From Free to Premium: A Dynamic Model of Product Usage and Conversion in Freemium Strategy

Research Proposal

Felix Nguyen

Wisconsin School of Business

University of Wisconsin - Madison

May 2019

1 Introduction

In his 2009 book "Free: The Future of a Radical Price", entrepreneur and former Editor-in-chief of WIRED magazine, Chris Anderson, talks about a new era of the business world. An era where savvy firms would meteorically become multi-billion dollar giants by riding the free business model (Anderson, 2009). This prediction soon became a reality, with the spectacular growths of companies such as Spotify, Slack, or Box. The business strategy Anderson discussed in his book has become widely known today as "freemium", a portmanteau of two seemingly opposite words: "free" and "premium".

Even though the term "freemium" was coined only around a decade ago, history of the concept itself goes back to the early 1980s, when software firms like Adobe started to offer a "lighter" version of their software, called "sharewares", or "cripplewares", with limited functionalities and were free of charge (Wagner et al., 2014). In contrast with its predecessor, the time-limited trial model, the free version offered in freemium model is *perpetually free*, which means there will always be a free segment, and the main goal is not converting all of customers to paying ones, but spreading product adoption and use the paid users to cover the costs of free users (Lee et al. 2017).

Since then, “freemium” has grown from a niche idea to help customers get used to specialty software to one of the most popular models of Internet-based product. In the early 2000s, email providers, such as Yahoo or AOL, offered users a free email service that was financed by advertisements. For a monthly fee, the service was advertisement-free and provided almost unlimited online storage space. More recently, Content-as-a-Service (CaaS) and Software-as-a-Service (SaaS) firms such as Spotify, Box, Skype... all rely on ‘freemium to quickly gain extensive userbases in short periods of time (Shi et al., 2019). Freemium also became particularly popular with mobile app market, with 80% top grossing iOS apps following this model (Lee et al., 2017). Within the last few years, we can see larger, more traditional tech giants are also gravitating towards this model, with LinkedIn offering Premium services to both job seekers and recruiters, Youtube launching Youtube Premium to watch videos without ads, or Amazon creating a new free tier for their Amazon Web Services. We can also see independent content creators started to embrace this model by offering free contents through their public social media channels and premium contents to their paid subscribers, through platforms such as Patreon. These ongoing trends are discernible validations of the success of freemium model.

Freemium, as its core, is a customer acquistion strategy. It is often used by firms because of its ability to attract a large amount of customers over a short period of time. This is particularly attractive to smaller sized firms, or startups, with limited resources to invest in marketing activities or building up large sales force. As a consequence, most of the users acquired through this model are non paying customers. A Techcrunch’s article, ”Should Your Startup Go Freemium?”, states that conversion rate from free to premium for most startups hovers around 1 to 10 % (Maltz & Barney, 2013), similar to Anderson’s predict of 95:5 ratio in his book (Anderson, 2009). There is also substantial heterogeneity in conversion rate, with Spotify averaging around 27% while only 4% of

DropBox and 0.5% of Google Drive users are fee-paying customers.

This generates a wide variety of oppositions and uncertainties around the freemium model. The main hurdle that businesses have to overcome when employing this model is achieving a certain premium to free ratio, so the revenue from fee paying customers could reasonably cover the cost incurred by free users. A Wall Street Journal's article, "When Freemium Fails", interviewed several entrepreneurs who moved away from the freemium model, calling it "a costly trap", forces a startup to sacrifices revenue and support freeloaders who will never become paying customer. In the same article, one entrepreneur noted: "freemium is like a Samurai sword: unless you're a master at using it, you can cut your arm off." (Needleman Loten, 2012). As user acquisition cost is rising fast for online and mobile product, with average cost per install of mobile games reaching $US 4.36 (Sifa et al., 2015), optimizing conversion rate is becoming more and more important for firms with freemium model.

