Incentivized actions in freemium games

Lifei Sheng[*] Christopher Thomas Ryan[†] Mahesh Nagarajan[‡]
Yuan Cheng [§] Chunyang Tong[¶]

July 2, 2020

Abstract

(1) **Problem definition: What is your research problem?**

Games are the fastest-growing sector of the entertainment industry. Freemium games are the fastest-growing segment within games. The concept behind freemium is to attract large pools of players, many of whom will never spend money on the game. When game publishers cannot earn directly from the pockets of consumers, they employ other revenue-generating content, such as advertising. Players can become irritated by revenue-generating content. A recent innovation is to offer incentives for players to interact with such content, such as clicking an ad or watching a video. These are termed incentivized (incented) actions. We study the optimal deployment of incented actions.

(2) **Academic/Practical relevance: How is your research problem relevant to the OM research/practice community?**

Removing or adding incented actions can essentially be done in real-time. Accordingly, the deployment of incented actions is a tactical, operational question for game designers.

(3) **Methodology: What is the underlying research method?**

We model the deployment problem as a Markov decision process (MDP). We study the performance of simple policies, as well as the structure of optimal policies. We use a proprietary data set to calibrate our MDP and derive insights.

(4) **Results: What are your key findings?**

Cannibalization – the degree to which incented actions distract players from making in-app purchases – is the key parameter for determining how to deploy incented actions. If cannibalization is sufficiently high, it is never optimal to offer incented actions. If cannibalization is sufficiently low, it is always optimal to offer. We find sufficient conditions for the optimality of threshold strategies that offer incented actions to low-engagement users and later remove them once a player is sufficiently engaged.

[*]College of Business, University of Houston-Clear Lake, E-mail: sheng@uhcl.edu

[†]Sauder School of Business, University of British Columbia, E-mail: chris.ryan@sauder.ubc.ca

[‡]Sauder School of Business, University of British Columbia, E-mail: mahesh.nagarajan@sauder.ubc.ca

[§]Corresponding author, School of Economics and Management, Tsinghua University, E-mail: chengyn@sem.tsinghua.edu.cn

[¶]School of Economics and Management, Tongji University, Email: tongchunyang@tongji.edu.cn

(5) **Managerial implications: How can academics/managers/decision-makers benefit from your study?**

This research introduces operations management academics to a new class of operational issues in the games industry. Managers in the games industry can gain insights into when incentivized actions can be more or less effective. Game designers can use our MDP model to make data-driven decisions for deploying incented actions.

1 Introduction

Video games are both the largest and fastest-growing segment of the entertainment industry.[1] Mobile games are the largest segment within video games,[2] also representing around 3/4 of total app store revenues in 2018.[3] Freemium games are free to download and play and earn revenue through advertising or selling game enhancements to dedicated players. When accessed in April 2020, Apple Inc.'s App Store showed all 50 of the top revenue-generating games in the United States were free to download.[4] The concept behind freemium is to attract large pools of players, many of whom might never make an in-app purchase. Players that pay out of pocket are said to *monetize*. In general, successful games have a monetization rate of between 2 and 10% percent, with the average much closer to 2%.[5] When game publishers cannot earn directly from the pockets of consumers, they turn to other sources of revenue. This is largely through third parties who pay for delivering advertising content and having players download other apps, fill out surveys, or apply for services, such as credit cards. This stream of revenue is less lucrative per conversion than in-app purchases. Typically, delivering a video earns pennies on the dollar compared to in-app purchases.

However, players can become irritated by advertising, especially when it interrupts the flow or breaks the fiction of a game. A recent innovation is to offer "incentives" for players to click on a banner ad, watch a video, or fill out a survey. These are collectively called *incentivized actions* or, more simply, *incented actions*. Incented advertising, in particular, is different from traditional banner or interstitial advertising. The player is presented with the option to view an advertisement in exchange for a benefit, whereas banner or interstitial advertisements are shown without a choice by the player. In this sense, incentivized advertising is far less "intrusive". In a recent survey, 71% of players preferred taking incented actions to "pay" for mobile games, over traditional advertising, in-app purchases, and premium pricing.[6]

To get a clearer sense of how incented actions work, we examine a concrete example. *Candy Crush Saga* is a puzzle game, published by King. In *Candy Crush Saga*, a player attempts to solve a progression of increasingly challenging puzzles. At the higher levels, players get stuck for extended periods on a single puzzle. Player progression is further hindered by a "lives" mechanic where each failed attempt at a puzzle consumes one of at most five total lives. Lives are regenerated either through waiting long periods of real-time or by purchasing additional lives with real money. Players may also buy items that enhance their chances of completing a puzzle.

Early versions of *Candy Crush Saga* had incented actions, including advertising. A player could take an incented action to earn lives or items without using real money. However, in June of 2013, six months after *Candy Crush Saga* was launched on Apple iOS, King decided to drop all forms of in-game advertising.[7] King's choice was surprising to many observers. What was the logic for removing a potential revenue stream? How did this move affect the monetization rate? This paper explores the logic of when to offer incented actions in mobile games. This can be an "always offer" or "never offer" decision, like in the case of *Candy Crush*. Our model, however, is more nuanced. It explores how incented actions can be used tactically to drive retention and monetization of players throughout their lifetimes. Our model can help answer questions like (i) when, in the lifetime of a player, is it best to offer incented actions? (ii) Should some players be offered incented actions and others not? (iii) What types of games are best suited to offering incented actions?

In this paper, we present an analytical model to explore the use of incented actions and attempt to answer the above questions. Our model emphasizes the connection of incented actions to two other useful concepts often discussed in the game industry – engagement and retention. Highly engaged players are more likely to make in-app purchases and less likely to quit. The longer a player is retained in the game, the more likely they are to engage and monetize. Analytically, player engagement levels are modeled as states in a Markov chain. Retention is the time a player stays in the system before being absorbed into a "quit" state.

Engagement is a common concept for both industry practitioners and academics who study the video game industry. Although all observers emphasize the importance of measuring player engagement, there is little consensus on how to measure it. Among practitioners (see, for instance, Lovell (2012)), a common measurement is the ratio of daily active users (DAU) over monthly active users (MAU). DAU measures the average number of users who play the game at least once per day, and MAU measures the average number of users who play the game at least once per month. This aggregate measure says relatively little about the behavior of an individual player. However, it is easy to calculate for publicly-traded game companies that typically publish their DAU and MAU numbers. Aggregate measures also have the virtue of being universal across many game designs where the mechanics of the game (being a puzzle, sports, or adventure game) may have little in common.

There is also substantial research on engagement in academic video game development journals. A recent study by Abbasi et al. (2017) describes a variety of quantitative and qualitative measurements of engagement including notions of immersion, flow, presence, etc. These are typically measured via questionnaires (such as the game-experience questionnaire (GEQ) proposed in Jennett et al. (2008)). These notions of engagement focus on the psychological and cognitive state of the player and thus presents challenges to operationalize.

For our interests, another view of engagement, which connects to progress made by the player in the game, is more salient. Players can invest significant time and resources to reach higher "levels" of progression. The association of engagement with levels has the benefit of having both a psychological component and a basis in raw game data. Progression suggests an investment of time and achievement of mastery that correlate well with psychological notions of engagement. On the other hand, progression can often be directly measured. For games like *Candy Crush Saga*, there are distinct *levels* that are reached by the player.

Not every game has a level-based design. Other measures of engagement are used in these games. In Section 4, we analyze a proprietary data set for the game *Survival of Primitive* (or *Primitive*, for short), developed by Ebo games and published in 2016. This game has a less linear notion of progression than *Candy Crush*, where "survival" depends on a variety of resource collection activities with many possible paths towards success. In this setting, we adopt Ebo's measure of engagement based on the duration of daily play. Our general engagement model (in Section 3) is general and can incorporate multiple interpretations of the concept of engagement.

Our results. Our results come in three forms. First, we formulate the incented actions deployment model as a Markov decision problem (MDP). This formulation allows game developers to numerically and efficiently compute the optimal provisioning of incented actions. according to engagement levels. We demonstrate how to populate the input of the MDP by studying real data from the *Primitive*. This leads us to a second series of data-driven results on the effectiveness of incented actions in *Primitive*. Currently, the developers of *Primitive* follow an "always offer" strategy. Our results show that this strategy can perform well, or extremely poorly, depending on a key parameter: cannibalization.

Cannibalization is the degree to which incentivized actions distract players from making in-app purchases, as opposed to quitting. Indeed, removing incented actions removes an option for progression. Players can react by either monetizing (to keep the path of progression they are on) or quit. If cannibalization is sufficiently high, it can be optimal to never offer incentivized actions. If cannibalization is sufficiently low, it can be optimal to always offer incentivized actions to ensure players are less likely to quit. In many cases, it is prudent to follow an intermediate strategy of offering incented actions at early levels of engagement and removing them as a player progresses. According to our experiments with the *Primitive* data, "always offer" and "never offer" strategies can be far from optimal for intermediate levels of cannibalization whereas threshold policies are always close to optimal.

In order to learn the degree of cannibalization, one needs to experiment with offering and not offering incented actions. A major policy recommendation of our work is for game developers to invest in understanding the effects of incented actions on monetization and retention of players through experiments with offering them intermittently.

Our third set of results analytically explore the tradeoffs associated with incented actions, and particularly in intermediate cannibalization regimes where our numerical results suggest that threshold strategies are near-optimal.

Analytical investigation of our MDP model for general engagement proves extremely challenging, so we make assumptions and simplifications in order to drive analysis. In particular, we isolate attention to level-based games (like *Candy Crush*) where engagement can be captured by the progression of the player (in levels or stages) and the amount of free play that has been accumulated. Even in this restricted setting, analytical challenges abound, driven by the nonstationarity of the transition data for the underlying Markov chain from engagement level to engagement level.

Accordingly, our analytical results fall into two categories, one where we assume data is stationary and another where parameters are state-dependent. In the former, we show that either "always offer" or "never offer" are optimal, depending on the degree of cannibalization. High cannibalization is associated with "never offer" and low cannibalization is associated with "always offer". This confirms our earlier empirical investigations with the *Primitive* data.

The boundary between what is considered "high" and "low" cannibalization depends on other model parameters. For instance, when the probability for a player to monetize increases relative to her probability to take an incented action (all else being equal) the "never offer" strategy is optimal over a smaller range of cannibalization levels. If the probability for a player to monetize increases relative to her probability to take an incented action, it is less important to remove incented actions to drive monetization, which is happening already at a higher organic rate. Thus, removing incented actions is less attractive in this setting.

One may ask what a game looks like where players organically monetize at a higher rate than taking incented actions. There is a combination of factors, of course, but one possibility is for games where the "rewards" for monetizing greatly exceed the "rewards" for taking an incented action. Consider a game where $1 gets the player a tool that makes passing the level 20% more likely, but an incented action offers a benefit that makes passing the level only 1% more likely. Players may be more likely to monetize than take an incented action for such "underpowered" rewards. Of course, this also depends on the player's time value of money. For a player with a high time value (such as a working adult), the time sink of taking the incented action (such as watching a video ad) can be costlier than for a younger player. Thus, in games whose focus audience are young players (possibly a cartoon-themed puzzle game), our model suggests offering incented actions more liberally is optimal. These considerations show that it is a combination of both the type of game and the nature of the player, that drives the optimal deployment of incented actions.

