Modeling User Engagement in Mobile Content Consumption with Tapstream Data and Field Experiment

Yingjie Zhang
Heinz College Carnegie Mellon University, yingjie@cmu.edu,

Beibei Li
Heinz College Carnegie Mellon University, beibeili@andrew.cmu.edu,

Xueming Luo
Temple University, luoxm@temple.edu,

Xiaoyi Wang
Zhejiang University, kevinwxy@zju.edu.cn,

September 15, 2016

Low engagement rates and high attrition rates have been formidable challenges for mobile apps and their long-term success, especially for those whose revenues come mainly from in-app purchases. To date, still little is known about how companies can comprehensively identify user engagement stages so as to improve business revenues. This paper proposes a structural econometric framework for modeling of consumer latent engagement stages that accounts for both the time-varying nature of engagement and consumer forward-looking consumption behavior. The present study analyzed a fine-grained mobile tapstream dataset on mobile users' continuous content consumption behavior in a popular mobile reading app. Our policy simulation enabled us to tailor, based on the model-detected engagement stages, an optimal pricing strategy to each consumer. Interestingly, we found that such an engagement-specific pricing strategy leads, simultaneously, to lower average prices for consumers and higher overall business revenues for the app. To further evaluate the effectiveness of our method, we conducted a randomized field experiment on a mobile app platform. Our experimental results provide more causal evidence that a personalized promotion strategy targeting user engagement stages can both decrease costs to app users and enhance overall business performance. Our structural-model- and field-experimentation-based findings are nontrivial and suggest, with respect to the crucial role of modeling user engagement, potential overall welfare improvements in the mobile app market.

Key words: User engagement, Mobile content consumption, App platforms, Hidden-state model, Forward-looking behavior, Structural econometric model, Field experiment

1. Introduction

With the increasing popularity of mobile technologies, consumers today spend significant amount of time and activities on mobile apps every day. A recent study (ComScore (2016)) showed that in 2015, mobile users spent approximately 68.2 hours per month on smartphone apps, compared to just 29.6 hours on tablets. The average mobile app usage time among the young (i.e., age 18-24) was even higher, approximately 90.6 hours per month. As indicated by the extant literature, researchers thus far have been attracted to the mobile market (e.g., Andrews et al. (2015), Ghose and Han (2014), Han et al. (2015b), Luo et al. (2013), Xu et al. (2016)). Given this growing trend, however, the in-app conversion rate is quite low. It has been reported that in February 2016, only 1.9% of all players paid for in-game content, and half of the revenues from all mobile game apps were contributed by only 0.19% of all players.[1] The average conversion rate of mobile apps was less than 2% in the U.S. in 2014.[2] Moreover, the attrition rate was high, 19% of mobile apps having been opened just once in 2015.[3] Indeed, low engagement and high attrition rates have been major challenges to the long-term success of mobile app companies, especially those whose revenues come mainly from in-app purchases. In order to deal with these challenges, most mobile app companies have started applying different targeting strategies such as freemium (e.g., time-based freemium, feature-based freemium, seat-limited freemium) by adversely selecting consumers into different types according to their business values. Prior work has shown that providing a free version does improve the sales of paid versions at the aggregate level (Ghose and Han (2014)). Besides, some mobile apps offer various plans or sales promotions to encourage users to commit to multiple purchases over the long run.

[1] http://venturebeat.com/2016/03/23/half-of-all-mobile-games-revenue-comes-from-only-0-19-of-players-report/

[2] http://info.localytics.com/blog/mobile-apps-whats-a-good-conversion-rate

[3] http://tech.thaivisa.com/app-retention/11662/

However, most of those targeting strategies are designed to be identical for all mobile users; they are not tailored. In other words, all users are facing the same products, prices and plans over time. Such non-personalized strategies are problematic, especially considering the significant variance in the app user population (e.g., active vs. non-active) over time. Thus, it is essential to tailor personalized targeting strategies.

Prior studies have paid scant attention to the issue of how companies can understand and improve user engagement and business revenues by designing effective in-app targeting strategies. Nonetheless, some have realized the importance of understanding user engagement in purchasing decision making. For example, Kim et al. (2013) explored this through surveys, asking mobile users about their engagement stages and the reasons they were continually engaging with mobile activities. Such studies yield interesting psychological findings, but are difficult to apply in the real world. Others have proposed Hidden Markov model to detect consumers' stages (Abhishek et al. (2012), Netzer et al. (2008)). In these studies, consumers were clustered into different "hidden" stage strata (e.g., awareness, interest, desire, and action) based on their behavioral patterns and willingness to pay. However, they focused only on consumers' *one-time* purchasing behavior. Such insights, unfortunately, cannot be generalized to the context of mobile apps, whose success often is a function of consumers' *repeated* purchasing and long-time loyalty. In the present study, we aimed to model consumer latent engagement stages by accounting for repeated purchasing. This is one major literature gap our study was undertaken, uniquely, to fill.

Another literature shortcoming is the assumption, when modeling hidden stages and purchase decisions, that consumers are myopic (i.e., not forward-looking). This assumption can be unrealistic, especially when modeling consumer behavior on mobile app platforms, because of consumers' repeated purchasing and long-term relationships with app platforms.

For example, mobile apps, such as Candy Crush, might offer both pay-per-use and subscription options. Thus, if a mobile user is running out of "lives," she can choose to pay for either a 2-hour unlimited lives (i.e., the subscription option) or per life (i.e., the pay-per-use option). Such a decision is based on the user's expectation of her future consumption. To retain long-term consumers then, it is essential for app platforms to understand users' forward-looking behavior in order to predict their current and future decisions and, therefore, proactively tailor targeting strategies for improved user engagement for better experience. However, from a methodological perspective, little knowledge has been gained with regard to how to model users' latent engagement stages while accounting for their forward-looking behavior over time. Our study fills this literature gap as well.

Finally, the issue of how businesses should leverage this knowledge in designing and measuring the effectiveness of personalized targeting strategies is, due to the potential for consumers' self-selection of engagement activities, rather challenging. For example, consumers who have stronger inherent preferences for the products or services on the app platform are more likely to become highly engaged, and also, meanwhile, to make purchases. Therefore, simply observing a positive relationship between engagement levels and purchase activities does not suggest a causal impact, nor does it indicate that companies should target consumers in the high-engagement stage to increase purchase rates. In the present study, we aimed not only to design effective personalized targeting strategies based on the detected user engagement levels, but also to evaluate the effectiveness of our approach from a causal perspective using a randomized field experiment. Randomization with exogenous promotion treatments enables clear identification of the sales impacts of different targeting strategies under various user engagement stages. The findings of prior studies with observational data based on correlational analyses can be confounded by self-selection and endogeneity biases, the correction of which by statistical instruments is onerous. This, then, is the third research gap filled by our work.

More specifically, this paper answers the following three research questions:

- How can we model the heterogeneous user engagement stages over time?
- In what ways can we leverage the detected user engagement stages in designing personalized targeting strategies to improve mobile app business revenues?
- How can we gauge the sales impacts of those personalized targeting strategies?

To address these questions, we proposed and estimated a structural model by combining the single-agent dynamic discrete choice model with the Hidden Markov Model (HMM). This framework can model consumer latent engagement stages by accounting for both the time-varying nature of engagement and consumer forward-looking consumption behavior. We validate our study by combining large-scale archival data analysis and randomized field experiment. Our fine-grained data contain the mobile "tapstream," the value of which is based on the fact that it is easier for mobile users to "browse, read, and buy from any device."[4] Our data contain 524,704 records on 10,970 consumers' approximately three months' worth of continuous content consumption behaviors in a popular mobile reading app in 2015.

