}(\cdot)$ sorts the states based on $[d_s]_{s\in\mathbb{S}}$ (churn deltas) and experiments in increasing order. Thus, users with the lowest churn delta are experimented on first, then the users with the second-lowest delta, and so on, until the experimentation budget is used up.

CM is not only intuitive but also a natural myopic policy to consider. To formalize this notion, we will show churn minimization is the experimentation policy that minimizes the one-step churn. In particular, given market state $\boldsymbol{\Lambda}^{(t)}$ during period t, observe that due to flow-balance Eq. (4), the total market size during the next period equals:

$$\mathbf{1}^\top \boldsymbol{\Lambda}^{(t+1)} = \mathbf{1}^\top \boldsymbol{\Lambda}^{(t)} + \underbrace{\mathbf{1}^\top \boldsymbol{\lambda}}_{\text{in-flow}} + \underbrace{(\text{number of churns after period } t)}_{\text{out-flow}}. \tag{7}$$

Given experimentation action $\boldsymbol{\pi}^{(t)}$ during period t, the number of churns (transitions to quit state q) from state $s \in \mathbb{S}$ after period t equals:

$$p_{s2q}\pi_s^{(t)} + p_{s1q}\left(\Lambda_s^{(t)} - \pi_s^{(t)}\right) = \underbrace{(p_{s2q} - p_{s1q})}_{=d_s}\pi_s^{(t)} + \underbrace{p_{s1q}\Lambda_s^{(t)}}_{\text{constant w.r.t. } \boldsymbol{\pi}^{(t)}}.$$

Therefore, given the budget constraint $\mathbf{1}^\top \boldsymbol{\pi}^{(t)} = \frac{\mathbf{1}^\top \boldsymbol{\Lambda}^{(t)}}{\eta}$, allocating the experimentation budget in increasing order of the churn delta $[d_s]_{s\in\mathbb{S}}$ minimizes the one-step churn, or equivalently, maximizes the market size $\mathbf{1}^\top \boldsymbol{\Lambda}^{(t+1)}$ in the next period.

Though CM is myopically optimal, it is unclear if CM maximizes long-term steady-state reward (as discussed above). In the remainder of this section, we analyze the steady-state performance of CM, placing emphasis on binary (Section 5.1) and funnel (Section 5.2) state spaces.

5.1. CM for the Binary State Space

In the proof of Proposition 1, we introduced CM as an alternative policy and showed it outperformed blind randomization. We will now show that CM is indeed optimal for the binary state space. To do so, we first establish that CM leads to a unique steady-state market.

Lemma 4 **(CM Steady State for Binary Space)**. *For the binary state space, under CM, the steady-state market* $\boldsymbol{\Lambda}^{\mathsf{CM}}$ *exists and is unique.*

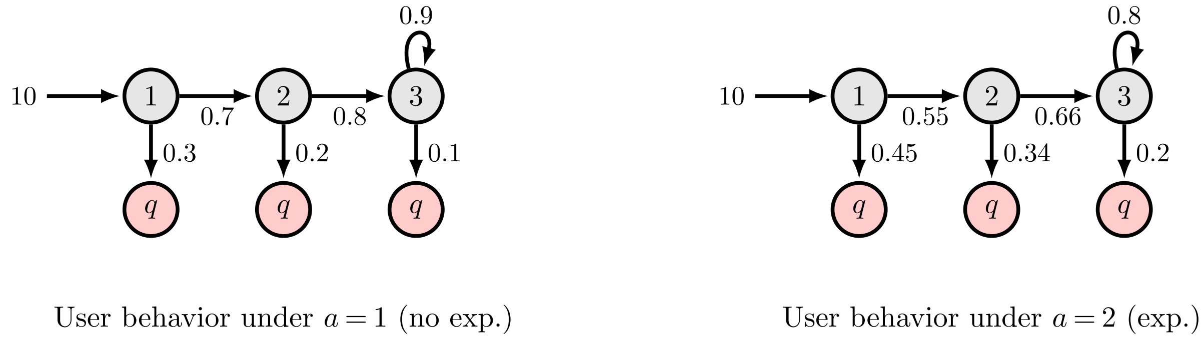
Having established the existence and uniqueness of the steady state, we can now evaluate the performance of CM for the binary state space. In Theorem 1, we show that the myopically optimal policy of CM is, in fact, long-term optimal for the binary state space.

Theorem 1 **(CM Optimal for Binary Space)**. *For the binary state space, CM is optimal.*

Our proof leverages the LP upper bound from Lemma 2 to evaluate the quality of the policy. We invoke strong duality for linear programming to solve for the optimal policy (Ch. 4 of Bertsimas and Tsitsiklis (1997)) and demonstrate that it coincides with CM in all instances. We also note our proof is constructive, and supplies the optimal experimentation policy in every instance.

As CM is interpretable and optimal for the binary state space, one might wonder whether the optimality of CM holds more generally. However, in some sense, Theorem 1 is sharp; the optimality of CM is no longer guaranteed even for simple extensions involving a third state, which we demonstrate in Example 1 next. (CM not being optimal in general is in fact in agreement with Lemmens and Gupta (2020).)

Figure 5 Three state funnel for Example 1.



Note. Depicted is a funnel state space model with $m = 3$. Users only arrive in the first state, and there is a self-loop in the third state. The transition probabilities are written on the arcs for the two possible actions, $a = 1$ on the left and $a = 2$ on the right. Note that for this instance we have $(d_1, d_2, d_3) = (0.15, 0.14, 0.10)$ and hence, $d_1 > d_2 > d_3$.

Example 1 (CM is not necessarily optimal beyond the binary state space). Consider the funnel state space model shown in Fig. 5. There are $m = 3$ states. Exogenous arrivals occur only to state 1 at rate $\lambda_1 = 10$. In addition to obeying "leakage", the instance obeys $d_1 > d_2 > d_3$. Thus, users become less sensitive to experimentation as they spend more time on the platform. However, solving[9] the LP from Lemma 2 with $\eta = 1.2$, we get a steady-state market $\mathbf{\Lambda}^{\mathsf{LP}} \approx (10, 5.61, 22.44)$ (and hence, total market size $\mathbf{1}^\top \mathbf{\Lambda}^{\mathsf{LP}} \approx 38.05$) with optimal experimentation policy $\boldsymbol{\pi}^{\mathsf{LP}} \approx (9.27, 0, 22.44)$ (all numbers reported in this example are rounded to 2 decimal places and hence, we use the "$\approx$" symbol). Note, this policy prioritizes experimentation in the following order: state 3 over *state 1 over state 2*, and thus is not CM. Moreover, when we simulate this policy, we do obtain the steady-state corresponding to $(\mathbf{\Lambda}^{\mathsf{LP}}, \boldsymbol{\pi}^{\mathsf{LP}})$, implying the LP steady-state is achievable by a feasible policy. For comparison, since $d_1 > d_2 > d_3$, CM prioritizes state 3 over state 2 over state 1, resulting in steady-state market $\mathbf{\Lambda}^{\mathsf{CM}} \approx (10, 6.44, 21.26)$ (and hence, market size $\mathbf{1}^\top \mathbf{\Lambda}^{\mathsf{CM}} \approx 37.70$) with experimentation policy $\boldsymbol{\pi}^{\mathsf{CM}} \approx (3.72, 6.44, 21.26)$. Though not by much (37.70 vs. 38.05), CM is strictly sub-optimal. $\square$

We note that Example 1 can be seen as a minimal extension beyond the setting of Theorem 1, with only one additional state, and while maintaining the sequential flow of users. To understand

[9] Details on our `MATLAB` implementation (for both optimization and simulation) are provided in Section B.1.

why CM is sub-optimal here, observe that though we have monotonicity between the corresponding churn deltas (i.e., $d_1 > d_2$), the ratio $d_2/d_1 = 0.14/0.15$ is "close" to 1. Since CM does not account for state structure and simply targets less sensitive users, it thus prioritizes experimenting on state 2 users even though state 1 users are nearly equivalent. In doing so, it ignores that state 1 users have lower future value since they are farther from becoming "regular", leading to sub-optimality.

Intuitively, we expect CM to perform best when there are large differences in churn deltas across states. As such, it seems desirable to explore whether there exists a more general setting than the binary state space for which CM remains optimal. In the next subsection, we will explore this idea and study the performance of CM for the funnel state space under some conditions on the churn deltas (i.e., that the ratio of churn deltas d_{s+1}/d_s is bounded away from 1 for all $s < m$).

5.2. CM for the Funnel State Space

The funnel state space (recall Fig. 3) generalizes the binary state space. In this section, we will explore CM for the funnel state space under the additional assumption that $d_s > d_{s+1}$ for $s \in \{1, \ldots, m-1\}$. This assumption implies that as users spend more time on the platform, their sensitivity to experimental content decreases, matching our observations in Section 2.

Under this assumption, it is straightforward to define CM for the funnel state space. The platform first prioritizes experimenting on the state m users, and the experimentation trickles down to states $m-1$, $m-2$, …, 1 until all the experimentation budget is used up. The existence and uniqueness of the corresponding steady-state market can be established using a similar analysis as in Lemma 4 and we skip it for brevity. We focus on understanding the performance of CM. We do so in Theorem 2 next, where we characterize the optimality of CM as a function of the churn deltas (recall the discussion after Example 1) and the return probabilities. (Recall $r_s := 1 - p_{s1q}$ denotes the return probability from state $s \in \mathbb{S}$ under no experimentation.)

Theorem 2 **(CM Optimality for Funnel)**. *For the funnel state space, CM is optimal if* $\frac{d_{s+1}}{d_s} < r_{s+1}$ *for* $s \in \{1, \ldots, m-2\}$.

Theorem 2 shows CM is an optimal policy for many instances of the funnel state space under one additional condition that strengthens $\frac{d_{s+1}}{d_s} < 1$ (as the return probability r_{s+1} is less than 1) to enforce some level of separation between the churn deltas. In light of this theorem, we can revisit Example 1 and verify that it violates the required condition: $d_2/d_1 = 0.14/0.15 \approx 0.93 > 0.8 = r_2$. As such, not only is the Theorem 2 condition sufficient for the optimality of CM, it is also necessary in the sense that under its absence, there exists a funnel setting where d_s is decreasing in s, but still, CM is strictly sub-optimal.

Remark 1 (Customer Lifetime Value). Intuitively, as alluded to above in our Ascarza (2018) and Lemmens and Gupta (2020) discussion, CM's optimality depends on the tension between the short-term gains from targeting less sensitive users, possibly by ignoring some long-term value that depends on the structure of the state space. In the condition for Theorem 2, the short-term gains are captured by the ratio of churn deltas appearing in the LHS. Furthermore, though not directly visible, the long-term value is accounted for by the return probability in the RHS. To see this, let f_s denote the expected number of visits before churning under no experimentation for a user in state s. We can think of f_s as the *future value* (or the *customer lifetime value*) of a user from state $s \in \{1, \ldots, m-1\}$, and note that it obeys the following recursion[10]:

$$f_s = 1 + r_{s+1} f_{s+1}. \tag{8}$$

That is, the future value from state s is today's visit ("1") plus the expected number of visits from $s+1$ assuming they return ("$r_{s+1} f_{s+1}$"). It follows from Eq. (8) that $r_{s+1} = \frac{f_s - 1}{f_{s+1}} \approx \frac{f_s}{f_{s+1}}$. Thus, the condition in Theorem 1 roughly maps to

$$f_s d_s > f_{s+1} d_{s+1} \tag{9}$$

for $s \in \{1, \ldots, m-2\}$. The tension between the short-term delta and the long-term value clearly pops out in Eq. (9). By prioritizing experimenting in state $s+1$ (over s), we benefit from the

[10] This is similar but not identical to the recursion in Eq. (2) ("survivor function") of Fader and Hardie (2007). In particular, we focus on the expected number of visits whereas Fader and Hardie (2007) consider the probability of survival. As such, we have an additional term ("1") that gets added in the RHS.

short-term delta gain (as we avoid d_s and $d_s > d_{s+1}$). However, we take a hit via the state $s+1$ long-term value f_{s+1} as opposed to the state s value f_s. As such, it makes sense that f_{s+1} should not be much bigger than f_s for CM to be optimal. In particular, if $f_{s+1} > \frac{d_s}{d_{s+1}} f_s$, then we end up violating Eq. (9). This finding provides a meaningful perspective to the views in Ascarza (2018) and Lemmens and Gupta (2020). As Ascarza (2018) points out, platforms must account for how sensitive customers are to the retention efforts (d_s); however, they can not ignore the long-term value of retaining the users (f_s) while doing so, as advocated by Lemmens and Gupta (2020).

It is of interest to connect Theorem 2 to more tangible models of user behavior, such as the ones discussed in Section 2. In this direction, we next discuss a parameterization of the funnel state space that is motivated by our data analysis (recall Eq. (1)) and show how Theorem 2 allows us to extract clean managerial insights.