Our last series of results concern settings where data is nonstationary. In such settings, analysis proves difficult, so we add additional data restrictions. These restrictions allow us to analytically isolate three key effects that drive the deployment of incented actions. First is the *revenue effect*. By offering incented actions, game publishers open up another channel of revenue. However, the net revenue of offering incented may nonetheless be *negative* if one accounts for the opportunity costs of players not making in-app purchases. This captures the possibility that a player would have made an in-app purchase if an incented action was not available. Second, the *retention effect* measures how effective an incented action is at keeping players from quitting. In other words, incented actions can delay a player's decision to quit the game. Third, the *progression effect* refers to the effectiveness of an incented action in deepening the engagement level of the player. It refers to an incented action's ability to increase the player's attachment to the game.

Due to nonstationarity, the key parameter here is the *rate of growth* in cannibalization as engagement level increases. Our main results give sufficient conditions on this rate of growth to establish the optimality of threshold policies where incented actions are offered at early engagement levels and later removed. The rate of growth in cannibalization impacts all three effects, and if sufficiently large, have all three effects favor removing incented actions.

The logic of threshold policies is as follows: by offering incented actions at early stages, the retention effect and progression effect keep the player in for longer by providing a non-monetizing option for progression. However, once a player is sufficiently engaged, the revenue effect becomes less beneficial and the retention effect less significant because highly engaged players are more likely to buy in-app purchases and keep playing the game. This suggests that it is optimal to remove incented actions and attempt to extract revenue directly from the player through monetization. Our sufficient conditions justify this logic, but we also explore settings where this basic intuition breaks down. For instance, the retention effect may remain a dominant concern even at higher engagement levels. Indeed, a highly engaged player may be quite likely to monetize, and so there is a strong desire on the part of the publisher to keep the player in the system for longer by offering incented actions to bolster retention.

The relative strengths of these three effects depend on the characteristics of the game that are reflected in the parameters of our MDP model. For instance, we show analytically that the more able players are at attracting their friends into playing the game, the greater should be the threshold for offering incented actions. This suggests that social games that include player interaction as part of their design should offer incented actions more broadly, particularly when the retention effect is strongly positive since keeping players in the game for longer gives them more opportunities to invite friends. Indeed, a common incented action is to contact friends in your social network or to build a social network to earn in-game rewards.

In summary, we provide the following tools and insights for game developers in assessing deployment of incented actions.

- We provide a general model for deciding the optimal deployment of incented actions as a function of engagement. This model can be instantiated with existing game data and solved numerically. We illustrate this procedure in the context of real data.
- Based on findings in real game data, we illustrate the importance of understanding how incented actions cannibalize in-app purchases. The most common incented action practices ("always offer" and "never offer") can be far from optimal under certain cannibalization regimes.
- Based on findings from real game data, policies that offer incented actions until some engagement threshold is met are near-optimal for a wide range of cannibalization behavior.
- Based on the previous two insights, we highly recommend game developers to experiment with offering and not offering incented actions to assess cannibalization behavior. This can

greatly assist in designing near-optimal threshold strategies or determine which of the simple "always offer" or "never offer" strategies are close to optimal.

- Analytical results confirm the importance of cannibalization, but also highlight thresholds for when cannibalization changes the optimal structure. This allows us to derive a several managerial insights, including that (i) games with players with a high monetary value for time (say among mature players) should offer fewer incented actions, while (ii) social games should offer more incented actions.

2 Related literature

As freemium business models have grown in prominence, so has interest in studying various aspects of freemium in the management literature. Papers in the marketing literature have mainly been empirical (see, for instance, Lee et al. (2017)), our work connects most directly to a stream of analytical studies in the information systems literature that explores how "free" is used in the software industry. Two important papers for our context are Niculescu and Wu (2014) and Cheng et al. (2015) that, together, establish a taxonomy of different freemium strategies and examine in what situations a given strategy is most advantageous.

Our work departs from this established literature in at least two dimensions. First, we focus on how to tactically implement a freemium strategy, in particular, when and how to offer incented actions to drive player retention and monetization. It does not so much argue the merits of freemium versus other models but describes how to implement freemium.

Second, games present a specific context that may be at odds with some common conceptualizations of a freemium software product. For a productivity-focused product, such as a PDF editor, a typical implementation of freemium is to put certain advanced features behind a pay-wall, such as the ability to make handwritten edits on files using a stylus. Once purchased, features are typically unlocked either in perpetuity or for a fixed duration by the paying player. Virtual items in games enhance the in-game experience, speed progression, or provide some competitive advantage. These virtual items are often *consumables*, meaning that they are depleted through use. This is true, for instance, of all purchases in *Candy Crush Saga*. Our model allows for a player to make repeated purchases and the degree of intensity of monetization to evolve.

Other researchers have examined the specific context offered by games, as opposed to general software products, and have adapted specialized theory to this specific context. Guo et al. (2019a) examine how the sale of virtual currencies in digital games can create a win-win scenario for players and publishers from a social welfare perspective. They make a strong case for the value created by games offering virtual currency systems. Our work adds a layer by examining how virtual currencies (and other in-game rewards) can be used to incentivize players to take actions that are profitable to the firm that does not involve a real-money exchange. Also, Guo et al. (2019a) develop a static model where players decide on how to allocate a budget between play and purchasing virtual currency. We relate a player's willingness to take incented actions or monetize as their engagement with the game evolves, necessitating the use of a dynamic model. This allows us to explore how a freemium design can respond to the actions of players over time.

Another recent paper by Huang et al. (2019) operationalizes the notion of engagement to drive decisions in game design. This work attempts to quantify engagement empirically using a hidden Markov model in order to improve in-game player matching approaches. In this paper, we connect the concept of engagement to the effectiveness of incented actions. Also using industry data, we use the company's definition of engagement to prescribe an optimal policy for incented action deployment.

The specific question of incentivized advertising in games has also been explored in the literature (concurrent with the writing of our paper). The effectiveness of incentivized ads is explored empirically in Chiong et al. (2017) and Lee and Shin (2017). The tradeoffs associated with incented ads (or reward ads, as they are sometimes known) were studied analytically in Guo et al. (2019b) in a static model. They examine conditions for when to offer incented ads along with in-app purchases to optimize revenues. Their model does not consider how players evolve and does not take up the question of the "quitting" behavior of players that is prevalent in many mobile games. The time dimension and possibility of quitting are distinguishing features of our study.

Finally, our modeling approach of using a discrete-time Markov decision process model in search of threshold policies is a standard-bearer of analysis in the operations management literature. Threshold policies have the benefit of being easily implementable and thus draw favor in studies of tactical decision-making. In our many conversations with some of the largest game developers in the world, this sentiment is very evident. The simplest type of threshold policies allows the system designer to keep track of nothing but the threshold (target) level and monitor the state of the system and take the appropriate action to reap the benefits of optimality. This is in contrast to situations where the optimal policy can be complex and have nontrivial state and parameter dependencies. Examples of effective use of this approach in dynamic settings include inventory and capacity management and control (Zipkin, 2000) and revenue management (Talluri and Van Ryzin, 2006).

3 General engagement model

We take the perspective of a game publisher who is deciding how to deploy incented actions optimally. Incented actions can be offered (or not) at different times during a player's experience with the game. For example, a novice player may be able to watch video ads for rewards during the first few hours of gameplay, only later to have this option removed.

Our model has two agents: the game publisher and a single player. This assumes that the game publisher can offer a customized policy to each player, or at least customized policies to different classes of players. In other words, the "player" in our model can be seen as the representative of a class of players who behave similarly. The publisher may need to decide on several different policies for different classes of players for an overall optimal design. An important challenge here is for the game designer to calibrate the data in the model given observations from a given class of players (this is discussed in our application to *Primitive* in Remark 3 below).

We assume that the player behaves stochastically according to the options presented to her by the game publisher. The hope is that this model allows for many interpretations of the meaning of "engagement", several of which were discussed in the introduction. In this section, we do not model down to the specifics of a particular game and instead provide what we feel is a robust approach to engagement. In Section 5, we assume more structure on the problem to provide a concrete interpretation of engagement.

The game publisher's decision can be modeled as a Markov Decision Problem (MDP), where the stochasticity is a function of the underlying player model, and the publisher's decision is whether or not to offer incented actions. We describe the player model in detail in the next subsection. The publisher's problem is detailed in Section 3.2.

3.1 Player model. We model player behavior as a Markov chain. We describe the components of the Markov chain (time epochs, states, and transitions) as follows.

A time epoch is the elapsed time between points when the player randomly selects from revenue-generating actions (described in more detail below) or quitting. We may think of these "decision" points where the player selects from these options as points when the player is fed up with her rate of progress in the game and are no longer interested in progressing organically via free play. That is, "free play" is considered a default action of the player at all other points in time.

The decision epochs where actions are undertaken occur when the player is assessing whether or not they want to continue playing the game. The real elapsed time between decision epochs is not constant since it depends on the behavior of the player between sessions of play. Some players frequently play, others play only for a few minutes per day. A player might be highly engaged but have little time to play due to other life obligations. This reality suggests that the elapsed time between decision epochs should not be a critical factor in our model.

The state of the Markov chain is the engagement level of the player. We model engagement abstractly as simply an ordered set of states. We restrict the set $E = \left\{-1, 0, 1, \ldots, \bar{E}\right\}$ of finitely many engagement levels where -1 denotes an absorbing "quit" state where the player no longer plays the game. We denote the engagement level at time epoch t by e_t.

Transitions between states in the Markov chain correspond to the player randomly selecting among three different actions at each time epoch. The first action is to *monetize* (denoted M) by making an in-app purchase with real money. The second is for the player to take an *incented action* (denoted I). Third, the player can quit (denoted Q). The set of available actions is determined by whether the publisher offers an incented action or not. We let $A_1 = \{M, I, Q\}$ denote the set of available actions when an incented action is offered and $A_0 = \{M, Q\}$ otherwise. We denote the action at time epoch t by a_t.

The probability that the player takes a particular action depends on her engagement level and what actions are available to her. We used the letter "p" to denote probabilities when an incented action is available and write $p_a(e)$ to denote the probability of taking action $a \in A_1$ at engagement level $e \in E$. We use the letter "q" to denote action probabilities when the incented action is unavailable and write $q_a(e)$ for the probability of taking action $a \in A_0$ at engagement level $e \in E$. By definition $p_M(e) + p_I(e) + p_Q(e) = 1$ and $q_M(e) + q_Q(e) = 1$ for all $e \in E$.

Remark 1. It is important to stress the implications of the Markovian assumption for our state space. The probability of monetizing or taking incented actions depends only on the level of engagement and not whether a player has previously taken an incented action or monetized. $\triangleleft$

There is a relationship between $p_a(e)$ and $q_a(e)$. When an incentivized action is not available, the probability $p_I(e)$ is allocated to the remaining two actions M and Q. For each $e \in E$, we assume that there exists a parameter $\alpha(e) \in [0, 1]$ such that

$$q_M(e) = p_M(e) + \alpha(e)p_I(e) \tag{1}$$

$$q_Q(e) = p_Q(e) + (1 - \alpha(e))p_I(e). \tag{2}$$

We call $\alpha(e)$ the *cannibalization parameter at level* e since $\alpha(e)$ measures the impact of removing an incented action on the probability of monetizing and thus captures the degree to which incented actions cannibalize demand for in-app purchases. A large $\alpha(e)$ (close to 1) implies strong cannibalization, whereas a small $\alpha(e)$ (close to 0) signifies weak cannibalization.

To turn random action choices into transition probabilities between states, it remains to consider how a player transitions from one engagement level to another after taking an action. In principle, we would need to determine individually each transition probability $\mathbb{P}(e_{t+1} = e'|e_t = e \text{ and } a_t = a)$. For actions $a \in \{M, I\}$, we will assume that transition probabilities are stationary and set $\mathbb{P}(e_{t+1} = e'|e_t = e \text{ and } a_t = a) = \tau_a(e'|e)$ for all times t, where τ is a $[0, 1]$-valued function such that $\sum_{e' \in E} \tau_a(e'|e) = 1$ for all $e \in E$ and $a \in \{M, I\}$. For the quit action, $\mathbb{P}(e_{t+1} = -1|e_t = e' \text{ and } a_t = Q) = 1$ for all times t and engagement levels e'. In other words, there are no "failed attempts" at quitting.