Our empirical analyses yielded some interesting findings. *First*, we detected four user engagement stages on the mobile reading app. Users at different engagement stages have different behavioral patterns. Promotions tailored to users from different mind stages can become more effective. *Second*, we found that, without any extra policy interventions, users in most of the engagement stages are likely to leave the app or become less engaged. On the other hand, users at the lower stages still have the potential to jump to a higher

[4] https://www.dynamicyield.com/2015/01/personalization-trends/

stage inherently. This finding is strong evidence for the importance of designing effective interventions to improve mobile app user engagement. *Third*, to evaluate our proposed model and to improve the current condition described in our second finding, we conducted several policy simulations to explore optimal pricing strategies. To that end, we designed a personalized pricing strategy based on users' real-time engagement stages. Our results suggest that total revenue increases relative to that earned with the non-personalized optimal pricing strategy. Interestingly, we found that such an engagement-specific pricing strategy leads to lower average prices for consumers *and* higher overall business revenues, thus suggesting the potential welfare improvement in the mobile app market. *Fourth*, by randomized field experiment, we further tested the causal impact and generalized the effectiveness of our engagement detection method. The results afforded causal evidence that personalized promotion strategies targeting user engagement stages could improve business performance and lower costs at the same time. All of these structural-model- and field-experiment-based findings are non-trivial; they suggest the high potential for welfare improvement in the mobile app market, particularly with respect to the crucial role of user engagement modeling.

This paper makes the following four key contributions. *First*, we propose a novel multi-stage structural econometric model for studies of individual-level mobile content consumption behavior. The model advances the existing literature by simultaneously accounting for both individual consumers' latent mind stages and their forward-looking behavior during the decision-making processes. The proposed modeling framework also can be generalized to study other online or offline contexts, such as consumers' purchasing behavior on durable goods with respect to their long-term expectations about future consumption. *Second*, our model for detection of user engagement stages is a feature dimensionality reduction method.

With this method, we can extract a one-dimensional indicator of user engagement from the multi-dimensional features of users' historical behavior. Compared with multi-dimensional features that were directly applied in previous studies, this single indicator will simplify the design and implementation of our methodology in practice, without losing tracks of consumer behavior information. *Third*, based on the detected mobile user engagement stages, we are able to design and evaluate user-engagement-based personalized targeting strategies. We show the strength of our framework in helping digital infrastructures and platforms to design personalized targeting strategies towards increasing the overall revenue of the business as well as consumer welfare. *Finally*, our study demonstrates the value of leveraging fine-grained mobile tapstream data by means of randomized field experiment for better understanding of user behavior on digital infrastructures. Because consumers often incur higher search costs on mobile devices than on personal computers (Ghose et al. (2012)), such individual-level mobile tapstream data would provide more valuable information in revealing users' economic preferences.

2. Literature Review

Our study draws from and builds on the following streams of literature.

2.1. User Engagement

Recently, the term "engagement" has been increasingly applied within the academic marketing field. Brodie et al. (2011) performed an exploratory analysis of its theoretical meaning and foundations. To the best of our knowledge, however, there are few studies empirically and effectively detecting the unobserved user engagement with a product to improve market efficiency. Kim et al. (2013) conducted a survey of mobile users' engagement stages and the reasons for their continually engaging with mobile activities. They found that engagement is the product of utilitarian, hedonic, and social motivations. Such survey data, however, can be subjective according to interviewees' differing standards applied in answering the same questions. Claussen et al. (2013) studied the effect of the Facebook *Reward* program on engaging users. But the proxy they used for user engagement was at the aggregated level for the number of active users on an app. As is noted in Hollebeek et al. (2014) and Vivek et al. (2012), consumer engagement stages can be reflected in multi-dimensional features. Although "engagement" provides enhanced predictive and explanatory power with regard to focal consumer behavior outcomes, without an effective detection method, we are hard-pressed to benefit from this term. In our study, we automatically detected "true" engagement using historical behavior data, which will prove both more convincing and more practical in future studies.

2.2. Business Strategies on Mobile Platforms

The freemium strategy has attracted attentions from both the mobile app industry and academics. Most of the previous empirical studies in the literature, though, have focused on the aggregated demand side, specifically determining, by reduced-form analysis, how the free version affects the paid version (Ghose and Han (2014), Lee et al. (2013), Liu et al. (2012), Wagner et al. (2013)). For example, Ghose and Han (2014) studied the impacts of different features of mobile apps on demand, finding that demand increased with the availability of the in-app purchase option. Liu et al. (2012) studied the Google Play platform, concluding that the sale rank and revenue of a paid version would increase if its trial version were offered. In addition to the mobile app market, the freemium strategy is widely used in the software industry, which is the underlying setting of many relevant theoretic papers (e.g., Faugère and Tayi (2007), Niculescu and Wu (2014)). Unlike the empirical work, these theoretical papers focus on the consumer-behavior side, especially their learning and awareness of product quality over time. In the software industry, each product can have true quality, which is the source of consumers' learning behavior. In the mobile app market by contrast, such "true" quality is subjective to each individual user, being dependent on their experience with the products. In other words, mobile users choose to pay mainly due to their high level of engagement with the product. Therefore, our empirical study cannot directly benefit from this well-established theoretic framework. In the present study then, we empirically estimated individual users' changes of behaviors over time and, additionally, detected their engagement with the product. With this framework, we were able to design a personalized freemium strategy based on user engagement detection. Our personalized targeting strategy, as distinguished from the commonly-used non-personalized freemium strategy, has the potential to simultaneously increase company revenues via discriminating prices and attracting more users.

As conditional on the paid version of a product, consumers might still face multiple contracts to choose from. The design of optimal pricing strategies is what tariff-choice-related papers are interested in (e.g., Goettler and Clay (2011), Han et al. (2015a), Lambrecht and Skiera (2006), Vigna and Malmendier (2006)). According to these papers, there exist two tariff-choice biases: flat-rate bias and pay-per-use bias, based on individual users' tendency to overestimate their participation with, and to underestimate of their usage of, a product, respectively. In the current study, we empirically built a model to explain how users make decisions among free versions and paid versions with different contracts. Our framework will help to clarify whether those biases exist and, if they do, how to address them with personalized targeting strategies.

2.3. Pricing in Online Markets

Pricing personalization has been widely studied in the literature stream (Bakos and Brynjolfsson (1999), Choudhary (2010), Choudhary et al. (2005), Huang and Sundararajan

(2011)). Most papers have addressed this issue using analytical frameworks developed for that purpose. This paper, by contrast, will provide solutions for personalized pricing strategies and empirically test their effectiveness. There is one working paper that aimed to design a personalized model for big data (Shiller et al. (2016)). Measuring the effects of pricing personalization, however, raises an endogeneity issue that was not addressed in that paper. Our study has the particular strength of combining an econometric structural model with a field-experimentation design, which provides causal-evidence explanatory power.

2.4. Methodology

From a methodological perspective, our work utilizes the framework of single-agent dynamic discrete-choice structural models (Hotz et al. (1994), Miller (1984), Rust (1987)) by incorporating HMM (MacDonald and Zucchini (1997)). The HMM has been widely used in the Machine Learning field (e.g., Punera and Merugu (2010), Laxman et al. (2008)) and was recently introduced into the marketing field (e.g.,Abhishek et al. (2012), Montgomery et al. (2004), Netzer et al. (2008), Kumar et al. (2011)). For example, Abhishek et al. (2012) applied the HMM to study consumers' stages in advertising analysis. The choice model embedded with the HMM in most marketing papers, however, is static without considering consumers' forward-looking behaviors. Identification and estimation issues have been proved theoretically in a recent working paper: (Connault (2014)). The current study applied such a combination as the major framework with which to empirically analyze our scenario.

3. Data and Industry Context

We initialized our analysis using a rich dataset from a top mobile reading app in China that offers more than 400,000 mobile books and attracts over 130 million users per month. This mobile app provides products very similar to Amazon Kindle but with specialized mobile-platform services.