Given this emerging challenge, the ability to define, detect, and predict users' behavioral attributes related to conversion as early as possible has become an important factor of success in modern internet and mobile markets (Hadiji et al., 2014; Runge et al., 2014). Understanding who could be converted to premium users, what affects their conversion choice, and when should a firm target them for conversion, are all steps in building an accurate customer's life time value prediction. This is the motivation for my main research goal in this paper: Understanding the relationship between users' behavior in a freemium environment and their conversion rate from free to fee-paying, premium tier. In particular, I am looking to capture the dynamics between individual customers' product or application usage within a time period and their potential conversion rate from free to premium, i.e. whether they are both affected by a common variable.

2 Literature Review

This paper intersects between several substantive research streams within marketing and information system literature: (1) *Product sampling*; (2) *Freemium*; (3) *Discrete-choice model*; and finally (4) *Temporal model of consumer behavior*.

2.1 Product Sampling

The first relevant stream of literature is product sampling, which includes freemium as well as product trial in digital products. Prior researches on digital goods have established that most of these are *experience goods*, and as such customers require time to obtain values from them, chiefly through trial processes (Lehmann and Esteban-Bravo, 2006; Heiman et al., 2001; Chellappa and Shivendu, 2005). Under product sampling, customers update their priors on added utilities of premium features based on their own experience with basic versions of the products. Built on this assumption, most of literature on product sampling focused on exploring how firms can improve adoption rate by educating customers of product values, either through demonstration (Heiman and Muller, 1996; Heiman et al., 2001), or free trial and sampling (Bawa and Shoemaker, 2004; Cheng and Tang, 2010; Dey et al., 2013). These studies, however, differ from the main context of this paper, since they mostly concern with how to educate customers through limited time free offering, instead of how to convert customers from perpetually free version to fee paying premium version.

2.2 Freemium

Recently, a new research stream is emerging to tackle issues specific to the freemium model, either through theoretical models (Niculescu and Wu, 2014; Kamada and Ory, 2019; Shi et al., 2019), or empirical research (Liu et al., 2014; Wagner et al., 2014; Lee et al., 2017). Kamada and Ory (2019) and Shi et al. (2019) examine the referral behavior of customers in a freemium context through micro-economics models, and investigate whether the network effects generates by freemium are efficient in generating Word-of-Mouth. The two papers also look at the trade-off between incremental Word-of-Mouth and cannibalization of premium version. Similarly, Niculescu and Wu (2014) examines the differences between two freemium implementations: free-limited freemium and uniform seeding. While these provide very useful managerial insights for firms to consider whether or not to employ freemium model, they look at the problem from a firm-side perspective and are not as helpful for firms who already follow such model, which, as mentioned above, account for up to 80% of the market.

There is still a significant paucity in studies, both theoretical and empirical, that focus on the consumer-side of freemium model, especially on how it affects consumer choice behavior and how to optimize conversion from free to premium. Earlier literature has shown that instead of simply picking a choice with the highest cost-benefit difference, people tend to perceive the benefits associated with free products as higher than the same benefits of paid products (Shampanier, Mazar, and Ariely, 2007: Niemand et al., 2015). Other studies look at how functional design of different versions influence consumer behavior, and signify the importance of functional fit between free version and premium version, as well as consistency in quality in both versions (Liu et al., 2014; Wagner et al., 2014). Liu et al. (2014) also found that freemium model could dilute the negative effect of product reviews in mobile app market.

However, these are still focusing on the advertising effect of freemium, and, as far as I know, the only study into conversion from free to premium is a working paper by Lee et al. (2017). The paper develops a structural model to study the design of freemium, with regards to a wide range of customer behaviors including adoption, upgrade, referral, and usage. This paper shares similar elements with Lee et al. (2017), however, instead of identifying the strategic balance between growth and monetization in freemium model, I focus on a more narrow issue of identifying conversion opportunity based on temporal usage data.

2.3 Temporal Model of Consumer Behavior

Methodically speaking, this paper is related to a growing stream of marketing literature: models of consumer behavior with inter-temporal dynamics. These dynamics often involves a latent structure that is difficult to capture with observational data, and prior marketing and management science literature has come up with various ways to model these dynamics, which could be divided to continuous and discrete state space structures.

