A Continuous Time Model of Product Usage: Measuring the Effect of Product Design and Rewards in Online Games

Yulia Nevskaya* Paulo Albuquerque†

October 3, 2012

Preliminary. Please do not cite without authors' permission.

Abstract

The paper proposes a demand model of product usage in continuous time. Our setting is flexible enough to simultaneously explain usage frequency, duration of usage, and consumer response to product features and firm's actions. Based on product usage literature, we define the main components of our model to include intrinsic motivations, product characteristics, external rewards provided by the firm, and past consumption. In our model, the influence of past consumer choices on decisions takes the form of a cue-based habit formation mechanism. The model is estimated on a novel dataset of online game usage where we observe the usage decisions of a large sample of individuals with a periodicity of 10 minutes. We provide managerial insights on product design and reward systems by testing different product configurations and measuring shirk with changes in reward frequency and product complexity. The proposed model can be applicable to a large number of product categories characterized by repeated product usage or content consumption.

Keywords: Product Usage, Dynamic Structural Discrete Choice Models, Continuous Time

* Ph.D. Candidate at the Simon Graduate School of Business, University of Rochester.

† Associate Professor of Marketing at the Simon Graduate School of Business, University of Rochester.

1 Introduction

To generate loyalty and repeat purchases, many firms rely on frequent incentives for continuous product usage: they either introduce stimulating product features that increase consumption enjoyment, or provide valuable reward systems that generate long-term goals for consumers. With each product usage occasion, consumers obtain a stream of utility with a hedonic component from the experience of using the product and also a goal-driven value from getting closer to and eventually obtaining a reward. Both intrinsic motivation resulting from a steady flow of feelings and emotions experienced while using a product and external rewards linked to the utilitarian uses of a product have been shown to be major drivers of product usage (Holbrook and Hirschman, 1982; Novak et al., 2003).

Additional factors and dimensions should also be considered to fully understand motivations for product consumption, including usage proficiency or mastery, cue-based consumption and habit formation, and complexity of tasks and information processing. First, usage proficiency or mastery development is present in productive and intellectual activities, sports, and other leisure activities, when a user has an opportunity to test one's abilities and conquer the environment in some way (London et al., 1977). Learning to use product features is a result of a number of usage occasions, and a level of experience or proficiency affects consumer enjoyment and derived utility (Lakshmanan et al., 2010). Second, literature on cue-based consumption and habit formation has shown that the presence of the cue raises marginal utility of consumption, and pairing a cue and consumption eventually creates cue-based complementarities (Laibson, 2001). Finally, complexity of tasks leads to changes in consumption patterns when products, experiences, and content lead to higher personal relevance or motivate a multiplicity of cognitive responses if there is a necessity to process information (Leavitt, Greenwald, and Obermiller 1981). Berlyne (1971) showed that there is a strong non-monotonic relation between stimulus complexity and hedonic value, arguing that to investigate the consumption experience, stimuli should be designed to vary in complexity over a range broad enough to permit the full relationship with enjoyment to appear.

Examples of product usage where these elements are common come from a wide variety of categories and many can be found among entertainment products such as computer games, television series, computer software in general, educational products, and leisure activities, such as hobbies and sports. Demanding a high level of expertise and information processing to use software or a computer game, or an increasing complexity of a story-line in a television series can reduce enjoyment and lead to shirking. Consumers who demonstrate low levels of enjoyment or interest reduce their frequency of consumption, for example, they watch an episode of a TV series or play sports only occasionally and not in a habitual pattern. These are also individuals who are more likely to abandon the product or activity. The sporadic usage with low enjoyment also correlates with low or negative consumer reviews for the products.

Most of the product usage drivers are at some level controllable by managers or influenced by marketing actions, and therefore, firms are inevitably interested in understanding the relation between the usage drivers and the long-term user participation, loyalty, and shirk rates. Understanding of the post-purchase dynamic is also an invaluable input into the firm's innovation agenda. For example, e-book vendors such as Amazon.com, Barnes&Noble and Kobo.Inc try to study behavior of users of their digital book reading devices. They track how far readers get into a particular book, how fast they read it, what lines they highlight, how readers of particular genres engage with books. Some of those companies share the knowledge with the publishers to help them create books that hold consumers' attention (The Wall Street Journal, 2012). In general, the "big data" trend of gathering massive amounts of information about consumer interactions with products along multiple dimensions, as well as many traditional products going digital, is bringing rich data on consumer post-purchase behavior that were not available before.

Our contribution to the marketing field is two-fold. First, we propose a model of product usage in continuous time, based on recent developments in dynamic structural discrete choice modeling (Doraszelski and Judd, 2011; Arcidiacono et al., 2012). We find that the flexibility of being able to model a large number of consumer decisions made with varying frequency and changes in the product features that occur sequentially, makes these methods very appealing for explaining product usage. In other areas, continuous time approaches are frequently used, for example, in the finance and macroeconomics fields, often yielding closed form solutions for problems that demand computationally intensive methods or simulations under discrete time. For example, pricing decisions and wage dispersion were analyzed using a continuous time approach (Calvo, 1983; Burdett and Mortensen, 1998). Doraszelski and Judd (2011) and Arcidiacono et al. (2012) contribute to these methods by developing estimation methods for dynamic discrete choice models in both single- and multi-agent problems. We extend the model in Arcidiacono et al. (2012) to accommodate consumer behavioral processes that are generally unobserved, such as habit development.

Our application is one of the first to fully utilize the advantages of a continuous time discrete choice model to address a marketing problem. Our approach (1) allows for a flexible decision schedule for consumers instead of using ad-hoc assumptions that individuals make usage decisions rigidly every day or every hour; (2) builds a consumer utility function that structurally reflects not only the instantaneous value of using the product that is present in discrete time models but also the duration of product usage; (3) permits simultaneous modeling of multiple types of consumer decisions, such as when and for how long to use the product, whether to respond to usage cues and rewards provided by the firm, and whether to abandon the product.

Additionally, our approach provides valuable insights into several marketing aspects related to product usage, especially product design, and depth and frequency of rewards. Table 1 lists examples of previous work and defines the place of this paper in this literature. Early literature (e.g., Holbrook and Hirshman, 1982; Unger and Kerman, 1983) developed the main theoretical body of work that described the main motivations and dimensions of usage. Subsequent papers applied this framework using surveys and experimental exercises to investigate the role of innovation and a state of experiential flow in product usage, explaining either frequency or duration decisions (e.g., Novak et al., 2003; Shih and Venkatesh, 2004; Huh and Kim, 2007; Lakshmanan et al., 2010). Increasing availability of data on consumer product usage makes it possible to implement studies based on revealed preferences (Albuquerque and Nevskaya, 2012; Huang et al., 2012; this paper), which have the advantage of observing multiple consumption actions of the same individual over time. While Albuquerque and Nevskaya (2012) focus on the relation between innovation and content consumption choices across alternatives and Huang et al. (2012) model consumer trade-off between short-term needs of hydration and mood pick-up and long-term health related goals in beverage consumption, we focus on the measurement of frequency and duration of usage decisions, modeling habit formation, and providing insights on product design, usage cues, and reward programs.

We empirically demonstrate the value of our approach using an individual-level panel dataset of consumers using an online video game *World of Warcraft* by Blizzard Entertainment. This category is very suitable for our analysis. First, the online games category is becoming one of the dominant ones in the entertainment industry, generating more revenues than the film industry (Wolf, 2006). Second, "gamification" of a number of products is now becoming very frequent, which makes it important to understand how consumers react to game-like product features and rewards. More specifically, in the game under study we observe a complex design created to motivate consumption and involvement. The game has multiple environments accessible to consumers, a wide variety of content that requires from a few seconds to many hours of playtime to consume, provides frequent feedback about accomplishments, and a complex reward system that motivates long-term and goal-driven usage.

Our model is used to measure the impact of changes in several components that explain usage. First, looking at the product design, we test alternative requirements for the amount of experience and mastery that needs to be accumulated by consumers to use advanced features of the product. Second, we change the frequency of the usage cues provided by the firm and measure its impact on habit formation and resulting frequency and duration of usage. Third, we provide findings on the depth and frequency of rewards obtained through product usage.

The rest of the paper is organized as follows. The next section describes the model. Section 3 provides information about the online games industry and details about the data used in our empirical application. In section 4 we discuss the estimation procedure. Section 5 discusses the results. We provide managerial implications of our results in section 6 and conclude in section 7.

2 A Continuous Time Model of Product Usage

In this section, we develop a model of product usage based on theory on experiential utility from consumption, valuation of rewards, and cue-based habit formation. As described in the previous section, these components have been shown to be essential drivers of product usage decisions.

2.1 A Framework for Product Usage Decisions

Consumers maximize their utility by dividing their time between using a product and engaging in other activities (the outside option), i.e. by making decisions about when to start using the product, for how long to use it until stopping, and whether to completely abandon the product. The decisions to start using a product are made in response to contextual cues that consumers receive from their environment that remind or prompt them to use a product (Wood and Neal, 2007, 2009; Laibson, 2001). For example, seeing an advertisement or having some free time may remind individuals of entertainment products and hobbies, leading them to a choice occasion. Note, that these cue arrivals can be controlled by the firm in certain cases through establishing communication channels with consumers.

In this setting, consumers are assumed to be in either active state $g = 1$ (using the product), or in idle state $g = 0$ (being engaged in some other activity), or in the absorbing shirk state $g = 2$ (having decided to abandon the product). At each choice opportunity driven by a cue arrival, an inactive consumer can decide to continue being idle (action $a = 0$) or start using the product (action $a = 1$). The action $a = 1$

will result in a transition from the idle state $g = 0$ to the active state $g = 1$. When a consumer is in the active state, she chooses between actions $a = 1$ (continue using the product), $a = 0$ (stop using), and $a = 2$ (abandon the product).

Once a consumer starts using a product, she obtains some utility for as long as she is actively engaged with the product. The utility provides an intrinsic motivation to use a product, and in some cases, consumers immerse themselves in the usage process and pay little attention to what happens around them, at least for some period of time. Early studies of product usage and leisure activities documented evidence of high involvement or total absorption in an activity, with consumers becoming so involved that they enter a "microcosm distinct from daily life" (Unger and Kerman, 1983; Gordon, Gaitz, and Scott 1976; Foote 1966). More recently, this high involvement also became known as a state of flow (Csíkszentmihályi, 1990). It has been shown that web-browsing and some other digital content consumption activities, e.g. video gaming, can lead to the state of "flow" (Cowley, 2008; Novak et al, 2003). The state of flow has limited duration, and upon the end of the flow state a consumer has an opportunity to make a decision.

Consumers gain experience with the product through the usage process. Consumer experience, or mastery of content, impacts product utility in two ways: by changing the intrinsic utility of product features that the consumer currently uses and by providing access to more advanced product features. We define the product mastery level by $l = 1, ..., L$. With prolonged usage, consumers may receive extrinsic rewards. Extrinsic rewards can be broadly defined as related to the goals of product usage and provide an additional value to consumers to use the product again. The firm decides on the frequency and depth of rewards and the amount of product usage experience necessary to obtain a reward. In our empirical application to online video games, the rewards take the form of access to additional content and virtual goods. The consumer valuation of these rewards might affect the utility derived from product usage.