3.1. Industry Context

Any mobile phone users can easily and freely download the reading app from app stores. They can then freely sign up for the app with their phone numbers. Every time the user finishes reading a content unit, the app can jump to the next content unit automatically. If the user chooses not to read the given content unit, the app will show her a new book automatically. In each book, the first several content units[5] are free for any users. After that, to continue reading, the user needs to either pay per content or subscribe to the app so that she can read all of the content units within a calendar month. At the beginning of a new calendar month, the subscription contract will continue automatically unless the user chooses to quit the current contract, which will end at the beginning of the next calendar month.

3.2. Tapstream Data

Similar to the clickstream data, which is used to understand PC-user behavior through mouse clicks, *tapstream* data records fine-grained information about individual behavioral trails on mobile platforms through finger taps. Our data contain a total of 524,704 records on 10,970 consumers' approximately three-month continuous content consumption behaviors in a popular mobile reading app for 2015. Each record includes user ID, time stamp and content information (i.e., name of the content unit, book name, book genre) as well as the payment option (i.e., free chapter, pay-per-use option, or subscription) the user chose.

In the raw tapstream data, books are pre-classified into 117 types by the app company. We grouped these types into three genres: "fiction", "casual," and "practical"(Li (2015)). Types within the same genre have similar reading purposes. For example, "casual" books mostly serve entertainment purposes, while "practical" books require in-depth reading or

[5] The numbers of free content units are determined by the mobile app company.

Table 1 Descriptive Statistics on Key Variables

Variable	Description	Mean	Minimum	Maximum	Std.dev.
T_i	Number of decision periods	183.7247	1	2650	329.4378
X_1	Popularity indicator	0.9562	0	1	0.2047
X_2	Fiction genre indicator	0.7146	0	1	0.4516
X_3	Casual genre indicator	0.2005	0	1	0.4003
sub	Indicator of subscription	0.8507	0	1	0.3564
F	Indicator of free content	0.8721	0	1	0.3340

even note-taking. Additionally, we counted the total number of records for each book and divided all of the books into two groups, ordinary and popular books, based on the mean number of records.[6] The limitation of our tapstream dataset is that it does not include any actual reading time information. To infer this information, we assume that if the time gap between two consecutive records is longer than 10 minutes, the user chose outside goods once in the gap. Otherwise, the time gap between two consecutive records is treated as the reading time of the given chapter.[7] The descriptive statistics on the key model variables are presented in Table 1.

3.3. Model Free Data Patterns

We observed two patterns in the tapstream data: the behavior variation across individuals, and the behavior trend of the same individual over time. Such behavior change, to some degree, can justify our claim about the variation of users' engagement stages.

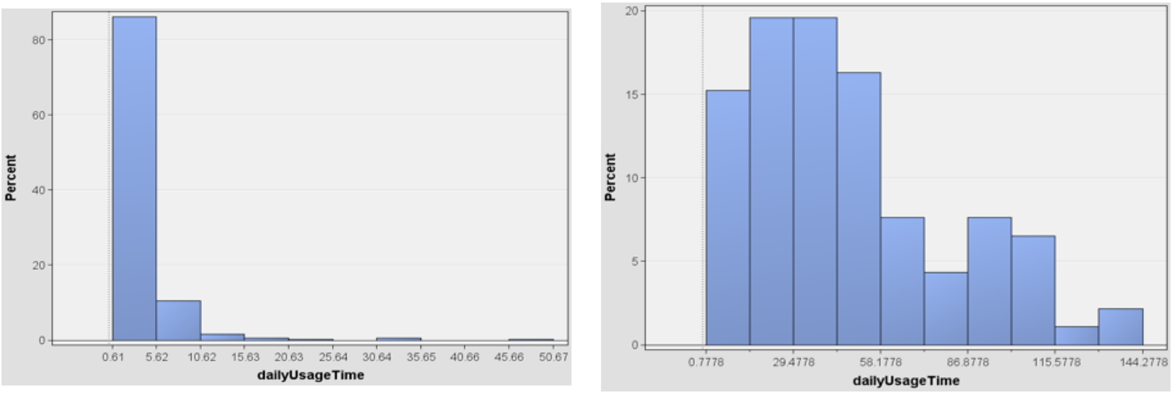
(1) Variation among all users

[6] We have tried various other definitions based on median value, or top/botoom 25 percentile as robustness tests. And the results show consistency.

[7] The data show that the majority of time gaps fall below 20 minutes. We also tried 5-min, 15-min and 20-min as different extraction criteria for reading time, and the results were qualitatively robust.

First, we extracted 23 variables that describe individuals' reading patterns at three levels: event-level, time-level, and book-level. Four variables are at event level, including number of days with more than one event, number of days under subscription, mean and std. dev. of daily number of events; nine variables are at book level, including total and daily number of read content-units/books/book genres, average number of units per book, and total number of free (paid) content units; ten variables are at time level, including total usage time (minutes), average daily usage time (hours), daily number of reading slots, standard deviation of daily reading time, mean and std. dev. of daily time gap between slots, mean and std. dev. of daily number of gap slots, average time gap per slot, and average usage time per event. We then utilized these features to segment mobile users using the K-means clustering algorithm (Hartigan and Wong (1979)). Our results show that users can be divided into five segments according to their reading-behavioral patterns and engagement with the reading app. For example, in Figure 1, we show the distribution of daily usage time of users in two of the five segments. The x-axis is individual's daily usage time in minutes and the y-axis is the frequency in percentage. We found that some users spent much more time on the reading app than the others. In the meantime, the distributions of the other features suggest that the segment of users with higher usage time tend to explore more book categories and make more pay-per-content payment options. This finding indicates that the engagement level varies among users. Similarly, Figure 2 shows that users in different segments present different volumes of reading activities per day. The x-axis is the number of content units an individual user reads daily, and the y-axis is the frequency in percentage. Figure 2 also suggests that it is possible that some users did not have any purchases (i.e., reading activities) during some periods.

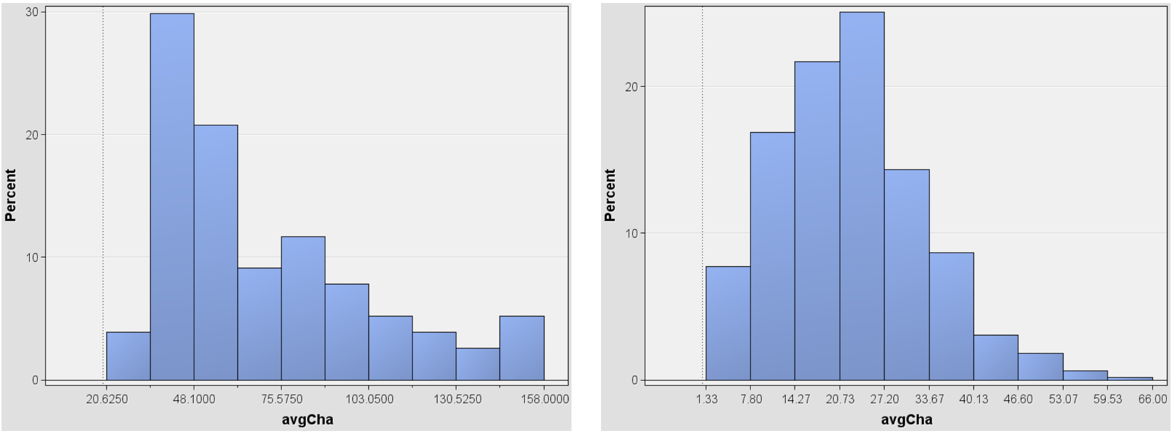
(2) Variation in individual users' behavior over time