A continuous state space structure is often employed if one could reasonably assume that the dynamic is gradual and smooth. A wide variety of techniques have been employed by researchers to model this continous state space, including time series autoregressive model (Dekimpe and Hanssens, 2000; Pauwels et al., 2004), or exponentially smoothed sum of past usages, which is widely used in choice modeling (Guadagni and Little, 1983; Srinivasan and Kevasan, 1976), or a conditional hazard model (Borle et al., 2008). A common feature of these approaches is that they all calculate the current value of variable of interest as a direct function of its past values.

While these approaches could be very efficient and accurate, the continunous state space model is adequate in capturing dynamics which evolve in a more discrete manner, such as instant shift in customer preferences (Netzer et al., 2008). For the context of this paper, one example could be a customer suddenly become more involved in a video game because his or her friends just joined in, or another customer lose interest in the game because they just switch to a different one. Therefore, in those scenarios, a model that allows customers to transition between discrete set of states is a more appropriate approach. The simplest implementation of this approach is the state-dependent model, whereby the observed previous choice is included in the model as the current state of customer (Heckman, 1981, McAlister et al., 1991). More sophiscated models could also be applied in the same manner, for example, Carraway (2000) captures dynamics in consumer lifetime value using a Markov model of observed purchase recency states.

However, these models are flawed because they often ignored the external sources of dynamics such as effects of marketing stimuli, and the states of the concerned dynamics may not be observable to researchers (Netzer et al., 2008). In recent year, more and more marketing studies have attempted to alleviate this through the usage of Hidden Markov Model (HMM). This is a stochastic model that describe transition between a set of latent discrete state, and at each discrete period, an observation is emitted based on the state the subject is currently in. A detailed overview and tutorial of HMM could be found in Rabiner (1989). A few examples of HMM's applications in marketing literature include: Montgomerry et al. (2004) studies web-path analysis using a time-continous HMM model; Netzer et al. (2008) uses HMM to model firm-customer relationship dynamics; Ascazar and Hardie (2013) employs a joint HMM of usage and churn in contractual setting, and so on. These new developments prove the usefulness of Hidden Markov Model in modeling temporal dynamics within consumer behavior.

3 The Model

For the model, the basic underlying assumptions are: (1) Observed usage and consumer choice to purchase premium features both reflect an individual's *level of interest* with the product or service, which itself is latent; (2) this latent variable is discrete in nature, and it stochastically evolves over time; (3) Our observed usage data reflects variation in interest in a specific moment in time; and (4) the choice to purchase premium add-ons also reflect this variation.

I employ a Hidden Markov Model to describe the above dynamics between product usage, customer's level of interest in that product, and the choice of purchasing premium features. This model includes four basic components as described below:

- *Initial State Distribution:* This is the probability of a customer i is in state s at time 1.

- *Inter-state Transition Probabilities:* A sequence of Markovian transition probabilities that express whether the level of interest of a customer i in a previous period *t-1* is strong enough to make them transition to the next state at time t.

- *State-dependent Product Usage:* This is our observed data - The time individual i spend using the product at a specific period t, conditional on state s.
- *State-dependent Premium Purchase:* This is the probability that customer i in state s would purchase a premium feature from a list of features at time t.

Next, I will go into more details about each of these components.

3.1 Markovian Transition Matrix

The transition matrix for an individual i at time t is a matrix of probabilities i would move between different states. I assume K number of hidden states of customer's interest in the product, and these states differ with respect to product usage and likelihood a customer would purchase premium add-ons. In order to model the variation in customer interests over time, I allow customers to transition between states when they move from time t to time *t+1*. From this assumption, the evolution of S_{it}, the state customer i is in at time t, could be described through matrix Q_{it} below:

$$Q_{it} = \begin{bmatrix} q_{(1,1)it} & q_{(1,2)it} & \cdots & q_{(1,K-1)it} & q_{(1,K)it} \\ q_{(2,1)it} & q_{(2,2)it} & \cdots & q_{(2,K-1)it} & q_{(2,K)it} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ q_{(K-1,1)it} & q_{(K-1,2)it} & \cdots & q_{(K-1,K-1)it} & q_{(K-1,K)it} \\ q_{(K,1)it} & q_{(K,2)it} & \cdots & q_{(K,K-1)it} & q_{(K,K)it} \end{bmatrix}$$

Where

$$P(S_{ii} = s' | S_{i(t-1)} = s) = q_{(s,s')it} \text{ for s, s' } \in \{1, ..., K\} \tag{1}$$

I assume that the transition from one state to another follows a Markov process; that is, the transitions between states are stochastically determined through the conditional transition probability to another state given the current state.