It is important to stress that a player can increase her level of engagement at any time between decision points. For example, a transition from e to e' after a monetization action can occur immediately after monetizing, or after some duration of free play. What is essential is that this change in engagement in the pursuing time epoch is impacted by the choice of action a (monetization or incented action) through the transition probabilities $\tau_a(e'|e)$. In particular, a change in engagement can occur due to free play, but the chance of this occurrence is modulated by the revenue-generating actions M and I. We make some of these engagement mechanics between decision epochs explicit when we study level-based games in Section 5.

If incented actions are available, the transition probability from engagement level e to engagement level e' is

$$\mathbb{P}_1(e'|e) := \begin{cases} p_M(e)\tau_M(e'|e) + p_I(e)\tau_I(e'|e) & \text{if } e, e' \in E \\ p_Q(e) & \text{if } e \in E, e = -1 \\ 1 & \text{if } e, e' = -1 \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

and if incented actions are not available

$$\mathbb{P}_1(e'|e) := \begin{cases} q_M(e)\tau_M(e'|e) & \text{if } e, e' \in E \\ q_Q(e) & \text{if } e \in E, e = -1 \\ 1 & \text{if } e, e' = -1 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

To help in analyzing the problem, we make the following assumption.

Assumption 1. No matter how engaged, there is always a positive probability that a player will quit; i.e., $p_Q(e), q_Q(e) > 0$ for all $e \in E$.

This acknowledges the fact that games are entertainment activities, and there are numerous reasons for a player to quit, even when engrossed in the game. This is an important technical assumption since it implies the publisher's problem is an absorbing Markov decision process.

3.2 The publisher's problem. We model the publisher's problem as an infinite horizon Markov decision process under a total reward criterion (for details, see Puterman (1994)). In our setting, the set of states is $\{-1\} \cup E$ and the set of controls $U = \{0, 1\}$ is independent of the state, where 1 represents offering an incented action and 0 not offering an incented action. The transition probabilities are given by (3) when $u = 1$ and (4) when $u = 0$. The reward depends on the action of the player. When the player quits, the publisher earns no revenue. When the player takes an incented action, the publisher earns μ_I, while a monetization action earns μ_M.

The expected reward in state e under control u is:

$$r(e, u) = \begin{cases} p_M(e)\mu_M + p_I(e)\mu_I & \text{if } e \in E \text{ and } u = 1 \\ q_M(e)\mu_M & \text{if } e \in E \text{ and } u = 0 \\ 0 & \text{if } e = -1. \end{cases}$$

Note that expected rewards do not depend on whether the player transitions to a higher engagement level and so the probabilities $\tau_a(e'|e)$ do not appear in $r(e, u)$.

A *policy* y is a mapping from E to U. On occasion, we will express a policy by the vector form of its image. That is, the vector $y = (1, 0, 1)$ denotes offering incented actions in engagement levels 1 and 3. Each policy y induces a stochastic process over rewards, yielding expected total reward:

$$W^y(e) := \mathbb{E}_e^y \left[\sum_{t=1}^{\infty} r(e_t, y(e_t)) \right] \tag{5}$$

where e is the player's initial level, and the expectation $\mathbb{E}_e^y[\cdot]$ derives from the induced stochastic process.

Remark 2. Rewards in (5) are not discounted. There are three reasons for this. Firstly, "time" is a tricky concept in our setting. There is the time playing the game and the time "outside" of the game experienced by the player. Should we discount rewards according to the elapsed time of playing the game or the "clock" time of the player? Second, our definition of time epochs includes varying amounts of "free" play, and the elapsed in-game or "clock" time are not captured. In Section 5, we account for the amount of "free play" more directly, but still do not fully capture "clock" time in any meaningful sense. Finally, time discounting is often assumed for analytical tractability. Total rewards can be unbounded where discounted rewards over an infinite horizon are summable. The absorbing nature of our MDP does not require a discounting assumption for analytical tractability. ◁

The game publisher chooses a policy to solve the following optimization problem. Given a starting engagement level e, solve:

$$\max_{y \in \{0,1\}^E} W^y(e). \tag{6}$$

This problem can be solved numerically using tools such as policy iteration (see, for instance, Puterman (1994)).

We believe numerical solutions to (6) are of interest to companies, as they allow them to decide in practice how they might deploy incented actions. The challenge, of course, in solving Section 3.2 is fitting data to the model. This is taken up in the next section. Later, we seek to further operationalize some aspects of our model and prove some analytical results.

4 Application: *Survival of the Primitive*

We apply our Markov decision process to the mobile game *Survival of Primitive* developed by Tianjin Ebo Digital Technology Co., Ltd. We now describe the data we received from Ebo and carefully describe how it fits our model.

4.1 Data. Our *Primitive* data set consists of the daily game behavior of 5000 randomly selected players starting from the release date of the game in July 2016 through 31 March 2017. For each of these 5000 players, we have daily entries for play time duration, the number of incented actions taken, and the number of in-app purchases made. The data used to build the model has a total of 1,529,577 player-day observations.

We use this data to define engagement levels and estimate the parameters of the publisher's problem in (6). We consulted with Ebo on how they defined engagement. Based on these conversations we tied engagement level to how "frequently" the player plays the game, as captured by daily gameplay durations. We computed the (moving) average of play durations trailing three days. Engagement levels consist of discrete buckets of play duration averages. For our main data analysis, we chose duration buckets of 0 to 30 minutes, 30 minutes to 60 minutes, 60 to 90 minutes, and 90 to 120 minutes. That is, if the average time the player spent is less than 30 minutes, her engagement level is 0; between 30 minutes and 1 hour, her engagement level is 1, etc. No player had a moving average duration greater than two hours, so there were four engagement levels overall.

After defining our notion of engagement, we implemented our notion of decision epochs. Recall, that we demarcate decision epochs by revenue-generating actions and quitting. There were naturally many days that a given player only engaged in free play. According to our model, this day should not be considered a decision epoch for that player. When we removed player-day observation that did not correspond to decision epochs in our model we were left with 123,117 observations. Over the 5000 players, an average of 4.7% of the 106 playing days had players monetizing, as compared to 17.4% of days had players taking incented actions.

We now describe how we used the data to compute $p_M(e), p_I(e), \tau_M(e'|e)$, and $\tau_I(e'|e)$. We identified the quitting decision as follows: if the last log-in of a player was greater than seven days before the end of the data collection period (31 March 2017), then we assume that the player took the quit action on the seventh data after her last login.

Given an engagement level e, let $TN(e)$ denote the total number of observations where the player was in engagement level e, let $MN(e)$ denote the number of these observations where a player monetized, and let $IN(e)$ denote the number of these observations where a player took an incentivized action. Then our estimate of $p_M(e)$ is the fraction $MN(e)/TN(e)$ and our estimate of $p_I(e)$ is the fraction $IN(e)/TN(e)$. The quit probability estimate is $p_Q(e) = 1 - p_M(e) - p_I(e)$. Table 1 captures the values we estimated from data.

Moving average play time (min)	Engagement Level (e)	p_I	p_M	p_Q	μ_M	μ_I
$[0, 5]$	1	0.7711	0.2176	0.0113	1	0.05
$(5, 30]$	2	0.7552	0.2143	0.0305		
$(30, 60]$	3	0.7576	0.2112	0.0312		
$(60, 90]$	4	0.7621	0.2064	0.0315		
$(90, 120]$	5	0.7534	0.2141	0.0325		

Table 1: Probabilities estimated using data from *Survival of Primitive*.

We assume that given an action a, the probability of changing engagement does not depend on the current engagement level (other than restricting that engagement remain in $\{0, 1, \ldots, \bar{E}-1, \bar{E}\}$). To compute these conditional transition probabilities, we first calculate the total number $MN = \sum_e MN(e)$ of player-days where a monetization action occurs and the total number $IN = \sum_e IN(e)$ of player days where an incented action is taken. Then we calculate the total number MN_{+i} of player-days where engagement increased by i (for $i = 0, 1, \ldots, \bar{E}$) engagement levels following a monetization action and the total number of MN_{-i} of player-days where monetization was followed by a decrease in engagement by i (for $i = 1, \ldots, \bar{E}$) engagement levels. The values IN_{+i} and IN_{-i} are defined analogously for incented actions. The resulting conditional transition probabilities, which are calculated by taking the appropriate ratio MN_{+i}/MN are captured in Table 2.

Remark 3. We estimated the parameters of our model using information from all 5000 players as a

	Changes in engagement level								
Action	-4	-3	-2	-1	0	1	2	3	4
$a = M$	0	0.0028	0.0123	0.1610	0.6303	0.1620	0.0203	0.0106	0.0006
$a = I$	0	0.0032	0.0145	0.1973	0.5496	0.1984	0.0231	0.0131	0.0008

Table 2: Conditional Transition Probabilities. Columns correspond to changes in engagement level: -4 means engagement level goes down by 4 from the current engagement level and 4 means engagement level goes up by 4 levels, etc.

single group. Of course, we could have divided up the data based on other covariates to potentially get more accurate transition probabilities for subgroups of players. We considered several player classifications, including by region (grouping into North America, Asia, Europe, and other). The resulting estimates of our transition probabilities largely resembled those in Tables 1 and 2. Our understanding, from discussion with Ebo, that this is also the company's common practice. $\triangleleft$

Ebo also did not provide information on the value of a monetization action μ_M or an incented action μ_I. Typically, these parameters depend on a variety of factors, including the nature of the action (for instance the length of a video ad) and the nature of contracts with third-party providers. We normalize $\mu_M = 1$ and set $\mu_I = 0.05$ (as reflected in Table 1). This choice of μ_I is reasonable, the value of an incented action is typically cents on the dollar when compared to monetization. Other choices for μ_I were considered, with the same qualitative findings as below.

Ideally, we would also like to provide an estimate of the cannibalization parameter α. However, incented actions were always available to players during the data collection period. Given the unavailability of data for α, we will often parameterize our results regarding the parameter α.

4.2 Policy comparisons. In this section, we numerically examine the performance of various policies for solving the publisher's problem (6) using the parameter values estimated in Tables 1 and 2. We examine four policies:

(A) "Always offer": always offer incented actions; that is, $y(e) = 1$ for all $e \in E$.

(N) "Never offer": never offer incented actions; that is, $y(e) = 0$ for all $e \in E$.

(T) "Optimal threshold": the policy that results from finding the optimal engagement level e^* such that $y(e^*) = 1$ for $e \leq e^*$ and $y(e) = 0$ otherwise.

(*) "Optimal": the policy that results from optimally solving the MDP in Section 3.2 using policy iteration.

To compare all four policies we assume constant cannibalization; that is, there exists a constant α such that $\alpha(e) = \alpha$ for all $e \in E$. Figure 1 compares the average value $\mathbb{E}(W) = \frac{\sum_e W(e)}{\bar{E}}$ under each of the four policies versus α ranging from 0 to 1. The optimal threshold policy tracks very closely to the optimal policy, dipping below optimality only for intermediate values of α. This suggests that optimal policies should be quite effective in practice, an idea we explore in more detail later in this section (and throughout the remainder of the paper).