(a) Segment of users with lower usage time

(b) Segment of users with higher usage time

Figure 1 Daily Usage time (minutes) between User Segments

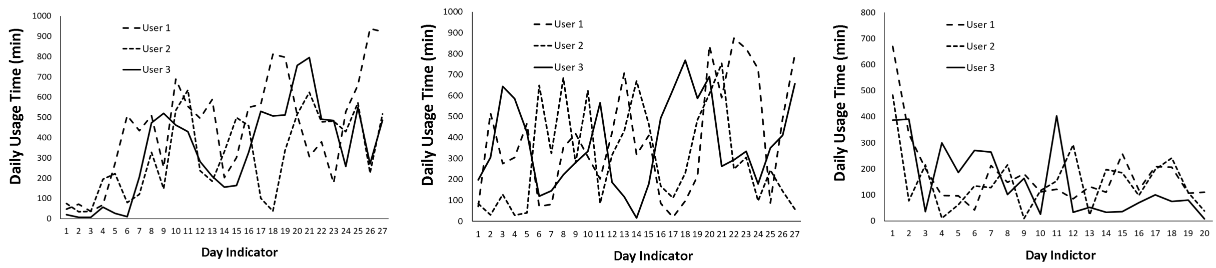


(a) Segment of users with more activities

(b) Segment of users with fewer activities

Figure 2 Average Daily Number of Read Content Units between User Segments

We calculated the individual users' daily usage time, and found that it varies significantly over time and across users. We sampled a number of mobile users in order to plot, as indicated in Figure 3, their usage paths over time. In the three plots, the x-axis is the time indicator, and the y-axis is the daily usage time in minutes. We divided those sampled users into three groups based on their different behavioral patterns. Respectively, the users in Figure 3a spent increasing time on the reading app, while those in Figure 3c spent decreasing time on it. Compared with the above two user groups, the reading behavior of users in Figure 3b varied over time. Clearly, we see significant heterogeneity in users' usage time over time. In some sense, a longer usage time indicates a higher level of user engagement with the app, which in turns will bring more potential revenues to the app company. Hence, detection of the time-varying engagement stages is helpful to the app company's efforts to apply corresponding targeting strategies and, thereby, improve their business revenues.



(a) Increasing Usage (b) Varied Usage (c) Decreasing Usage

Figure 3 Individual samples' daily usage time over time

4. Model

We constructed our framework by combining both the single-agent dynamic discrete choice model (Rust (1987)) and the HMM (MacDonald and Zucchini (1997)).

The HMM is a stochastic process model in which the states are unobserved but can affect the observed outcome. In our framework, we modeled individual users' engagement with the reading app as a hidden state in the HMM. We present a schematization of our proposed framework in Figure 4. Users would be at different engagement stages in different phases. The engagement stage will affect their period utility, which is used to form the expectation about future values. Finally, the decision is made based on the lifetime expected utility. A high level of engagement would, similarly to the purchase funnel concept, lead to a high probability of purchasing if other factors remain constant. In the following sub-sections, we will discuss the model in detail.

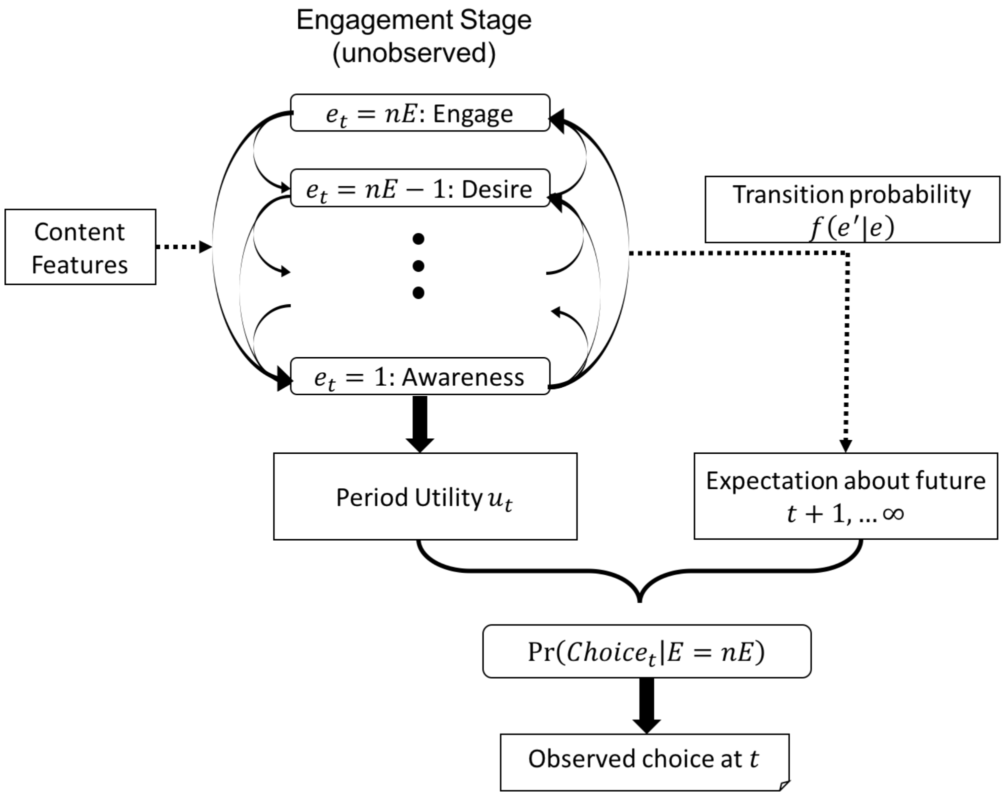


Figure 4 **Model Framework**

4.1. Model User Decisions

In each time period, $t \in \{1, \ldots, T_i\}$ (the total number of time periods T_i can vary among users), a new content unit $j \in \{1, \ldots, J\}$ will be shown on the mobile reading app. Then, a mobile user $i \in \{1, \ldots, I\}$ has the following three options.[8]

1. $d_{ijt} = 0$, the user chooses outside goods (i.e., leaves the mobile reading app platform) instead of the given content unit without any payment. The expected utility of choosing outside goods is normalized to zero.

[8] In our model, we do not consider users' un-subscribing decision. On the mobile app platform, by default the subscription contract will continue without any extra action. Also, tapstream data do not include the unsubscribing action. But this action can be captured in our model by assuming that the users chose outside options and had no choice on the app furthermore.

2. $d_{ijt} = 1$, the user chooses to continue reading with the pay-per-use option. But she needs to pay the fee, P_C, if the content unit is not free and she is not under any subscription contract. Note that if the user has been under a subscription contract and chooses to read a content unit at time t, the option she chooses, d_{ijt}, is equal to one, but with free payment (i.e., $P_C = 0$) after the subscription.

3. $d_{ijt} = 2$, the user chooses the subscription option with payment P_s. Then, up to the expiration of the subscription content, she has free access to all content units.

Note that we do not model mobile users' decision-making processes with respect to the content selection. In other words, if t is given, content unit j is given by the reading app itself. If the user has not finished one book, the reading app will automatically display the next content unit of the same book when the user completes reading the current content unit. If the user finishes one book, the reading app will recommend the first content unit of one new book for her decision-making with regard to whether to continue reading or not. Due to data limitation, we ignore the situation wherein mobile users self-select a new book either through the search option or through some other means such as first-page recommendation.[9] Also, from the reading app perspective, the firm is more interested in engaging consumers in the platforms than promoting a specific book. Thus we will ignore subscript j in the following discussions.