As a result of continued interaction with the product, consumers might develop a habit for product usage. Following psychology and neuroscience literatures, this paper defines habit "as a psychological disposition to repeat past behavior" (Wood and Neal, 2007, 2009). The habit can also be understood as a general state of high consumer involvement with the product. We denote the habit strength by discrete levels $h = 0, 1, .., \bar{H}$. Consumer develops a strong habit for product usage when she consistently responds to cues by using the product. We assume that the level of habit strength increases with every session of product usage, thereby h goes up to $h+1$. Habit might lead to the change in the utility derived from product usage, as well as in the way contextual cues are interpreted. Consumers with stronger habit might interpret certain cues as prompting to use the product; the same cues would have not been interpreted this way if the habit were absent. The habit strength can be reduced and, eventually, the habit can be completely broken by consistently abstaining from product usage in response to a cue arrival. This process works as follows. If upon a cue arrival a consumer decides not to use a product, the habit strength decreases with probability ϕ. In other words, in order for the habit strength to decrease, a consumer has to deny herself using the product when prompted by a cue, and she might have to repeat this pattern of "receive cue, do not use the product" several times before the habit strength subsides and, eventually, the habit is erased completely. Finally, we assume that consumers are forward-looking about the habit formation process.

2.2 Consumer Utility from Product Usage

A consumer receives intrinsic utility from product usage while she is in active state $g = 1$ and utility from an outside activity when she is in idle state $g = 0$. We define the intrinsic utility as follows:

$$U_{intrinsic}(s) = \begin{cases} \alpha_1 + \alpha_2 f_2(l) + \alpha_3 f_3(h), & if\ g = active, \\ 0, & otherwise. \end{cases} \tag{1}$$

In the above, $s = (g, l, h)$ is a consumer state at time t, where g is a usage state, l is the level of experience with the product, and h is a habit state. The intrinsic utility of product usage depends on some base preference for the product measured by α_1; the return of the level of product mastery on consumer intrinsic utility is measured by the term $\alpha_2 f_2(l)$; and the change in consumer intrinsic utility due to habit is measured by $\alpha_3 f_3(h)$. The $f_2(\cdot)$ and $f_3(\cdot)$ are functions to be specified. The intrinsic utility of being in the inactive state $g = 0$ or the shirk state $g = 2$ are normalized to zero for identification purposes.

Because the consumer receives intrinsic utility at all instants of time, the total utility over some time period of duration τ can be represented as:

$$U^{\tau}_{intrinsic}(s) = \int_0^{\tau} e^{-\rho t} U_{intrinsic}(s) dt. \tag{2}$$

This formulation implies that a consumer exponentially discounts future payoffs at a rate ρ. Without loss of generality, we define $[0, \tau)$ to be the time interval when no change in the consumer state occurs. Therefore, during this period the consumer makes no product usage decisions: during an idle period $g = 0$ no cues arrive prompting her to use a product; during an active period $g = 1$ the consumer is in a state of flow and is not able to make usage decisions as well. Also, during time interval $[0, \tau)$, the consumer does not change her level of product mastery l. At time τ, the state of flow ends (if $g = 1$) or a cue arrives (if $g = 0$), both leading to a choice opportunity. Alternatively, at time τ the change in the product experience level l occurs if $g = 1$.

While the intrinsic utility is specified as being received continuously over time, the utility from extrinsic rewards is defined as an instantaneous utility that consumer obtains at the moment when the reward is provided. The instantaneous utility of extrinsic rewards is defined as:

$$U_{extrinsic}(s) = \beta_1 + \beta_2 f_e(l, X_e). \tag{3}$$

In the above expression, β_1 defines a base utility derived from getting an extrinsic reward; the term $\beta_2 f_e(l, X_e)$ makes the valuation of the reward be a function of the product experience level l and the reward type X_e. The firm decides on the timing and design of the rewards by choosing at which level of experience consumers receive the reward and the characteristics of those rewards.

Finally, we assume that the consumer may also get some instantaneous utility (in some cases, disutility) from the act of starting or finishing the usage session itself, e.g. from the setup needed to get the product ready to use. Formally, we define the instantaneous utility of action $a = 0, 1, 2$ as follows:

$$U_a(s) = \psi_a(s) + e_a. \tag{4}$$

The term e_a is a random shock to consumer utility of action a. It follows Type II extreme value distribution and it is identically and independently distributed across time and consumer actions.

2.3 Evolution of Consumer States

A consumer makes decisions about product usage with certain frequency. In the idle state $g = 0$ the environmental cues prompting product usage arrive via a Poisson process defined by the rate λ_h^{cue}. The rate of cue arrivals is made a function of the consumer habit state h because consumers with a strong usage habit may have a higher disposition to interpret the contextual signals as usage cues. Upon a cue arrival, the consumer makes a decision about product usage.

If a consumer starts using a product, she gets into the "flow" state for some period of time. In the "flow" state she is completely absorbed in the product usage activity and does not think about stopping it. The length of the "flow" period follows an exponential distribution with the rate λ_h^{flow}. When the "flow" state ends, the consumer has the opportunity to make a decision to continue or stop using the product[1].

Additionally, while using the product, changes in the consumer mastery of the product occur[2]. We assume that those changes follow a Poisson process, with the arrival rate λ_l^{exp}. The arrival rate of product mastery changes is influenced by two factors. First, usage allows consumers to gain knowledge about the product. Second, and perhaps more important, the firm's decisions on product complexity and on requirements for access to advanced product features strongly impact how fast a consumer masters a product. In cases where advanced features require a lot of knowledge and experience, or if some content is made available only to individuals that have been using the product for a long time, the arrival rate of gains from product usage will be considerably slower.

2.4 Consumer Decisions

Given the randomness of the cue arrival process, the "flow" durations, and the uncertainty in the timing of product mastery accumulation, the consumer has uncertainty about the timing of her next decision and the timing of the change to her state l. It is also generally true for many experiential products, that the consumer does not know for sure, which event will happen next: opportunity for him to make a decision about product usage provided by the cue arrival / the "flow" state interruption or the arrival of the exogenous event that would change her level of product usage. For example, in many video games progression happens rather fast, and a consumer is able to pass through several levels of the game during one period of "flow". We denote the probability that the next event is the opportunity to make a choice

[1]The products with very short instances of usage do not place consumers into the "flow" state. The model still applies but there is no need to model the duration of the usage session.

[2]The mastery of the product usage can be understood broadly as a level of product usage. The level of product usage can be influenced by the user herself or by the firm. For example, in some online video games progression through the levels is exogenous and its speed is defined by the game developer: the gamer cannot deliberately select the level to play at, instead all the gamer can do to reach a higher level is to invest more time in the game. For some products, such as a video camera, the level of product usage is endogenous: consumers can use an automatic setting or use more advanced controls. Both endogenous decision and exogenous changes in the product usage levels can be easily accommodated by the model.

as $\pi_a(s)$ and the probability that the next event is the change in the product usage level as $\pi_l(s)$. These probabilities are derived based on the properties of exponential distribution[3].

In our empirical application to the online video games the change in the product usage level l occurs only while the user is active ($g = 1$). Thus, in our case the probabilities $\pi_l(s)$ and $\pi_a(s)$ are defined as:

$$\pi_l(s) = \begin{cases} \dfrac{\lambda_l^{exp}}{\lambda_l^{exp} + \lambda_h^{flow}}, & if\ g = 1, \\ 0, & if\ g = 0. \end{cases} \tag{5}$$

and

$$\pi_a(s) = \begin{cases} \dfrac{\lambda_h^{flow}}{\lambda_l^{exp} + \lambda_h^{flow}}, & if\ g = 1, \\ 1, & if\ g = 0. \end{cases} \tag{6}$$

Consumers are forward-looking, and hence take into account both the current enjoyment from product usage and the consequences of the present choices for the future utility. Let $v_a(s)$ denote a continuation value received by the consumer after making a choice a in state $s = (g, l, h)$. All consumer choices lead to a deterministic resulting activity state g, however there might be uncertainty about the resulting habit state h. Also, by choosing the terminal action $a = 2$ consumer decides to never use the product again, and we set the continuation value of this choice to zero. Hence, formally we define the continuation value as:

$$v_a(g, l, h) = \begin{cases} \phi \cdot V\left(s = (0, l, max\left[0, h-1\right])\right) + (1-\phi) \cdot V\left(s = (0, l, h)\right), & if\ g = 0\ and\ a = 0, \\ V(s = (1, l, h)) & if\ g = 0, 1\ and\ a = 1, \\ V(s = (0, l, min\left[h+1, \bar{H}\right])) & if\ g = 1\ and\ a = 0, \\ 0 & if\ g = 1\ and\ a = 2. \end{cases} \tag{7}$$

The term $V(s)$ above stands for the value function that assigns to each state s the present discounted value of all future utility obtained from starting in that state and behaving optimally from there on.

[3] If $X \sim exp(A)$ and $Y \sim exp(B)$ then $\min(X, Y) \sim exp(A + B)$. Then it follows, the probability that a realization of X is smaller than a realization of Y is $\dfrac{A}{A+B}$.

The value function for state s is defined recursively as follows:

$$V(s) = E\left[U^{\tau}_{intrinsic}(s) + e^{-\rho\tau}\left(\pi_l(s)\left[U_{extrinsic}(s) + V(s')\right] + \pi_a(s)\max_{a\in A(s)}\left(\psi_a(s) + e_a + v_a(s)\right)\right)\right], \quad (8)$$

where s' is new consumer state at the end of the period $[0, \tau)$, with $s' = (g, l', h)$. The expectation E in the value function is taken with respect to the duration of time τ during which the consumer state s does not change. The three elements in the expectation E that compose the value function represent the following: first, over the duration τ, the consumer in state s gets the integrated intrinsic utility, $U^{\tau}_{intrinsic}(s)$; second, upon the end of that period, a change in level l to l' will occur with probability $\pi_l(s)$ and value of the new state s' is $V(s')$; third, an opportunity to make a decision will arrive with probability $\pi_a(s)$ and she will obtain the value from the best action a that belongs to the action set $A(s)$. The second and third elements in expectation E represent the expected future value received from the end of duration τ onward and are exponentially discounted by discount rate ρ.

The assumption about the extreme value distribution of the error terms in the utility gives rise to the conditional choice probabilities:

$$P(a\,|s) = \frac{exp\left[(v_a(s) + \psi_a(s)) - (v_0(s) + \psi_0(s))\right]}{\sum\limits_{a'\in A_g} exp\left[(v_{a'}(s) + \psi_{a'}(s)) - (v_0(s) + \psi_0(s))\right]}. \quad (9)$$

3 Data

In this section, we describe the industry and product setting of our empirical application. Our model is general and can be applied to a large number of categories, but here we use our model in the context of online video games. Hence, we clearly define what the dimensions of product usage and reward outcomes are in the video game industry. We provide the generalization of our application and expected results to other categories when we discuss managerial implications in section 6.