We can model each elements of matrix Q_{it}, using a multinominal logit model (Ascarza, 2018), as follow:

$$q_{(s,s')it} = \frac{exp(\mu(s,s')_i + \mathbf{x^q_{it}}\beta^{\mathbf{q}}_{\mathbf{s,s'}})}{1 + \sum_{j=1}^{K-1} exp(\mu(s,j)_i + \mathbf{x^q_{it}}\beta^{\mathbf{q}}_{\mathbf{s,j}})'} \tag{2}$$

$$q_{(s,K)it} = \frac{1}{1 + \sum_{j=1}^{K-1} exp(\mu(s,j)_i + \mathbf{x^q_{it}}\beta^{\mathbf{q}}_{\mathbf{s,j}})'} \tag{3}$$

The parameter $\mu(s,s')_i$ denotes individual specific propensity to move from state s to s', $\mathbf{x^q_{it}}$ is a vector of time-varying covariates that could potentially affect the transition process, and $\beta^{\mathbf{q}}_{\mathbf{s,s'}}$ is the vector of corresponding coefficients of these covariates.

In order to account for customer heterogeneity in their propensities to transition between states, as customers may be different in their "stickiness" to an interest state, and some maybe more susceptible to changes than others, we can allow the transition propensity to vary across individuals:

$$\mu(s,s')_i = M(s,s') + \delta_{(s,s')i} \ , \tag{4}$$

in which $M(s,s')$ is the average transition propensity from s to s', and $\delta_{(s,s')i}$ is individual heterogeneous adjustment. We then have the following hyperparameters:

$$\mathbf{M} = [M(1,1), ...M(1,K-1), ..., M(K,1), ..., M(K,K-1)] \tag{5}$$

$$\delta_{\mathbf{i}} = [\delta_{(1,1)i}, ..., \delta_{(1,K-1)i}, ..., \delta_{(K,1)i}, ..., \delta_{(K,K-1)i}] \tag{6}$$

3.2 State-Depedent Product Usage

Due to the nature of digital products such as mobile games or streaming services, it is reasonable to expect that there would be a significant portion of our observed data where a customer i have a product usage of zero at time t. Many customers would play a game only a few times a week or a month, for example. To account for the potentiality of two underlying processes, one determines whether customers use the product at all at time t, and one determines the usage level, I assume that product usage follows a Poisson distribution with parameter:

$$\lambda_{it}|[S_{it} = s] = \pi_s \gamma_i exp(\omega^{\mathbf{z}}\mathbf{z_i} + \omega^{\mathbf{u}}\mathbf{u_{it}}) \tag{7}$$

The variable γ_i represents heterogeneity across population in product usage, allowing two customers with same interest level and same set of covariates varies in level of usage. Following Ascarza (2013), this variable is assumed to follow a lognormal distribution with average = 0 and standard deviation σ_γ for computational convenience purpose.

Next, the vector $\boldsymbol{\pi} = [\pi_1, \pi_2, ..., \pi_K]$ is a vector of state-specific mean usage parameters, represents the change in usage behavior stemmed from the transition between states. As λ_{it} must > 0, π_s is constrained to be > 0 as well. Furthermore, as usage behavior should positively correlated with change in level of interest, $\pi_1, \pi_2, ..., \pi_K$ should monotonically increasing, i.e. $0 < \pi_1 < \pi_2 < ... < \pi_K$.

Equation (7) also include covariates that may affect customer usage behavior of a specific time period, denoted by $\mathbf{z_i}$, the vector of time invariant covariates, such as customers demographics information, and $\mathbf{u_{it}}$, vector of time-varying covariates, such as the firm's marketing activities at time t. Last, $\boldsymbol{\omega^z}$ and $\boldsymbol{\omega^u}$ are parameters vector of these covariates.