On the other hand, "always offer" and "never offer" can perform quite poorly, depending on

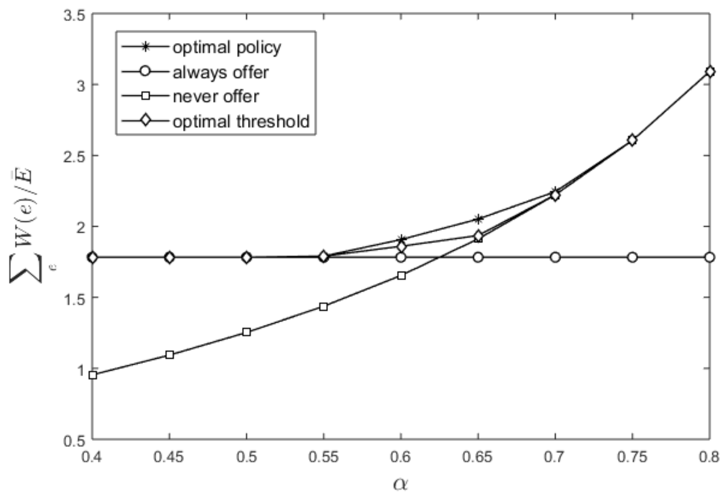


Figure 1: Comparison of four policies when α increases.

the parameter α. When cannibalization is high (large values of α), (A) performs quite poorly. Conversely, when cannibalization is low (small values of α) (N) performs poorly, players are driven to quit in high proportions due to a lack of incented actions to help them progress. Notice that (A) is optimal for low cannibalization regimes and (N) is optimal for high cannibalization regimes.

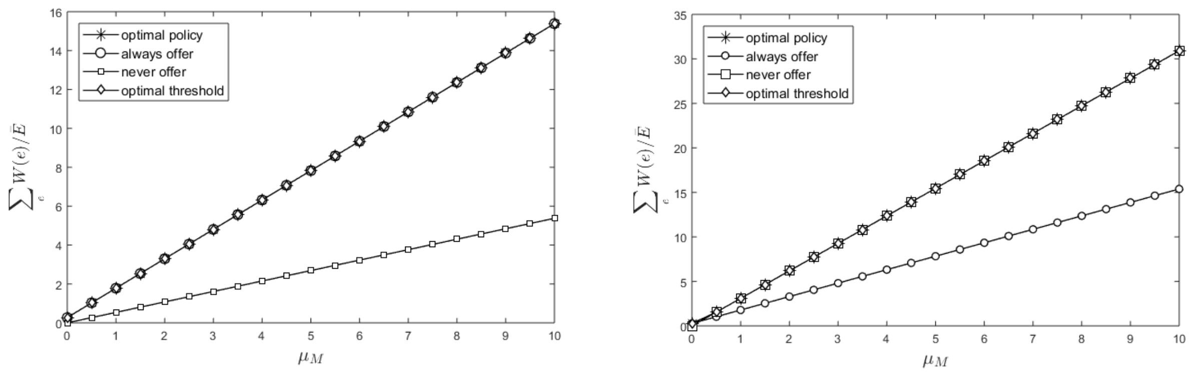
Currently, Ebo follows an "always offer" strategy. According to our calculations, this strategy is likely to be suboptimal if α is sufficiently large. Below we calculate the worst-case loss (over the choice of α) for taking a suboptimal policy. The worst-case losses for the three strategies are

$$\sup_{\alpha\in[0,1]} \frac{\mathbb{E}(W^*)-\mathbb{E}(W^A)}{\mathbb{E}(W^*)} = 77.12\%$$
$$\sup_{\alpha\in[0,1]} \frac{\mathbb{E}(W^*)-\mathbb{E}(W^N)}{\mathbb{E}(W^*)} = 85.37\%$$
$$\sup_{\alpha\in[0,1]} \frac{\mathbb{E}(W^*)-\mathbb{E}(W^T)}{\mathbb{E}(W^*)} = 6.00\%.$$

This result starkly underscores the importance for the company of experimenting with not offering incented actions in order to estimate α. The worst-case optimality gaps of (A) and (N) worsen as μ_M gets larger (see Figure 2 for an illustration). However, the optimal threshold policy performs quite well; its worst-case optimality gap over α is 6%. If we allow μ_M to be any value below 10 dollars, the (worst-case) optimality gap will increase, but within control:

$$\sup_{\alpha\in[0,1],\mu_M\in[0,10]} \frac{\mathbb{E}(W^*)-\mathbb{E}(W^T)}{\mathbb{E}(W^*)} = 6.42\%.$$

These results suggest that restricting attention to optimal threshold policies is a robust and near-optimal strategy. We examine this takeaway further by examining more complex functional forms of the cannibalization parameter $\alpha(e)$ (other than constant). We consider three different constructions of the cannibalization parameter $\alpha(e)$ as a function of e: (i) linear, $\alpha(e) = \alpha_{\text{initial}} + \alpha_{\text{step}}(e-1)$; (ii) quadratic, $\alpha(e) = t_1e^2 + t_2e$; and (iii) square-root, $\alpha(e) = \sqrt{ke}$, where α_{initial},α_{step}, t_1, t_2, and



(a) $\alpha = 0.2$

(b) $\alpha = 0.8$

Figure 2: Comparison of four policies when μ_M increases (with $\mu_I = 0.05$ fixed).

k are parameters that we enumerate in our numerical experiments.

Figure 3 shows the optimal policy for a linear α as we change α_{initial} and α_{step}. Threshold policies are optimal in a large regime in the linear case. The gray area indicates the region where the optimal policy is not a threshold.

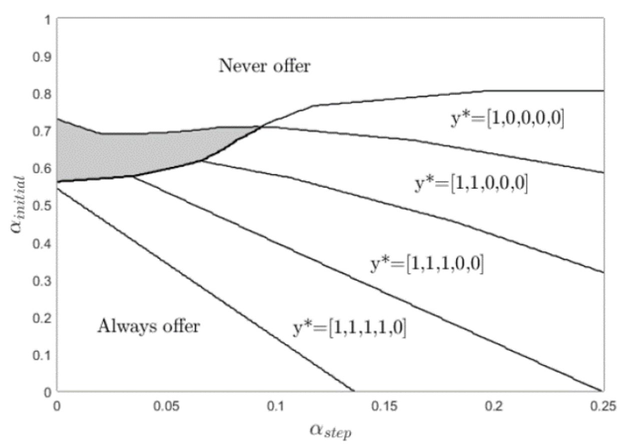
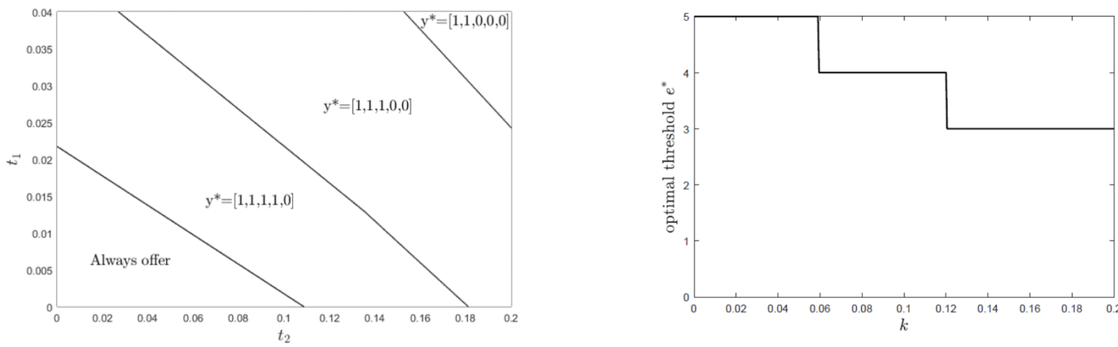


Figure 3: Threshold policies for linear α when changing α_{initial} and α_{step}. (The gray area indicates where no threshold policy is optimal)

Figure 4 shows the optimal policy for a quadratic α when we change the first- and second-order coefficients t_1 and t_2; and for a square root α when we change k. We observe that the threshold policy is always optimal in the quadratic and square-root cases. This underscores the fact that the optimality of threshold policies is not driven by the underlying linearity of the data. We did similar calculations for defining engagement using 5-day and 7-day moving averages and the results were similar (for brevity, we only show the figures for the 3-day moving average case).

4.2.1 Data-driven sensitivity analysis. Next, we conduct sensitivity analyses to the revenue parameters μ_I and μ_M that were estimated above. What matters is the ratio μ_I/μ_M since one of the parameters can always be normalized. Figure 5 illustrates a typical result for two different



(a) quadratic, $\alpha(e) = t_1 e^2 + t_2 e$.

(b) square-root, $\alpha(e) = \sqrt{ke}$.

Figure 4: Robustness check for threshold policy with quadratic α and square-root α.

choices of α. Observe that the threshold increases in the ratio μ_I/μ_M. This result is intuitive. As incented actions yield more revenue (relative to monetization), it becomes more attractive to offer them more extensively. Conversely, if monetization yields more revenue (relative to incented actions) then incented actions should be used less often. This provides clear guidance for Ebo. In later sections, we reproduce these results analytically under some technical conditions.

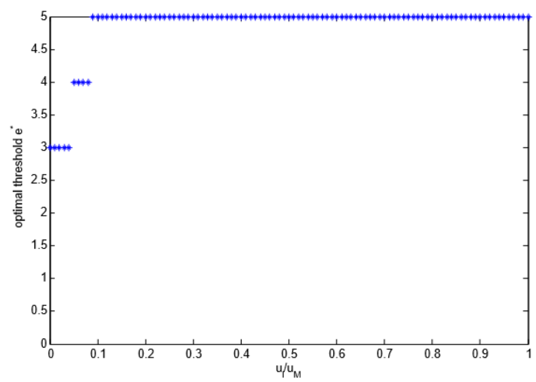


Figure 5: The effect of μ_I/μ_M on the optimal threshold for $\alpha = [0.127, 0.632, 0.815, 0.906, 0.913]$.

5 Model for level-based games

Our findings from the *Primitive* dataset suggest that threshold policies may often be optimal. However, an attempt to directly analyze the model presented in Section 3 to establish the optimality of threshold policies quickly proves hopeless. Part of the challenge is the lack of structure in how the model treats engagement levels. Our first step is to operationalize engagement in more concrete settings. In this section, we aim to provide analytical proofs of the optimality of threshold policies under certain technical and interpretable conditions.

The concrete setting we discuss here is inspired by the puzzle genre of games, typified by *Candy Crush*. These are level-based games where the player attempts a sequence of puzzles or discrete stages. The mechanics of level-based games allow us to operationalize the notion of engagement level. Roughly speaking, the engagement level of the player now consists of the tuple (ℓ, f) where ℓ is the current level the player is attempting and f is the total number of free attempts made at the current level. This becomes the state of a new Markov chain describing player behavior. The set L of levels is a finite set $\{-1, 0, 1, 2, \ldots, \bar{L}\}$ where $\bar{L}$ represents the highest level of the game. Again, level -1 captures a quit state. If a player reaches the highest level, it does not mean that the player quits the game. We will model players at the highest level to be in some holding pattern where there is some probability they will quit, but will otherwise stay connected to the game. Although we do not model it here, additional levels are released over time by game publishers, and players at the highest levels can wait to play these newly released levels.

The second state is the cumulative number f of "free plays" taken at a given level. Recall that the time epochs record points in time where the player selects one of M, I, and Q. The state f counts the number of free plays made between these time epochs at the given level. The set F of free play counts is also a discrete set denoted $\{0, 1, 2, \ldots, \bar{F}\}$ where $\bar{F}$ is the maximum effective free play accumulation. Once $\bar{F}$ free plays have been accumulated, further free plays do not impact state transitions.