4.2. Period Utility Function

The mobile user's decision-making process is not based solely on the period utility but also on inter-temporal trade-offs, which means that she is a forward-looking consumer. The determined part of the utility function consists of two components: utility of money and

[9] So far, the mobile reading app has the largest market demand in China, thus it is not providing any strategic recommendations for its users.

utility of reading. Utility of money is a linear function of the price the user needs to pay at time t for option d_{it}. Utility of reading indicates the benefits of reading from the current content unit. We model this part as a user-engagement-specific constant. Mathematically,

$$
\begin{aligned}
&U(d_{it}=d,\mathtt{sub}_{it},F_t,e_{it}=e,\varepsilon_{it};\Theta)\\
&=\underbrace{\alpha(P_C\cdot\mathbb{1}\{d=1\}\cdot\mathbb{1}\{F_t=0\}+P_s\cdot\mathbb{1}\{d=2\})\cdot\mathbb{1}\{\mathtt{sub}_{it}=0\}}_{\text{Utility of money}}+\underbrace{\tilde{\omega_e}\cdot\mathbb{1}\{d\neq 0\}}_{\text{Utility of reading}}+\varepsilon_{it}(d_{it}),
\end{aligned}
\tag{1}
$$

where $\mathtt{sub}_{it}$ is the indicator of whether the mobile user i is under subscription contract or not. Mathematically, $\mathtt{sub}_{it}$ if and only if $\exists n\in[0,t], d_{i,t-n}=2$. With this indicator, our dynamic modeling framework can capture the fact that mobile users can gain more benefits from reading additional content after subscription, even though their period utility might be lower (i.e., the subscription price is higher than the per-content price) than that with the non-subscription option. F_t indicates whether the content unit read at time t is free (i.e., if the content unit is free, $F_t=1$). e_{it} denotes the user's engagement. We treat it as a hidden state used to predict users' probability of purchase. The number of engagement stages can be empirically tested, which will be discussed in the results section. α is the price coefficient, which is identical among users. Based on the engagement stages, we define $\tilde{\omega_e}$ as an engagement-specific parameter vector (Netzer et al. (2008)). Due to the identification issue, we do not include a constant term in the utility; otherwise, we cannot simultaneously estimate $\tilde{\omega}$ and the constant term. In our framework, the evolution of the hidden state can be affected by the features ($\boldsymbol{X_{it}}$) of the content read by the user at time t. Thus, the state variable vector in this model is $S_{ijt}=(e_{it},\mathtt{sub}_{it},F_t,\boldsymbol{X_{it}})$.[10] The utility function indicates that the utility of outside goods is normalized to zero. ε_{it} is the

[10] Note that price is not included, because both the pay-per-content price and the subscription price are constant over time.

idiosyncratic choice-specific shock, which is assumed to follow the Type I Extreme Value distribution (i.i.d). It is not content-specific, because we assume that j is fixed given t. This error term brings uncertainty to the model, because it captures the unobserved factors that would affect users' utility. Without this stochastic term, reading free content (or any content after subscription) would be consistently better than choosing outside goods. Examples of unobserved factors that might affect this stochastic term include promotion or advertisement of outside goods, social influence from friends, and so on.

To ensure identification of the hidden state, we assume the choice probabilities to be non-decreasing with an increasing engagement state value. Mathematically, this assumption is operationalized as

$$
\begin{aligned}
&\tilde{\omega_1} = \omega_1, \\
&\tilde{\omega_2} = \tilde{\omega_1} + exp(\omega_2), \\
&\cdots \\
&\tilde{\omega_{nE}} = \tilde{\omega_{nE-1}} + exp(\omega_{nE}),
\end{aligned}
$$

where ω is the estimate from data. This assumption ($\tilde{\omega_1} < \tilde{\omega_2} < \cdots < \tilde{\omega_{nE}}$) has been commonly used in previous HMM-related work (Abhishek et al. (2012), Ascarza and Hardie (2013), Netzer et al. (2008)).

4.3. State Evolution

The state variable vector in this model is $S_{ijt} = (e_{it}, \mathtt{sub}_{it}, F_t, \boldsymbol{X_{it}})$. We assume that all of the elements in the state vector are independent of each other as conditional on the given state values and decisions; therefore, the state transition probability f_S can be expressed as a multiplication of the four elements' transition probabilities $(f_e, f_F, f_{\mathtt{sub}}, f_{\boldsymbol{X}})$:

$$f_S(S'|S,d) = f_e(e'|S,d) \cdot f_F(F'|S,d) \cdot f_{\mathtt{sub}}(\mathtt{sub}'|S,d) \cdot f_{\boldsymbol{X}}(\boldsymbol{X'}|S,d). \tag{2}$$

According to the mobile reading app from which our data were collected, for each book, the first N content units are free. But the number of free content units in each book is determined by the app company. To estimate the transition probability $(f_F(F'|S,d))$ of the free content units, we follow the design of the app and assume that the probability of reading a free content unit at time t given last period's state is fixed.

Regarding the subscription option, the mobile user can benefit from it within the subscription contract period. Mathematically,

$$\texttt{sub}' = \begin{cases} 1 & \text{if } \texttt{sub} = 1 \text{ or } d = 2 \\ \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Therefore, as described in the above two equations, $f_F(F'|S,d)$ and $f_{\texttt{sub}}(\texttt{sub}'|S,d)$ are both determined once the conditional states and actions are fixed.

$\boldsymbol{X_{it}}$ is a vector of observed features describing the content units read by the user at time t. We include the following elements: book popularity indicator and book genre indicators. Similarly to the case for the above two state variables, we empirically estimate the transition probability of $\boldsymbol{X_{it}}$ from data.

The user engagement measurement (e_{it}) is hidden (i.e., not observed from data) in our model. Like Netzer et al. (2008), we model the transition between the engagement stages as a threshold model, wherein a discrete transition occurs if the transition propensity passes a threshold level. We assume that the transition probability is book-specific. In other words, the observed features of the content units $\boldsymbol{X_{it}}$ form the transition matrix of the hidden engagement stages. For example, if the user is provided with content units from a popular book, her transition propensity is likely to be shifted above the threshold to a higher state; otherwise, her engagement is transited to a lower state, because the estimated transition propensity is below the threshold. In addition, we allow the weights of the observed features to be heterogeneous when measuring the transition probability. That is, even with the same content units, users at different stages show diverse transition probabilities. Let nE denote the total number of engagement stages. Assuming that the unobserved shock of the transition propensity follows the Type I Extreme Value distribution (i.i.d), we model the non-homogeneous transition probabilities as the following ordered logit model (note we have one constraint $\sum_{e'=1}^{nE} f_e(e'|S)=1$):

$$
\begin{aligned}
&f_e(e'=1|S)=\frac{\exp(h(1,e)-\boldsymbol{\delta_e X})}{1+\exp(h(1,e)-\boldsymbol{\delta_e X})}\\
&f_e(e'|S)=\frac{\exp(h(e',e)-\boldsymbol{\delta_e X})}{1+\exp(h(e',e)-\boldsymbol{\delta_e X})}-\frac{\exp(h(e'-1,e)-\boldsymbol{\delta_e X})}{1+\exp(h(e'-1,e)-\boldsymbol{\delta_e X})}, e'\in\{2,\ldots,nE-1\} \quad (4)\\
&f_e(e'=nE|S)=1-\frac{\exp(h(nE-1,e)-\boldsymbol{\delta_e X})}{1+\exp(h(nE-1,e)-\boldsymbol{\delta_e X})}
\end{aligned}
$$

where $\boldsymbol{\delta_e}$ is a vector of the response coefficients, and $h(e',e)$[11] is the e' ordered logit threshold in state e.

4.4. Dynamics in Mobile Users' Decisions

Mobile users make decisions by maximizing the sum of discounted future period utilities as

$$\max_{D_i=\{d_{i1},d_{i2},\ldots,d_{iT_i}\}} E[\sum_{t=1}^{\infty}\beta^{t-1}U(d_{it},S_{it},\varepsilon_{it};\Theta)|S_0,\varepsilon_{i0}] \tag{5}$$

Note that in our setting given t, j is fixed (i.e., no strategic book recommendation or selection). But for a different user i, different j might be faced at the same t because different choices are made over time. To simplify, we remove notation j in the following discussion. The solution to the above dynamic programming problem is the same as that to the Bellman equation: we rewrite the utility from choosing $d_t=k, k\in\{0,1,2\}$ in state $S_t=(e,F,\texttt{sub})$, with $\varepsilon_t(d_t)=\varepsilon_{kt}$ as the additive structure

$$U(S_t,d_t=k)=u_k(S_t)+\varepsilon_{kt}. \tag{6}$$

[11] To guarantee all the transition probabilities are within the range $[0,1]$, we need to include an additional assumption: $h(e',e)\geq h(e'-1,e), \forall e'=2,\ldots,nE$.