From the above, we can have the likelihood function of product usage of a customer i with T_i observatory periods and a state sequence $S_i = [S_{i1}, S_{i2}, ..., S_{i(T_i)}]$ as follow:

$$\mathcal{L}_i^{use}(\boldsymbol{\gamma_i}, \boldsymbol{\pi}, \boldsymbol{\omega^z}, \boldsymbol{\omega^u} | S_i = s_i, data) = \prod_{t=1}^{T_i} P(Y_{it}^{use} = y_{it}^{use} | S_{it} = s_{it}, \boldsymbol{\gamma_i}, \boldsymbol{\pi}, \boldsymbol{\omega^z}, \boldsymbol{\omega^u}, data) \tag{8}$$

$$= \prod_{t=1}^{T_i} \frac{(\gamma_i \pi_{s_{it}} e^{\omega^z z_i + \omega^u u_{it}})^{y_{it}^{use}} e^{-(\gamma_i \pi_{s_{it}} e^{\omega^z z_i + \omega^u u_{it}})}}{y_{it}^{use}!} \tag{9}$$

3.3 State Dependent Premium Purchase

The next component in this paper's model is "State Dependent Premium Purchase", or the probabilities that a customer i in state s will purchase a premium add-ons Pre_n at time t. Since a customer must be using the product at time t in order to purchase premium features, I assume consumer's purchase choice is conditional on their usage at t is different from 0.

Assuming that the freemium product offers a list of N different premium features for purchase $\mathbf{PRE} = [Pre_1, Pre_2, ..., Pre_N]$, and these options differ by multiple covariates such as cost or effect duration, we can model the utility of purchase premium feature Pre_n as:

$$U_{n,it|s} = \eta_{n,it|s} + \epsilon_{n,it|s}$$

with $\eta^n_{it|s}$ as the systematic component and $\epsilon^n_{it|s}$ as the random component. The probability a customer i in state s purchase premium feature Pre_n, assuming the error term is i.i.d. following a Gumbel Type I extreme distribution, could be modelled through a conditional multinominal logit model:

$$\begin{cases} p_{n,it|s} = 0, & \text{if } Y^{use}_{it} = 0\text{, else} \\ p_{n,it|s} = \frac{exp(\eta^n_{it|s})}{1+\sum_{j=1}^{N-1} exp((\eta^n_{it|s}))'}, & \\ p_{N,it|s} = \frac{1}{1+\sum_{j=1}^{N-1} exp((\eta^n_{it|s}))'} & \end{cases} \tag{10}$$

Where,

$$\eta_{n,it|s} = \alpha_{n,is} + \mathbf{x^p_{it}}\theta_\mathbf{s} + \nu_n\zeta_s \tag{11}$$

Variable $\alpha_{n,is}$ represent the heterogeneity in utility across population and state, and this could be broken down into two components:

$$\alpha_{n,is} = \psi_{n,s} + \chi_{n,i}$$

With $\psi_{n,s}$ represent the average utility of Pre_n in state s, and $\chi_{n,i}$ is the heterogeneous effect across different individuals. By splitting α into a state-specific and an individual-specific component, I ensure the identification of the unobserved individual parameters in both the transition probabilities and observed behavior.

Vector $\mathbf{x^p_{it}}$ represents the vector of covariates that could potentially affect customer's choice of purchasing premium feature, and θ_s is the state-specific vector of parameters corresponds to these covariates.

Last, vector ν_n is a vector of feature-specific covariates, such as price, popularity..., which differs across all offered premium features. Similar to above, corresponding to these covariates are state-specific vectors ζ_s of parameters.

3.4 Stationary Initial State Distribution

As this is a HMM model with time-variant and invariant covariates, the stationary distribution could be calculated by solve the equation $\boldsymbol{\tau} = \boldsymbol{\tau}\bar{\mathbf{Q}}$, conditional by sum of $\boldsymbol{\tau}$ equals one, whereby $\bar{\mathbf{Q}}$ is the transition matrix with the estimated parameters. Given that the data is not left-censored, i.e. we can observe the first period since customers just download the application, $\boldsymbol{\tau}$ could be calculated with covariates set to zero. Following Netzer et al. (2008), if all estimated transition probabilities are positive, we can confirm a unique stationary distribution.