The probability that the player takes a particular action now depends on the state (ℓ, f) and on what actions are available to her. Again, we use the letter "p" to denote probabilities when an incented action is available and write $p_a(\ell, f)$ to denote the probability of taking action $a \in A_1$ at state (ℓ, f). Again, we use the letter "q" to denote action probabilities when the incented action is unavailable and write $q_a(\ell, f)$ for the probability of taking action $a \in A_0$ at state (ℓ, f). By definition, $p_M(e, f) + p_I(e, f) + p_Q(e, f) = 1$ and $q_M(e, f) + q_Q(e, f) = 1$ for all (e, f). Finally, we can define the cannibalization parameter as follows:

$$\begin{aligned} q_M(e, f) &= p_M(e, f) + \alpha(e, f) p_I(e, f) \\ q_Q(e, f) &= p_Q(e, f) + (1 - \alpha(e, f)) p_I(e, f). \end{aligned}$$

The action-specific state transition probabilities are denoted $\tau_a(\ell', f' | \ell, f)$.

The rest of the setup of the model is precisely the same as before. We get aggregate transition probabilities from state to state, depending on whether incented actions are available or not. If incented actions are available, the transition probabilities are

$$\mathbb{P}_1(\ell', f' | \ell, f) := \begin{cases} p_M(\ell, f) \tau_M(\ell', f' | \ell, f) + p_I(\ell, f) \tau_I(\ell', f' | \ell, f) & \text{if } \ell' \in \{\ell, \ell + 1\}, \ell' \neq -1 \\ p_Q(\ell, f) & \text{if } \ell = -1 \\ 1 & \text{if } \ell, \ell' = -1 \\ 0 & \text{otherwise,} \end{cases}$$

and if incented actions are not available,

$$\mathbb{P}_0(\ell', f'|\ell, f) := \begin{cases} q_M(\ell, f)\tau_M(\ell', f'|\ell, f) & \text{if } \ell' \in \{\ell, \ell+1\}, \ell' \neq -1 \\ q_Q(e, f) & \text{if } \ell' = -1 \\ 1 & \text{if } \ell, \ell' = -1 \\ 0 & \text{otherwise.} \end{cases}$$

Assumption 2. There is always a positive probability that a player will quit; i.e., $p_Q(\ell, f), q_Q(\ell, f) > 0$ for all $\ell \in L$ and $f \in F$.

Assumption 3. The benefit (to the game designer) of monetizing is greater than the benefit of incented actions: $\mu_M > \mu_I$.

Assumption 2 again yields an absorbing Markov chain with a total reward criterion, where the final optimization problem is:

$$\max_{y \in \{0,1\}^{L \times F}} W^y(\ell, f)$$

where $y(\ell, f) = 1$ means offer incented actions in state (ℓ, f) and $y(\ell, f) = 0$ means do not offer.

In its full complexity, this model proves no easier to analyze than the model introduced in Section 3. Indeed, the problem now has a two-dimensional state. The goal in this section is not to tackle this model in its full complexity, but to add sufficient structure to derive analytical insights. This work is carried out over the next two subsections.

A first step in structuring this model is to constrain levels to go up one at a time; that is,

Assumption 4. For any f, f', $\tau_a(\ell', f'|\ell, f) = 0$ if $\ell' \geq \ell + 2$.

This assumption enforces the condition that a player *cannot* advance more than one level at a time on free play alone. That is, every level of the game the player has a decision point where she either monetizes, takes an incented action, or quits. The general model in Section 3 had no such restriction, engagement can (and did in the case of *Primitive*) jump around significantly. The requirement here is stricter, but we believe justifiable in practice.

Consider the case of *Candy Crush Saga*. As of January 2020, there were 6,230 distinct stages across 416 episodes of ten to fifteen stages each. Our notion of "level" here need not correspond precisely to each stage or even episodes of stages. Our notion of "level" can correspond to a sufficiently large collection of stages and episodes where we can be quite confident that the average player either monetizes, takes an incented action, or quits. Indeed, freemium games often have "pinch points" in their design to challenge players to start taking revenue-generating actions (Luton, 2013). In this interpretation, our level can represent the number of "pinch points" passed by the player along a linear progression of "pinch points". We believe this to be typical of many level-based freemium designs.

In this setting, we conduct numerical experiments on synthetic data to test the efficacy of threshold policies. Of course, we need to be clear what we mean by threshold policies in this setting since the state is now two-dimensional. We study *rectangular-threshold* policies where there

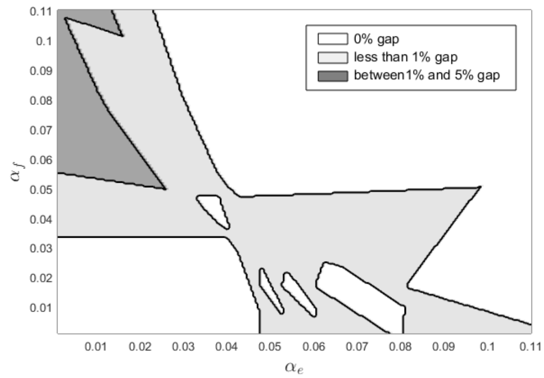


Figure 6: Revenue gap with optimal threshold policy versus optimal policy. The data specification is follows: $\mu_M = 1$, $\mu_I = 0.05$, $p_M(\ell, f) = 0.02 + 0.03\ell + 0.01f$, $p_I(\ell, f) = 0.4 - 0.02\ell + 0.04f$, $\tau_M(\ell+1, f'|\ell, f) = 0.8[0.05 + 0.04(f - f')]$, $\tau_M(\ell, f'|\ell, f) = 0.2[0.1 + 0.2(f - f')]$, $\tau_I(\ell+1, f'|\ell, f) = 0.3[0.05 + 0.04(f - f')]$, $\tau_O(\ell, f'|l, f) = 0.7[0.1 + 0.2(f - f')]$

exists a threshold (ℓ^*, f^*) such that $y(\ell, f) = 1$ if $\ell \leq \ell^*$ and $f \leq f^*$ and $y(\ell, f) = 0$ otherwise.

Our investigations suggest that threshold policies are still often optimal, or close to optimal (see Figure 6). In a wide variety of parameter specifications, we find that threshold policies do no worse than 5% loss from optimal, and very often, are optimal. This is analogous to the results we uncovered in Section 4.2 for *Primitive*.

Despite these positive results, finding conditions for the optimality of threshold policies for the general level-based setting proves very difficult. One of the major challenges is handling the nonstationarity of the transition probabilities as engagement changes. Only when we assume that some of the nonstationarity is "controlled" in some way, can we establish optimal threshold policy results under various restrictions.

5.1 Special case: Stationary transition probabilities. A natural first direction is to assume that the data of the problem is stationary and attempt to establish the optimality of threshold policies. This is achieved in the following result.

Theorem 1. Suppose the parameters are stationary, i.e. p_M, p_I, α are constant. Moreover, suppose the transition probabilities $\tau_a(\ell', f'|\ell, f)$ is stationary for $\ell' \geq \ell$ and $f' \geq f$ (where at least one inequality is strict) and $a = M, I$ (that is, when engagement level (strictly) goes up). Then, there exists an α^* such that "always offer" is optimal when $\alpha < \alpha^*$ and "never offer" is optimal when $\alpha \geq \alpha^*$, where specifically

$$\alpha^* = \tfrac{p_M \mu_M + (1-p_M)\mu_I}{p_I \mu_I + (1-p_I)\mu_M}. \tag{7}$$

In particular, there is never an optimal solution where incented actions are offered at some engagement levels and not at others.

The result that the optimal policy is either "always offer" or "never offer" is intuitive because of the stationarity of the data, a policy choice that is good at one engagement level should be just as good at another. Of more interest is discovering a clean boundary in α for when to switch between "always offer" and "never offer". High cannibalization (large α) is associated with "never offer" and low cannibalization (small α) is associated with "always offer". The intuition here is simple (and discussed earlier in Section 4), when α is large then there is little additional risk of players quitting when incented actions are removed and they are more likely to monetize, so it makes sense to not offer incented actions.

The boundary value α^* (given in (7)) that distinguishes "always offer" and "never offer" also has an interesting structure. For instance, if μ_I increases relative to μ_M then one can show that α^* increases. That is, the range of cannibalizations for "never offer" narrows because incented actions are relatively more valuable (cf. Proposition 3 below where a similar result is discussed in a nonstationary setting). Notice also that if p_M increases relative to p_I, then α^* increases (in light of Assumption 3) and so "always offer" remains optimal for even higher values of α. This is because if p_I decreases, there is less value to removing incented actions since players take incented actions less frequently. All of this confirms with intuition.

What does a game look like where p_M is relatively higher than p_I? As discussed in the introduction, this can reflect a combination of factors related to the relative strength of an incented action and the player's own time value of money. Games where incented actions warrant high rewards will have higher p_I, moving down the threshold α^*.

Another intuitive aspect, given the stationary nature of the problem, is that τ does not appear in the formulation of α^*. Indeed, if the problem is stationary, there is no real benefit to progressing since each state is the same probabilistically. Proposition 4 below shows that in a nonstationary setting, the dependence on τ plays an important role.

Regarding the stationarity assumption, in light of our data in Table 1, the stationarity of $p_M(\ell, f)$ and $p_I(\ell, f)$ does not seem unreasonable. However, the stationarity of $\alpha(\ell, f)$ and τ are more problematic. In our later analysis, we could not handle the nonstationarity of τ and α simultaneously, so our conditions focus on the impact of the nonstationarity of α.[8]

5.2 Special case: Cumulative free play is irrelevant. The rest of the paper focuses on establishing an analytical condition to show the optimality of threshold policies when p_I, p_M, and α are no longer stationary. Handling the complexity of this case drives several assumptions of varying degrees of strength. We will examine each of these assumptions carefully and justify them to the extent possible. Many of the assumptions will arise to control the speed of change in α (relative to p_M and p_I) as engagement increases (see, for instance, Assumption 8).

5.2.1 Additional assumptions. The most important assumption we need to make is in how free play affects (or more precisely, does not affect) state transition.

Assumption 5. The state variable f capturing the cumulative amount of free play has a cap of $\bar{F} = 0$. In other words, the accumulation of free plays does not impact state transitions.

With this assumption, we may simplify our notation to only involve the state ℓ. That is, we will write $p_M(\ell)$, $p_I(\ell)$, $q_M(\ell)$, and $\alpha(\ell)$. We simplify the transition probabilities $\tau_M(\ell'|\ell)$ and $\tau_I(\ell'|\ell)$ even further, in concert with Assumption 4:

Assumption 6. We assume that τ is stationary, given the choice of action; that is, $\tau_M(\ell'|\ell) = \tau_M$ if $\ell' = \ell + 1$ and 0 otherwise, while $\tau_I(\ell'|\ell) = \tau_I$ if $\ell' = \ell + 1$ and 0 otherwise.

In other words, if incented actions are available, our transition probabilities are now:

$$\mathbb{P}_1(\ell'|\ell) = \begin{cases} p_M(\ell)\tau_M + p_I(\ell)\tau_I & \text{if } \ell' = \ell + 1 \text{ and } -1 < \ell < \bar{L} \\ p_M(\ell)(1-\tau_M) + p_I(\ell)(1-\tau_I) & \text{if } -1 < \ell' = \ell < \bar{L} \\ 1 & \text{if } \ell = \ell' = \bar{L} \\ p_Q(\ell) & \text{if } \ell = -1 \\ 0 & \text{otherwise} \end{cases}$$

and if incented actions are not available:

$$\mathbb{P}_0(\ell'|\ell) = \begin{cases} q_M(\ell)\tau_M & \text{if } \ell' = \ell + 1 \text{ and } -1 < \ell < \bar{L} \\ q_M(\ell)(1-\tau_M) & \text{if } -1 < \ell' = \ell < \bar{L} \\ 1 & \text{if } \ell = \ell' = \bar{L} \\ q_Q(\ell) & \text{if } \ell = -1 \\ 0 & \text{otherwise.} \end{cases}$$

Accordingly, we may also restate the publisher's problem as:

$$\max_{y \in \{0,1\}^L} W^y(\ell)$$

where $W^y(\ell) := \mathbb{E}_\ell^y\left[\sum_{t=1}^{\infty} r(\ell_t, y(\ell_t))\right]$ as analogously defined in Section 3.