Table 2 Summary of Notations

Notation	Description
i, j, t	Indices of mobile user, content unit, and period
I, J, T_i	Total numbers of users, chapters, periods (for user i)
$n^{\{S\}}, n^{\{E\}}, n^{\{D\}}$	Total numbers of states, engagement stages, and decision choices
d_{ijt}	User i's choice at time t (for content unit j)
P_c, P_S	Per-content price (0.12 CNY) and subscription price (5 CNY)
$\texttt{sub}_{it}, F_t$	Indicator of subscription, and free content unit
e_{it}	User i's engagement stage at time t
α	Price coefficient in utility function
$\tilde{\omega_e}$	Engagement-specific coefficient of reading in period utility function
$(f_e, f_F, f_{\texttt{sub}}, f_{\boldsymbol{X}})$	Transition probabilities
$\boldsymbol{X_{it}}$	Feature vector of content units read by user i at time t
$h(e', e)$	e'-ordered logit threshold in state e
$\boldsymbol{\delta_e}$	Vector of engagement-specific response coefficients of $\boldsymbol{X_{it}}$
β	Discount factors in lifetime utility function (assumed as 0.99)

Then, by following the assumptions in Rust (1987), and assuming that ε_{kt} is i.i.d. across actions and time periods, we obtain the associated value function

$$\begin{aligned}\nu(S,\varepsilon) &= \max_{k\in\{0,1,2\}}\{u_k(S)+\varepsilon_k+\beta E[\nu(S',\varepsilon')|S,d=k]\}\\ &= \max_{k\in\{0,1,2\}}\{u_k(S)+\varepsilon_k+\beta\int\nu(S',\varepsilon')f_S(S'|S,k)g(\varepsilon')d(S',\varepsilon')\}.\end{aligned} \tag{7}$$

A summary of all variables and notations is presented in Table 2.

4.5. Identification and Estimation

As described in An et al. (2013), our proposed model can theoretically identify the transition probability of unobserved states and the conditional probability of outcomes given those states. According to the identification theorem discussed in Magnac and Thesmar (2002), given the distribution of error term (assumed as Type I Extreme Value distribution), the transition matrix (identified from the above assumption), discount factor (fixed to 0.99[12]) and utility of outside options normalized to 0, the utility function can be identified for all states and all decision choices.

Empirically, as shown in Equation (1), ω_e is the engagement-specific coefficient of reading utility. The repeated reading activities of the same user or reading activities from similar users at the same engagement stage, conditional on the same price (i.e., the same free content after subscription, the same pay-per-content unit price before subscription, or the same bundling price upon subscription decision), can help us to identify the reading utility coefficient at each engagement stage. More specifically, conditional on the same engagement stage and price, if we observe that users choose to read the content rather than switch to outside options, we can infer that the reading utility coefficient ω_e is high.

Conditional on the reading utility, we can then identify the price coefficient α through users' repeated purchases (reading activities). Specifically, a high magnitude of α indicates a higher reading frequency in the post-subscription period, meaning that users are more price sensitive and that more reading activities can minimize the waste of money on subscriptions. Moreover, if there is subscription behavior, we observe the user's repeated reading activities before and after the subscription. On that basis, we can also identify the price coefficient with the change in the frequency of repeated reading activities from the same user before and after subscription. For example, if we observe a significant increase in the frequency of repeated reading activities from users after subscription, this could indicate that they are relatively price sensitive (with a high magnitude of α). Additionally, the dynamics in the mobile users' forward-looking behavior also can help us identify the price elasticity. For example, if we observe a high frequency of repeated pay-per-content reading activities from users with no subscription, this might indicate that those users are less price elastic (with a low magnitude of α).

[12] The qualitative nature of the results is robust to several other values of the discounted factors.

Table 3 Comparison of HMM Models

Model	# of states	Log-likelihood	AIC	BIC	# of Variables
HMM	Three	-10327.3	-10415.7	-10365.3	19
	Four	-9524.7	-9582.7	-9659.6	29
	Five	-10277.8	-10359.8	-10468.5	41
	Six	-12314.0	-12424.0	-12569.8	55

Note: The best model in each column is shown in bold.

We follow the nested pseudo maximum likelihood procedure that was used in Aguirregabiria and Mira (2007), as well as Huang et al. (2015), to solve the hidden-state single-agent dynamic problem. The detailed estimation procedure is provided in the online appendix.

5. Results

The first stage in estimating our model is to determine the number of engagement stages (i.e., the hidden state). We compared several alternative models with different numbers of engagement stages (varying from 3 to 6). With respect to the AIC and BIC criteria, the results indicated that the best solution is to identify *four* engagement stages (Table 3). We herein label the four stages as "*aware*," "*familiar*," "*active*," and "*addicted*."

Table 4 reports the estimated parameters in the utility function. The second and third columns are the values obtained from our estimation procedure. Note that the price coefficient is user inherent attribute and is common across all stages. Whereas, the reading utility coefficient is stage-specific. With the assumption of non-decreasing reading utility values, we converted the estimated ω_e to the true reading utility coefficient ($\tilde{\omega_e}$) reported in the fourth column of Table 4. As Equation (1) shows, the utility function has two components: money utility and reading utility. Assuming that everything else is constant, a lower reading utility has a lower probability of purchase. We found that users can gain

Table 4 Estimated Utility Function Parameters

Engagement stage	Price coefficient α	ω_e	Reading utility coefficient ($\tilde{\omega_e}$)
$e = 1$ (aware)		0.7906***	0.7906
		(0.0009)	
$e = 2$ (familiar)		-1.3614***	1.0469
	-0.3778***	(0.0808)	
$e = 3$ (active)	(0.0000)	-2.0107***	1.1808
		(0.0026)	
$e = 4$ (addicted)		-1.7664***	1.3508
		(0.0009)	

Notes:
***$p < 0.001$; ** $p < 0.01$; * $p < 0.05$.
Standard errors are shown in parentheses.

32.42% more utility from reading when they move up from an "aware" mind stage to a "familiar" stage. Table 5 reports the estimated parameters in the transition function of the engagement stages. The threshold $h(e', e)$ (shown in rows 2-4 of Table 5) is the e' ordered logit threshold in state e, meaning that if a user's propensity from reading experience is above the high threshold, she is very likely to move forward to a higher engagement stage. The results show that for "aware" and "active" stages, the highest threshold is much larger than the lowest threshold, indicating that the mobile users at these stages are not stable and more likely to move down than to move up.