3.5 Likelihood Function

Finally, we can combine all submodels above to derive the final likelihood function of the full model. Due to the Markovian structure of the model, the likelihood function of $\mathbf{y_{it}} = [y_{it}^{use}, y_{it}^{Pre1}, ..., y_{it}^{PreN}]$, the vector of all observe behaviors of customer i is calculated through all the paths of hidden states that an individual consumer could take over time. Therefore, the joint likelihood is given by the sum over all possible routes for each sequence:

$$\begin{aligned} &\mathcal{L}_i(\boldsymbol{M}, \boldsymbol{\delta_i}, \boldsymbol{\omega}, \boldsymbol{\pi}, \boldsymbol{\gamma_i}, \boldsymbol{\psi}, \boldsymbol{\chi_i}, \boldsymbol{\beta}, \boldsymbol{\zeta} | Data) \\ = \sum_{S_{i1}=1}^{K} \sum_{S_{i2}=1}^{K} \dots \sum_{S_{iT_i}=1}^{K} &[P(S_{i1} = s_{i1}) \prod_{t=2}^{T_i} P(S_{it} = s_{it} | S_{i(t-1)} = s_{i(t-1)}) \\ &\times \prod_{t=1}^{T_i} P(\boldsymbol{Y_{it}} = \boldsymbol{y_{it}} | S_{it} = s_{it})] \end{aligned}$$

Using the notation in the submodels, the above likelhood function could be re-written as:

$$\begin{aligned} &\mathcal{L}_i(\boldsymbol{M}, \boldsymbol{\delta_i}, \boldsymbol{\omega}, \boldsymbol{\pi}, \boldsymbol{\gamma_i}, \boldsymbol{\psi}, \boldsymbol{\chi_i}, \boldsymbol{\beta}, \boldsymbol{\zeta} | Data) \\ &= \sum_{S_{i1}=1}^{K} \sum_{S_{i2}=1}^{K} \dots \sum_{S_{iT_i}=1}^{K} [\tau_{s1} \prod_{t=2}^{T_i} q_{(s_{i(t-1)}, s_{it})it} \prod_{t=1}^{T_i} h_{it|s_{it}}] \end{aligned}$$

With $h_{it|s_{it}} = P(\boldsymbol{Y_{it}} = \boldsymbol{y_{it}} | S_{it} = s_{it})$.

Following Zucchini and MacDonald (2009), we can write the function in matrix form as:

$$\mathcal{L}_i(\boldsymbol{M}, \boldsymbol{\delta_i}, \boldsymbol{\omega}, \boldsymbol{\pi}, \boldsymbol{\gamma_i}, \boldsymbol{\psi}, \boldsymbol{\chi_i}, \boldsymbol{\beta}, \boldsymbol{\zeta} | Data) = \boldsymbol{\tau H_{i1} Q_{i2} H_{i2} ... Q_{iT_i} H_{iT_i} I_K} \tag{12}$$

Where $\boldsymbol{H_i t} = diag(h_{it|1}, ..., h_{it|K})$ is diagonal matrix of size $K \times K$ which diagonal elements as the probabilities of observed behaviors at each state, and $\boldsymbol{I}_K$ is a $1 \times K$ vector of ones. We can estimate the model parameters using a hierarchical Bayesian framework similar to Ascarza (2013) and Netzer et al. (2008).

4 Data

4.1 Data Requirements

In order to empirically apply the proposed model, I would use the data of a video game application that is popular on AppStore or Google Playstore. The requirements for this data are as follow:

- The game should follow a freemium model where customers can download a free version of the game free of charge, and this version should be *perpetually free* and always available to customers even if they do not move to the fee-paying segments.

- The game should offer premium features instead of a premium version, since I want to test a multinominal logit model. A popular example is free games with "shop" where customers could purchase items or "power-booster" using real world money. With this model, customers could have multiple premium options to choose from.

- The data should include the whole lifetime product usage data of customers, i.e. since they first download the game, so we can calculate the initial transition matrix $\boldsymbol{\tau}$.