A few words of justification for Assumption 5. Following the statement of Assumption 4, we described an interpretation of levels as successive "pinch points" in a game. Therefore, the random behavior of players depended on how many "pinch points" the player has passed and the number of free play attempts the player has made. We are now simplifying this further to say that the game company is only interested in the number of "pinch points" passed by the player. This interpretation makes it possible to give clean interpretations of our transition probabilities and their behavior.

Assumption 7. We make the following additional assumptions:

(A7.1) $p_M(\ell)$ and $q_M(\ell)$ increase in ℓ,

(A7.2) $p_Q(\ell)$ and $q_Q(\ell)$ decrease in ℓ,

(A7.3) $\tau_M > \tau_I$, and

(A7.4) $\alpha(\ell)$ is increasing in ℓ.

Assumption (A7.1) and Assumption (A7.2) ensure that players at higher levels are more likely to make in-app purchases and less likely to quit. The concept here is that more engaged players are more likely to spend and less likely to quit. An astute reader may observe that item (A7.2) is violated by the data in Table 1 and thus not every conception of engagement may support this assumption. Note that we do not make an assumption about the monotonicity of $p_I(\ell)$ in ℓ, but of course, it is affected by changes in $p_M(\ell)$ and $p_Q(\ell)$.

Assumption (A7.3) implies that a player is more likely to progress a level when monetizing than when taking an incented action. Again, the rewards for incented actions are typically less powerful than what can be purchased with real money and so monetizing is more likely to lead to increased engagement. Finally, (A7.4) implies that a greater share of the probability of taking an incented actions when offered is allocated to monetization when an incented action is removed (see (1)). As a player moves higher up in levels, the monetization option becomes relatively more attractive than quitting once the incented action is removed. Indeed, quitting has the player walking away from a potentially significant investment of time and mastery captured by a high level in the game.

5.2.2 Revenue, retention, and progression effects. Additional assumptions are needed for our analysis, but they are harder to state and interpret without discussion of additional concepts. Let $y^1_{\bar{\ell}}$ be a given policy with $y^1_{\bar{\ell}}(\bar{\ell}) = 0$ for some level $\bar{\ell}$. Consider a local change to a new policy $y^2_{\bar{\ell}}$ where $y^2_{\bar{\ell}}(\bar{\ell}) = 1$ but $y^2_{\bar{\ell}}(\ell) = y^1_{\bar{\ell}}(\ell)$ for $\ell \neq \bar{\ell}$. We call $y^1_{\bar{\ell}}$ and $y^2_{\bar{\ell}}$ *paired policies with a local change at* $\bar{\ell}$. Analyzing local changes at the target level $\bar{\ell}$ gives insight into the effect of starting to offer an incented action at a given level. For ease of notation, let $W^1(\ell) = W^{y^1_{\bar{\ell}}}(\ell)$ and $W^2(\ell) = W^{y^2_{\bar{\ell}}}(\ell)$.

Our goal is to understand the change in expected revenue moving from policy $y^1_{\bar{\ell}}$ to policy $y^2_{\bar{\ell}}$ where the player starts (or has reached) level $\bar{\ell}$. Indeed, because the player's level does not decrease, if she reaches level $\bar{\ell}$, the result is the same as if the player just started at level $\bar{\ell}$ by the Markovian property of the player model. Understanding when, and for what reasons, this change has a positive impact on revenue provides insights into the value of incented actions.

It is timely to exploit the special structure offered by absorbing Markov chains to rewrite the total reward for a policy as follows:

$$W^y(\ell) = \sum_{\ell' \in L} n^y_{\ell,\ell'} r(\ell', y(\ell')) \tag{8}$$

where $n^y_{\ell,\ell'}$ is the expected number of visits to level ℓ' starting in level ℓ. We derive closed-form expressions for $n_{\ell,\ell'}$ that facilitate analysis. For details, see Section A1 of the e-companion.

The change in total expected revenue from the policy change $y^1_{\bar{\ell}}$ to $y^2_{\bar{\ell}}$ at level $\bar{\ell}$ is:

$$W^2(\bar{\ell}) - W^1(\bar{\ell}) = \underbrace{n^2_{\bar{\ell},\bar{\ell}} r(\bar{\ell}, 1) - n^1_{\bar{\ell},\bar{\ell}} r(\bar{\ell}, 0)}_{(C(\bar{\ell}))} + \underbrace{\sum_{\ell > \bar{\ell}} (n^2_{\bar{\ell},\ell} - n^1_{\bar{\ell},\ell}) r(\ell, y(\ell))}_{(F(\bar{\ell}))}. \tag{9}$$

Term $C(\bar{\ell})$ is the change of revenue accrued from visits to the current level $\bar{\ell}$. We may think of $C(\bar{\ell})$ as denoting the *current* benefits of offering an incented action in state $\bar{\ell}$, where "current" means the current level. Term $F(\bar{\ell})$ captures the change due to visits to all other levels. We may think of $F(\bar{\ell})$ as denoting the *future* benefits of visiting higher ("future") levels. We can give explicit formulas for $C(\bar{\ell})$ and $F(\bar{\ell})$ for $\ell < \bar{L}$ (after some work detailed in Section A4 of the e-companion) as follows:

$$C(\bar{\ell}) = \frac{p_M(\bar{\ell})\mu_M + p_I(\bar{\ell})\mu_I}{1-p_M(\bar{\ell})(1-\tau_M)-p_I(\bar{\ell})(1-\tau_I)} - \frac{q_M(\bar{\ell})\mu_M}{1-q_M(\bar{\ell})(1-\tau_M)} \tag{10}$$

$$= \frac{p_I(\bar{\ell})\{\mu_I - \alpha(\bar{\ell})\mu_M + [p_M(\bar{\ell})+\alpha(\bar{\ell})p_I(\bar{\ell})][(1-\tau_I)\mu_M-(1-\tau_M)\mu_I]\}}{[1-p_M(\bar{\ell})(1-\tau_M)-p_I(\bar{\ell})(1-\tau_I)][1-p_M(\bar{\ell})(1-\tau_M)-\alpha(\bar{\ell})p_I(\bar{\ell})(1-\tau_M)]} \tag{11}$$

and

$$F(\bar{\ell}) = \{\frac{p_M(\bar{\ell})\tau_M + p_I(\bar{\ell})\tau_I}{1-p_M(\bar{\ell})(1-\tau_M)-p_I(\bar{\ell})(1-\tau_I)} - \frac{q_M(\bar{\ell})\tau_M}{1-q_M(\bar{\ell})(1-\tau_M)}\}\{\sum_{\ell'>\bar{\ell}} n^{y^1}_{\bar{\ell}+1,\ell'} r(\ell', y(\ell'))\} \tag{12}$$

$$= \frac{p_I(\bar{\ell})\{\sum_{\ell'>\bar{\ell}} n^{y^1(\bar{\ell}+1)}_{\bar{\ell}+1,\ell'} r(\ell',y(\ell'))\}\{\tau_I - \alpha(\bar{\ell})\tau_M + [p_M(\bar{\ell})+\alpha(\bar{\ell})p_I(\bar{\ell})][(1-\tau_I)\tau_M-(1-\tau_M)\tau_I]\}}{[1-p_M(\bar{\ell})(1-\tau_M)-p_I(\bar{\ell})(1-\tau_I)][1-p_M(\bar{\ell})(1-\tau_M)-\alpha(\bar{\ell})p_I(\bar{\ell})(1-\tau_M)]}. \tag{13}$$

One interpretation of the formula $C(\bar{\ell})$ is that the two terms are conditional expected revenues associated with progressing to level $\bar{\ell}+1$ conditioned on the event that the player does not stay in level ℓ (by either quitting or advancing). Thus, $C(\bar{\ell})$ is the change in conditional expected revenue from offering incented actions. There is a similar interpretation of the expression $\frac{p_M(\bar{\ell})\tau_M + p_I(\bar{\ell})\tau_I}{1-p_M(\bar{\ell})(1-\tau_M)-p_I(\bar{\ell})(1-\tau_I)} - \frac{q_M(\bar{\ell})\tau_M}{1-q_M(\bar{\ell})(1-\tau_M)}$ in the definition of $F(\bar{\ell})$. Both terms are conditional probabilities of progressing from level $\bar{\ell}$ to level $\bar{\ell}+1$ conditioned on the event that the player does not stay in level $\bar{\ell}$ (by either quitting or advancing). Thus, $F(\bar{\ell})$ can be seen as the product of a term representing the increase in the conditional probability of progressing to level $\bar{\ell}$ and the sum of revenues from expected visits from state $\bar{\ell}+1$ to the higher levels.

We provide some intuition behind what drives the benefits of offering incented actions by isolating three effects of incented actions that were discussed informally in the introduction. To this end, we introduce the notation $\Delta_r(\ell|\bar{\ell}) := r(\ell, y^2_{\bar{\ell}}(\ell)) - r(\ell, y^1_{\bar{\ell}}(\ell))$, which expresses the change in the expected revenue per visit to level ℓ and $\Delta_n(\ell|\bar{\ell}) := n^2_{\bar{\ell},\ell} - n^1_{\bar{\ell},\ell}$, which expresses the change in the number of expected visits to level ℓ (starting at level $\bar{\ell}$) before quitting. Note that $\Delta_r(\ell|\bar{\ell}) = 0$ for $\ell \neq \bar{\ell}$ since we are only considering a local change in policy at level $\bar{\ell}$. On the other hand,

$$\Delta_r(\bar{\ell}|\bar{\ell}) = -(q_M(\bar{\ell}) - p_M(\bar{\ell}))\mu_M + p_I(\bar{\ell})\mu_I = p_I(\bar{\ell})[\mu_I - \alpha(\bar{\ell})\mu_M]. \tag{14}$$

The value is called the *revenue effect*. It expresses the change in the revenue per visit to the starting level $\bar{\ell}$. The *retention effect* is the value $\Delta_n(\bar{\ell}|\bar{\ell})$ and expresses the change in the number of visits to $\bar{\ell}$. Lastly, we refer to the value $\Delta_n(\ell|\bar{\ell})$ for $\ell > \bar{\ell}$ as the *progression effect at level ℓ*.

At first blush, it may seem possible for the progression effect to have a different signs at different levels, but the following result shows that the progression effect is uniform in sign.

Proposition 1. Under Assumptions 2 to 7, the progression effect is uniform in sign; that is, either $\Delta_n(\ell|\bar{\ell}) \geq 0$ for all $\ell \neq \bar{\ell}$ or $\Delta_n(\ell|\bar{\ell}) \leq 0$ for all $\ell \neq \bar{\ell}$.

Because of the consistency in sign, we may refer to the *progression effect* generally (without reference to a particular level).

If both the revenue effect and retention effects are positive, $C(\bar{\ell})$ in (9) is positive, and there is a net increase in revenue due to visits to level $\bar{\ell}$. Similarly, if both effects are negative, then $C(\bar{\ell})$ is negative. When one effect is positive and the other is negative, the sign of $C(\bar{\ell})$ is unclear. By contrast, the retention effect has a consistent sign.

Policy	$W^y(0)$	$W^y(1)$
$y = (0,0)$	0.723	1.13
$y = (1,0)$	0.691	1.13
$y = (1,1)$	0.701	1.15
$y^* = (0,1)$	0.727	1.15

Table 3: Total expected profit for Example 1.