3.1 Patterns of Product Usage

With about \$15 billion in sales in 2010 and additional sales of virtual goods in excess of \$1 billion, the online video gaming industry is becoming one of the dominant areas in entertainment (Playlogic Entertainment Inc, 2010), especially benefiting from the new technologies that allow almost permanent online connectivity. We use data from the online game *World of Warcraft*, launched by Blizzard Entertainment, a division of Activision Blizzard, in 2004. Using a description of the game's website, *World of Warcraft* is a "Massively Multiplayer Online Role-Playing Game (MMORPG), set in the high-fantasy universe centered around persistent online personae". In other words, a player chooses a character, develops him over time while gaming and uses him to explore the environment created by the game developer. It was the bestselling PC game of 2005 and 2006 and had more than 10.2 million subscribers worldwide in 2011.[4]

The firm designed the game in a way that applies all the motivation elements described previously in section 1. First, intrinsic motivation is created by the large simulated environment where consumers enjoy a storyline by making their chosen characters do tasks in the game. Second, the player advancement in the game is built as a complex reward system: a consumer gains access to more game content and features the more she plays. The virtual goods are also provided as a recompense for the user's accomplishments at specific stages of the game. We note that players can at any time terminate a gaming session without any penalty, with the restarting point being the last location and level that they had when they finished, which allows for rewards based on the entire time played in the game. With a specific speed chosen by the firm, players "level up" in the game after having earned a required number of experience points.[5] Additionally, the firm creates challenges that reset daily or weekly and can be repeated by users for additional rewards. The firm's website, containing news and game statistics, as well as frequent content updates about the storyline emailed at regular intervals to users, constitute two of the types of cues that the firm uses to prompt engagement. Eventually, as one progresses through the storyline, the game becomes complex and challenging, adding another motivational dimension.

The firm charges a fixed price for buying the software that needs to be installed on each player's computer and an additional fee for a 24 hours/7 days a week access to the server (game is not available on Tuesdays for about 3 hours when server maintenance is done). Most of the users pay these subscription fees using a weekly automated payment plan, and the price ranges from about 40 cents to 50 cents per day, depending on different payment schedules. The prices for gaming have not changed in the entire analysis period in our dataset. Users can buy additional virtual goods to use in the game, but these goods do not help advance in the game in any way, possessing a prestige nature.

[4] For more details about the game go to http://us.battle.net/wow/en/.

[5] There were 60 levels of the game at the time of our analysis.

The data is collected by a special program that logs into the game server every 10 minutes and records the ids of the gamers playing at the moment, as well as their current level, content area where consumers are playing, and affiliation with a game community, if applicable. The data set and the collection process are described in detail in Lee et al. (2011), a paper in computer science. We use a random sample of 956 individuals who started gaming in March-October of 2006. Additionally, the official game website provides detailed information about the reward system used in the game, as well as any other virtual good offerings. We next describe some of the patterns in the data.

Table 2 shows some descriptive statistics regarding product usage. Individuals on average play for about one hour and a half, and the idle periods between sessions last about 30 hours. We observe that the average number of days that players stay loyal to the game, from their initial starting date to the date when they completely abandon the game[6] is about 25 days. On average, consumers reach level 10 (out of a total of 60 levels), so there is a considerable percentage of users that drop out of the game at initial stages. Finally, across all these measures, the standard deviation shows that there is considerable heterogeneity across the user population.

Figure 1 shows additional details of product usage across sessions. The panel on the left shows the distribution of the duration of idle time between two game sessions, measured in hours, across all players. Conditional on a decision to start a game session, we observe that about 50% of the sessions to follow happen during the same day (about 8000 sessions), with less than 4 hours of idle time in between, and about 30% of them happen in consecutive days. This result is evidence that content consumption mostly occurs with high frequency, either daily or twice daily, on a schedule similar to a habitual activity.

The panel on the left of figure 1 shows the distribution of the length of gaming sessions, in minutes. We observe that about 30% of the sessions last less than 20 minutes, and about 60% last less than one hour. The remaining 40% can be considered long duration sessions, where consumers are likely to become involved in a number of different content activities, exploring the environment and accomplishing a number of tasks in the game. The longer sessions are also more likely to be associated with experienced players who have advanced to higher levels, where rewards are less frequent and progression slows down, while the game content becomes more complex.

[6] Defined as user not logging in for the rest of the time period.

3.2 Product Design and Reward Systems in Online Video Games

The main characteristics of video games include graphics, background and setting, duration of game, advancement rate, use of humor, control options, winning and losing features, character development, brand assurance, and multi-player features (Wood et al., 2004). In more general terms, these different aspect can be grouped into product design characteristics (e.g., sound, graphics, use of humor, brand assurance), which directly influence the intrinsic experience and enjoyment of playing the game, and stages and outcomes of a reward system (e.g., advancement rate, wining and losing features, character development), which motivate players by giving them goals that provide external rewards or benefits from long-term usage and progression in the game.

To describe player advancement and complexity requirements in more detail, Figure 2 shows the time required, in hours, to move up a level, for each of the 60 levels in the game. We see that initial levels are simple, easy to complete, and consequently advancement happens at a very fast pace. This is also a stage where consumers receive a lot of encouragement from the game developer, in the form of notices and messages in the game that provide guidance and motivation to keep consumers engaged. As the players progress in the game, they take longer at each level. On average, the last two levels take about 15 hours of playing time to be completed.

We note that this figure is a result of two different components. First, the firm has implemented a schedule in the product design such that the time to reach the next level of the game gradually increases. Second, consumers can spend time using the product, but doing activities that are not beneficial to advancement. An interesting aspect to note is that although the firm has implemented monotonically increasing time requirements, the figure shows some spikes and dips, a result of consumer usage behavior. For example, the firm has designed the last level to be the most demanding in terms of time to complete, but we see that in practice consumers are more efficient and focus on activities that help completing that level faster. In general, levels where we see a dip in the time required by users are also levels with large rewards from the game. For example, after completing the last level, level 60, users gain access to a large number of content areas and virtual goods that are not available at lower levels. These patterns illustrate forward-looking behavior about rewards and the influence of goal-oriented expectation on product usage.

Figure 3 provides additional details about some of the rewards included in the game, as well as an aggregate measure of the drop-outs at each level. We see that the firm has designed the product to have a clear schedule of rewards to encourage consumer loyalty and mastery of the game. Usually, at the levels indexed by the numbers that are multiples of 10, consumers have a strong reward, such as access to more content or a virtual good, but users also get additional rewards at other levels, depending on the firm's reward system design. A majority of players stop early in the game, mostly at the levels 2 to 15. This higher quitting rate in earlier stages happens because consumers are likely to be still trying to understand the game and if it is a good match for their preferences. After that, any decisions are likely to be more related to product usage.

4 Estimation

We now describe in more detail how our model is translated to the video games setting.

4.1 Implementation Details

The state of a gamer i ($i = 1, ..., N$) at any time t, $t \in [0, T]$, is described by the following: 1) the gaming status g; 2) the level in the game l[7]; 3) his habit, or involvement, level h. We now describe in more details the meaning of each of these state variables in the video game environment.

The choice set of the gamer is $A_0 = \{0, 1\}$, where $a = 1$ stands for "play", $a = 0$ denotes "stop", if the player is in the idle state $g = 0$, and $A_0 = \{0, 1, 2\}$ if he is in the active state $g = 1$, where $a = 2$ means "abandon the game". The actions results in the gaming states $g = a$ in a straightforward way. The difference between $g = 0$ and $g = 2$ is that $g = 2$ is an absorbing state, while $g = 0$ is just a temporal state of being "inactive", where leveling up in the game is paused until the gamer decides to "play".

There are $L = 60$ levels of the game that are available to a gamer to gain experience, enjoy different content, and master the different elements of the game. Progression to the next level in the game is exogenous to consumer decisions, since she cannot deliberately start playing at a higher level. Instead, the consumer has to earn the progression by investing her time into playing the game. Getting to the next level brings extrinsic rewards, both tangible and intangible. The tangible rewards can include

[7] In order to keep track of the inter-level progress (since some levels take long time to progress to) we also introduce an auxiliary state m, $m = 1, ..., 4$. The intra-level progress state m provides measure on the number of experience points accumulated by the gamer since the beginning of the game at level l. The number of points required to be accumulated at level l to progress to level $l + 1$ is normalized to 100 for any l. Within each level, and since the number of the experience points is not observed by the researcher, we assume that the points are earned at a constant rate across gaming time (for a given level l). With this assumption, the time that the gamer has already spent "active" ($g = 1$) at level l becomes an indication of the number of accumulated experience points. We discretize the stock of accumulated points into four groups. Thus, if by some time t the gamer has "actively" spent at level l less than 25% of the total time he "actively" spent at level l in total, he is assigned to be in the state $m = 0$ at time t. For clarity of explanation will further refer to state l only.

gaining new abilities, privilege to explore new territories, right to own certain property. Progressing to the next level per se, without getting any of the above mentioned tangible benefits, is defined to be an intangible reward. Every even level of the game provides tangible rewards, and every 10th level provides substantial tangible rewards. We mark the rewarding levels by the dummies "even level" and "10th" and estimate the coefficients on them to measure the value of those rewards. The value of intangible rewards provided by "empty" leveling is estimated as an intercept in the extrinsic utility function described earlier. Thus, the extrinsic utility is specified as $U_{extrinsic}(s) = \beta_1 + \beta_2 X_{even} + \beta_3 X_{10th}$. The intrinsic utility of gaming is specified as a linear function of a product usage level and the habit strength, i.e $U_{intrinsic}(s) = \alpha_1 + \alpha_2 l + \alpha_3 h$.

4.2 A Markov Chain in Continuous Time

Formally, there are three continuous-time processes evolving in the model: 1) the product usage process $G = \{G(t),\ t \geq 0\}$, 2) the game progression process $L = \{L(t),\ t \geq 0\}$, and 3) the habit process $H = \{H(t),\ t \geq 0\}$. These three processes constitute a multivariate continuous-time finite-state homogeneous Markov chain $S = (G, L, H) = \{G(t),\ L(t),\ H(t),\ t \geq 0\}$. Figure 4 shows a graphical representation of the chain. The process G has states $g = 0, 1, 2$; process L has state space $l = 1, ..., 60$; and process H has states $h = 0, ..., \bar{H}$.

The processes evolve as follows. When the product usage process G is in the state $g = active$, the game progression process L is allowed to evolve through its states; otherwise, the process L is on hold. In other words, consumers can reach a higher level only while playing the game. In its turn, the state l of the progression process L affects the state transition probabilities of the product usage process G. The reason is that the forward-looking consumers might have a higher propensity to play the game when they approach an especially rewarding level. The state transitions of the usage process G lead to the state transitions of the habit process H. That is, the product usage choices being made by a consumer build or break her habit. In its turn, the state h of the habit process affects the state transition probabilities of the usage process G, since a stronger/weaker habit results in more/less active product usage.