4.2 Data Structure

The general structure of required data should be a time series data of customer usage by day, $Usage_{it}$ as well as whether customers purchase a premium feature in a given day, denoted by a vector $[Pre1_{it}, Pre2_{it}, ..., PreN_{it}]$, which take value of 0 if customer i does not purchase that feature at day t and 1 if they do. Beside these main variables, potential covariates that could be used for the model are as follow:

- *Covariates that affect transition probabilities*: These could include variables such as (1) Current in-game level, (2) Number of friends in the game, (3) Number of other games in their phone, (4) General popularity of the game, measured by ranking in marketplace, (5) Marketing activities by game publisher

- *Time invariant covariates that affect Product Usage:* (1) Age, (2) Gender, (3) Location, (4) Type of device.

- *Time variant covariates that affect Product Usage:* (1) Day of the week, (2) Season, (3) Marketing activities, (4) Whether there is a new update with new content, (5) Special in-game events.

- *Individual-specific covariates that affect Premium purchase:* (1) Type of device, (2) Number of days since last purchase, (3) Prior investment in the game, (4) Demographic variables.

- *Feature-specific covariates that affect Premium purchase:* (1) Price, (2) Popularity, (3) Order of listing.

The above list, however, is not exhaustive, and more covariates could be utilized, dependent on the nature of available data as well as the design of the game we can get data from.

5 Expected Results

5.1 Choosing number of states

A 80:20 split could be applied to the data, in order to set aside 20% of the observation as hold out. The model would be estimated using various number of state, incremental by one from one state and up. We can compute the log predictive density, the Watanabe–Akaike information criterion (WAIC), and the meansquare error (MSE) and mean absolute percentage error(MAPE) between the observed and predicted product usage and premium feature purchase. These benchmarks could be used together in order to choose the best number of states, in the same manner as Netzer et al. (2008).

5.2 Results report

We could continue the result section with using the average observed variables for each state to come up with a general, characterizing description of each state. We could then report the transition matrix, as well as reporting the average product usage and premium feature purchase probabilities for each state. The main assumption is that both Product Usage and Premium Purchase likelihood would differs across states, and thus proves that there are underlying dynamics between the two, through an “interest” or “relationship” state variable.

Next, we could investigate the effects of covariates on state-dependent product usage and premium feature purchase, with each section for each type of covariates listed out above. Some expected results could be: Special in-game event increase Product Usage and Premium Purchase for customers with high interest, but does not affect customers with low interest, for example.

Next, we perform State Recovery using equation (12), and smoothing approach similar to Zucchini and MacDonald (2009) to compute the probability of each customer belonging to each state at any time period. This could create good managerial insights on how customer base evolves over time. Combining with the effects of covariates, we can map out which marketing activities or product adjustment are necessary in each period to maximize the premium purchases from customers.

5.3 Benchmarking

We could conduct benchmarking of model predictive power by comparing it with some other popular models, such as:

- Lagged covariates: A model in which Product Usage and Premium Purchase behaviors at a specific time period are modelled as functions of these behaviors at previous periods.

- Recency & Frequency: This is another popular model, in which the observed behaviors are modelled as function of recency (number of periods since last perform the action) and frequency (the proportion of performing the action in previous periods).

- Trajectory Analysis: Trajectory analysis is a cluster-based approach for identification of clusters of individual trajectories within a given population (Nagin 2005). This method models the linkage between time and behavior by allowing for polynomial relationships.

6 Limitations and Extensions

A potential limitation of this paper is that it does not account for the scenario where customers stop using the product, or "churn". Since the main purpose of this paper is exploring the dynamics between usage and premium purchases, this could be possibly be ignored. However, in real life, a large number of customers would transition into a state "zero" and stop using the game altogether. The model could be expanded to include this, by incorporating app deletion data and model the probability of deleting as a state-dependent function.

Furthermore, in this paper I do not include other important aspects of freemium model, such as network effects and utility from referrals. These aspects could heavily influence managerial strategic decision in relation to freemium strategy (Niculescu and Wu, 2014). Similar to above, these could also be incorporated to the model as sub processes, all dynamically linked together through the state latent variable. For example, if data is available, we could easily model the referral probability, as well as including network effects from user's friends into proposed submodels.