To better interpret the results, in Table 6, we provide the estimated engagement state transition. We first computed the book-specific transition matrices by incorporation with the different values of the three feature indicators. Table 6 provides the weighted averages of those matrices. First, it shows that users at most of the engagement stages are highly likely to switch to the lowest stage. This downward trend can be more significant with non-popular books. This finding suggests that the mobile app company could lose their mobile

Table 5 Estimated Transition Matrix Parameters

Variable	Notation	$e = 1$ (aware)	$e = 2$ (familiar)	$e = 3$ (active)	$e = 4$ (addicted)
Threshold 1 (lowest)	$h(1, e)$	-0.9535***	1.3354***	0.3856***	-1.1501***
		(0.0019)	(0.1593)	(0.0902)	(0.0053)
Threshold 2	$h(2, e)$	-0.8925***	-1.5416***	0.4211***	-1.0876***
		(0.0008)	(0.0754)	(0.0202)	(0.0618)
Threshold 3 (highest)	$h(3, e)$	4.6663***	1.5585***	3.5767***	-1.0640***
		(0.0324)	(0.0357)	(0.0083)	(0.0738)
Popularity indicator	X_1	-0.2827***	1.8529***	-1.1012***	2.3624***
		(0.0012)	(0.2378)	(0.1119)	(0.0087)
Fiction genre indicator	X_2	-1.3257***	1.0932*	-0.5990**	2.4137***
		(0.0011)	(0.1391)	(0.2045)	(0.0000)
Casual genre indicator	X_3	-0.8003***	0.5227**	0.4919**	0.3134***
		(0.0062)	(0.2609)	(0.2252)	(0.0575)

Notes:
***$p < 0.001$; ** $p < 0.01$; * $p < 0.05$.
Standard errors are shown in parentheses.

users since they are highly likely to leave the market or become less engaged without any additional policy intervention. The only exception is the highest stage. On average, users at the "addicted" mind stage behave stably by staying at the same stage with the highest probability. On the other hand, however, our results indicate that users at the lower stages still have potential business value. For example, "aware" or "familiar" users might jump to a higher stage (i.e., the "active" or "addicted" stage) with a probability similar to that of being at the lowest stage. This finding indicates that users are heterogeneous with different inherent engagement stages and preferences. Therefore, a targeting strategy corresponding to their engagement stages would be effective.

Table 6 Estimated Transition Matrix of Engagement Stages

$f(e \mid e', \bar{X})$	$e=1$ (aware)	$e=2$ (familiar)	$e=3$ (active)	$e=4$ (addicted)
$e'=1$	0.5725	0.0142	0.4103	0.0030
$e'=2$	0.4047	0.0407	0.0034	0.5512
$e'=3$	0.7585	0.0059	0.2210	0.0147
$e'=4$	0.0432	0.0024	0.0009	0.9535

6. Policy Simulations

Given the estimated parameters, we numerically solve for the reading app's optimal pricing strategy. For each individual mobile user, we first calculate her choice probability at each period given the same state values as we observe from the data. To measure the effects of policy intervention, we use the users' total payment *per period*. Specifically, we first compute the users' decision probability at each period given the same state variable values we observe from data;[13] we then calculate the expected payment amount using the decision probability and average to the expected payment per period; finally, we aggregate all users to compute the overall expected revenue for one single period.

We first examine the optimal per-content price (or subscription price) by fixing the current subscription price (or per-content price). The estimation shows that given the current subscription price (5 CNY), the optimal per-content price is approximately 4.3 CNY. Based on this new per-content pricing strategy, the overall per-period revenue increases from 2820.17 CNY to 8864.89 CNY (column five "Total revenue" in Table 7). The results

[13] The simulated decision probability is computed for each period. In each period, we observe mobile users' past reading behavior as their state variable values. The simulated decision probability is still based on users' forward-looking behavioral patterns and the same state variables, but with different policy interventions (e.g., different prices). An alternative computation method is to compute users' complete decision sequences from the first period using the initial state variable values. This requires us to compute the sequential state variable values as well, which can incur more uncertainty in the predicted decision probability.

are shown in the third row (i.e., With identical optimal per-content price) in Table 7. On the other hand, given the current per-content price (0.12 CNY), the optimal subscription price is approximately 3.1 CNY. Therefore, this mobile reading app company should be advised to narrow the price gap between the pay-per-content and subscription options. Even though the above numerical results are specific to the mobile reading app company, our model framework can be applied to any companies (or industries) with similar contexts to help them explore their optimal pricing strategy.

Secondly, we design a third-degree price discrimination by assigning different prices to users at different engagement stages. To examine the effectiveness of the personalized intervention, we conduct the following policy simulation to compare the total revenues before and after the personalization: first, we compute the optimal engagement-specific prices of paid content; then we use the four optimal prices to compute the total revenues under personalized (engagement-specific) per-content pricing.[14] The results are shown in the last row (i.e., With personalized optimal per-content price) in Table 7. There are two interesting findings here. (1) From the current pricing strategy (i.e., row two "Without intervention") to the non-personalized optimal per-content price (i.e., row three "With identical optimal per-content price"), both pay-per-content revenue and subscription revenue increase significantly. This is due mainly to the fact that the demand curve is inelastic, according to which, an increase in price leads only to a small decrease in demand. On the other hand, the significant increase in price drives a steep increase in pay-per-content revenue. Meanwhile, the price gap between the pay-per-use and subscription options is narrowed, leading some users to switch from paying per content to subscribing. The greater subscription demand explains

[14] With this simulation mechanism, we assume that mobile users are aware of the facts that the reading app can detect their real-time engagement stages and that the per-content price varies with them.

Table 7 Revenue Comparison (Unite: CNY)

	(Average) per-content price	Pay-per-content revenue	Subscription revenue	Total revenue
Without intervention	0.12	371.36	2448.81	2820.17
With identical optimal per-content price	4.30	4683.97	4180.92	8864.89
With personalized optimal per-content price	3.40	4687.72	4286.33	8974.05

the increase in subscription revenue. (2) From the non-personalized (i.e., row three) to personalized pricing strategy (i.e., row four "With personalized optimal per-content price"), both pay-per-content revenue and subscription revenue increase. The main reason is that the third-degree price discrimination strategy extracts all of the consumer surplus. When we translate the revenue increase into percentage improvement compared with the original revenue, we find an additional 3.87% increase from the non-personalized strategy to the personalized pricing strategy. This finding highlights our engagement detection method by showing the potential value of the engagement-based personalization strategy. Interestingly, the total revenue increases even though the average pay-per-content price is lower than the non-personalized optimal price (i.e., 3.40 vs. 4.30). This finding is interesting, because in the simulation, we simply assume that at each period, the mobile users' state values stay the same as what we observe from the data, while in the real world, a lower price can attract even more users and encourage them to move forward to a more engaged level. Therefore, we can consider this 3.87% as a lower bound for revenue improvement from the non-personalized pricing to a stage-specific pricing strategy.

7. Randomized Field Experiment

In order to further examine the effectiveness of different targeting strategies under different user engagement levels, we designed and conducted a field experiment to measure the effects from a causal perspective. The field experiment was conducted from September 28 to December 12 2015. Our overall goal is to examine the impact of potential pricing strategy on user purchase behavior under different engagement stages. To control for the potential effect from the design of the ad copy, we considered two forms of pricing promotion design at the same price discount level: price discount and free content. In particular, users were randomly assigned to three groups: two treatment groups and one control group. The first treatment was price promotion, whereby users were provided with discount vouchers (0.60 CNY) for reading any content unit in the app; the second treatment was free content promotion, whereby users were provided with five free-content-unit vouchers (with a total worth of also 0.60 CNY) for reading any content unit in the app. In the control group, users were not provided with any promotion but rather with a reminder notification messages (i.e., non-pricing advertisement), in order to identify the incremental effects (lift) for the two treatments. To compare the performances between the treatment groups and the control group, we determined whether the pricing promotions or non-pricing advertisement could achieve better performance in terms of sales lift. In other words, the two treatments were designed to test whether users show more attention to money or to products in marketing promotions. The pre-treatment period was from September 28 to October 27 2015; the treatment period was from October 28 to November 8 2015; and the post-treatment period was from November 9 to December 12 2015. In each group, we used the tapstream data from a random sample of 1,000 active users to conduct the following analysis.[15]

[15] Initially, we assigned 20,000 users to each group; however, not all of them responded. We first removed the non-active users who did not have any record during the three months. Then, to guarantee comparability among the three groups, we randomly selected 1,000 users from each. We checked our results with the different random samples and found them to be consistent.