Not all three processes are equally well observed by the researcher. The progression process L is fully observed, since the information about the consumer's level in the game at any point in time is available. The product usage process G is partially observed. Finally, the habit process H is an underlying process that is not observed.

The reason that process H and process G are not fully observed is the existence of product usage choices that are inherently unobservable to the researcher. Consider a consumer being in the active state $g = 1$ for some period of time. Upon the end of the "flow", the consumer makes a choice between continuing the game session, stopping it, or abandoning the game for good. However, the researcher does not observe when the "flow" ends, as well as how many "flow" periods the consumer experienced during the period. If the consumer decides to continue the session, he will remain in the state $g = 1$. Therefore, the researcher will not observe that choice and the timing of the resulting state self-transition. Any other consumer choice would lead to a transition to a state $g \neq 1$, and thus the choice would be observable to the researcher. The same logic applies to the decisions and self-transitions in the idle state $g = 0$. Here, the consumer makes a choice upon the arrival of the cue. The cue arrival is unobservable to the researcher, and thus the decision to remain idle will lead to the unobserved choice and a state self-transition.

It is necessary to model the above unobservable choices for two reasons. First, having their structure in place allows to estimate how frequently consumers make product usage decisions, instead of enforcing the assumption about that frequency. The decision frequency has clear implications for the correct estimation of the consumer utility. Suppose, the assumed frequency is higher than the true one, then the consumer appears making choices against product usage (purchase) more often than he truly does, and as a result the utility is underestimated. Second, the unobserved product usage choices that are made during the idle period are crucial for the habit breaking process. The habit can become weaker when a consumer, prompted by an unobserved cue, decides to further abstain from using the product. The alternative approach to modeling the habit breaking process would be making it a function of time spent without using the product. However, that approach would be purely descriptive and thus unsuitable for the counterfactual analysis of habit formation.

4.2.1 Transition Rates During an Active Period

The state transition in mastery level, $(g, l, h) \rightarrow (g, l' = l + 1, h)$, is guided by the rate of the Poisson process described by the rate λ_l^{exp}, as discussed earlier. To specify the rates of the state transitions changing the usage state g and habit level h, we use the structure of the consumer decision-making process. While in state $g = 1$, a consumer engaged with the product for a duration that follows an exponential distribution with the rate parameter λ_h^{flow}. In other words, the end of the active state arrives with the rate λ_h^{flow} and the consumer then makes a choice $a \in \{0,1,2\}$ with probability $P(a|s)$. A choice $a = (1\,|g = 1, l, h\,)$ of continuing to use a product results in a self-transition to prolonging state $g = 1$. In most applications including ours, researchers do not observed the action $a = (1\,|g = 1, l, h\,)$ of remaining in the active state. The other two choices, $a = (0\,|g = 1, l, h\,)$, and $a = (2\,|g = 1, l, h\,)$, which are commonly observed by researchers, lead to the transitions to states $g = 0$ and $g = 2$, respectively.

Because of the unobserved action $a = (1\,|g = 1, l, h\,)$, the arrival rates of state transitions that we take to the data need to account for the probability that that action occurs, $P(a = 1\,|g = 1, l, h\,)$. Following Arcidiacono et. al. (2010), it can be shown that the waiting time until a transition from state $g = 1$ to each of the usage states $g = 0, 2$ through respective actions $a = 0, 2$, is distributed exponentially with the following rate parameters $\lambda_a(s)$:

$$\lambda_0(s) = P(a = 0|g = 1, l, h)\,(1 - P\,(a = 1|g = 1, l, h))\,\lambda_h^{flow}, \tag{10}$$

$$\lambda_2(s) = P(a = 2|g = 1, l, h)\,(1 - P\,(a = 1|g = 1, l, h))\,\lambda_h^{flow}. \tag{11}$$

Finally, the overall rate of leaving state $s = (g = 1, l, h)$, either because of a change in level l or because of action $a = 0$ or $a = 2$, is the negative of $\lambda(s)$ that is defined as:

$$\lambda(s) = \lambda_l^{exp} + \lambda_0(s) + \lambda_2(s) = \lambda_l^{exp} + (1 - P\,(a = 1|s))\,\lambda_h^{flow}. \tag{12}$$

4.2.2 Transition Rates During an Idle Period

If a consumer is not using the product and has also not chosen to abandon it completely, she is in state $s = (g = 0, l, h)$. In this state, she receives cues prompting her to use a product and make a choice to continue to be idle or to resume using the product, $a \in \{0, 1\}$.[8] If following a cue the consumer chooses $a = (1\,|g = 0, l, h\,)$, she goes into the active state. With a choice of $a = (0\,|g = 0, l, h\,)$, unobserved by the researcher, the consumer remains idle but as a consequence, there is a decrease in the consumer's habit strength with probability ϕ. Formally, the possible state transitions to other states from the idle state with habit level h are:

[8] For identification purposes, we assume that the consumer cannot abandon the product from the idle stage. He can do it from the active stage only.

1. $(g = 0, l, h) \rightarrow (g = 0, l, h' = h - 1)$, if the consumer chooses to remain idle and $h > 0$;[9]

2. $(g = 0, l, h) \rightarrow (g = 1, l, h)$, if the consumer chooses to go into the active state.

Note, that the state h is a function of unobserved decisions, and consequently the state transition in the case (1) above is not observed by the researcher. As before, the derivation of the state transition rates is based on the structure of the consumer decision-making process. A consumer receives cues via the Poison process with the arrival rate λ_h^{cue}. At that stage, he will either decide to continue to be idle and transition to a lower habit state $(h - 1)$ with probability ϕ or move to the active state $g = 1$, with the following respective rates $\mu_a(s)$:

$$\mu_0(s) = \lambda_h^{cue} P\left(a = 0|g = 0, l, h\right) \times \phi, \; with \, h > 0 \tag{13}$$

$$\mu_1(s) = \lambda_h^{cue} P\left(a = 1|g = 0, l, h\right). \tag{14}$$

The overall rate of leaving state $s = (g = 0, l, h)$ is the negative of $\mu(s)$ that is defined as:

$$\mu(s) = \mu_0(s) + \mu_1(s) = \lambda_h^{cue} \left(1 - P\left(a = 0|g = 0, l, h\right)\left(1 - \phi I\left(h > 0\right)\right)\right). \tag{15}$$

As noted above, while being inactive, a consumer may choose to continue to be idle multiple times until eventually becoming active again (these decisions are unobserved to the researcher). The duration of time until moving to the active state, accounting for any unobserved decisions that arrive at rate λ_h^{cue} during that period, follows the Coxian distribution. The Coxian distribution is a phase-type distribution that describes the time until reaching an absorbing state in a Markov chain. In this case, the idle state $g = 0$ constitutes the non-absorbing state and the active state constitutes the absorbing state of the Markov sub-chain that evolves during the idle time period. We note that although based on a Markov process, the Coxian distribution does not possess the "memoryless" property. The absence of the "memoryless" property is highly desirable here, since observasionally consumer propensity to start a new usage session depends on the amount of time that passed since the last usage session. In particular, we illustarte the dynamic effect of time and habit on the probability of consumers becoming active in the results section.

[9] If $h = 0$, and the consumer chooses to continue to be idle, then she would stay in the same state $(g = 0, l, h)$

4.3 The Density of the Observed Processes

As mentioned above, the processes in the model are of the three types: observed, partially observed, and unobserved. We address the challenge presented by the presence of the unobserved habit formation process following the approach suggested in Mark and Ephraim (2011).

Based on the structure of the partially observed usage process we derived the transition rates of the observed usage process in the preceding sections. We denote the observed usage process by G'. The state space of the process G' is exactly the same as the state space of the original usage process G, i.e. $g = \{0, 1, 2\}$. Thus, now all the processes are either observed (the progression process L and the observed usage process G') or unobserved (the habit process H). Hence, the observed process is two-dimensional and we denote it as $Z(t) = (G'(t), L(t))$ with a generic state $z = (g, l)$ and the state space $\bar{Z}$, $g = 0, 1, 2$ and $l = 1, ..., 60$. Thus, the state of the Markov chain can be written as $s = (z, h)$. Using the notations just introduced, the intensity matrix can be represented as a block matrix, $R = \{R_{zz'},\ z, z' \in \bar{Z}\}$, where $R_{zz'} = \{r_{zz'}(hh'),\ h, h' = 0, ..., \bar{H}\}$ are $\bar{H} \times \bar{H}$ matrices. The matrix $R_{zz'}$ contains the rates of transition between two observed states z and z' and all the unobserved states h. The intensity matrix R is shown in appendix B.

The density of the observed process $Z = (G', L)$ can be obtained from the transition density of the Markov chain. Following Mark and Ephraim (2011), it can be shown that the density of the observed process for one consumer can be computed as follows:

$$P\left(Z\left(t\right), 0 \leq t \leq T\right) = \boldsymbol{h_0} \left\{ \prod_{k=1}^{K_n} \boldsymbol{f}^{\boldsymbol{\tilde{z}_{k-1}\tilde{z}_k}} \left(\triangle_k\right) \right\} \boldsymbol{f}^{\boldsymbol{\bar{z}_K}} \left(T - t_K\right) \mathbf{1} \tag{16}$$

In the expression above, the bold font denotes matrices or vectors. In particular, $\boldsymbol{h_0}$ is a row vector of the initial habit state probabilities, $\bar{z}_k = (g, l)$ represents the realization of the observed processes of G' and L at the k-th jump of the observed Markov chain ($k = 1, ..., K$), $\triangle_k$ is the duration of time between the $(k-1)$-th and the k-th jumps, T is the total duration of time the processes were observed for, t_K is the time of the last jump of the chain, and $\mathbf{1}$ is a column vector of ones. The transition density matrix $\boldsymbol{f}^{\boldsymbol{\tilde{z}_{k-1}\tilde{z}_k}}\left(\triangle_k\right)$ of the k-th jump in the observed process is defined as follows:

$$\boldsymbol{f}^{\boldsymbol{\tilde{z}_{k-1}\tilde{z}_k}}\left(\triangle_k\right) = e^{R_{\tilde{z}_{k-1}\tilde{z}_{k-1}}\triangle_k} R_{\tilde{z}_{k-1}\tilde{z}_k}, \tag{17}$$

where $\bar{z}_{k-1} \neq \bar{z}_k$. The transition probability matrix over the last duration with no jump in the observed process $\boldsymbol{f}^{\bar{\boldsymbol{z}}_{\boldsymbol{K}}}(T-t_K)$ is defines as:

$$\boldsymbol{f}^{\bar{\boldsymbol{z}}_{\boldsymbol{K}}}(T-t_K)=e^{R_{\tilde{z}_{k-1}\tilde{z}_{k-1}}(T-t_K)}. \tag{18}$$

The matrix exponentials above are computed in the closed form for our empirical application, since the sub-intensity matrices $R_{zz'}$ are diagnosable. We describe computation of the transition densities in appendix C. We use forward-backward recursion in order to evaluate the above density[10]. The need arises due to the unobservable paths that the habit process H is taking.