Proposition 2. Under Assumptions 2 to 7, the retention effect is nonnegative; i.e., $\Delta_n(\bar{\ell}|\bar{\ell}) \geq 0$.

Since the sign of $F(\bar{\ell})$ is determined by the direction of the progression effect, its sign is less straightforward to determine and is unlikely to have a consistent sign.

5.2.3 The logic of the optimality of threshold policies. Using the terminology of the "three effects", we can give an intuition for why threshold policies are likely to be optimal. When players start out playing a game, their level is low and they are likely to quit. The following result shows that high levels are of higher value to the designer.

Lemma 1. Under Assumptions 2 to 7, $W(\ell)$ is a nondecreasing function of ℓ.

Indeed, Lemma 1 says we get more value out of players at higher levels. Hence, retaining players and progressing them to higher levels is important for overall revenue. Proposition 2 shows the retention effect of offering incented actions is always positive, and intuitively, the revenue and progression effects are largest at low levels because players are unlikely to monetize early on and the benefits derived from progressing players are likely to be at their greatest. This suggests it is optimal to offer incented actions at low levels. However, once players are sufficiently engaged, it makes sense to remove incented actions to focus their attention on the monetization option. If sufficiently engaged and $\alpha(\ell)$ is sufficiently large, most of the probability of taking the incented action shifts to monetizing, which drives greater revenue.

Despite this appealing logic, the next example shows that our current assumptions are insufficient to guarantee the existence of an optimal threshold policy, even under all of our assumptions.

Example 1. Consider the following two level example. Assume $\mu_M = 1$, $\mu_I = 0.05$, $\tau_M = 0.5$, $\tau_I = 0.4$. At level 0, $p_M(0) = 0.05$, $p_I(0) = 0.65$, $\alpha(0) = 0.5$ and thereby $q_M(0) = 0.375$. At level 1, $p_M(1) = 0.2$, $p_I(1) = 0.6$, $\alpha(1) = 0.55$ and thereby $q_M(1) = 0.53$.

We solve the optimal policy by backward induction. At level 1, $W(1, y = 1) = \frac{p_M(1)\mu_M + p_I(1)\mu_I}{1 - p_M(1) - p_I(1)} = 1.15$ while $W(1, y = 0) = \frac{q_M(1)\mu_M}{1 - q_M(1)} \approx 1.13$. Therefore, $y^*(1) = 1$ and $W(1) = 1.15$. At level 0, $W(0, y = 1) = 0.701$ and $W(0, y = 0) = 0.727$ hence $y^*(0) = 0$ and $W(0) = 0.727$.

Next, we show that $y^* = (0, 1)$ is the only optimal policy. In fact, we compute $W^y(0)$ and $W^y(1)$ under all possible policies in the following table. We observe that none of $(0, 0)$, $(1, 0)$ and $(1, 1)$ are optimal. This implies y^* is the only optimal policy. Since y^* is not a forward threshold policy, this implies there is no optimal forward threshold policy. $\triangleleft$

This example illustrates a break in the logic set at the outset of this subsection. It is optimal to offer incented actions at the higher level because of the dramatic reduction in the quitting probability when offered, reducing the quitting probability compared to a 0.47 quitting probability when not offering incented actions. Although the expected revenue per period the player stays at the highest level is lower when incented actions are offered (0.23 as compared to 0.47), the player will stay longer and thus earn additional revenue. However, at the lowest level, the immediate reward of not offering incented actions (0.462 versus 0.141) outweighs losses due to a lower chance of advancing to the higher level.

The previous example shows how the *growth rate* of α plays a key role in determining whether a threshold policy is optimal or not. When incentives actions are removed the probability $p_I(\ell)$ is distributed to the monetization and quitting actions according to $\alpha(\ell)$. The associated increase in the probability of monetizing from $p_M(\ell)$ to $q_M(\ell)$ makes removing incented actions attractive since the player is more likely to pay. However, the quitting probability increases from $p_Q(\ell)$ to $q_Q(\ell)$, a downside of removing incented actions. If $\alpha(\ell)$ grows sufficiently quickly, the benefits will outweigh the costs of removing incented actions. From Assumption (A7.4) we know that $\alpha(\ell)$ increases, but this alone is insufficient. Just how quickly we require $\alpha(\ell)$ to grow to ensure a threshold policy requires careful analysis. This analysis results in lower bounds on the growth of $\alpha(\ell)$ that culminates in Theorem 2 below. Here are the assumptions we require on α:

Assumption 8. The following hold:

(A8.1) $\alpha(\bar{L}) = 1$; that is, $q_Q(\bar{L}) = p_Q(\bar{L})$ and $q_M(\bar{L}) = p_M(\bar{L}) + p_I(\bar{L})$,

(A8.2) $\alpha(\ell+1) - \alpha(\ell) > q_M(\ell+1) - q_M(\ell)$ for all $\ell \in L$., and

(A8.3) $1 - \alpha(\ell+1) \leq (1 - \alpha(\ell)) \frac{p_M(\ell+1)\tau_M + p_I(\ell+1)\tau_I}{p_Q(\ell+1) + p_M(\ell+1)\tau_M + p_I(\ell+1)\tau_I}$ for $\ell = 1, 2, \ldots, \bar{L} - 1$.

It is straightforward to see that (A8.1) implies that it is never optimal to offer incented action at the highest level. Players at the highest level are no more likely to quit when the incented action is removed. Assumption (A8.2) provides a lower bound on the growth of $\alpha(\ell)$; it says that $\alpha(\ell)$ must grow faster than the probability $q_M(\ell)$ of monetizing when the incented action is not offered. Assumption (A8.2) provides a second lower bound on the growth of $\alpha(\ell)$. The fractional term in the assumption is the probability of advancing from level $\ell+1$ to $\ell+2$ conditioned on leaving level $\ell+1$ and is thus less than one. Which bound in (A8.2) or (A8.3) is tighter depends on the data specifications that arise from specific game settings.

Theorem 2. Suppose Assumptions 2 to 8 hold. Then there exists an optimal threshold policy with threshold level ℓ^*. That is, there exists an optimal policy y^* with $y^*(\ell) = 1$ for any $\ell \leq \ell^*$ and $y^*(\ell) = 0$ for any $\ell > \ell^*$.

In Section A7 of the e-companion we show that if we drop (A8.2) then a threshold policy may no longer be optimal. Section A8 shows that the same is true if (A8.3) is dropped. As we see in some examples in the next subsection, our assumptions are sufficient but not necessary conditions for an optimal threshold policy to exist.

To simplify matters further, we also take the convention that when there is a tie in Bellman's equation

$$\begin{aligned}W(\ell) &= \max\{W(\ell, y=1), W(\ell, y=0)\}\\ &= \max\{r(\ell,1) + \mathbb{P}_1(\ell|\ell)W(\ell) + \mathbb{P}_1(\ell+1|\ell)W(\ell+1),\\ &\qquad r(\ell,0) + \mathbb{P}_0(\ell|\ell)W(\ell) + \mathbb{P}_0(\ell+1|\ell)W(\ell+1)\}\end{aligned}$$

whether to offer an incented action or not, the publisher always chooses not to offer. This is consistent with the fact that there is a cost to offering incented actions. Although we do not model costs formally, we will use this reasoning to break ties. Under this tie-breaking rule, there is a *unique* optimal threshold policy guaranteed by Theorem 2.

5.2.4 Game design and optimal use of incented actions. So far we have provided a detailed analytical description of the possible benefits of offering incented actions and the optimality of threshold policies under certain conditions. There remain questions surrounding what impacts the choice of threshold, and how this threshold is determined by attributes of the game.

We first consider how differences in the revenue parameters μ_I and μ_M affect ℓ^*. Observe that only the revenue effect in (14) is impacted by changes in μ_I and μ_M, the retention and progression effects are unaffected. This suggests the following result:

Proposition 3. The optimal threshold ℓ^* is a nondecreasing function of the ratio μ_I/μ_M.

Note that the revenue effect is nondecreasing in the ratio μ_I/μ_M. Since the other effects are unchanged, this implies that the benefit of offering incented actions at each level is nondecreasing in μ_I/μ_M, thus establishing the monotonicity of ℓ^* in μ_I/μ_M.

To interpret this result, we consider what types of games have a large or small ratio μ_I/μ_M. From the introduction in Section 1, we know that incented actions typically deliver far less revenue to the publisher than in-app purchases. This suggests that the ratio is small, favoring a lower threshold. However, this conclusion ignores how players in the game may influence each other. In many cases, a core value of a player to the game publisher is the word-of-mouth a player spreads to their contacts. In cases where this "social effect" is significant, this can change our revenue considerations. Let δ be the revenue attributable to the word-of-mouth or network effects of a player, regardless of whether the player takes an incented actions or monetizes, then the ratio of interest is $\frac{\mu_I+\delta}{\mu_M+\delta}$, under the assumption that whenever the player is in the game, it generates extra revenue δ for the publisher.

This new ratio comes from writing the expected reward in state ℓ as

$$r(\ell, y) = \begin{cases} p_M(\ell)\mu_M + p_I(\ell)\mu_I + (p_M(\ell) + p_I(\ell))\delta, & \text{if } \ell \in L \text{ and } y = 1\\ q_M(\ell)\mu_M + q_M(\ell)\delta, & \text{if } \ell \in E \text{ and } y = 0\\ 0, & \text{if } \ell = -1.\end{cases}$$

The larger is δ, the larger is the ratio $\frac{\mu_I+\delta}{\mu_M+\delta}$, and according to Proposition 3, the larger is the optimal threshold.

This analysis suggests that games with a significant social component should offer incented actions more broadly than in games with weak social aspects. In a social game, it is important to have a large player base to create positive externalities for new players to join, and so having players quit is of greater concern than in non-social games. Hence, it is best to offer incented action until higher levels are reached. All of this intuition is confirmed by Proposition 3.

Besides the social nature of the game, other factors can greatly impact the optimal threshold. Genre, intended audience, and structure of the game affect the other parameters of our model; particularly, τ_I, τ_M, and $\alpha(\ell)$. We first examine the progression probabilities τ_I and τ_M. As we did in the case of the revenue parameters, we focus on the ratio τ_I/τ_M. This ratio measures the relative probability of advancing through incented actions versus monetization. By (A7.3), $\tau_I/\tau_M \leq 1$ but its precise value can depend on several factors. As discussed in the previous section, the ratio has a number of possible interpretations.

Proposition 4. The optimal threshold ℓ^* is a nondecreasing function of the ratio τ_I/τ_M.

One interpretation of this result is that the more effective an incented action is at increasing levels of the player, the longer the incented action should be offered. This is consistent with our discussion of Theorem 1, but now in a nonstationary setting.

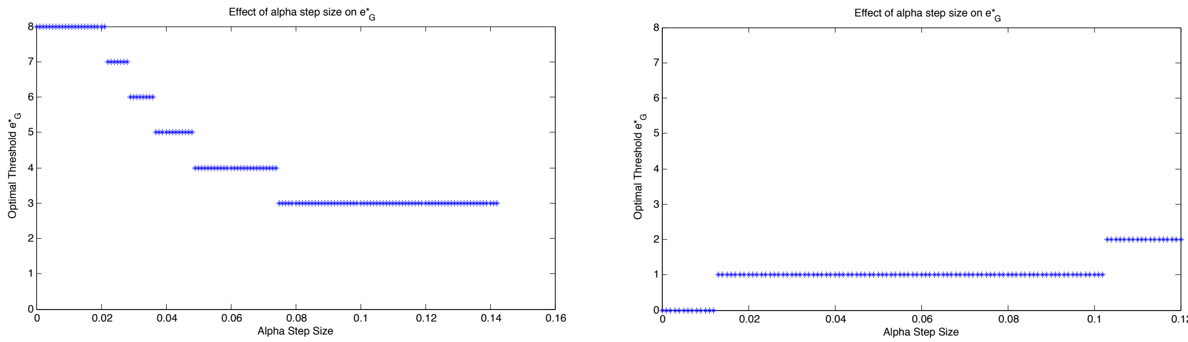
This leads us to investigate how changes in the degree of cannibalization between incented actions and monetization impacts the optimal threshold. For the sake of analysis, we assume that $\alpha(\ell)$ is an affine function of ℓ with $\alpha(\ell) = \alpha_{\text{initial}} + \alpha_{\text{step}}\ell$ where α_{initial} and α_{step} are nonnegative real numbers (as before our results are robust to nonlinear instantiations of the function $\alpha(e)$).