The Exponential Decay Parameterization. We now analyze the funnel under a specific parameterization of the transition probabilities. In particular, we study a parameterization that matches our observed churn probabilities in Fig. 2, and where both the churn probability (under no experimentation) and the churn delta decay exponentially in the state number (frequency):

$$1 - r_s = \gamma_0 e^{-\gamma_1 \times (s-1)} \ \forall s \in \mathbb{S} \tag{10a}$$

$$d_s = \theta_0 e^{-\theta_1 \times (s-1)} \ \forall s \in \mathbb{S}. \tag{10b}$$

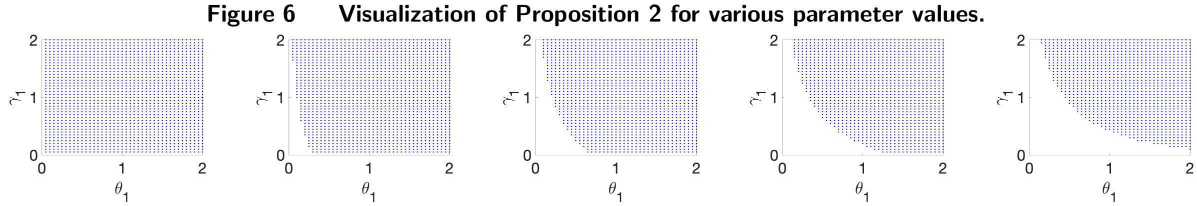
Note that $1 - r_s$ denotes the churn probability p_{s1q} from state s (under no experimentation), parameter $\gamma_0 \in (0,1)$ equals the churn probability from state 1, and $\gamma_1 > 0$ denotes the rate of decay (as the user becomes more regular). Similarly, $\theta_0 \in (0,1)$ denotes the churn delta from state 1, and $\theta_1 > 0$ the corresponding rate of decay. Given $\theta_1 > 0$, this model obeys $d_s > d_{s+1}$ for $s < m$.

Such a parameterization has a practical appeal as it captures user behavior via just four parameters $(\gamma_0, \gamma_1, \theta_0, \theta_1)$, and can do so quite well (cf. Section 2.2). Thus, it is of interest to understand if the optimality of CM can be expressed in terms of these primitives. We do so next.

Proposition 2 **(CM Optimality for Exponential Decay)**. *For the funnel state space with the exponential decay parameterization, CM is optimal if* $1 - e^{-\theta_1} > \gamma_0 e^{-\gamma_1}$.

This result equips a practitioner with an easy-to-use rule for deciding when to use CM. We visualize Proposition 2 in Fig. 6, where we highlight the parameter regions that obey the condition. A higher θ_1 drives the ratio of churn deltas further away from 1, and hence, a higher θ_1 favors CM (consistent with our discussion around short-term/long-term trade-off). Similar intuition holds for γ_1 since a higher γ_1 results in the ratio of long-term values being closer to 1. Mapping this to the data analysis in Section 2 ($\gamma_0 \approx 0.5$, $\gamma_1 \approx 0.1$, $\theta_0 \approx 0.05$, and $\theta_1 \approx 1$), CM appears to be optimal at NetEase. Furthermore, it adds a non-trivial value over BR. For instance, with $m = 5$ and $\eta = 5$, the market size under BR over that under CM equals 0.9870, implying a value-add of approx. 1.3%.

Figure 6 Visualization of Proposition 2 for various parameter values.



(a) $\gamma_0 = 0.05$ (b) $\gamma_0 = 0.25$ (c) $\gamma_0 = 0.5$ (d) $\gamma_0 = 0.75$ (e) $\gamma_0 = 0.95$

Note. For each $\gamma_0 \in \{0.05, 0.25, 0.5, 0.75, 0.95\}$, we plot the condition as θ_1 and γ_1 vary over $\{0.05, 0.1, \ldots 2\}$. A blue dot means $1 - e^{-\theta_1} > \gamma_0 e^{-\gamma_1}$ and hence, implies CM is optimal by Proposition 2.

Finally, it is of interest to understand the performance of CM outside the regimes studied so far. Though it is possible to extend Theorem 2 to more general settings (e.g., incorporating self-loops or momentum), the resulting optimality conditions are rather opaque. As such, in Section 6, we shift our focus to numerical simulations, where we leverage the LP upper bound as a benchmark. Remarkably, we find CM to be optimal in a wide spectrum of settings.

6. Numerical Study

In this section, we numerically study (a) how much value CM adds over BR and (b) the quality of CM relative to the LP upper bound, over a broad class of parameter values within and beyond the funnel. We find that (a) CM significantly outperforms BR in all regimes and (b) CM's optimality is robust to various changes in the underlying environment. From a managerial point-of-view, our findings here provide a positive support for the proposal in Ascarza (2018), at least within the class of state spaces we study: funnel (Section 6.1), self-loops (Section 6.2), and momentum (Section 6.3). Note that our numerics are implemented in `MATLAB` (details in Section B.1).

6.1. Funnel Numerics

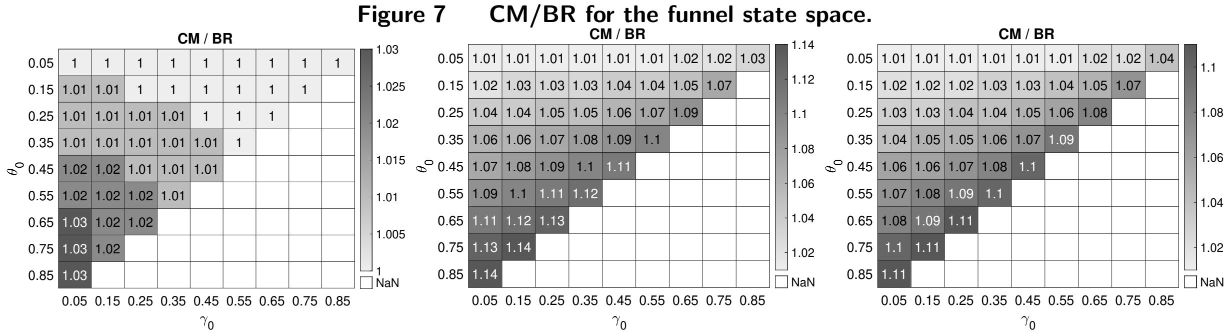
For the funnel state space (Fig. 3), we parameterize the transition probabilities via the exponential decay model (10). We set $\lambda_1 = 100$ wlog (since it only has a scaling effect) and experiment with the remaining parameters η, m, and $(\gamma_0, \theta_0, \gamma_1, \theta_1)$ as follows. We set $\eta = 10$ in our baseline numerics, implying the platform experiments on $1/10 = 10\%$ of the users in every period. We perform sensitivity analysis by varying $\eta \in \{2, 10, 50\}$ (reported in Section B.2). The number of states m is set to 5, and we find results to be similar for other values of m. We vary $\gamma_0 \in \{0.05, 0.15, \ldots, 0.85\}$ and $\theta_0 \in \{0.05, 0.15, \ldots, 0.85\}$ such that $\gamma_0 + \theta_0 < 1$.[11] Finally, for conciseness, we set the two decay rates γ_1 and θ_1 to be equal to each other and vary them in $\{0.05, 1, 2\}$ (slow, medium, and fast decay), which correspond to the bottom-left, middle, and top-right points in Fig. 6 (and hence, cover regimes where the condition of Proposition 2 does not hold).

In Fig. 7, we visualize the value added by CM over BR. The ratio is never below 1, implying CM is always at least as good as BR. In fact, depending on the parameter regime, CM can add a non-trivial value over BR. When the decay rates (γ_1, θ_1) are small (Fig. 7a), the increase in objective value is modest (3% at most). This makes intuitive sense since a small decay rate means little heterogeneity in the state-specific churn deltas and future values. For medium and large decay rates (Figs. 7b and 7c), the gains can be over 10%. Naturally, one would expect the gains to be bigger when the churn deltas are bigger. This is what we see in Fig. 7 as the value-add increases with θ_0 (recall θ_0 is the churn delta at state 1). A similar logic explains the increase of ratios with γ_0 (recall γ_0 is the churn probability at state 1).

We also evaluate how far CM is from optimality by benchmarking it against the LP upper bound in Fig. 8. Remarkably, for all instances tested, CM equals the upper bound, implying optimality.

[11] Some parameter combinations are infeasible as they result in transition probabilities outside $(0, 1)$. In the funnel, it suffices to ensure $r_s - d_s > 0$ for all s, which is implied by $r_1 - d_1 > 0$ (since r_s is increasing and d_s is decreasing in s). Observing $r_1 - d_1 = 1 - \gamma_0 - \theta_0$ (cf. (10)), we only enumerate (γ_0, θ_0) pairs such that $\gamma_0 + \theta_0 < 1$.

Figure 7 **CM/BR for the funnel state space.**

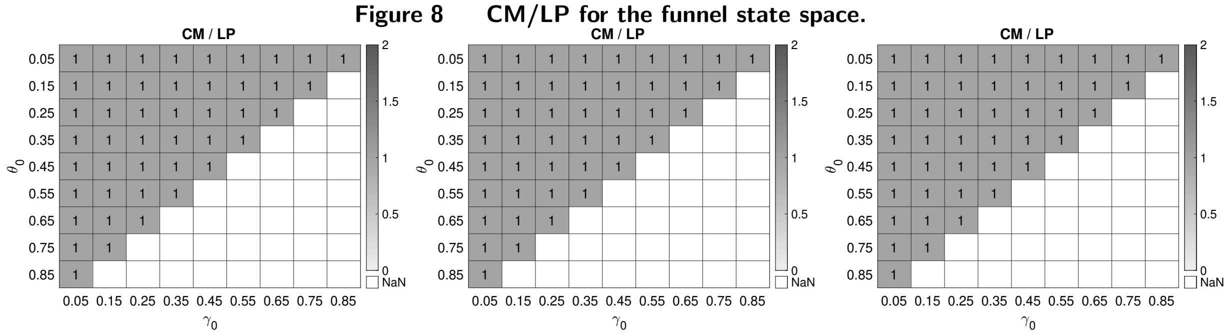


(a) $\gamma_1 = \theta_1 = 0.05$ (b) $\gamma_1 = \theta_1 = 1$ (c) $\gamma_1 = \theta_1 = 2$

Note. Market size under CM over that under BR for the funnel setup ($m = 5$, $\eta = 10$) discussed in Section 6.1. Recall that only (θ_0, γ_0) values obeying $\gamma_0 + \theta_0 < 1$ are valid.

We note that though the user model is a funnel in this subsection, the parameter settings do not necessarily obey the Proposition 2 condition, highlighting CM's robustness. Of course, it is of interest to test CM beyond the funnel state space, which we do next.

Figure 8 **CM/LP for the funnel state space.**



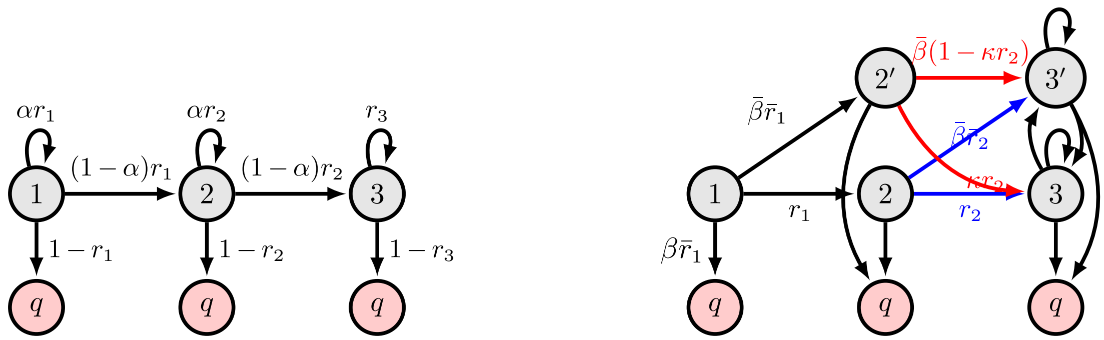
(a) $\gamma_1 = \theta_1 = 0.05$ (b) $\gamma_1 = \theta_1 = 1$ (c) $\gamma_1 = \theta_1 = 2$

Note. Market size under CM over that under LP for the funnel setup ($m = 5$, $\eta = 10$) discussed in Section 6.1.

6.2. Self-loops Numerics

Recall in the definition of the funnel state space we assumed there was only a self-loop in the terminal state. Here we augment the funnel model by introducing a self-loop at each state with a corresponding parameter $\alpha \in [0, 1]$ (see left panel of Fig. 9). These self-loops model the behavior of users who do not necessarily become more regular in each visit but incur some inertia, which is quantified by α. For instance, with $m = 2$ (binary state space), a new user might not become

Figure 9 Model of user behavior with self-loops (left panel) and momentum (right panel).