4.4 Optimization

We use the mathematical programming with equilibrium constrains approach (MPEC) to estimate the model. In this approach, the density of the observed process is maximized subject to the constraint formed by the consumer value function. This approach allows to significantly economize on the time required to estimate the model since it allows to avoid the nested fixed point iterations on the value function (Dube et al, 2001).

The integral defining intrinsic utility in Section 2.4 has a closed form solution:

$$U^{\tau}_{intrinsic}(s)=\int_{0}^{\tau(s)}e^{-\rho t}U_{intrinsic}(s)dt=-\frac{U_{intrinsic}(s)}{\rho}\left(e^{-\rho\tau(s)}-1\right) \tag{19}$$

The value function is approximated using Monte Carlo integration by drawing $R=250$ draws of τ from the estimated exponential distribution of the length of the time intervals when consumer state does not change.

We employ a discrete representation of unobserved consumer heterogeneity and set the number of consumer segments to be $M=2$. The segments differ in the parameters of their utility functions. Let $P_{ij}\left(Z_i\left(t\right),0\le t\le T_i\right)$ be the density of the observed processes for consumer i $(i=1,...,N)$, conditional on individual i belonging to segment j, $(j=1,...,M)$. We denote the probability of consumer i belonging to segment j conditional on the observed data as $p(i\in j)$. Also, let Θ denote the parameters of the

[10] The forward density is defined as a row vector $L(k)$, where $L(k)=h_0$ and $L(k)=L(k-1)\boldsymbol{f}^{\tilde{\boldsymbol{z}}_{\boldsymbol{k-1}}\tilde{\boldsymbol{z}}_{\boldsymbol{k}}}(\triangle_k)$. The backward recursion is defined as $R(K+1)=\boldsymbol{f}^{\bar{\boldsymbol{z}}_{\boldsymbol{K}}}(T-t_K)\,\mathbf{1}$, $R(k)=\boldsymbol{f}^{\tilde{\boldsymbol{z}}_{\boldsymbol{k-1}}\tilde{\boldsymbol{z}}_{\boldsymbol{k}}}(\triangle_k)\,R(k+1)$. Then the density of the observed process is $P\left(Z\left(t\right),0\le t\le T\right)=L(k)R(k+1)$, as described in Mark and Ephraim (2011).

model to be estimated. Then, we write the maximization problem as

$$\max \sum_{i=1}^{N}\sum_{j=1}^{M} p(i \in j) log\left[P_{ij}\left(\{Z_i\left(t\right), 0 \le t \le T_i\}; \Theta\right)\right] \quad subject\ to:$$

$$V_i(s) \approx \frac{1}{R}\sum_{s=1}^{R}\left[U_{i,intrinsic}^{\tau_r}(s) + e^{-\rho\tau_r(s)}\left(\pi_l(s)\left[U_{i,extrinsic}(s) + V_i(s')\right] + \pi_a(s) ln\left(\sum_{a' \in A_g} exp(\psi_{i,a}^{s} + v_{i,a}(s))\right)\right)\right]. \quad (20)$$

In the approximated value function above, the $ln(\cdot)$ term is the closed form solution for the integral over the utility shocks e_a that follow extreme value distribution.

We use the habit state space $H = \{0, 1\}$ and, as mentioned above, specify the structure of the habit formation process that assumes that at the end of each game session a consumer has his habit reinforced through product usage, i.e. $h = 1$. Hence, the state of the habit process remains unobserved at any point in time except when a consumer ends his game session[11].

The EM algorithm is used to get the estimates of the sizes of consumer segments and the segment-specific parameters of consumer utility function. We use the KNITRO solver for non-linear optimization to estimate the model[12].

5 Results

In this section, we start by analyzing the parameter estimates and their implications in terms of consumer product usage behavior. Then we provide three counterfactual experiments that make use of our model to provide insights for the product design decisions and reward programs.

[11] With this assumption, we do not need to employ the Expectation-Maximization (EM) algorithm to estimate the transition probabilities of the unobserved process (in our case, the habit process), as described in Mark and Ephraim (2011).

[12] The KNITRO does not allow for drawing random variables within the optimization routine. However, the exponential distribution is closed under the scalar transformation, i.e. if $\tau \sim Exp(b)$ then $k\tau \sim Exp\left(\frac{b}{k}\right)$, where k is some positive scalar. Thus, before the start of the optimization, τ' are drawn from the exponential distribution with the rate parameter of 1, and the draws used for estimation are $\tau = (\hat{b})\tau'$, where$\hat{b}$ is the estimate of the rate.

5.1 Parameter Estimates

We discuss the parameter estimates, which are displayed in Table 3.[13] We divided the parameters into groups defined by the motivation components that explain choice: intrinsic motivation, extrinsic rewards, cue effects, and habit formation. The model was estimated with two discrete segments to control for unobserved heterogeneity. The choice of two segments was based on model fit, number of parameters, and computational requirements, when compared to alternative models with one and three segments.

We start by looking at the parameters related to the intrinsic motivation that measure the level of flow utility derived from playing the game for some duration of time. Comparing the segments, we find that segment 1 has a lower base value of intrinsic motivation to play the game than segment 2, but has a higher preference for later stages of content, with a value for the utility gain per level of 0.066 versus 0.018 of segment 2. Segment 2 enjoys the game more across all levels of content, and consumers in this group are likely to develop a strong habit for the game, which we analyze later in this section.

The parameters related to extrinsic rewards show that segment 1 draws more utility from obtaining rewards than segment 2, although segment 2 has a stronger preference for smaller and more frequent rewards. Combining this result and the findings described in the previous paragraph, we conclude that segment 1, with about 80% of the market, has a lower intrinsic utility but is goal-oriented and motivated by rewards, and could be named "reward seekers". Consumers in this segment use the product because they want to reach the next level, with a steeper increases for utility resulting from advancement and significantly more enjoyment from receiving the rewards. On the other hand, the remaining 20% of users in segment 2 enjoy the content more and are prone to habit formation, playing the game for its own sake. For gamers in segment 2, intrinsic motivations dominate over the impact of external rewards. For a more intuitive analysis, this segment is denoted as "content enthusiasts" for the remainder of this section.

Looking now at the parameters related to the frequency of consumer usage decisions, we see a clear difference between the rates in the active state and the idle state .[14]We computed the implied decision frequency based on these parameters, and table 4 shows the results. If the consumer is in a habit state $h = 1$, she gets a cue that prompts her to play every 7 hours, while a consumer in habit state $h = 0$

[13]With this set of parameters, the fit of the model was measured in several ways and for all of them, the model performed very well. For example, both product usage frequency in figure 1 and the drop out rates shown in figure 3 are predicted well.

[14]We chose not to introduce unobserved heterogeneity in this parameters since in our setting the decision opportunity arrival depends in part on the firm's actions.. However, it is possible that some consumer segments pay more attention to the usage cues originating from the firm or look for more cues than others.

receives a cue every 65 hours. This seems to be a reasonable result. Habit leads consumers to think more often about the game and to pay more attention to information relevant to the game, by visiting the game website or signing up for newsletters and email subscriptions, which would lead to more frequent occasions to make the decision to play or not to play.

Additionally, we also obtain estimates of how frequently consumers make decisions about stopping a game session while playing. We interpret the durations between two decision points as the time spent by players in the state of flow (Cowley, 2008). We find that consumers with a habit for gaming have shorter "flow", with the duration of 9 minutes, compared to 11 minutes for users with no habit. This suggests that habituated consumers are less immersed into the game, since for them the game demands less concentration given their frequent past experience, justifying a shorter time during which these consumers do not think about stopping their game session.[15]

To better interpret the parameter estimates related to habit formation, we look at the estimated probability of consumers maintaining the higher habit level after having finished a game session. The decay in habit happens when consumer decides not to play the game after having received cues prompting her to do so. The left panel of Figure 5 shows the evolution of the probability of moving to the no-habit state over time during the idle period, for each of the two segments. We observe that "content enthusiasts" have a significantly higher probability of staying in the high habit state for longer, by about 6 hours. On average, for the "reward seekers", it takes about 38 hours for the habit to fully disappear, and about 50 hours for content enthusiasts. This is an important finding for managers in related industries, since it provides input for the frequency of communications and cue provisions to the users to keep them engaged and with the additional motivation generated by the positive impact of habit.

The right panel in Figure 5 shows the probability of starting a new gaming session after stopping the game. We note that there is a rapid increase in the probability of having another gaming session in the first 8 to 24 hours after the last gaming session was ended by the gamer, especially for the "content enthusiasts". However, as time progresses, the probability of another session flattens out. This pattern occurs because the longer consumers remain idle, the higher the probability that they have moved from the higher to the lower habit state. This transition indicates that the positive impact of habit disappears and it will not affect consumer decisions at all after 38 and 50 hours for segment 1 and segment 2, respectively. Additionally, without habit, the cue arrival interval while not playing also

[15]This interpretation is similar to products that generate addiction. As individuals become more addicted, the effect of the addiction is reduced with additional usage.

increases considerably. Hence, our results show that losing the consumer's higher involvement created by habit has a double jeopardy effect: lower enjoyment from the game and lower frequency of occasions for deciding to return to play. We note that a priori assumptions about the frequency of decisions commonly used in discrete time models would miss this relevant insight.

5.2 Managerial Implications

Our analysis continues with the use of our approach to provide insights into three important managerial problems: (1) the frequency of cues to be implemented by the firm, (2) the level of complexity and effort required from players to obtain access to additional features and content, and (3) the choice of rewards programs, in terms of frequency and depth.

5.2.1 Usage Cues

One of the ways that firms can influence the engagement of consumers and increase usage of their product is through the management of cues. The cues prompt the decision to start a gaming session and lead to higher or more frequent engagement. In our empirical application, Blizzard Inc. can modify game content to require daily activities,[16] increase the frequency of availability of content, send emails and newsletters with reminders, or create forums with frequently updated game and player statistics.

Using our approach, we tested the effect of the additional cues on the length of idle time between two sessions. To do so, we created a counterfactual situation where we increased the rate of cue arrivals by 50%[17], and therefore, consumers are faced with more frequent opportunities to decide on starting a playing session. We computed both the counterfactual and actual transition probabilities between the idle and the active consumer states, which are shown in figure6. We can see that consumers react to additional cues by accelerating the timing of the next gaming session, and that this effect is especially pronounced for the "reward seekers", who are less intrinsically motivated to play the game and are less habit-prone. We also note that the impact of cues on the game participation is somewhat limited: even with an increase of 50% in the frequency of cues, the probabilities of starting playing session increased by only about 5% for segment 1 and 3% for segment 2. Even with more decision occasions, consumers are going to make their final usage choice based on the remaining motivations for usage, namely intrinsic

[16]The use of daily activities is also present in a large number of gaming apps in Facebook, for example, in the popular game FarmVille.