Another potential extension of the model is the inclusion of another latent state sequence representing the relationship between customeres and the platform itself. This is not substantial for mobile games, but for many other applications, such as PC-based video games purchasing through Steam platform, or purchasing of premium contents of a specific channel of Youtube, the underlying dynamics may come from customer relationship with both the specific product or channel and the relationship with the publishing platform. In this case, a factorial Hidden Markov Model (Ghahramani and Jordan, 1997) would allow us to model observed behaviors as being dependent on more than one state variable.

Last but not least, Hidden Markov Model assumes that each state last only one period of time, transitions could happen at the end of each period, and observed behaviors are independent of time spent in each state. However, this may not fully reflect reality, as states may varies in length, and the time customers spent in each state could greatly influence their observed behavior. Therefore, this model could be extended using a Hidden Semi Markov Model, or non-stationary Hidden Markov Model, to account for these variations. A good theoretical overview of this model could be found in Yu (2010).

7 References

1. Anderson S, Celik L (2014) Product line design. Working paper, Centre for Economic Policy Research, London.

2. Ascarza E, Hardie BGS (2013) A joint model of usage and churn in contractual settings. Marketing Sci. 32(4):570–590.

3. Coussement K, Van den Poel D (2008) Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. Expert Systems Appl. 34(1):313–327.

4. Chintagunta PK (1998) Inertia and variety seeking in a model of brand-purchase timing. Marketing Sci. 17(3):253–270.

5. Dey D, Lahiri A, Liu D (2013) Consumer learning and time-locked trials of software products. J. Management Inform. Systems 30(2):239–268.

6. Ghahramani, Z., Jordan (1997) Factorial Hidden Markov Models. M.I. Machine Learning. 29: 245.

7. Guadagni PM, Little JD (2008) A logit model of brand choice calibrated on scanner data. Marketing Sci. 27(1):29–48.

8. Kamada Y, Ory A (2019) Encouraging word of mouth: Free contracts, referral programs, or both? Working paper, Cowles Foundation for Research in Economics, Yale University, New Haven, CT.

9. Kumar V, Zhang X, Luo A (2014) Modeling customer opt-in and optout in a permission-based marketing context. J. Marketing Res. 51(4):403–419.

10. Kumar V (2014) Making “freemium” work. Harvard Bus. Rev. (May): 701–703.

11. Lee C, Kumar V, Gupta S (2017) Designing freemium: Strategic balancing growth and monetization. Working paper, Samuel Curtis Johnson Graduate School of Management, Cornell University, Ithaca, NY. Available at SSRN: https://ssrn.com/abstract =2767135.

12. Maltz J., Barney D. (2012) Shoud Your Startup Go Freemium. TechCrunch, https://techcrunch.com/2012/11/04/should-your-startup-go-freemium.

13. Montgomery, A.L., S. Li, K. Srinivasan, J. Liechty (2004) Modeling Online Browsing and Path Analysis Using Clickstream Data. Marketing Science. 23 (4): 579-595.

14. Montoya R, Netzer O, Jedidi K (2010) Dynamic allocation of pharmaceutical detailing and sampling for long-term profitability. Marketing Sci. 29(5):909–924.

15. Needleman S, Loten A (2012) When freemium fails. Wall Street Journal (August 22), https://www.wsj.com/articles/SB1000087239639 0443713704577603782317318996.

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

17. Niculescu M, Wu D (2014) Economics of free under perpetual licensing: Implications for the software industry. Inform. Systems Res. 25(1):173–199.

18. Rabiner, L. R., Juang, B. H. (1986) An Introduction to hidden Markov models. IEEE Acoustics, Speech & Signal Processing Magazine, 3, 4–16.

19. Shi, Z., Zhang, K., Srinivasan, K. (2019) Freemium as an Optimal Strategy for Market Dominant Firms. Marketing Science. 38. 10.1287/mksc.2018.1109.

20. Wagner TM, Benlian A, Hess T (2014) Converting freemium customers from free to premium—the role of the perceived premium fit in the case of music as a service. Electronic Markets 24(4): 259–268.

21. Shun-Zheng Yu (2010) Hidden semi-Markov models. Artificial Intelligence. 174(2): 15-243.

22. Zucchini W, MacDonald IL (2009) Hidden Markov Models for Time Series: An Introduction Using R (CRC Press, Boca Raton, FL)