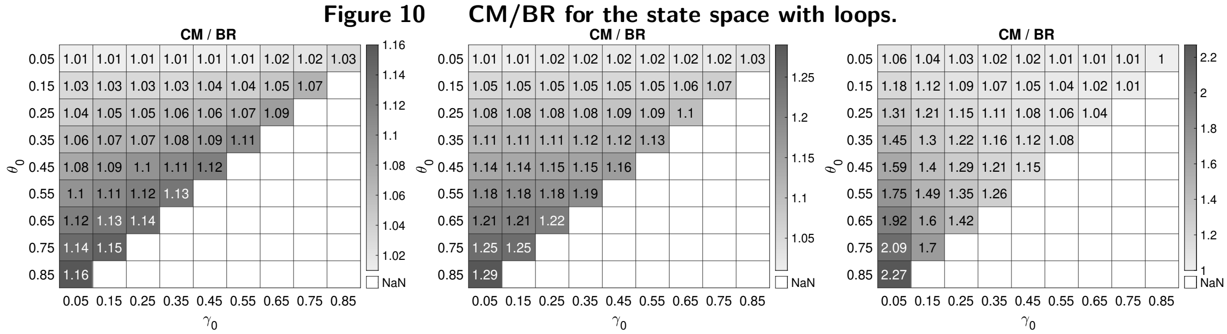


Self-loops (under $a = 1$) Momentum (under $a = 1$)

Note. We show the transition probabilities under no experimentation ($a = 1$) and note that similar to the funnel state, under experimentation ($a = 2$), we simply replace r_s by $r_s - d_s$ everywhere for all s. Furthermore, for simplicity, the visuals here are for $m = 3$ states and it is straightforward to generalize. We use the notation $\bar{\beta} := 1 - \beta$ and omit some transition probabilities in the right panel (due to space considerations), which are discussed in Section 6.3.

regular after one visit but do so eventually. As in Section 6.1, we parameterize $[r_s]_s$ and $[d_s]_s$ via the exponential decay model (10), setting $\lambda_1 = 100$, $\eta = 10$, $m = 5$, vary $\gamma_0 \in \{0.05, 0.15, \ldots, 0.85\}$ and $\theta_0 \in \{0.05, 0.15, \ldots, 0.85\}$ such that $\gamma_0 + \theta_0 < 1$, and set $\gamma_1 = \theta_1$. Though we generated results for $\gamma_1 \in \{0.05, 1, 2\}$, we only show the ones for $\gamma_1 = 1$ for conciseness and note that the other two sets of results are qualitatively similar. Finally, we vary $\alpha \in \{0.1, 0.5, 0.9\}$ (low, medium, high).

Figure 10 CM/BR for the state space with loops.



(a) $\alpha = 0.1$ (b) $\alpha = 0.5$ (c) $\alpha = 0.9$

Note. Market size under CM over that under BR for the loops setup discussed in Section 6.2.

In Fig. 10, we visualize the value added by CM over BR. CM continues to dominate BR in all instances, and the performance gap increases for larger decay rates. In fact, CM becomes even more valuable in the presence of self-loops, which is evident when we (a) compare the numbers in Fig. 10 to those in Fig. 7b and (b) observe the value-add becomes bigger as α increases from 0.1 to 0.5 to 0.9.[12] As we did in Fig. 8, we can evaluate the (sub-)optimality of CM. We do not show the corresponding figure and note that it is identical to Fig. 8, meaning CM remains optimal.

6.3. Momentum Numerics

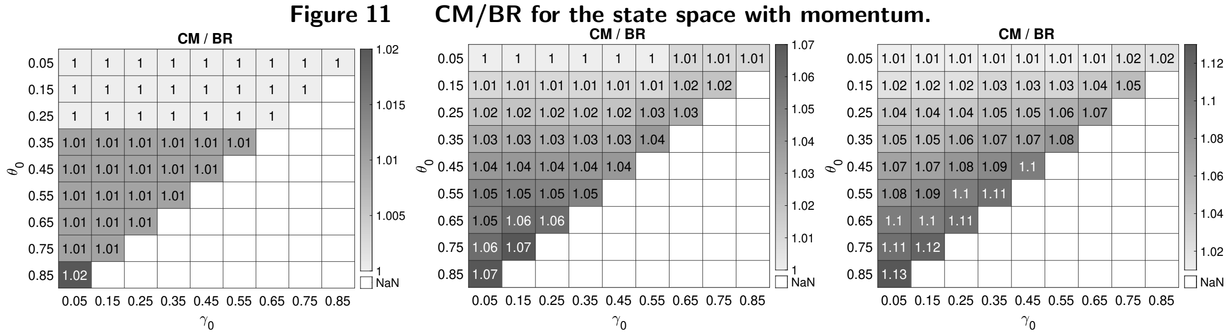
As alluded to in Section 3, we now incorporate momentum in the user model where we allow the user behavior to vary as a function of their second most experience with the platform (in addition to the most recent experience that is already captured via the churn delta). In particular, for each state $s \in \{2, \ldots, m\}$, we introduce a second dimension that tracks whether the user has a positive or a negative experience in the previous visit. We show the state space in Fig. 9 (right panel) where states 2 and 3 denote states with positive experience, and $2'$ and $3'$ denote negative experience. Under no experimentation[13], the return probability from state s to $s+1$ remains the same as before (r_s), and we decompose $1 - r_s$ via the parameter $\beta \in [0, 1]$, which determines the split between churn ($\beta(1 - r_s)$) and bad experience ($(1 - \beta)(1 - r_s)$). For a bad-experience state s', the return probability (to the good-experience state $s + 1$) is scaled by κ, i.e., κr_s, where $\kappa \in [0, 1]$ captures the dampening due to bad experience; and similar to above, we decompose $1 - \kappa r_s$ via β, which determines the split between churn ($\beta(1 - \kappa r_s)$) and bad experience ($(1 - \beta)(1 - \kappa r_s)$). Such a parameterization ensures a user with a positive experience is (a) more likely to return and (b) and less likely to churn (than a user with a negative experience). We vary $\beta \in \{0.1, 0.5, 0.9\}$ (low, medium, high) and set $\kappa = 0.9$, noting that results for $\kappa \in \{0.1, 0.5\}$ are similar. All other parameters are identical to those in Section 6.2.

As seen in Fig. 11, CM still adds a non-trivial value over BR, with the value-add increasing in β. We also evaluate the (sub-)optimality of CM (as in Fig. 8). Instead of showing the plot, we note that it is identical to Fig. 8, highlighting the robustness of CM to environments beyond the funnel.

[12] Note that $\alpha = 0$ is equivalent to not having loops.

[13] The transition probabilities under experimentation are identical to the ones under no experimentation except that we replace r_s by $r_s - d_s$ everywhere.

Figure 11 CM/BR for the state space with momentum.



(a) $\beta = 0.1$ (b) $\beta = 0.5$ (c) $\beta = 0.9$

Note. Market size under CM over that under BR for the momentum setup discussed in Section 6.3.

7. Concluding Remarks

We propose a modeling framework that integrates a discrete-time RFM-style *behavioral* model of user behavior (capturing heterogeneous churning) with a *prescriptive* model to determine the optimal experimentation policy (under heterogeneous supply-side quality), with the goal of maximizing long-term value. Our main analytical result establishes that the simple policy of churn minimization (experimenting on users who are least sensitive to experimental content) is optimal under certain interpretable conditions on the primitives, bridging the perspectives of Ascarza (2018) and Lemmens and Gupta (2020). Our numerical simulations provide evidence that this simple policy is optimal even for settings beyond the ones covered in our theory.

Though we grounded ourselves in the content recommendation context, our framework is quite general and applicable to other two-sided markets such as labor (Upwork, TaskRabbit), transportation (Uber, Lyft), food and grocery delivery (DoorDash, Instacart), lodging (Airbnb, Vrbo), and relationships (Match, Hinge). In particular, our RFM-style stochastic model of user behavior can be apt for markets where users interact with the platform in discrete time periods and exhibit heterogeneous churning behavior, and our supply-side experimentation model is apt for platforms that have supply of heterogeneous quality and the platform needs to learn the underlying quality. For instance, in labor markets such as Upwork, the supply corresponds to workers that complete the job and the demand side corresponds to the platform users who submit job requests. In transportation markets, the drivers form the supply side whereas the riders form the demand side. Naturally, users in such markets are sensitive to service quality and this sensitivity can be heterogeneous depending on how long they have been using the platform. Further, the quality of service can vary across providers, which the platform needs to learn as new providers join.

Our work opens the door for multiple possible avenues of future research. On the optimization side, it is of interest to understand if we can analytically establish the optimality of CM outside the conditions we develop, and to also solve for more complicated state spaces. In terms of the user behavior model, while the model presented in Section 3 is general, it does abstract away some peculiarities of the real world. It is of interest to study the extensions that capture such peculiarities. For instance, we model the reward parameters (e.g., click probability of a card) as being independent of the user's state. This allowed us to decouple the objective into two terms (recall Lemma 1) and facilitated a closed-form analysis of the resulting optimization problem. However, it is possible that users who have visited more spend more time during a visit, e.g., see more content in a recommendation platform. Such behavior can be captured via a cascade model (Craswell et al. 2008) where a user sees a sequence of cards during a visit and the number of cards she sees depends on her state. However, this will require a careful analysis of the corresponding optimization problem. In addition, as alluded to in Section 3, it is of interest to study the continuous-time variant of our model. Finally, in this work, we focused on the behavior of the users as a function of the platform's experimentation policy and assumed the content generation process to be exogenous. Understanding the effects of the platform's actions on content creators' motivation to publish new content and capturing such feedback loop in the model itself would be useful. We hope to pursue some of these directions in future research.

References

Agnew, Julie R, Hazel Bateman, Christine Eckert, Fedor Iskhakov, Jordan Louviere, Susan Thorp. 2018a. First impressions matter: An experimental investigation of online financial advice. *Management Science* **64**(1) 288–307.

Agnew, Julie R, Hazel Bateman, Christine Eckert, Fedor Iskhakov, Jordan J Louviere, Susan Thorp. 2018b. Learning and confirmation bias: Measuring the impact of first impressions and ambiguous signals. *UNSW Business School Research Paper, Wharton Pension Research Council Working Paper* (2018-08).

Alexander, Julia. 2020. Recommendation is one of the biggest issues facing streamers like Netflix, HBO Max, and more URL `https://www.theverge.com/2020/1/9/21058599/netflix-streaming-farewell-recommendations-lulu-wang-hbo-max-quibi`.

Ascarza, Eva. 2018. Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research* **55**(1) 80–98.

Azevedo, Eduardo M, Jacob D Leshno. 2016. A supply and demand framework for two-sided matching markets. *Journal of Political Economy* **124**(5) 1235–1268.

Bastani, Hamsa, Pavithra Harsha, Georgia Perakis, Divya Singhvi. 2022. Learning personalized product recommendations with customer disengagement. *Manufacturing & Service Operations Management* **24**(4) 2010–2028.

Bennett, James, Stan Lanning, et al. 2007. The Netflix Prize. *Proceedings of KDD cup and workshop*, vol. 2007. Citeseer, 35.

Bertsimas, Dimitris, John N Tsitsiklis. 1997. *Introduction to linear optimization*, vol. 6. Athena scientific Belmont, MA.

Chen, Mingliu, Adam N Elmachtoub, Xiao Lei. 2021. Matchmaking Strategies for Maximizing Player Engagement in Video Games. *Available at SSRN 3928966* .

Craswell, Nick, Onno Zoeter, Michael Taylor, Bill Ramsey. 2008. An experimental comparison of click position-bias models. *Proceedings of the 2008 international conference on web search and data mining*.

Dredge, Stuart. 2023. NetEase Cloud Music now has 206.7m active users in China URL `https://musically.com/2023/08/25/netease-cloud-music-now-has-206-7m-active-users-in-china/`.

Fader, Peter S, Bruce GS Hardie. 2007. How to project customer retention. *Journal of Interactive Marketing* **21**(1) 76–90.

Fader, Peter S, Bruce GS Hardie. 2010. Customer-base valuation in a contractual setting: The perils of ignoring heterogeneity. *Marketing Science* **29**(1) 85–93.

Fader, Peter S, Bruce GS Hardie, Ka Lok Lee. 2005. RFM and CLV: Using iso-value curves for customer base analysis. *Journal of marketing research* **42**(4) 415–430.

Fader, Peter S, Bruce GS Hardie, Jen Shang. 2010. Customer-base analysis in a discrete-time noncontractual setting. *Marketing Science* **29**(6) 1086–1108.

Freund, Daniel, Chamsi Hssaine. 2021. Fair Incentives for Repeated Engagement. *arXiv preprint arXiv:2111.00002* .

Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, Donald B Rubin. 2013. *Bayesian data analysis*. CRC press.

Glenski, Maria, Tim Weninger. 2017. Predicting user-interactions on reddit. *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*.

Gupta, Sunil, Donald Lehmann. 2005. *Managing customers as investments the strategic value of customers in the long run*. Wharton School Publishing.

Gurobi Optimization, LLC. 2023. Gurobi Optimizer Reference Manual. URL `https://www.gurobi.com`.

IFPI. 2022. IFPI Global Music Report 2022 URL `https://www.ifpi.org/Global_Music_Report_2023_State_of_the_Industry`.

Kemeny, John G, James Laurie Snell. 1976. *Finite Markov chains*. Springer.

Kveton, Branislav, Csaba Szepesvari, Zheng Wen, Azin Ashkan. 2015. Cascading bandits: Learning to rank in the cascade model. *International Conference on Machine Learning*. PMLR, 767–776.

Lemmens, Aurélie, Sunil Gupta. 2020. Managing churn to maximize profits. *Marketing Science* **39**(5) 956–973.

Lindgaard, Gitte, Gary Fernandes, Cathy Dudek, Judith Brown. 2006. Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & information technology* **25**(2) 115–126.

Liu, Guimei, Tam T Nguyen, Gang Zhao, Wei Zha, Jianbo Yang, Jianneng Cao, Min Wu, Peilin Zhao, Wei Chen. 2016. Repeat buyer prediction for e-commerce. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 155–164.

MATLAB. 2021. *Version 9.10.0 (R2021b)*. The MathWorks Inc., Natick, Massachusetts.

McInerney, James, Benjamin Lacker, Samantha Hansen, Karl Higley, Hugues Bouchard, Alois Gruson, Rishabh Mehrotra. 2018. Explore, exploit, and explain: personalizing explainable recommendations with bandits. *Proceedings of the 12th ACM conference on recommender systems*. 31–39.