[17]In our model, the cue arrival rate is an exogenous measure to consumers and so we can test different values for the estimate parameter in counterfactual scenarios.

motivations and the reward valuation, which are not changed between the two scenarios.

5.2.2 Product Design: Complexity

In the second counterfactual, we explore a possibility of the firm changing the requirements necessary to access more advanced features and content. Our model accounts for this aspect of the product through the time needed to advance across the different levels of the game and reach higher, and more rewarding, content. We construct a counterfactual scenario where the advancement characteristic of the game design is modified: the time required to move from one level to the next is doubled, for all 60 levels. For example, in the actual game a consumer needs an average of 2.5 hours of gaming time to get from level 9 to l0, while in the counterfactual scenario the time required to do the same is 5 hours. This could be implemented by the firm by increasing the complexity of more advanced product features or game content.

Figure 7 displays a comparison between the drop-out rates by level for the two designs. The levels are combined in the groups of five for ease of exposition. We see an increase in the number of drop-outs from the very early levels in the game, especially in the 1-5 levels, while most of the remaining levels maintain very similar drop-out rates between the two scenarios. In the first five levels, the increase in the drop-outs leads to a loss of 80 players out of a total of 350 players, about 20% loss in participation. This result is a consequence of the fact that the new game design negatively influences the reward seekers, who have lower intrinsic motivation and need to have frequent rewards and advancement in the game to be motivated to stay engaged. With the new design, they are required to wait longer for each level to advance and to obtain the rewards, and thus they drop out earlier in the game than before. The segment of content enthusiasts who have high intrinsic motivation for gaming and habit are almost not affected by this change in the game design.

5.2.3 Reward System

Finally, we use our approach to test the impact of reward scheduling on product usage. As an example, we test the counterfactual situation where the reward at level 10 is removed from the game. In the game environment, rewards included at level 10 are an opportunity to have a virtual pet and participate in more challenging player versus player activities. Figure 1 shows the shirk rates of consumers in the actual and counterfactual scenarios. We observe interesting changes for the levels close to level 10, especially levels 6, 9, and 10. The largest increases in drop-out rates occur in level 6, which is the last level to offer a reward before level 10, and at the immediately preceding level 9. These larger increases in shirking in the absence of the reward have a reasonable explanation. Without a goal to look for at level 10, some reward seekers decide to abandon the game when they reach a good reward at level 6. Additionally, not having a goal to look for at level 9 also makes consumers drop out since they do not have the incentive that would be arriving soon, at the end of the next level. Matching the literature on intrinsic motivation and external rewards, we find that some of the later levels have slightly lower levels of shirk. This is because consumers that remain in the game are likely to play not just because of short-term goals but instead derive substantial enjoyment from playing, thus leading to lower incidence of shirking.

6 Conclusion

Our paper focused on explaining product usage and investigating how managers can use product features, cues and reward programs to increase consumption and reduce consumer shirking. We contribute to the marketing literature in two ways. First, we propose a dynamic structural model of product usage in continuous time that allows for the main drivers examined in the previous literature: extrinsic rewards, intrinsic utility, cue-based habit formation, and complexity and access to advanced features and content.

Second, we present a number of substantive findings relevant to managers of products where usage is critical to maintain revenues in the long-run: (1) we show that intrinsic and extrinsic rewards drive the product usage, but differently across consumer segments: a large consumer segment of “reward seekers” is motivated by specific goals and rewards in the game, while a smaller segment of “content enthusiasts” is driven by higher intrinsic motivation and is habit-prone; (2) we are able to estimate the frequency of cue arrivals that prompt consumers to decide whether to use a product, and find that habit plays a very important role in reducing idle time; (3) we test a number of different scenarios regarding marketing decisions that influence product usage, such as the use of cues to motivate participation, change in product complexity, and alternative reward programs.

References

1. Albuquerque, Paulo and Yulia Nevskaya (2012), "The Impact of Innovation on Product Usage: A Dynamic Model with Progression in Content Consumption", Working Paper.

2. Alter, Alexandra, "Your E-Book is Reading You". The Wall Street Journal, June 29, 2012, p. D1.

3. Arcidiacono, Peter, Patrick Bayer, Jason R. Blevins, and Paul B. Ellickson (2010), "Estimation of Dynamic Discrete Choice Models in Continuous Time," Economic Research Initiatives at Duke (ERID) Working Paper No. 50.

4. Berlyne, Daniel E. (1971), *Aesthetics and Psychobiology*, New York: Appleton- Century-Crofts.

5. Burdett, Kenneth and Dale T. Mortensen, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review* Vol. 39, No. 2 (May), pp. 257-273.

6. Calvo, G. A. (1983), "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics* 12, 383–398.

7. Csikszentmihalyi, Mihalyi (1990)*, Flow: The psychology of optimal experience.* New York: Harper and Row.

8. Cowley, B., Charles, D., Black, M., & Hickey, R (2008)*, "Toward an understanding of flow in video games*", Computers in Entertainment, 6(2), 1-27.

9. Doraszelski, U. and K. L. Judd (2008), "Avoiding the curse of dimensionality in dynamic stochastic games," Working paper, Harvard University.

10. Foote, Nelson N. (1966), "The Time Dimension and Consumer Behavior," in *On Knowing the Consumer,* ed. J. W. Newman, New York: John Wiley, 38-46.

11. Gordon, Chad, Charles M. Gaitz, and Judith Scott (1976), "Leisure and Lives: Personal Expressivity Across the Life Span", in *Handbook of Aging and the Social Sciences*, eds. Robert H. Binstock and Ethel Shanas, New York: Van Nostrand Reinhold, 310-341.

12. Holbrook, M. B. and E. C. Hirschman (1982), "The Experiential Aspects of Consumption: Consumer Fantasies, Feelings, and Fun", *Journal of Consumer Research*, Vol. 9, No. 2 (Sep.), pp. 132-140.

13. Huang, Guofang, Ahmed Khwaja, and K. Sudhir (2012), “Short Run Needs and Long Term Goals: A Dynamic Model of Thirst Management,” Working Paper.

14. Huh, Y. E. and S.-H. Kim (2008), “Do early adopters upgrade early? Role of post-adoption behavior in the purchase of next-generation products”, *Journal of Business Research*, Vol. 61, Issue 1, January, Pages 40–46.

15. Laibson, David I. (2001), “A Cue-Theory of Consumption”, *Quarterly Journal of Economics*, 116(1): 81-119.

16. Lakshmanan, Arun, Charles D. Lindsey, and H. Shanker Krishnan (2010), “Practice Makes Perfect? When Does Massed Learning Improve Product Usage Proficiency?,” *Journal of Consumer Research*, Vol. 37, No. 4 (December), pp. 599-613.

17. Leavitt, Clark, Anthony G. Greenwald, Carl Obermiller (1981), “What is Low Involvement Low In?,” *Advances in Consumer Research*, Volume 08, Association for Consumer Research, pp. 15-19.

18. Lee, Yeng-Ting Lee, Kuan-Ta Chen, Yun-Maw Cheng, and Chin-Laung Lei, “World of Warcraft Avatar History Dataset,” MMSys’11, February 23–25, 2011, San Jose, California, USA.

19. London, Manuel, Rick Crandall, and Dale Fitzgibbons (1977), "The Psychological Structure of Leisure: Activities, Needs, People," *Journal of Leisure Research* 9(4), 252-263.

20. Mark, Brian L. and Yariv Ephraim (2011), “An EM Algorithm for Continuous-time Bivariate Markov Chains”, George Mason University, Working Paper.

21. Novak et al. (2003), “The Influence of Goal-Directed and Experiential Activities on Online Flow Experiences,” *Journal of Consumer Psychology*, 13(1&2), 3–16

22. Shih, C.-F. and A. Venkatesh (2004), “Beyond Adoption: Development and Application of a Use-Diffusion Model,” *Journal of Marketing*, Vol. 68, No. 1 (Jan.), pp. 59-72.

23. Su, C.-L. and K. Judd (2011), “Constrained Optimization Approaches to Estimation of Structural Models,” *Econometrica*, Forthcoming.

24. Unger, L. and J. Kernan (1983), “On the Meaning of Leisure: An Investigation of Some Determinants of the Subjective Experience,” *Journal of Consumer Research*, Vol. 9, No. 4 (March), pp. 381-392

25. Wolf, Mark J. P., *Game Studies and Beyond Games and Culture*, January 2006 1: 116-118.

26. Wood, Wendy, and Neal, David T. (2007), "A new look at habits and the habit-goal interface," *Psychological Review*, Vol. 114, No. 4, 843–863.

27. Wood, W. and David T. Neal (2009), "The Habitual Consumer", *Journal of Consumer Psychology*, 19, 579-592.

Appendix A. Tables and Graphs

Paper	Approach and Data	Dependent Variable	Intrinsic Motivation	External Rewards	Usage Cues	Habit or Involvement	Innovation	Decisions Scheduling
Holbrook & Hirshman (1982)	Theory	Instant,Duration	Yes	Yes	Yes	Yes	No	N/A
Unger & Kernan (1983)	Survey	Instant	Yes	Yes	No	Yes	No	N/A
Laibson (2001)	Theory	Instant	Yes	No	Yes	Yes	No	N/A
Novak et al. (2003)	Survey	Duration	Yes	Yes	No	No	No	N/A
Shih & Venkatesh (2004)	Survey	Instant	Yes	No	Yes	No	Yes	Discrete
Huh & Kim (2007)	Survey	Instant,Duration	Yes	No	No	No	Yes	Discrete
Lakshmanan et al. (2010)	Experiment	Instant	Yes	No	No	No	No	N/A
Albuquerque & Nevskaya (2012)	Revealed	Instant	Yes	Yes	No	No	Yes	Discrete
This Paper	Revealed	Instant,Duration	Yes	Yes	Yes	Yes	No	Continuous

Table 1: Summary of literature.

	Mean	Std. Deviation
Game session duration, in minutes	87.8	109.1
Duration of the idle period between the two game sessions, in hours	31.5	130.1
Number of days spent playing the game	25.5	44.0
Level reached in the game	10.3	12.1

Table 2: Descriptive statistics for product usage patterns in the game.

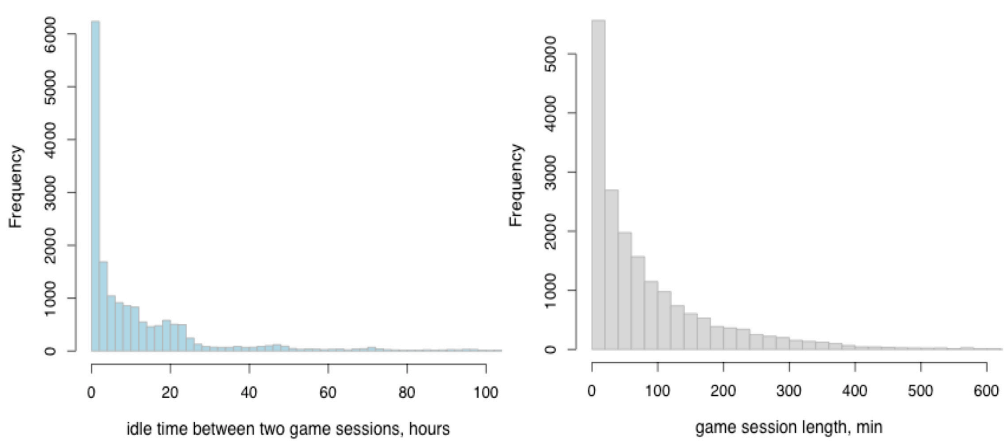

Figure 1: Duration of gaming sessions (right) and idle periods between game sessions (left).