Analysis of how different values for α_{initial} and α_{step} impact the optimal threshold is not straightforward. This is illustrated in the following two examples. The first considers the sensitivity of the optimal threshold to α_{step}.

Example 2. Consider the following example with nine levels and the following data: $\mu_m = 1$, $\mu_I = 1, 0.05$, $\tau_M = 0.8$, $\tau_I = 0.4$, $p_M(\ell) = 0.0001 + 0.00005\ell$ and $p_I(\ell) = 0.7 - 0.00001\ell$ for $\ell = 0, 1, \dots, 8$. We have not yet specified $\alpha(\ell)$. We examine two scenarios: (a) where $\alpha_{\text{initial}} = 0$ and we vary the value of α_{step} (see Figure 7a) and (b) where $\alpha(0) = 0.16$ and we vary the value of α_{step} (see Figure 7b). The vertical axis of these figures is the optimal threshold of the unique optimal threshold policy for that scenario. What is striking is that the threshold ℓ^* is nonincreasing in α_{step} when $\alpha_{\text{initial}} = 0$ but nondecreasing in α_{step} when $\alpha_{\text{initial}} = 0.16$.

One explanation of the difference in patterns between Figures 7a and 7b concerns whether it is optimal to include incented actions initially or not. In Figure 7a the initial degree of cannibalization is zero, making it costless to offer incented actions initially. When α_{step} is very small cannibalization is never an issue, and incented actions are offered throughout. However, as α_{step} increases, the degree of cannibalization eventually makes it optimal to stop offering incented actions to encourage monetization. This explains the nonincreasing pattern in Figure 7a.

By contrast, in Figure 7b the initial degree of cannibalization is already quite high, making it optimal to start by offering for low values of α_{step}. However, when α_{step} is sufficiently large,



(a) Initial value $\alpha_{\text{initial}} = 0$.

(b) Initial value $\alpha_{\text{initial}} = 0.16$.

Figure 7: Sensitivity of the optimal threshold to changes in α_{step}.

there are benefits to encouraging the player to advance. Recall, $\alpha(\ell)$ affects both the probability of monetization and the probability of quitting. In the case where α_{step} is sufficiently high, there are greater benefits to the player progressing, making quitting early more costly. Hence it can be optimal to offer incented actions initially to discourage quitting and encourage progression. This explains the nondecreasing pattern in Figure 7b. $\triangleleft$

As the following example illustrates, adjusting for changes in α_{initial} reveals a different type of complexity.

Example 3. Consider the following two-level example. Assume $\mu_M = 1$, $\mu_I = 0.0001$, $\tau_M = 0.01$, $\tau_I = 0.009$. At level 0, $p_M(0) = 0.05$, $p_I(0) = 0.68$. At level 1, $p_M(1) = 0.3$, $p_I(1) = 0.65$. We set α step size be 0.6, i.e. $\alpha(1) = \alpha_{\text{initial}} + 0.6$. Figure 8 captures how changes in α_{initial} leads to different optimal thresholds (details on how the figure is derived is suppressed for brevity, but follows the same pattern as previous examples). The striking feature of the figure is that the optimal threshold decreases, and then increases, as α_{initial} becomes larger. This "U"-shaped pattern reveals competing effects associated with changes in α_{initial}. As α_{initial} increases, the benefit of increasing retention (at the cost of harming retention) weakens. This contributes to downward pressure on the optimal threshold. On the other hand, increasing α_{initial} also increases $\alpha(1)$. This increases the attractiveness of reaching a higher level and dropping the incented action. Indeed, $W^y(1)$ is increasing in $\alpha(1)$ when $y(0) = 1$. This puts upward pressure on the optimal threshold. This latter "future" benefit is weak for lower levels of $\alpha(0)$, where it may be optimal to offer an incented action in the last period. This provides justification for the "U"-shaped pattern. $\triangleleft$

The scenarios in the above two examples provide a clear illustration of the complexity of our model. At different levels, and with different prospects for the value of future benefits, the optimal strategy can be influenced in nonintuitive ways. This is particularly true for changes in $\alpha(\ell)$ as it impacts all three effects – revenue, retention, and progression. In some sense, cannibalization is the core issue in offering incented actions. This is evident in our examples and a careful examination of the conditions in Assumption 8 – the parameter $\alpha(\ell)$ is ubiquitous.

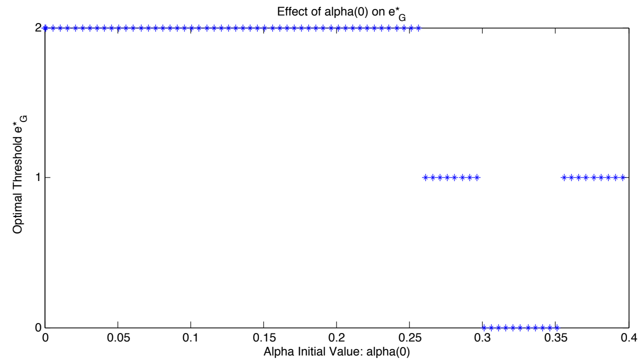


Figure 8: Sensitivity of the optimal threshold to changes in α_{initial}.

6 Conclusion

In this paper, we investigated the use of incented actions in mobile games, a popular strategy for extracting additional revenue from players in freemium games where the vast majority of players are unlikely to monetize. We built an MDP model to make tactical decisions for offering incented actions and showed how to implement the MDP model in a concrete case study. Our approach of using an MDP has some direct benefit to practitioners. With player data and relevant game parameters that companies have access to in the age of big data, validating our model and using it to derive insights on the impact of certain policies is plausible.

Moreover, we discussed the intuitive reasons for offering incented actions and built an analytical model to assess the associated tradeoffs. This understanding lead us to define sufficient conditions for the optimality of certain specially structured policies, which we later analyzed to provide managerial insights into what types of game designs are best suited to offering incented actions.

Our analytical approach was to devise a parsimonious stylized model that abstracts a fair deal from reality and yet maintained the salient features needed to assess the impact and effects of offering incented actions. For instance, we assume the publisher has complete knowledge about the player's transition probabilities and awareness of the level. In the setting where transition probabilities are unknown, some statistical learning algorithms and classification of players into types would be required. Moreover, in the situation where engagement is difficult to define or measure, a partially observed Markov decision process (POMDP) model would be required, where only certain signals of the player's underlying engagement can be observed.

There is also the question of micro-founding the player model that we explore, asking what random utility model could give rise to the transition probabilities that we take as given in our model. All these questions are outside of our current scope but could nonetheless add realism to our approach. Of course, the challenge of establishing the existence of threshold policies in these extensions is likely to be prohibitive. Indeed, discovering analytical properties of optimal policies of any form in a POMDP is challenging (Krishnamurthy and Djonin, 2007). These extensions would likely produce studies that are even more algorithmic and numerical, whereas in the current study we worked hard to deliver analytical insights.

Finally, the current study ignores an important actor in the case of games hosted on mobile platforms – the platform holder. In the case of the iOS App Store, Apple has made several interventions that either limited or more closely monitored the practice of incented actions.[9] The platform holder and game publisher have misaligned incentives when it comes to incented actions. Typically, the revenue derived from incented actions is not processed through the platform, whereas in-app purchases are. We feel that investigation of the incentive misalignment problem between platform and publisher, possibly as a dynamic contracting problem, is a promising area of future research. The model developed here is a building block for such a study.

Acknowledgments

We thank the associate editor and two anonymous reviews for their very helpful feedback, which greatly improved the paper. We also thank the technical staff at Ebo Games for their able assistance. All authors contributed equally in this research. The work of Dr. Tong was supported by the National Natural Science Foundation of China [Grands 71772137, 71531011 and 71890973].

Notes

[1]https://www.reuters.com/sponsored/article/popularity-of-gaming

[2]https://www.newzoo.com/globalgamesreport

[3]https://www.businessofapps.com/data/app-revenues/

[4]https://www.appannie.com/en/apps/ios/top/united-states/games/iphone/

[5]http://www.gamesbrief.com/2011/11/conversion-rate/

[6]http://response.unity3d.com/in-game-advertising-the-right-way-monetize-engage-retain-whitepaper

[7]http://techcrunch.com/2013/06/12/king-quits-advertising-since-it-earns-so-much-on-candy-crush-purchases/

[8]Notice however that p_M and τ_M play complementary roles in determining the transition probabilities of the underlying Markov chain (see (3) and (4)). Thus, by analyzing the case where p_M is nonstationary, we can capture the dynamics even while assuming τ_M and τ_I do not depend on the state.

[9]http://venturebeat.com/2014/06/21/apples-crackdown-on-incentivizing-app-installs-means-marketers-need-new-tric

References

A.Z. Abbasi, D.H. Ting, and H. Hlavacs. Engagement in games: Developing an instrument to measure consumer videogame engagement and its validation. *International Journal of Computer Games Technology*, 2017, 2017.

H.K. Cheng, S. Li, and Y. Liu. Optimal software free trial strategy: Limited version, time-locked, or hybrid? *Production and Operations Management*, 24(3):504–517, 2015.

Khai Chiong, Richard Chen, and Sha Yang. Incentivized advertising: Treatment effect and adverse selection. *arXiv preprint arXiv:1709.00197*, 2017.

Hong Guo, Lin Hao, Tridas Mukhopadhyay, and Daewon Sun. Selling virtual currency in digital games: Implications for gameplay and social welfare. *Information Systems Research*, 2019a.

Hong Guo, Xuying Zhao, Lin Hao, and De Liu. Economic analysis of reward advertising. *Production and Operations Management*, 28:2413–2430, 2019b.

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

C. Jennett, A.L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9):641–661, 2008.

V. Krishnamurthy and D.V. Djonin. Structured threshold policies for dynamic sensor scheduling — A partially observed Markov decision process approach. *IEEE Transactions on Signal Processing*, 55(10):4938–4957, 2007.

Clarence Lee, Vineet Kumar, and Sunil Gupta. Designing freemium: Strategic balancing growth and monetization. *Available at SSRN 2767135*, 2017.

Joowon Lee and Dong-Hee Shin. Positive side effects of in-app reward advertising: Free items boost sales: A focus on sampling effects. *Journal of Advertising Research*, 57(3):272–282, 2017.

N. Lovell. *Design Rules for Free-to-Play Games*. GAMESbrief, 2012.

W. Luton. *Free-to-Play: Making Money From Games You Give Away*. New Riders, 2013.

M.F. Niculescu and D.J. Wu. Economics of free under perpetual licensing: Implications for the software industry. *Information Systems Research*, 25(1):173–199, 2014.

M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 1994.

K.T. Talluri and G.J. Van Ryzin. *The Theory and Practice of Revenue Management*. Springer, 2006.

P.H. Zipkin. *Foundations of Inventory Management*, volume 20. McGraw-Hill New York, 2000.