Melville, Prem, Vikas Sindhwani. 2010. Recommender systems. *Encyclopedia of machine learning* **1** 829–838.

Misra, Kanishka, Eric M Schwartz, Jacob Abernethy. 2019. Dynamic online pricing with incomplete information using multiarmed bandit experiments. *Marketing Science* **38**(2) 226–252.

Netzer, Oded, James M Lattin, Vikram Srinivasan. 2008. A hidden Markov model of customer relationship dynamics. *Marketing science* **27**(2) 185–204.

Padilla, Nicolas, Eva Ascarza. 2017. First impressions count: Leveraging acquisition data for customer management.

Rabin, Matthew, Joel L Schrag. 1999. First impressions matter: A model of confirmatory bias. *The quarterly journal of economics* **114**(1) 37–82.

Schwartz, Eric M, Eric Bradlow, Peter Fader, Yao Zhang. 2011. 'Children of the HMM': Modeling Longitudinal Customer Behavior at Hulu. com. *Com (August 1, 2011)* .

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

Stewart, Gilbert W. 1998. *Matrix algorithms: volume 1: basic decompositions*. SIAM.

Sutton, Richard S, Andrew G Barto. 2018. *Reinforcement learning: An introduction*. MIT press.

Watson, Amy (November-14). 2018. Digital Music - Statistics and Facts URL `https://www.statista.com/topics/1386/digital-music/`.

Wei, Kangning, Jinghua Huang, Shaohong Fu. 2007. A survey of e-commerce recommender systems. *2007 international conference on service systems and service management*. IEEE, 1–5.

Xu, Boya, Yiting Deng, Carl Mela. 2022. A scalable recommendation engine for new users and items. *arXiv preprint arXiv:2209.06128* .

Zhang, Boxun, Gunnar Kreitz, Marcus Isaksson, Javier Ubillos, Guido Urdaneta, Johan A Pouwelse, Dick Epema. 2013. Understanding user behavior in Spotify. *2013 Proceedings IEEE INFOCOM*. IEEE, 220–224.

Zhang, Dennis J, Ming Hu, Xiaofei Liu, Yuxiang Wu, Yong Li. 2020. NetEase Cloud Music Data. *Manufacturing & Service Operations Management* .

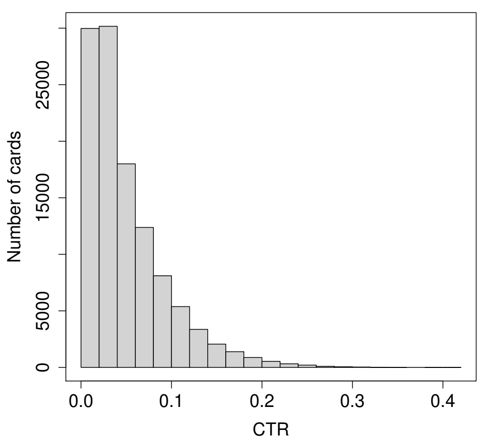
Churning While Experimenting (Electronic Companion)

Appendix A: Further Details on the Data Analysis

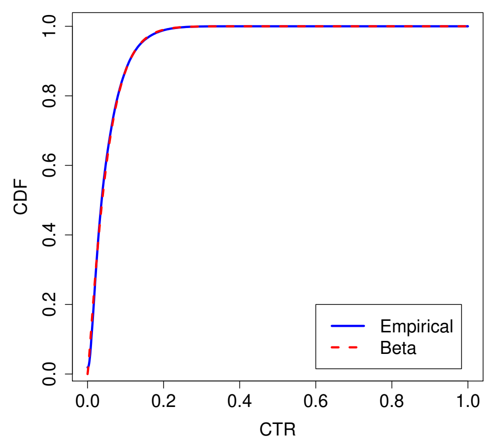
A.1. Empirical Distribution of CTR

To understand card heterogeneity, a key quantity to understand is the distribution over CTRs. In Fig. EC.1a, we show the empirical histogram obtained via a sample estimate of each card's CTR in the NetEase dataset that was shown at least 100 times (around 110,000 such cards).

Figure EC.1 CTR distribution at NetEase.



(a) Empirical histogram of CTR



(b) Beta fit to the empirical distribution

Note. For each card, we estimate the CTR (subplot (a)) as the proportion of times it was clicked. The sample mean equals $\mu_0 = 0.0507$, and the sample standard deviation equals $\sigma_0 = 0.0441$. To fit the Beta distribution (subplot (b)), we calibrate to the first two moments of the data, as in Eq. (EC.1).

Visually, the histogram for the empirical CTR resembles a Beta distribution. In fact, as we show in Fig. EC.1b, the Beta distribution explains the data quite well. We obtain the fit by calibrating the first two moments of the $\text{Beta}(\alpha_0, \beta_0)$ distribution to the data-driven sample mean ($\mu_0 = 0.0507$) and standard deviation ($\sigma_0 = 0.0441$). In particular, for a $\text{Beta}(\alpha_0, \beta_0)$, we have $\mu_0 = \frac{\alpha_0}{\alpha_0+\beta_0}$ and $\sigma_0^2 = \frac{\alpha_0\beta_0}{(\alpha_0+\beta_0)^2(\alpha_0+\beta_0+1)}$. Rearranging gives us the following closed-form expressions for α_0 and β_0:

$$\alpha_0 = \left(\frac{1-\mu_0}{\sigma_0^2} - \frac{1}{\mu_0}\right)\mu_0^2, \qquad \beta_0 = \alpha_0\left(\frac{1}{\mu_0} - 1\right). \tag{EC.1}$$

A.2. Further Details on the First Impression Effect

In this appendix, we discuss the churn rate of "regular" users. Recall from Section 2.2 that for a new user, the churn rate is around 47% if they click and around 52% if they do not click, implying a "churn delta" of around 5 percentage points. For a "regular" user, we found the churn rate to be around 5% if they click, with a churn delta of around 1 to 2 percentage points. For illustrative purposes, we defined a regular user as follows. We considered the set of users who have been on the platform for at least six months. Intuitively, a regular user is someone who visits the platform frequently. To enforce this, we considered day 10 in the data and filtered for users (six months plus) who visited the platform on both days 9 and 10. There were 36,429 such users. We analyzed their interaction on day 10 and whether they churned (i.e., never returned) as a function of whether they clicked or not on day 10. Out of the 36,429 users, 6161 clicked during their visit on day 10, with 238 (out of 6161) churning, i.e., a churn rate of around 3.86% given a click. Of the remaining 30,268 users who did not click, 1705 churned (5.63%), implying a delta of 1.77 percentage points. As a robustness check, in addition to day 10, we repeated the calculation for days 7, 8, and 9 and found the numbers to be similar: $(3.56\%, 5.46\%)$, $(5.01\%, 6.26\%)$, and $(5.80\%, 6.90\%)$, respectively. For each day, there were around 40,000 users (high sample size). Note that our definition of a regular user is just one possibility, and it is possible to construct alternative "types" of users (perhaps with an even lower delta and hence, a stronger first impression effect). Our model and analysis in Sections 3 to 5 are flexible enough to allow for an arbitrary number of user types with corresponding churn rates and deltas.

A.3. Evidence of Blind Randomization

To understand whether the NetEase dataset supports the hypothesis of blind randomization put forth in Section 2.3, we consider the new cards in the dataset (cards published in November, the month for which the data is released). Though there are ∼100,000 such cards, it is possible that the first impression of these cards was to a user outside the 2 million users in the dataset. Accordingly, in order to be sure we only use the new cards for which we know the first impression, we focus on the cards that were shown exactly once in November (over *all* users at NetEase and not just the 2 million users in the released dataset). We inferred this using the `mlog_stats.csv` file provided in the dataset. We found 1024 such cards.

We conduct a Bayesian analysis with a null hypothesis that blind randomization exists. Mathematically, given a new card and two user types (regular and new), blind randomization is equivalent to:

$$\mathbb{P}\{\text{show new card} \mid \text{user is new}\} = \mathbb{P}\{\text{show new card} \mid \text{user is regular}\}. \tag{EC.2}$$

Under the null hypothesis, both probabilities in (EC.2) equal $\mathbb{P}\{\text{show new card}\}$ (i.e., experimentation is type-independent), which we denote by p_{blind}. For inference, we consider the following ratio as a test statistic:

$$\frac{\mathbb{P}\{\text{show new card} \mid \text{user is new}\}}{\mathbb{P}\{\text{show new card} \mid \text{user is regular}\}} = \frac{\mathbb{P}\{\text{user is new} \mid \text{show new card}\}}{\mathbb{P}\{\text{user is regular} \mid \text{show new card}\}} \times \frac{\mathbb{P}\{\text{user is regular}\}}{\mathbb{P}\{\text{user is new}\}} \tag{EC.3}$$

The equality follows Bayes' theorem. Under the null hypothesis, our test statistic should be close to 1 whereas a value less (more) than 1 indicates new users receive less (more) experimentation than regular users.

We classify a user to be "new" if they registered on the platform within x months of November and we vary $x \in \{0,1,2\}$ as a robustness check. We estimate $\mathbb{P}\{\text{user is new} \mid \text{show new card}\}$ as the empirical ratio $\sum_{c=1}^{1024} \frac{\mathbb{I}\{\text{first impression of card } c \text{ shown to a new user}\}}{1024}$, where the sum is over the 1024 new cards discussed above. Trivially, $\mathbb{P}\{\text{user is regular} \mid \text{show new card}\} = 1 - \mathbb{P}\{\text{user is new} \mid \text{show new card}\}$. Similarly, our estimate of $\mathbb{P}\{\text{user is new}\}$ equals the empirical proportion $\sum_{i=1}^{57{,}750{,}012} \frac{\mathbb{I}\{\text{impression } i \text{ shown to a new user}\}}{57{,}750{,}012}$, where the sum is over all the 57,750,012 impressions in the dataset. $\mathbb{P}\{\text{user is regular}\} = 1 - \mathbb{P}\{\text{user is new}\}$. We plug these estimates to evaluate our test statistic from (EC.3) and show results in Table EC.1.

x	0	1	2
Test statistic	0.72	1.82	1.24

Table EC.1 Test statistic as in (EC.3) **corresponding to the 1024 cards with exactly one impression. Recall that $x \in \{0,1,2\}$ defines the class of new users (x months since registration). In terms of the raw quantities, number of impressions to new users equals ∼0.4M, 2.4M, and 3.8M for $x = 0,1,2$, respectively. In addition, the number of new cards (out of 1024) with first impression to a new user equals 5, 76, and 82 for $x = 0,1,2$, respectively.**

The estimate of the test statistic for $x \in \{1,2\}$ is bigger than 1 whereas for $x = 0$, the estimate equals 0.74. By themselves, these numbers are hard to interpret. To quantify their statistical significance, we use the concept of Bayesian p-values (Gelman et al. 2013). As mentioned above, under the null hypothesis, the experimentation probability is independent of the type and we denote it by p_{blind}. Under the null hypothesis, it seems reasonable to estimate p_{blind} as the empirical ratio $\frac{\text{number of new cards}}{\text{number of impressions}} = \frac{1024}{57{,}750{,}012}$. Further, as before, we use the ratio $\frac{\text{number of impressions to new users}}{\text{number of impressions}} = \frac{n_{\text{new}}}{57{,}750{,}012}$ to capture $\mathbb{P}\{\text{user is new}\}$. Accordingly, we get

$$\frac{\mathbb{P}\{\text{user is regular}\}}{\mathbb{P}\{\text{user is new}\}} = \frac{n_{\text{regular}}}{n_{\text{new}}}, \tag{EC.4}$$

where $n_{\text{regular}} = 57,750,012 - n_{\text{new}}$ denotes the number of impressions to the regular users. Recall from the caption of Table EC.1 that n_{new} equals ∼0.4 million, 2.4 million, and 3.8 million for $x = 0,1,2$, respectively.

Under the null hypothesis, number of new cards (out of 1024) with first impression to a new user equals

$$y \sim \text{Binomial}(p_{\text{blind}}, n_{\text{new}}). \tag{EC.5}$$

It is possible for y to exceed 1024 (though unlikely given our estimates). If so, we cap it at 1024. Then, the number of new cards (out of 1024) with first impression to a regular user equals $1024 - y$. This implies

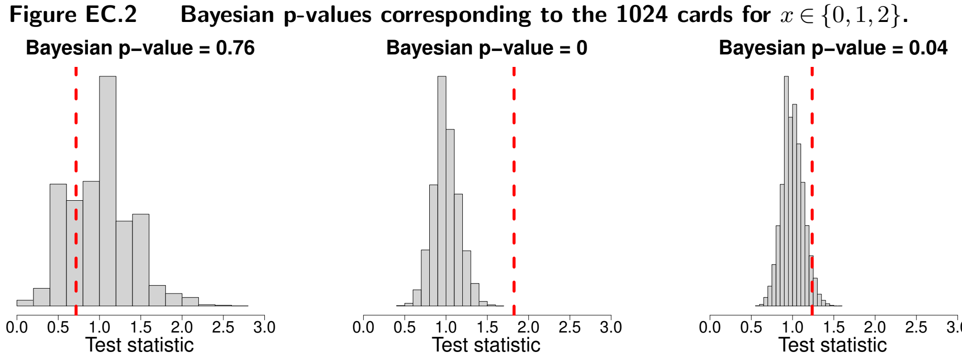
$$\frac{\mathbb{P}\{\text{user is new} \mid \text{show new card}\}}{\mathbb{P}\{\text{user is regular} \mid \text{show new card}\}} = \frac{y}{1024 - y}. \tag{EC.6}$$

Plugging (EC.4) and (EC.6) in (EC.3) gives us the distribution of our test statistic under the null hypothesis:

$$\frac{y}{1024 - y} \times \frac{n_{\text{regular}}}{n_{\text{new}}} \text{ where } y \text{ as in (EC.5).} \tag{EC.7}$$

We use Monte-Carlo simulation to understand this distribution (by simulating y). Furthermore, we know the value of the test statistic in real-data (see Table EC.1). Hence, we can estimate the Bayesian p-value for each $x \in \{0, 1, 2\}$. In Fig. EC.2, we report these Bayesian p-values. For $x = 0$, even though the test statistic is below 1, the Bayesian p-value equals 0.76, indicating the data is consistent with the null hypothesis. Moreover, for $x \in \{1, 2\}$, the Bayesian p-value is close to zero, indicating more experimentation on new users than suggested by the null hypothesis.