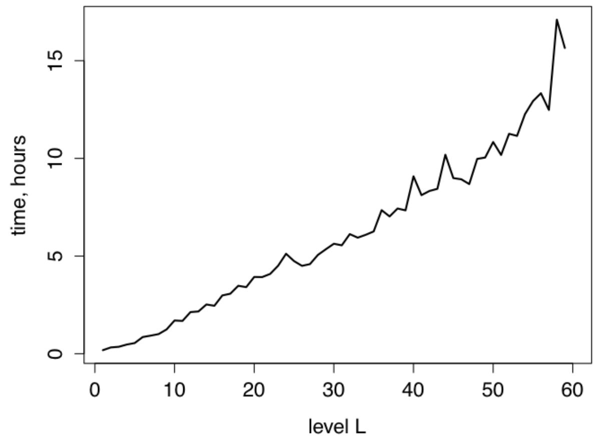

Figure 2: Average amount of gaming time observed before progression to the next level.

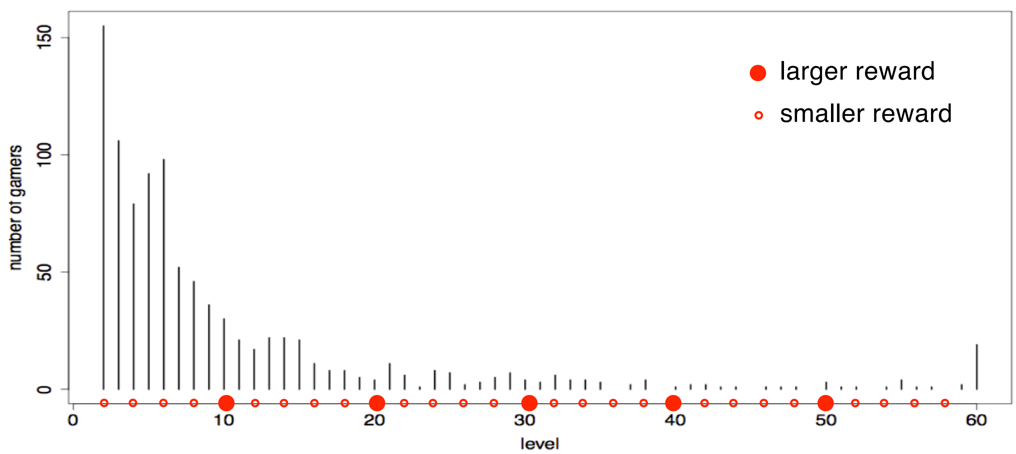

Figure 3: Histogram of the number of individuals reaching each of the 60 levels of the game and occurrence of large and small reward.

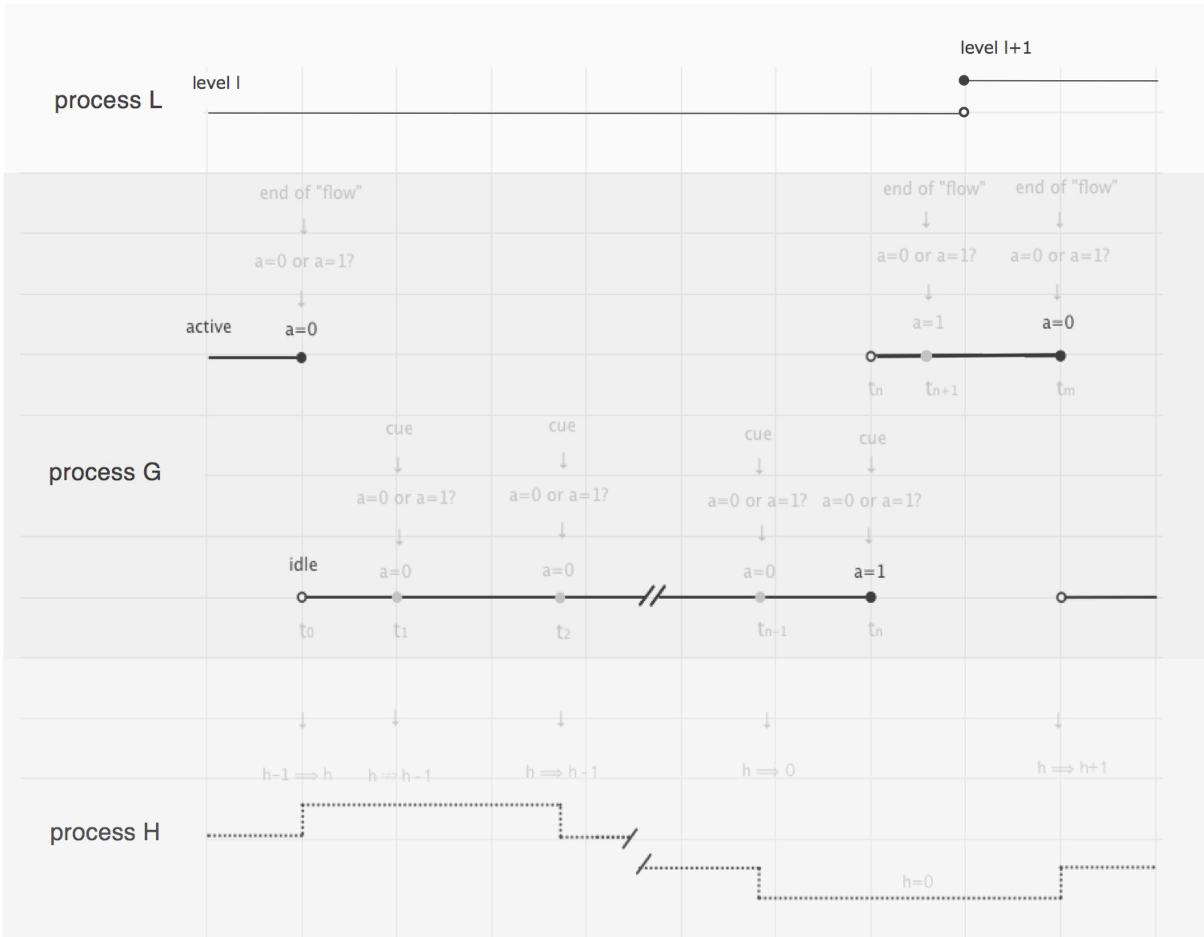

Figure 4: The three processes of the model: usage (G), progression (L), and habit formation (H).

		Segment 1	Segment 2
Intrinsic Motivation	Base	2.023*	5.970*
	Additional per Level	0.066*	0.018*
	Habit “Bump”	0.035*	0.420*
Extrinsic Rewards	Base	724.950*	251.349*
	Even level (small)	30.321**	154.972**
	Every 10th Level (large)	77.465*	93.172*
Cue for Decisions	Not playing, no habit	0.129**	0.129*
	Playing, no habit	46.696**	46.696**
	Not playing, habit	1.251**	1.251**
	Playing, habit	55.078**	55.078**
Habit	Probability of drop in habit	0.606*	0.606*
Segment Size		0.783**	0.217**

Table 3: Parameter estimates; “**”:significant at 2.5%; “*”significant at 5%.

Arrival Rates of Usage Decisions	Estimate	
During an Idle Period	with habit	1 every 7 hours
	without habit	1 every 65 hours
During a Game Session	with habit	1 every 11 min
	without habit	1 every 9 min

Table 4: Parameter estimates - Frequency of usage decisions. "**":significant at 2.5%; "*"significant at 5%.

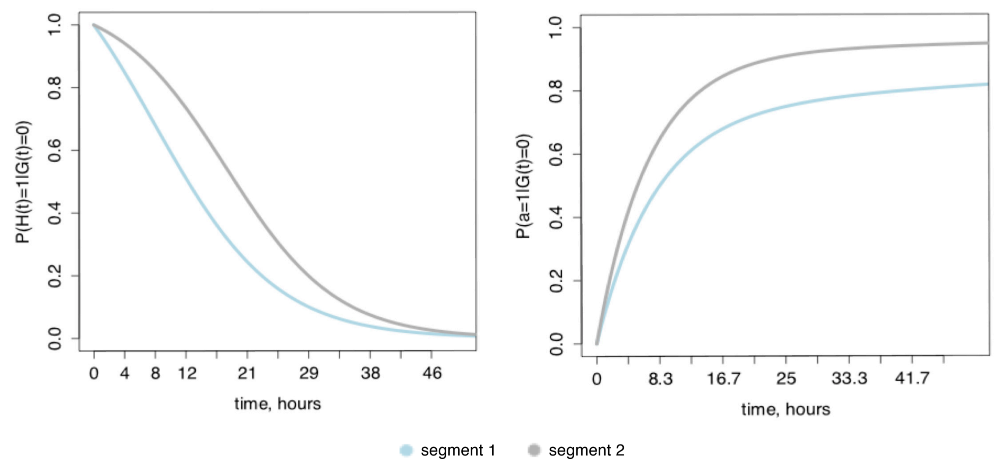

Figure 5: Habit decay (left) and probability to start a new gaming session (right) during idle period.