Figure EC.2 Bayesian p-values corresponding to the 1024 cards for $x \in \{0, 1, 2\}$.



(a) $x = 0$ (b) $x = 1$ (c) $x = 2$

Note. For each $x \in \{0, 1, 2\}$, we compute the distribution of our test statistic under the null hypothesis (as in (EC.7)) via Monte-Carlo simulation (10,000 samples of y as in (EC.5)). Then, we check where the ratio computed using real-data (recall Table EC.1) lies on this distribution (dotted red line) and compute the area to its right (Bayesian p-value). By definition, Bayesian p-value lies between 0 and 1. A value around 0.5 suggests the data is consistent with the null hypothesis whereas a value closer to 0 (1) suggests more experimentation on new (regular) users.

Appendix B: Further Details on the Numerics

B.1. Implementation Details

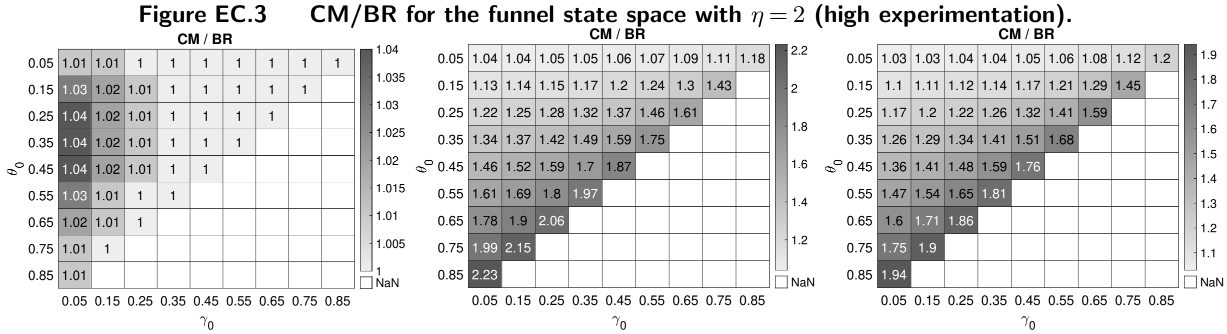
Our code is implemented in `MATLAB` (MATLAB 2021) on a MacBook Pro with an M1 Max chip and 64 GB memory. At a high-level, we perform three computations for a given problem instance: (1) LP upperbound, (2) steady-state of BR, and (3) steady-state of CM. We discuss them sequentially and note that the compute time for each of these three computations (for a given parameter setting) was less than a second.

We compute the LP upperbound by solving the underlying linear program using `gurobi` (Gurobi Optimization, LLC 2023). Computing the steady-state of BR is straightforward as it involves solving a system of linear equations (cf. Lemma 3). To compute the steady-state of CM, we perform fixed-point iteration. We start at a market state of $\mathbf{\Lambda}^{(1)} = (\lambda_1, 0, \ldots, 0)$, compute the allocation $\Pi^{\mathsf{CM}}(\mathbf{\Lambda}^{(1)})$, followed by the computation of the market state $\mathbf{\Lambda}^{(2)}$ in the next period (via the flow-balance equation (4)). We repeat this process until convergence, i.e., until $\mathbf{\Lambda}^{(t)} \approx \mathbf{\Lambda}^{(t+1)}$. Our numerical check for this is $\|\mathbf{\Lambda}^{(t)} - \mathbf{\Lambda}^{(t+1)}\|_2 < \epsilon$, with $\epsilon = 0.01$. In fact, we can do so to simulate any given policy, e.g., the LP policy in Example 1.

Note that given fluid transitions, there is no randomness. However, it is easy to relax the fluid transitions assumption in our numerics by drawing the stochastic transitions using the multinomial distribution and simulating the Markov process long enough to compute the steady state. In fact, as a robustness check against the fluid transitions assumption, we did implement this idea in our code and found the corresponding results to be nearly identical to the ones we report.

B.2. Sensitivity Analysis

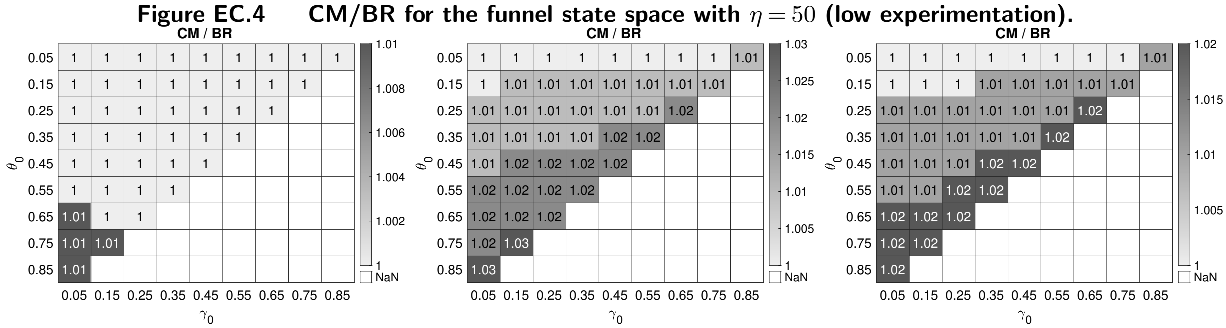
Figure EC.3 CM/BR for the funnel state space with $\eta = 2$ (high experimentation).



(a) $\gamma_1 = \theta_1 = 0.05$ (b) $\gamma_1 = \theta_1 = 1$ (c) $\gamma_1 = \theta_1 = 2$

Note. Market size under CM over that under BR for the Section 6.1 setup but with $\eta = 2$.

Figure EC.4 CM/BR for the funnel state space with $\eta = 50$ (low experimentation).



(a) $\gamma_1 = \theta_1 = 0.05$ (b) $\gamma_1 = \theta_1 = 1$ (c) $\gamma_1 = \theta_1 = 2$

Note. Market size under CM over that under BR for the Section 6.1 setup but with $\eta = 50$.

Appendix C: Omitted Proofs

C.1. Omitted Proofs from Section 4

Proof of Lemma 3. Plugging Definition 1 ($\pi_s = \frac{\Lambda_s}{\eta}\ \forall s$) into the steady-state Eq. (5) gives us

$$\Lambda_s = \lambda_s + \sum_{s'=1}^{m} \left\{ \frac{\Lambda_{s'}}{\eta}(p_{s'2s} - p_{s'1s}) + \Lambda_{s'} p_{s'1s} \right\}\ \forall s \in \mathbb{S}$$

$$\Longrightarrow \Lambda_s \left(1 - p_{s1s} - \frac{p_{s2s} - p_{s1s}}{\eta}\right) - \sum_{s' \neq s} \Lambda_{s'} \left(p_{s'1s} + \frac{p_{s'2s} - p_{s'1s}}{\eta}\right) = \lambda_s\ \forall s \in \mathbb{S}.$$

It suffices to show this system of linear equations admits a unique solution (w.r.t. $\boldsymbol{\Lambda}$). Transforming it into the standard form of $\boldsymbol{A}\boldsymbol{\Lambda} = \boldsymbol{b}$, we get the vector $\boldsymbol{b}$ equals $\boldsymbol{\lambda}$ and the matrix $\boldsymbol{A}$ admits the following structure:

$$\boldsymbol{A} = \begin{bmatrix} 1 + \left(\frac{1}{\eta} - 1\right) p_{111} - \frac{1}{\eta} p_{121} & \left(\frac{1}{\eta} - 1\right) p_{211} - \frac{1}{\eta} p_{221} & \cdots & \left(\frac{1}{\eta} - 1\right) p_{m11} - \frac{1}{\eta} p_{m21} \\ \left(\frac{1}{\eta} - 1\right) p_{112} - \frac{1}{\eta} p_{121} & 1 + \left(\frac{1}{\eta} - 1\right) p_{212} - \frac{1}{\eta} p_{222} & \cdots & \left(\frac{1}{\eta} - 1\right) p_{m12} - \frac{1}{\eta} p_{m21} \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{1}{\eta} - 1\right) p_{11m} - \frac{1}{\eta} p_{12m} & \left(\frac{1}{\eta} - 1\right) p_{21m} - \frac{1}{\eta} p_{22m} & \cdots & 1 + \left(\frac{1}{\eta} - 1\right) p_{m1m} - \frac{1}{\eta} p_{m2m} \end{bmatrix}.$$

If $\boldsymbol{A}^\top$ is non-singular, then so is $\boldsymbol{A}$. Hence, it suffices to show $\boldsymbol{A}^\top$ is non-singular. To do so, we use the following fact: if $\lim_{n\to\infty}(\boldsymbol{I} - \boldsymbol{A}^\top)^n = 0$, then $\boldsymbol{A}^\top$ is non-singular (p. 55 of Stewart (1998)). Observe that $\boldsymbol{I} - \boldsymbol{A}^\top = \left[\left(1 - \frac{1}{\eta}\right) p_{s1s'} + \frac{1}{\eta} p_{s2s'}\right]_{(s,s') \in \mathbb{S}\times\mathbb{S}}$ is a Markov chain transition matrix over $\mathbb{S}$ under the following randomized policy: do not experiment ($a = 1$) w.p. $1 - \frac{1}{\eta}$ and experiment ($a = 2$) w.p. $\frac{1}{\eta}$ (blind randomization). Due to the "leakage" assumption, it follows that $\lim_{n\to\infty}(\boldsymbol{I} - \boldsymbol{A}^\top)^n = 0$. This is because the (i, j)-th entry of the matrix $(\boldsymbol{I} - \boldsymbol{A}^\top)^n$ corresponds to the n-step probability of being in state j given a starting state of i, and leakage implies every user quits w.p. 1 as n goes to ∞ (Ch. 3 of Kemeny and Snell (1976)). Hence, $\boldsymbol{A}$ is non-singular, which means the steady-state $\boldsymbol{A}^{-1}\boldsymbol{b}$ exists uniquely. $\square$

Proof of Proposition 1. We consider a family of binary state space instances parameterized by ϵ and show that as $\epsilon \to 0$, BR yields arbitrarily poor performance. Fixing $\epsilon > 0$, the user model we consider is described in Fig. 4. The instance in Fig. 4 obeys the "leakage" assumption since $p_{saq} > 0$ for all s and a, as long as $\epsilon > 0$. Under BR, by Corollary EC.1 (see the text following this proof), the steady-state market size is:

$$\Lambda_1^{\mathsf{BR}} = \lambda_1,$$

$$\Lambda_2^{\mathsf{BR}} = \Lambda_1^{\mathsf{BR}} \frac{p_{112} - \frac{p_{112} - p_{122}}{\eta}}{1 - p_{212} + \frac{p_{212} - p_{222}}{\eta}} = \lambda_1 \frac{(1-2\epsilon) - \frac{(1-2\epsilon)-0}{\eta}}{1-(1-\epsilon) + \frac{(1-\epsilon)-(1-\epsilon)}{\eta}} = \frac{\lambda_1 (1-2\epsilon)\left(1 - \frac{1}{\eta}\right)}{\epsilon}.$$

Now, consider a policy that experiments first on users in state 2 (e.g., regular users), then experiments on users in state 1 (e.g., new users) only when it runs out of state 2 users to experiment on. We call this policy "churn minimization (CM)", the reasons for which will become clear in Section 5. Assuming there are enough state 2 users to experiment on, CM experiments as follows: $\pi_1^{\mathsf{CM}} = 0$ and $\pi_2^{\mathsf{CM}} = \frac{\Lambda_1^{\mathsf{CM}} + \Lambda_2^{\mathsf{CM}}}{\eta}$. It then follows from Eq. (5) that the steady-state market size under CM equals: $\Lambda_1^{\mathsf{CM}} = \lambda_1$ and $\Lambda_2^{\mathsf{CM}} = \frac{\lambda_1(1-2\epsilon)}{\epsilon}$. The "enough state 2 users" assumption for CM holds when there are enough users in state 2 to experiment on, i.e., $\pi_2^{\mathsf{CM}} \le \Lambda_2^{\mathsf{CM}}$. As we show in Corollary EC.2 (Section C.2) where we establish the steady-state of CM for the binary state space, this assumption is equivalent to $\eta \ge \frac{p_{112} + d_2 + p_{21q}}{p_{112}} = \frac{p_{112} + 1 - p_{222}}{p_{112}}$, which for the current instance simplifies to $\eta \ge \frac{1-\epsilon}{1-2\epsilon}$. Setting $\eta = \frac{1-\epsilon}{1-2\epsilon}$,

$$\frac{\Lambda_1^{\mathsf{BR}} + \Lambda_2^{\mathsf{BR}}}{\Lambda_1^{\mathsf{CM}} + \Lambda_2^{\mathsf{CM}}} = \frac{\epsilon + (1-2\epsilon)\left(1 - \frac{1}{\eta}\right)}{\epsilon + (1-2\epsilon)} = \frac{2\epsilon^2}{1-\epsilon} \to 0 \text{ as } \epsilon \to 0.$$

Hence, BR can achieve no constant factor performance of another feasible policy (CM). □

In the proof of Proposition 1 above, we invoke the following result, which characterizes the steady-state market size for BR under the binary state space and directly follows Lemma 3.

COROLLARY EC.1. *For the binary state space, BR results in the following steady-state market:*

$$\Lambda_1^{\mathsf{BR}} = \lambda_1, \qquad \Lambda_2^{\mathsf{BR}} = \Lambda_1^{\mathsf{BR}} \frac{p_{112} - \frac{p_{112} - p_{122}}{\eta}}{1 - p_{212} + \frac{p_{212} - p_{222}}{\eta}}.$$

C.2. Omitted Proofs from Section 5

Proof of Lemma 4. Recall the churn deltas are denoted by $[d_s]_s$ and the return probability (under no exp.) from state $s \in \mathbb{S}$ by $r_s := 1 - p_{s1q}$. We show that for $s \in \{1, 2\}$ (binary state space), the sequence $(\Lambda_s^{(t)})_t$ increases monotonically and is bounded, which implies that it has a finite limit (cf. monotone convergence theorem). For state 1, observe that $\Lambda_1^{(t)} = \lambda_1$ for $t \in \{1, 2, \ldots\}$. Hence, we only need to analyze $(\Lambda_2^{(t)})_t$.

First, to establish $\Lambda_2^{(t)}$ is upper bounded for all t, consider a hypothetical market with no experimentation (i.e., $\eta=\infty$) and denote the corresponding market state evolution by $(\bar{\boldsymbol{\Lambda}}_t)_t$. Observe that $\Lambda_2^{(t)} \le \bar{\Lambda}_2^{(t)}$ for all t. This is because experimentation leads to more churn (due to the churn deltas being non-negative by assumption). Furthermore, it is easy to see that $(\bar{\Lambda}_2^{(t)})_t$ is monotonically increasing and converges to $\bar{\Lambda}_2$, where $\bar{\Lambda}_2 = \lambda_1 r_1 + \bar{\Lambda}_2 r_2$ (flow-balance), i.e., $\bar{\Lambda}_2 = \frac{\lambda_1 r_1}{1-r_2}$. Due to the "leakage" assumption, $r_2<1$ and hence, $\bar{\Lambda}_2$ is finite. Second, to show $(\Lambda_2^{(t)})_t$ increases monotonically, we perform induction. For the base case, observe that $\Lambda_2^{(1)} \le \Lambda_2^{(2)}$. This holds because $\Lambda_2^{(1)}=0$ and $\Lambda_2^{(2)} \ge 0$. Now, suppose $\Lambda_2^{(t-1)} \le \Lambda_2^{(t)}$. It suffices to show $\Lambda_2^{(t)} \le \Lambda_2^{(t+1)}$. We proceed in two cases: (1) $d_1 \ge d_2$ and (2) $d_1 < d_2$.

Case 1: $d_1 \ge d_2$***.*** In this case, state 2 users are prioritized over state 1 users to receive experimentation. We examine four mutually exclusive and exhaustive sub-cases in terms of $(\Lambda_2^{(t-1)}, \Lambda_2^{(t)})$. First, consider $\Lambda_2^{(t-1)} \ge \frac{\lambda_1+\Lambda_2^{(t-1)}}{\eta}$ and $\Lambda_2^{(t)} \ge \frac{\lambda_1+\Lambda_2^{(t)}}{\eta}$. That is, in both periods $t-1$ and t, there are enough state 2 users so that the experimentation does not trickle down to state 1. Flow balance along with $\Lambda_2^{(t)} = \Lambda_2^{(t-1)} + (\Lambda_2^{(t)} - \Lambda_2^{(t-1)})$ implies $\Lambda_2^{(t+1)} = \frac{\lambda+\Lambda_2^{(t)}}{\eta}(r_2-d_2) + \left(\Lambda_2^{(t)} - \frac{\lambda+\Lambda_2^{(t)}}{\eta}\right) r_2 = \left(\frac{\lambda+\Lambda_2^{(t-1)}}{\eta} + \frac{\Lambda_2^{(t)}-\Lambda_2^{(t-1)}}{\eta}\right)(r_2-d_2) + \left(\Lambda_2^{(t-1)} - \frac{\lambda+\Lambda_2^{(t-1)}}{\eta} + (\Lambda_2^{(t)} - \Lambda_2^{(t-1)})\left(1-\frac{1}{\eta}\right)\right) r_2 \ge \frac{\lambda+\Lambda_2^{(t-1)}}{\eta}(r_2-d_2) + \left(\Lambda_2^{(t-1)} - \frac{\lambda+\Lambda_2^{(t-1)}}{\eta}\right) r_2 = \Lambda_2^{(t)}$. For the inequality, we invoke $\Lambda_2^{(t)} \ge \Lambda_2^{(t-1)}$ (induction hypothesis). The second ($\Lambda_2^{(t-1)} < \frac{\lambda_1+\Lambda_2^{(t-1)}}{\eta}$ and $\Lambda_2^{(t)} \ge \frac{\lambda_1+\Lambda_2^{(t)}}{\eta}$) and third ($\Lambda_2^{(t-1)} < \frac{\lambda_1+\Lambda_2^{(t-1)}}{\eta}$ and $\Lambda_2^{(t)} < \frac{\lambda_1+\Lambda_2^{(t)}}{\eta}$) sub-cases can be handled similarly via first principles. The fourth sub-case ($\Lambda_2^{(t-1)} \ge \frac{\lambda_1+\Lambda_2^{(t-1)}}{\eta}$ and $\Lambda_2^{(t)} < \frac{\lambda_1+\Lambda_2^{(t)}}{\eta}$) is impossible. This is because $\Lambda_2^{(t-1)} \ge \frac{\lambda_1+\Lambda_2^{(t-1)}}{\eta}$ implies $\Lambda_2^{(t-1)} \ge \frac{\lambda_1}{\eta-1}$ and $\Lambda_2^{(t)} < \frac{\lambda_1+\Lambda_2^{(t)}}{\eta}$ implies $\Lambda_2^{(t)} < \frac{\lambda_1}{\eta-1}$, resulting in $\Lambda_2^{(t-1)} > \Lambda_2^{(t)}$, which contradicts the induction hypothesis.

Case 2: $d_1 < d_2$***.*** In this case, state 1 users are prioritized over state 2 users to receive experimentation. We again examine four mutually exclusive and exhaustive sub-cases in terms of $(\Lambda_2^{(t-1)}, \Lambda_2^{(t)})$. First, consider $\lambda_1 \ge \frac{\lambda_1+\Lambda_2^{(t-1)}}{\eta}$ and $\lambda_1 \ge \frac{\lambda_1+\Lambda_2^{(t)}}{\eta}$. That is, in both periods $t-1$ and t, there are enough state 1 users so that the experimentation does not trickle up to state 2. Since $\Lambda_2^{(t)} \ge \Lambda_2^{(t-1)}$, an additional $\frac{\Lambda_2^{(t)}-\Lambda_2^{(t-1)}}{\eta}$ users receive experimentation in period t (compared to period $t-1$) and all the experimentation users are in state 1 (by the definition of this sub-case). This increased experimentation results in a lower number of users moving from state 1 to state 2; $\frac{\Lambda_2^{(t)}-\Lambda_2^{(t-1)}}{\eta} d_1$ to be precise. However, there is an increase of $(\Lambda_2^{(t)} - \Lambda_2^{(t-1)}) r_2$ users from state 2 to state 2 transitions. Hence, the net change equals $(\Lambda_2^{(t)} - \Lambda_2^{(t-1)})\left(r_2 - \frac{d_1}{\eta}\right) \ge 0$. The inequality is true because $r_2 - d_2 \ge 0$ and $d_1 < d_2$ imply $r_2 - d_1 \ge 0$, which implies $r_2 - \frac{d_1}{\eta} \ge 0$ (as $\eta > 1$). The second ($\lambda_1 \ge \frac{\lambda_1+\Lambda_2^{(t-1)}}{\eta}$ and $\lambda_1 < \frac{\lambda_1+\Lambda_2^{(t)}}{\eta}$) and third ($\lambda_1 < \frac{\lambda_1+\Lambda_2^{(t-1)}}{\eta}$ and $\lambda_1 < \frac{\lambda_1+\Lambda_2^{(t)}}{\eta}$) sub-cases can be handled similarly via first principles. The fourth sub-case ($\lambda_1 < \frac{\lambda_1+\Lambda_2^{(t-1)}}{\eta}$ and $\lambda_1 \geq \frac{\lambda_1+\Lambda_2^{(t)}}{\eta}$) implies $\Lambda_2^{(t-1)} > \Lambda_2^{(t)}$, which contradicts the induction hypothesis. Hence, this sub-case is impossible. $\square$

Proof of Theorem 1. We consider four cases depending on whether one or both states are used by CM, and which state is used first. In each case, we use the LP upper bound from Lemma 2 to demonstrate the optimality of CM. As a preliminary, we simplify the LP formulation under the binary state space.

Let (π_1, π_2) be an arbitrary steady-state experimentation action and recall $d_s := p_{s2_q} - p_{s1_q}$ for all $s \in \{1,2\}$, and that for the binary state space, we assume arrivals are only to the first state and self-loops only in the terminal state. Under (π_1, π_2), the number of steady-state users in state 1 and state 2 by Eq. (5) is:

$$\Lambda_1 = \lambda_1, \qquad \Lambda_2 = \frac{\lambda_1 p_{112} - d_1\pi_1 - d_2\pi_2}{p_{21_q}}. \tag{EC.9}$$

LP maximizes $\Lambda_1 + \Lambda_2$, which is equivalent to maximizing $-d_1\pi_1 - d_2\pi_2$ by Eq. (EC.9). In fact, Eq. (EC.9) allows us to eliminate the (Λ_1, Λ_2) variables from the LP, resulting in the following simplified LP (primal):

$$\max_{(\pi_1,\pi_2)\geq 0} \; -d_1\pi_1 - d_2\pi_2 \tag{EC.10a}$$

$$\text{s.t. } \pi_1 \leq \lambda_1 \tag{EC.10b}$$

$$d_1\pi_1 + (p_{21_q} + d_2)\pi_2 \leq \lambda_1 p_{112} \tag{EC.10c}$$

$$(p_{21_q}\eta + d_1)\pi_1 + (p_{21_q}\eta + d_2)\pi_2 = (p_{112} + p_{21_q})\lambda_1. \tag{EC.10d}$$

Eq. (EC.10b) and Eq. (EC.10c) enforce $\pi_s \leq \Lambda_s$ for $s \in \{1,2\}$, and Eq. (EC.10d) enforces adequate experimentation, i.e., $\pi_1 + \pi_2 = (\Lambda_1 + \Lambda_2)/\eta$. This linear program admits the following dual:

$$\min_{(\theta_1,\theta_2)\geq 0, z} \; \lambda_1\theta_1 + \lambda_1 p_{112}\theta_2 + \lambda_1(p_{112} + p_{21_q})z \tag{EC.11a}$$

$$\text{s.t. } \theta_1 + d_1\theta_2 + (p_{21_q}\eta + d_1)z \geq -d_1 \tag{EC.11b}$$

$$(p_{21_q} + d_2)\theta_2 + (p_{21_q}\eta + d_2)z \geq -d_2. \tag{EC.11c}$$

We will use strong duality for linear programming to solve for the optimal policy (Ch. 4 of Bertsimas and Tsitsiklis (1997)), and demonstrate that it is always CM. Our proof will proceed in four cases depending on whether the solution uses one state or both for experimentation.