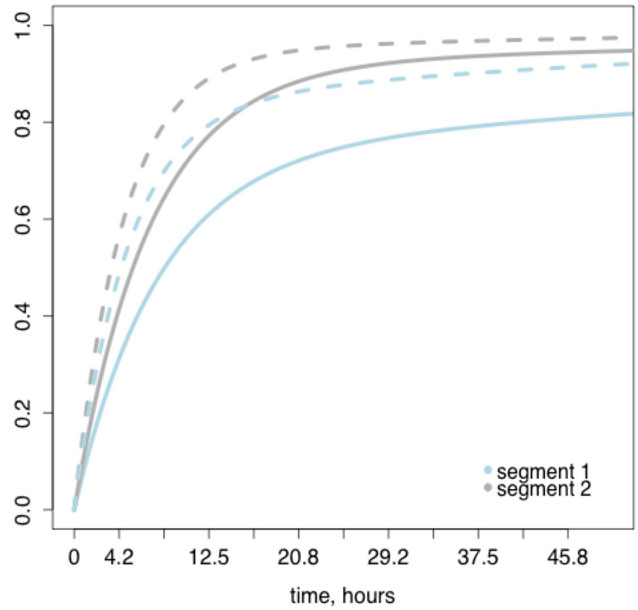

Figure 6: Probabilities of starting a new gaming session for two alternative scenarios: actual cue frequency (solid lines) and an hypothetical increase of 50% of cue arrivals (dashed lines)

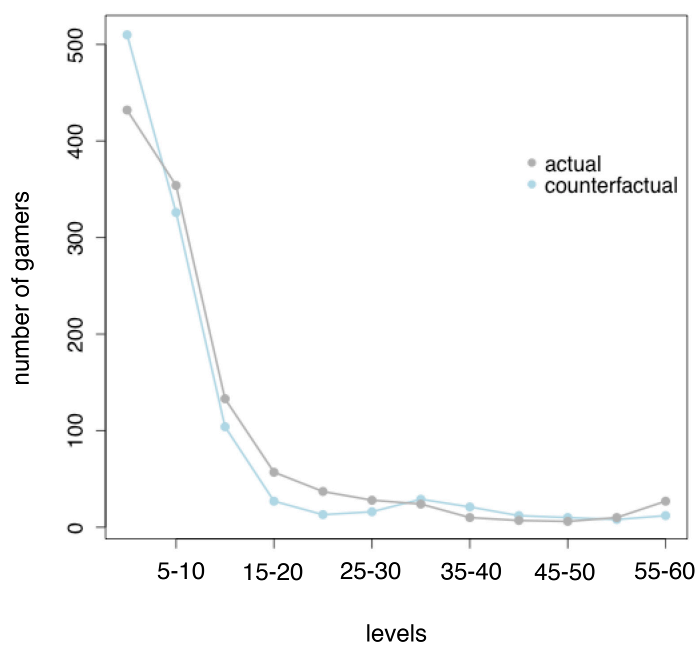

Figure 7: Drop-out rates for the actual setting and a counterfactual case where the complexity in increased.

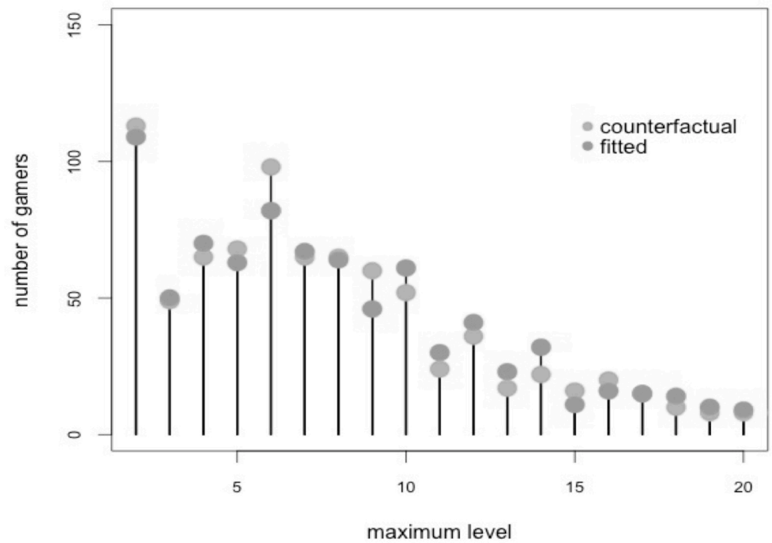

Figure 8: Drop-out rates of players per level with a reward (darker circles) and without a reward (lighter circles) at level 10.

Appendix B

The Intensity Matrix R

$s=(z,h)$		$z_1=(g=0,l=1)$		$z_2=(g=0,l=2)$		$z_3=(g=1,l=1)$		$z_4=(g=1,l=2)$		$z_5=(g=2,l=1)$		$z_6=(g=2,l=2)$	
		$h_1=0$	$h_2=1$	$h_1=0$	$h_2=1$	$h_1=0$	$h_2=1$	$h_1=0$	$h_2=1$	$h_1=0$	$h_2=1$	$h_1=0$	$h_2=1$
$z_1=(g=0,l=1)$	$h_1=0$	$-\mu_1(z_1,h_1)$	0	0	0	$\mu_1(z_1,h_1)$	0	0	0	0	0	0	0
	$h_2=1$	$\mu_h(z_1,h_1)$	$-\mu(z_1,h_2)$	0	0	0	$\mu_1(z_1,h_2)$	0	0	0	0	0	0
$z_2=(g=0,l=2)$	$h_1=0$	0	0	$-\mu_1(z_2,h_1)$	0	0	0	$\mu_1(z_2,h_1)$	0	0	0	0	0
	$h_2=1$	0	0	$\mu_h(z_2,h_1)$	$-\mu(z_2,h_2)$	0	0	0	$\mu_1(z_2,h_2)$	0	0	0	0
$z_3=(g=1,l=1)$	$h_1=0$	0	$\lambda_0(z_3,h_1)$	0	0	$-\lambda(z_3,h_1)$	0	0	$\lambda_{l=1}^{exp}$	0	$\lambda_2(z_3,h_1)$	0	0
	$h_2=1$	0	$\lambda_0(z_3,h_2)$	0	0	0	$-\lambda(z_3,h_2)$	$\lambda_{l=1}^{exp}$	0	0	$\lambda_2(z_3,h_2)$	0	0
$z_4=(g=1,l=2)$	$h_1=0$	0	0	0	$\lambda 0(z_4,h_1)$	0	0	$-\lambda(z_4,h_1)$	0	0	0	0	$\lambda_2(z_4,h_1)$
	$h_2=1$	0	0	0	$\lambda_0(z_4,h_2)$	0	0	0	$-\lambda(z_4,h_2)$	0	0	0	$\lambda_2(z_4,h_2)$
$z_5=(g=2,l=1)$	$h_1=0$	0	0	0	0	0	0	0	0	0	0	0	0
	$h_2=1$	0	0	0	0	0	0	0	0	0	0	0	0
$z_6=(g=2,l=2)$	$h_1=0$	0	0	0	0	0	0	0	0	0	0	0	0
	$h_2=1$	0	0	0	0	0	0	0	0	0	0	0	0

Note: We show the intensity matrix R for $h=0,1$ and $l=1,2$ for the sake of brevity, but extension to a larger state space is straightforward.

Appendix C. Transition Density Matrices

Estimation of the model requires computation of transition density matrices, as indicated in Section 4.3. We provide an example of computing one of the transition matrices needed for the model. Others can be computed analogously.

Suppose, the $(k-1)$-th observed event for consumer i was ending a game session at some time t while at level 2 and the k-th event happened in 47 minutes when the consumer started another game session. Thus, we observe a transition from observed state $\tilde{z}_{k-1} = z_2 = (g=0, l=2)$ to observed state $\tilde{z}_k = z_4 = (g=1, l=2)$ with the transition duration being $\triangle_k = 47$. The subscripts on the z state correspond to the ordering of states z in the intensity matrix R (see appendix B). Note, that the habit state h is unobserved, and a consumer can be in any state $h = 0, ..., \bar{H}$ at time t, as well as it might had been changing during those 47 minutes that the consumer was inactive with the game. We define $\bar{H} = 1$ in this application.

The probability density for the described transition is:

$$\boldsymbol{f}^{\boldsymbol{z_2 z_4}}(\triangle) = e^{\triangle R_{z_2 z_2}} R_{z_2 z_4},$$

where the intensity sub-matrix $R_{z_2 z_2}$ describes the transition rates of the unobserved habit state h while the observed state $z = z_2$, and the sub-matrix $R_{z_2 z_4}$ describes the rates of change in both the observed state z and the unobserved state h. The two matrices are:

$R_{z_2 z_2}$:

$s:$	(z_2, h_1)	(z_2, h_2)
(z_2, h_1)	$-\mu_1(z_2, h_1)$	0
(z_2, h_2)	$\mu_h(z_2, h_1)$	$-\mu(z_2, h_2)$

$R_{z_2 z_4}$:

$s:$	(z_4, h_1)	(z_4, h_2)
(z_2, h_1)	$\mu_1(z_2, h_1)$	0
(z_2, h_2)	0	$\mu_1(z_2, h_2)$

The matrix exponential e^X, where X is a square matrix, can easily be computed in a closed form provided that the dimensions of X are not prohibitively high and X is diagonizable. The matrix X is diagonizable if it can be represented as $X = Y \Sigma Y^{-1}$, where Σ is a diagonal matrix with the diagonal elements being the distinct eigenvalues of matrix X and the off-diagonal elements being 0, and Y is a matrix which columns are the eigenvectors corresponding to the eigenvalues at the diagonal of matrix Σ. In this case, the matrix exponential is:

$$e^X = Y e^{\Sigma} Y^{-1}.$$

Note, that if X' is a $d \times d$ diagonal matrix with all the off-diagonal elements being 0, then:

$$e^{X'} = \begin{bmatrix} e^{x'_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{x'_d} \end{bmatrix}$$

For example of the transition above, the eigenvalues of matrix $R_{z_2 z_2}$ are $\xi_1 = -\mu_1(z_2, h_1)$ and $\xi_2 = -\mu(z_2, h_2)$. The corresponding eigenvectors are of the form:

$$\xi_1\colon \quad \begin{bmatrix} 0 & w_1 \end{bmatrix} \qquad \xi_2\colon \quad \begin{bmatrix} w_2 & \dfrac{-\mu_h(z_1, h_1) \cdot w_2}{-\mu(z_1, h_2) + \mu_1(z_1, h_1)} \end{bmatrix},$$

where w_1 and w_2 can be any numbers. For convenience, we let $w_1 = \mu_h(z_2, h_1)$ and $w_2 = -\mu(z_1, h_2) + \mu_1(z_1, h_1)$ and therefore specify the eigenvectors to be:

$$\xi_1\colon \quad \begin{bmatrix} 0 & \mu_h(z_2, h_1) \end{bmatrix} \qquad \xi_2\colon \quad \begin{bmatrix} -\mu(z_1, h_2) + \mu_1(z_1, h_1) & -\mu_h(z_1, h_1) \end{bmatrix},$$

Then the matrices Σ and Y are:

$$\Sigma = \begin{bmatrix} e^{-\mu_1(z_2, h_1)} & 0 \\ 0 & e^{-\mu(z_2, h_2)} \end{bmatrix} \qquad Y = \begin{bmatrix} 0 & -\mu(z_1, h_2) + \mu_1(z_1, h_1) \\ \mu_h(z_2, h_1) & -\mu_h(z_1, h_1) \end{bmatrix}$$

and the matrix exponential $e^{\triangle R_{z_2 z_2}}$ becomes:

$$e^{\triangle R_{z_2 z_2}} = \left(Y e^{\Sigma} Y^{-1}\right)^{\triangle} = \begin{bmatrix} e^{-\mu_1(z_2, h_1)\triangle} & 0 \\ \dfrac{\mu_h(z_2, h_1)\left[e^{-\mu_1(z_2, h_1)\triangle} - e^{-\mu(z_2, h_2)\triangle}\right]}{-\mu(z_1, h_2) + \mu_1(z_1, h_1)} & e^{-\mu(z_2, h_2)\triangle} \end{bmatrix}$$

Given the matrices Y, Σ, and $R_{z_2 z_4}$ above, the transition density $\boldsymbol{f}^{\boldsymbol{z_2 z_4}}(\triangle)$ is computed.