Case 1: $p_{112} + p_{21_q} > \frac{p_{21_q}\eta + d_2}{p_{21_q} + d_2} p_{112}$**,** $d_1 > d_2$**.** In this case, both states are used by CM. Now, consider the following pair of primal and dual solutions: $(\pi_1, \pi_2) = \left(\frac{\lambda_1(d_2 + p_{112} - \eta p_{112} + p_{21_q})}{d_1 - d_1\eta + \eta(d_2 + p_{21_q})}, \frac{\lambda_1(d_1 - \eta p_{112})}{d_1(\eta - 1) - \eta(d_2 + p_{21_q})}\right)$ and

$(\theta_1,\theta_2,z)=\left(0,\frac{\eta(d_2-d_1)}{d_1(\eta-1)-\eta(d_2+p_{21q})},\frac{d_1}{d_1(\eta-1)-\eta(d_2+p_{21q})}\right)$. First, we claim that this pair satisfies strong duality. To see this, observe the primal objective Eq. (EC.10a) equals the dual objective Eq. (EC.11a): $-d_1\pi_1-d_2\pi_2=-d_1\frac{\lambda_1(d_2+p_{112}-\eta p_{112}+p_{21q})}{d_1-d_1\eta+\eta(d_2+p_{21q})}-d_2\frac{\lambda_1(d_1-\eta p_{112})}{d_1(\eta-1)-\eta(d_2+p_{21q})}=\frac{\lambda_1(d_2\eta p_{112}+d_1(p_{112}-\eta p_{112}+p_{21q}))}{d_1(\eta-1)-\eta(d_2+p_{21q})}=\frac{\lambda_1 p_{112}\eta(d_2-d_1)}{d_1(\eta-1)-\eta(d_2+p_{21q})}+\frac{\lambda_1(p_{112}+p_{21q})d_1}{d_1(\eta-1)-\eta(d_2+p_{21q})}=\lambda_1p_{112}\theta_2+\lambda_1(p_{112}+p_{21q}z)$. Now, we will show that (π_1,π_2) is primal-feasible. First, Eq. (EC.10b) holds since $\pi_1\le\lambda_1 \iff \frac{d_2+p_{112}-\eta p_{112}+p_{21q}}{d_1-d_1\eta+d_2\eta+p_{21q}\eta}\le 1 \iff 0\le(\eta-1)(d_2-d_1+p_{112}+p_{21q}) \iff 0\le(\eta-1)(p_{22q}+p_{122})$, where the final implication is true since $\eta>1$ and the transition probabilities are non-negative. The next two primal constraints, Eqs. (EC.10c) and (EC.10d), are tight by construction, i.e., π_1 and π_2 are chosen to satisfy them with equality. Last thing to check for primal feasibility is that these solutions are non-negative. We can rearrange the condition for this case into an upper bound on η: $\eta<\frac{(p_{112}+p_{21q})(p_{21q}+d_2)}{p_{112}p_{21q}}-\frac{d_2}{p_{21q}}=\frac{d_2+p_{112}+p_{21q}}{p_{112}}$. Note both expressions for π_1 and π_2 are non-negative when $\eta=1$, and monotone in η since:

$$\begin{aligned}
\frac{d}{d\eta}\pi_1&=\frac{\lambda_1(d_2+p_{21q})(d_1-d_2-p_{21q}-p_{112})}{(-d_1(\eta-1)+d_2\eta+\eta p_{21q})^2}=\frac{\lambda_1(d_2+p_{21q})(-p_{22q}-p_{122})}{(-d_1(\eta-1)+d_2\eta+\eta p_{21q})^2}<0\\
\frac{d}{d\eta}\pi_2&=\frac{\lambda_1 d_1(d_2-d_1+p_{21q}+p_{112})}{(\eta(d_2-d_1)+d_1+\eta p_{21q})^2}=\frac{\lambda_1 d_1(p_{22q}+p_{122})}{(\eta(d_2-d_1)+d_1+\eta p_{21q})^2}>0.
\end{aligned}$$

Since the derivative is positive for π_2, it is always non-negative in this case. Since the derivative is negative for π_1, to check for non-negativity, we need only check π_1 at the maximal value for η. Plugging in for π_1 we get, $\pi_1|_{\eta=\frac{d_2+p_{112}+p_{21q}}{p_{112}}}=0$. Thus, the solution is primal feasible. Now, we will check dual feasibility. For the dual solution, Eq. (EC.11b) and Eq. (EC.11c) are satisfied with equality (by construction). Since z is unconstrained, we need only check that $\theta_2\ge0$. Note that $d_1>d_2$ implies $\theta_2=\frac{\eta(d_2-d_1)}{d_1(\eta-1)-\eta(d_2+p_{21q})}\ge0 \iff d_1(\eta-1)-\eta(d_2+p_{21q})<0$. Now, the strict inequality holds when $\eta=1$, it's minimal (inf to be precise) value. Further, since it is a linear equation in η, we need only check the sign at the maximal value of η from the condition for this case: $d_1(\eta-1)-\eta(d_2+p_{21q})|_{\eta=\frac{d_2+p_{112}+p_{21q}}{p_{112}}}=\frac{(d_2p_{21q})(d_1-d_2-p_{112}-p_{21q})}{p_{112}}=\frac{(d_2p_{21q})(-p_{22q}-p_{122})}{p_{112}}<0$. Thus, the proposed solution is optimal, and both primal and dual feasible. It also experiments as much as possible in the second state, and since $d_1>d_2$, it is CM. We conclude CM is optimal in this case.

Case 2: $p_{112}+p_{21q}\le\frac{p_{21q}\eta+d_2}{p_{21q}+d_2}p_{112}$**,** $d_1>d_2$**.** Consider the following pair of primal and dual solutions: $(\pi_1,\pi_2)=\left(0,\frac{\lambda_1(p_{112}+p_{21q})}{p_{21q}\eta+d_2}\right)$ and $(\theta_1,\theta_2,z)=\left(0,0,\frac{-d_2}{p_{21q}\eta+d_2}\right)$. We claim that this pair satisfies strong duality. To see this, first observe the primal objective Eq. (EC.10a) equals the dual objective Eq. (EC.11a): $-d_1\pi_1-d_2\pi_2=-d_2\frac{\lambda_1(p_{112}+p_{21q})}{p_{21q}\eta+d_2}=\lambda_1\theta_1+\lambda_1p_{112}\theta_2+\lambda_1(p_{112}+p_{21q})z$. Second, observe that (π_1,π_2) is primal-feasible. Both π_1 and π_2 are non-negative, Eq. (EC.10b) holds since $\pi_1=0\le\lambda_1$, Eq. (EC.10c) holds since

$d_1\pi_1 + (p_{21_q} + d_2)\pi_2 = \frac{p_{21_q}+d_2}{p_{21_q}\eta+d_2} \times \lambda_1(p_{112} + p_{21_q}) \le \frac{p_{21_q}+d_2}{p_{21_q}\eta+d_2} \times \frac{p_{21_q}\eta+d_2}{p_{21_q}+d_2}\lambda_1 p_{112} = \lambda_1 p_{112}$, and Eq. (EC.10d) holds since $(p_{21_q}\eta + d_1)\pi_1 + (p_{21_q}\eta + d_2)\pi_2 = (p_{21_q}\eta + d_2)\frac{\lambda_1(p_{112}+p_{21_q})}{p_{21_q}\eta+d_2} = \lambda_1(p_{112} + p_{21_q})$. Third, observe that (θ_1, θ_2, z) is dual-feasible. Both θ_1 and θ_2 are non-negative and Eq. (EC.11b) holds since $\theta_1 + d_1\theta_2 + (p_{21_q}\eta + d_1)z = -(p_{21_q}\eta + d_1)\frac{d_2}{p_{21_q}\eta+d_2} \ge -d_1$. This inequality holds because $d_1 \ge d_2 \implies d_1 p_{21_q}\eta \ge d_2 p_{21_q}\eta \implies d_1 p_{21_q}\eta + d_1 d_2 \ge d_2 p_{21_q}\eta + d_1 d_2 \implies d_1(p_{21_q}\eta + d_2) \ge d_2(p_{21_q}\eta + d_1) \implies d_1 \ge (p_{21_q}\eta + d_1)\frac{d_2}{p_{21_q}\eta+d_2} \implies -d_1 \le -(p_{21_q}\eta + d_1)\frac{d_2}{p_{21_q}\eta+d_2}$. Eq. (EC.11c) holds since $(p_{21_q} + d_2)\theta_2 + (p_{21_q}\eta + d_2)z = -(p_{21_q}\eta + d_2)\frac{d_2}{p_{21_q}\eta+d_2} = -d_2 \ge -d_2$. Finally, observe that the primal solution maps to CM since it only experiments on state 2 users.

Case 3: $p_{112} + p_{21_q} \ge d_1 + \eta p_{21_q}$, $d_1 \le d_2$. This case is similar to case 1 in that both states are used, now with state 1 being the one that is used maximally (which is CM since $d_1 \le d_2$). Consider the following pair of primal and dual solutions: $(\pi_1, \pi_2) = \left(\lambda_1, \frac{\lambda_1(-d_1+p_{112}+p_{21_q}-\eta p_{21_q})}{d_2+\eta p_{21_q}}\right), (\theta_1, \theta_2, z) = \left(\frac{(-d_1+d_2)\eta p_{21_q}}{d_2+\eta p_{21_q}}, 0, \frac{-d_2}{d_2+\eta p_{21_q}}\right)$. It is easily checked that this pair of solutions satisfies strong duality with objective value $\frac{-\lambda_1\left(d_1\eta p_{21_q}+d_2(p_{112}+p_{21_q}-\eta p_{21_q})\right)}{d_2+\eta p_{21_q}}$. Thus, we need only check primal and dual feasibility. First, we show that (π_1, π_2) is primal-feasible. To this end, note Eq. (EC.10b) and Eq. (EC.10d) both hold with equality (by construction). To check Eq. (EC.10c), $d_1\pi_1 + (p_{21_q} + d_2)\pi_2 \le \lambda_1 p_{112} \iff \frac{(\eta-1)(d_1-d_2-p_{112}-p_{21_q})p_{21_q}}{d_2+\eta p_{21_q}} \le 0$, which holds since $\eta > 1$ and $d_1 \le d_2$ by assumption of the case. Lastly, we must check the primal solution is non-negative. π_1 is clearly non-negative and so we need only check π_2 is non-negative. Consider the derivative of π_2 as a function of η: $\frac{d}{d\eta}\pi_2 = \frac{-\lambda_1 p_{21_q}(d_2-d_1+p_{21_q}+p_{112})}{(d_2+\eta p_{21_q})^2} < 0$. Furthermore, since $\pi_2|_{\eta=1} = \lambda_1\frac{-d_1+p_{112}}{d_2 p_{21_q}} > 0$, to show non-negativity, we need only check π_2 at the maximal value of η. Namely, $\pi_2|_{\eta=\frac{p_{112}+p_{21_q}-d_1}{p_{21_q}}} = 0$. Thus, the solution is primal feasible. Now, we will check dual feasibility. For the dual solution, Eq. (EC.11b) and Eq. (EC.11c) are satisfied with equality (by construction). Since z is unconstrained and $\theta_2 = 0$, we need only check that $\theta_1 \ge 0$, which follows immediately from the fact that $d_2 \ge d_1$ in this case. Thus, the proposed solution is optimal, and primal and dual feasible. It also experiments as much as possible in the first state, which in this case means it is CM.

Case 4: $p_{112} + p_{21_q} \le d_1 + \eta p_{21_q}$, $d_1 \le d_2$. Consider the following pair of primal and dual solutions: $(\pi_1, \pi_2) = \left(\frac{\lambda_1(p_{112}+p_{21_q})}{\eta p_{21_q}+d_1}, 0\right)$ and $(\theta_1, \theta_2, z) = \left(0, 0, \frac{-d_1}{\eta p_{21_q}+d_2}\right)$. We claim that this pair satisfies strong duality. To see this, first observe the primal objective Eq. (EC.10a) equals the dual objective Eq. (EC.11a): $-d_1\pi_1 - d_2\pi_2 = \frac{-d_1\lambda_1(p_{112}+p_{21_q})}{\eta p_{21_q}+d_1} = \lambda_1\theta_1 + \lambda_1 p_{112}\theta_2 + \lambda_1(p_{112} + p_{21_q})z$. Second, observe that (π_1, π_2) is primal-feasible. Both π_1 and π_2 are non-negative, Eq. (EC.10b) holds since $\pi_1 = \frac{\lambda_1(p_{112}+p_{21_q})}{\eta p_{21_q}+d_1} \le \frac{\lambda_1\left(d_1+\eta p_{21_q}\right)}{\eta p_{21_q}+d_1} = \lambda_1$, Eq. (EC.10c) holds since $d_1\pi_1 + (p_{21_q} + d_2)\pi_2 = d_1\frac{\lambda_1(p_{112}+p_{21_q})}{\eta p_{21_q}+d_1} \le d_1\frac{\lambda_1\left(d_1+\eta p_{21_q}\right)}{\eta p_{21_q}+d_1} = \lambda_1 d_1 = \lambda_1(p_{12_q} - p_{11_q}) =$

$\lambda_1(1-p_{122}-(1-p_{112}))=\lambda_1(p_{112}-p_{122})\leq\lambda_1 p_{112}$, and Eq. (EC.10d) holds since $(p_{21q}\eta+d_1)\pi_1+(p_{21q}\eta+d_2)\pi_2=(p_{21q}\eta+d_1)\frac{\lambda_1(p_{112}+p_{21q})}{\eta p_{21q}+d_1}=\lambda_1(p_{112}+p_{21q})$. Third, observe that (θ_1,θ_2,z) is dual-feasible. Both θ_1 and θ_2 are non-negative and Eq. (EC.11b) holds since $\theta_1+d_1\theta_2+(p_{21q}\eta+d_1)z=-(p_{21q}\eta+d_1)\frac{d_1}{\eta p_{21q}+d_2}\geq -d_1$. Eq. (EC.11c) holds since $(p_{21q}+d_2)\theta_2+(p_{21q}\eta+d_2)z=-(p_{21q}\eta+d_2)\frac{d_1}{\eta p_{21q}+d_2}=-d_1\geq -d_2$. Finally, observe that the primal solution maps to CM since it only experiments on state 1 users. □

Theorem 1 allows us to characterize the steady-state market size under CM for binary state spaces. In Corollary EC.2 , we zoom in on the $d_1\geq d_2$ case as it is used in the proof of Proposition 1.

Corollary EC.2. *For the binary state space with $d_1\geq d_2$, CM results in the following steady-state market: $\Lambda_1^{\mathsf{CM}}=\lambda_1$ and*

$$\Lambda_2^{\mathsf{CM}}=\begin{cases}\frac{\lambda_1(\eta p_{112}-d_2)}{d_2+\eta p_{21q}} & \text{if } \eta\geq\frac{p_{112}+d_2+p_{21q}}{p_{112}}\\ \frac{\lambda_1(d_1-\eta p_{112})}{d_1(\eta-1)-\eta(d_2+p_{21q})} & \text{if } \eta<\frac{p_{112}+d_2+p_{21q}}{p_{112}}.\end{cases}$$

Proof. The two cases correspond to cases 1 and 2 in the proof of Theorem 1. In each case, the corresponding policy (π_1,π_2) is given. Plugging those policies into Eq. (EC.9) gives the result. □

Proof of Theorem 2. Recall for the funnel state space, we assume arrivals are only to the first state, and self-loops occur only in the terminal state. Now, observe that for the funnel, the steady-state market $\boldsymbol{\Lambda}$ can be characterized as a linear function of the steady-state experimentation action $\boldsymbol{\pi}$:

$$\Lambda_s=\lambda_1\prod_{\ell=1}^{s-1}r_\ell-\sum_{\ell=1}^{s-1}\pi_\ell d_\ell\prod_{k=\ell+1}^{s-1}r_s \text{ for } s\in\{1,\dots,m-1\} \tag{EC.12a}$$

$$\Lambda_m=\frac{1}{1-r_m}\left(\lambda_1\prod_{\ell=1}^{m-1}r_\ell-\sum_{\ell=1}^{m}\pi_\ell d_\ell\prod_{k=\ell+1}^{s-1}r_s\right). \tag{EC.12b}$$

An "empty" product is defined to be 1, i.e., a product with the start index greater than the end index, and recall the churn deltas are denoted by $[d_s]_s$ and for the funnel state space, the return probability (under no exp.) from state $s\in\mathbb{S}$ is $r_s:=1-p_{s1q}$. Denoting by $\boldsymbol{\Lambda}(\boldsymbol{\pi})$ the linear relationship in Eq. (EC.12), an immediate corollary is that $\boldsymbol{\Lambda}(\boldsymbol{\pi})$ is continuous and decreasing in $\boldsymbol{\pi}$. We will use these observations in the proof.

Now, let $(\boldsymbol{\pi}^*,\boldsymbol{\Lambda}^*)$ denote any optimal solution to LP (6). We show it coincides with CM. First, suppose $\pi_m^*<\Lambda_m^*$ but $\pi_s^*>0$ for some $s\in\{1,\dots,m-1\}$. We show the existence of an alternative feasible solution with a higher objective value, which contradicts the optimality of $(\boldsymbol{\pi}^*,\boldsymbol{\Lambda}^*)$. To construct the alternative solution, increase π_m^* by ϵ for some $\epsilon\in(0,\Lambda_m^*-\pi_m^*]$. Then, Eq. (EC.12) implies $(\Lambda_1,\dots,\Lambda_{m-1})$ remains constant but Λ_m decreases by $\frac{1}{1-r_m} d_m \epsilon$. To offset this decrease in Λ_m, decrease π_s by ϵ' so that Λ_m increases by $\frac{1}{1-r_m} d_s \prod_{k=s+1}^{m-1} r_k \epsilon'$ (cf. Eq. (EC.12)). For the net change in Λ_m to be positive, we need

$$\frac{1}{1-r_m} d_s \prod_{k=s+1}^{m-1} r_k \epsilon' > \frac{1}{1-r_m} d_m \epsilon \iff \epsilon' > \frac{d_m \epsilon}{d_s \prod_{k=s+1}^{m-1} r_k}. \tag{EC.13}$$

When we decrease π_s by ϵ', $(\Lambda_{s+1}, \dots, \Lambda_{m-1})$ increases as well (cf. Eq. (EC.12)). Putting everything together, under Eq. (EC.13), the net change in the total market size $\mathbf{1}^\top \boldsymbol{\Lambda}$ is positive. The change in the amount of experimentation $\mathbf{1}^\top \boldsymbol{\pi}$ equals $\epsilon - \epsilon'$, which is positive[14] if $\epsilon' < \epsilon$. Putting this together with (EC.13), we get

$$\frac{d_m \epsilon}{d_s \prod_{k=s+1}^{m-1} r_k} < \epsilon' < \epsilon. \tag{EC.14}$$

Hence, for ϵ' to exist, we need $\frac{d_m}{d_s \prod_{k=s+1}^{m-1} r_k} < 1$, which is equivalent to $d_s \prod_{k=s+1}^{m-1} r_k > d_m$. Observe that this holds as long as $\frac{d_{s+1}}{d_s} < r_{s+1}$ for $s \in \{1, \dots, m-2\}$:

$$d_s \prod_{k=s+1}^{m-1} r_k = d_s r_{s+1} \prod_{k=s+2}^{m-1} r_k > d_{s+1} \prod_{k=s+2}^{m-1} r_k = d_{s+1} r_{s+2} \prod_{k=s+3}^{m-1} r_k > d_{s+2} \prod_{k=s+3}^{m-1} r_k > d_{m-2} r_{m-1} > d_{m-1} > d_m.$$

Hence, under the alternative solution, both $\mathbf{1}^\top \boldsymbol{\Lambda}$ and $\mathbf{1}^\top \boldsymbol{\pi}$ undergo a positive increase (compared to $\mathbf{1}^\top \boldsymbol{\Lambda}^*$ and $\mathbf{1}^\top \boldsymbol{\pi}^*$, respectively). To ensure feasibility (recall the constraints in LP), we need to obey $\mathbf{1}^\top \boldsymbol{\pi} = \frac{\mathbf{1}^\top \boldsymbol{\Lambda}}{\eta}$, $\boldsymbol{\pi} \le \boldsymbol{\Lambda}$, and $(\boldsymbol{\pi}, \boldsymbol{\Lambda}) \ge 0$. Let's consider the last two constraints first ($\boldsymbol{\pi} \le \boldsymbol{\Lambda}$ and $(\boldsymbol{\pi}, \boldsymbol{\Lambda}) \ge 0$). It suffices to ensure $\epsilon' \le \pi_s^*$. We can ensure this by setting $\epsilon = \min\{\pi_s^*, \Lambda_m^* - \pi_m^*\}$ so that picking an ϵ' as per Eq. (EC.14) works. Let's consider the $\mathbf{1}^\top \boldsymbol{\pi} = \frac{\mathbf{1}^\top \boldsymbol{\Lambda}}{\eta}$ constraint now. The (ϵ, ϵ') tweak might result in the alternative solution violating this equality. However, as discussed above, both $\mathbf{1}^\top \boldsymbol{\Lambda}$ and $\mathbf{1}^\top \boldsymbol{\pi}$ undergo a positive increase due to the (ϵ, ϵ') tweak. Combining this with the fact that $\boldsymbol{\Lambda}$ is continuous and decreasing in $\boldsymbol{\pi}$, it follows that we can restore this equality while ensuring the net change in $\mathbf{1}^\top \boldsymbol{\Lambda}$ is positive. In particular, if the alternative solution results in $\mathbf{1}^\top \boldsymbol{\pi} < \frac{\mathbf{1}^\top \boldsymbol{\Lambda}}{\eta}$, then we can increase $\boldsymbol{\pi}$ (while ensuring $\boldsymbol{\pi} \le \boldsymbol{\Lambda}$) until we obtain the desired equality ($\mathbf{1}^\top \boldsymbol{\pi} = \frac{\mathbf{1}^\top \boldsymbol{\Lambda}}{\eta}$). On the other hand, if the alternative solution results in $\mathbf{1}^\top \boldsymbol{\pi} > \frac{\mathbf{1}^\top \boldsymbol{\Lambda}}{\eta}$, then we can decrease $\boldsymbol{\pi}$ (while ensuring $\boldsymbol{\pi} \ge 0$) until we obtain the desired equality.

Second, suppose $\pi_s^* < \Lambda_s^*$ but $\pi_{s'}^* > 0$ for some $s \in \{2, \dots, m-1\}$ and $s' < s$. As above, we show the existence of an alternative feasible solution with a higher objective value, which contradicts the optimality of $(\boldsymbol{\pi}^*, \boldsymbol{\Lambda}^*)$. To construct the alternative solution, increase π_s^* by ϵ for some $\epsilon \in (0, \Lambda_s^* - \pi_s^*]$. Then, Eq. (EC.12) implies $(\Lambda_1, \dots, \Lambda_s)$ remains constant but $(\Lambda_{s+1}, \dots, \Lambda_m)$ decreases

[14] The reason for why we need it be positive will become clear below.

as follows: Λ_{s+1} decreases by $d_s\epsilon$, Λ_{s+2} decreases by $d_s r_{s+1}\epsilon$, $\ldots$, Λ_{m-1} decreases by $d_s\prod_{k=s+1}^{m-2} r_k\epsilon$, and Λ_m decreases by $\frac{1}{1-r_m} d_s\prod_{k=s+1}^{m-1} r_k\epsilon$. To offset this decrease in $(\Lambda_{s+1},\ldots,\Lambda_m)$, decrease $\pi_{s'}$ by ϵ' so that Λ_{s+1} increases by $d_{s'}\prod_{k=s'+1}^{s} r_k\epsilon'$, Λ_{s+2} increases by $d_{s'}\prod_{k=s'+1}^{s} r_k r_{s+1}\epsilon'$, $\ldots$, Λ_{m-1} increases by $d_{s'}\prod_{k=s'+1}^{s} r_k\prod_{k=s+1}^{m-2} r_k\epsilon'$, and Λ_m increases by $\frac{1}{1-r_m} d_{s'}\prod_{k=s'+1}^{s} r_k\prod_{k=s+1}^{m-1} r_k\epsilon'$. For the net change in $(\Lambda_{s+1},\ldots,\Lambda_m)$ to be positive (element-wise), we need

$$d_{s'}\prod_{k=s'+1}^{s} r_k\epsilon' > d_s\epsilon \iff \epsilon' > \frac{d_s\epsilon}{d_{s'}\prod_{k=s'+1}^{s} r_k}. \tag{EC.15}$$

When we decrease $\pi_{s'}$ by ϵ', $(\Lambda_{s'+1},\ldots,\Lambda_s)$ increases as well (cf. Eq. (EC.12)). Putting everything together, under Eq. (EC.15), the net change in the total market size $\mathbf{1}^\top\mathbf{\Lambda}$ is positive. The change in the amount of experimentation $\mathbf{1}^\top\boldsymbol{\pi}$ equals $\epsilon-\epsilon'$, which is positive if $\epsilon'<\epsilon$. Putting this together with Eq. (EC.15), we get

$$\frac{d_s\epsilon}{d_{s'}\prod_{k=s'+1}^{s} r_k} < \epsilon' < \epsilon. \tag{EC.16}$$

Hence, for ϵ' to exist, we need $\frac{d_s}{d_{s'}\prod_{k=s'+1}^{s} r_k} < 1$, which is equivalent to $d_{s'}\prod_{k=s'+1}^{s} r_k > d_s$. Observe that this holds as long as $\frac{d_{s+1}}{d_s} < r_{s+1}$ for $s\in\{1,\ldots,m-2\}$:

$$d_{s'}\prod_{k=s'+1}^{s} r_k = d_{s'}r_{s'+1}\prod_{k=s'+2}^{s} r_k > d_{s'+1}\prod_{k=s'+2}^{s} r_k = d_{s'+1}r_{s'+2}\prod_{k=s'+3}^{s} r_k > d_{s'+2}\prod_{k=s'+3}^{s} r_k > d_{s-1}r_s > d_s.$$

Hence, under the alternative solution, both $\mathbf{1}^\top\mathbf{\Lambda}$ and $\mathbf{1}^\top\boldsymbol{\pi}$ undergo a positive increase (compared to $\mathbf{1}^\top\mathbf{\Lambda}^*$ and $\mathbf{1}^\top\boldsymbol{\pi}^*$, respectively). To ensure feasibility (recall the constraints in LP), we need to obey $\mathbf{1}^\top\boldsymbol{\pi} = \frac{\mathbf{1}^\top\mathbf{\Lambda}}{\eta}$, $\boldsymbol{\pi}\le\mathbf{\Lambda}$, and $(\boldsymbol{\pi},\mathbf{\Lambda})\ge 0$. This can be done in a manner analogous to the one above. This completes the proof. □

Proof of Proposition 2. Theorem 2 implies CM is optimal if $d_s r_{s+1} > d_{s+1}$ for $s\in\{1,\ldots,m-2\}$. Under the exponential decay parameterization, this condition is equivalent to

$$\theta_0 e^{-\theta_1(s-1)}\left\{1-\gamma_0 e^{-\gamma_1 s}\right\} > \theta_0 e^{-\theta_1 s} \iff s > -\frac{1}{\gamma_1}\log\left(\frac{1-e^{-\theta_1}}{\gamma_0}\right).$$

Since $s\ge 1$, it suffices to have $-\frac{1}{\gamma_1}\log\left(\frac{1-e^{-\theta_1}}{\gamma_0}\right) < 1$, which is equivalent to $1-e^{-\theta_1} > \gamma_0 e^{-\gamma_1}$. □