7.1. Model-free Evidence

First, we calculated the daily average number of read content units per user and plotted the change over time (Figure 5). The treatment period was from day 30 to day 43. The plot indicated that over time, users read less and less but that the promotions could alleviate this negative effect. We then plotted similar trends by engagement stages in Figure 6. Specifically, we used the estimated coefficients from our structural model to identify each mobile user' engagement stage at the beginning of treatment. Then, we divided all of the users into four groups based on their engagement stages and then plotted the corresponding behavior trends. The four plots showed that the positive effects brought by promotions varied according to the users' engagement stages. For example, while the improvement between the treatment groups and the control group was significant when users were at the "familiar" mind stage, users at the "active" stage showed insignificant changes for either the price promotion or free content promotion.[16]

7.2. Difference-in-difference Analysis

Furthermore, we quantified the effects using a Difference-in-Difference analysis, applying the equation

$$Y_{it} = \alpha_0 + \alpha_1 \texttt{Test}_t + \alpha_2 \texttt{Treat1}_i \times \texttt{Test}_t + \alpha_3 \texttt{Treat2}_i \times \texttt{Test}_t + \xi_j + \varepsilon_{it} \tag{8}$$

where Y_{it} is the outcome measure of user i at time t, and using the daily number of read content units as the measure. $\texttt{Treat1}_i$ indicates whether user i is in the first treatment: price promotion; $\texttt{Treat2}_i$ indicates whether user i is in the second treatment: free content

[16] Note that the mobile user segments are based on the users' engagement stages at the beginning of treatment. These might change after treatment. For example, it is possible that a user at the "active" stage segment would move to other stages during or after treatment.

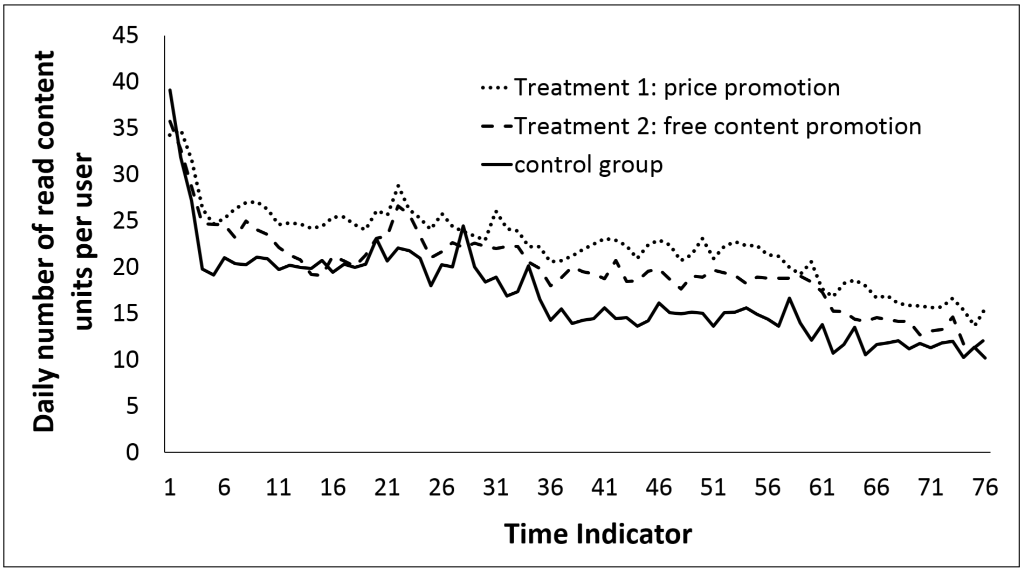
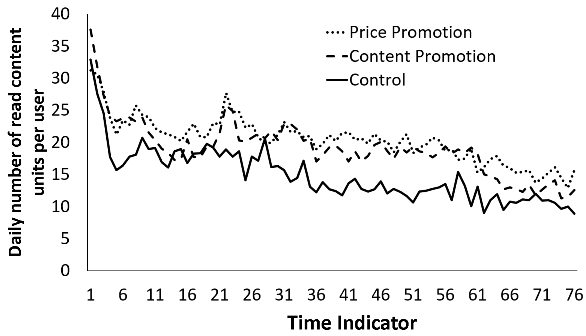
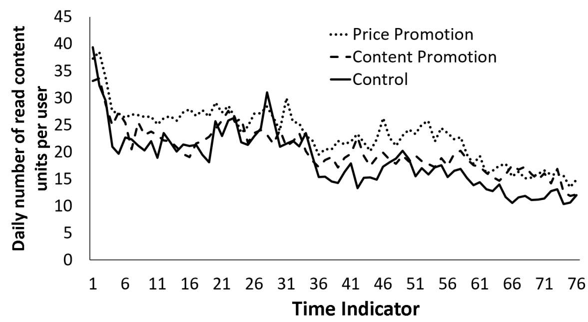


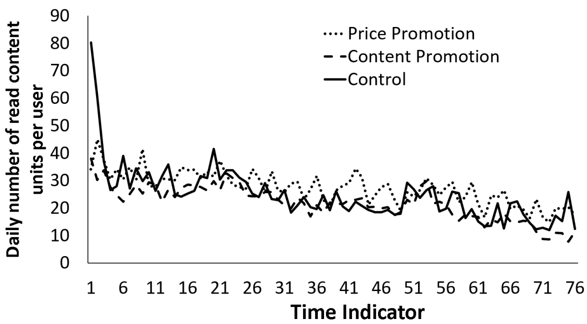
Figure 5 **Overall Effects of Promotions.**



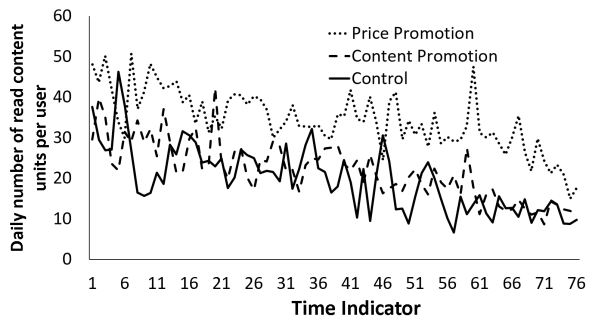
(a) $e = 1$ Aware stage



(b) $e = 2$ Familiar stage



(c) $e = 3$ Active stage



(d) $e = 4$ Addicted stage

Figure 6 **Engagement-specific Effects of Promotions**

promotion; and $\texttt{Test}_t$ indicates whether period t is in the during-treatment period. ξ_j, finally, represents the group-level fixed effects.

The results are shown in Table 8. *First*, the negative coefficients of Test variable suggest that over time, the mobile reading app market has a dismal customer attrition picture, but that mobile users' engagement stages and personalized targeting are key to alleviating such attrition. *Second*, the comparison among the interaction terms (`Treat1`, `Treat2`, and `Test`) indicates that the overall effect of free content promotion (i.e., 1.4265) is slightly better than that of price promotion (i.e., 1.3639). However, the effectiveness of treatments in alleviating the slippery slope of app consumption are different among the four engagement stages. On the whole, the treatment effects are mainly driven by users at the familiar and addicted stages. Interestingly, the optimal promotions differ among users: familiar users prefer price promotion to free content promotion, while addicted users show the opposite. Quantitatively, price promotion encourages familiar users to read 2.33 more content units per day. One managerial implication from this finding is that the app company should focus more on choosing the right promotion strategies to target familiar and addicted users, whereas they should avoid using the same promotion strategies for users at the other two stages (i.e., new users and active users) when providing personalized strategies.

The findings from the field experiment suggest that the effects of app notifications are dependent on the right mix of data analytics (user engagement modeling) and ad copy creativity (messages emphasizing free content promotion or price discount promotion). Because over 50% of app users find app notifications annoying (Localytics (2016)), firms should tap into their user engagement analytics to create personalized notifications. Such app notifications are what specific user segments want to receive; thus, they can generate substantially larger sales impacts than broadcasted notifications.

Table 8 Field Experiment Analysis

	Overall	$e = 1$ (aware)	$e = 2$ (familiar)	$e = 3$ (active)	$e = 4$ (addicted)
`Treat1` $\times$ `Test`	1.3639**	0.4330	2.3282**	-0.3430	1.8796**
	(0.4030)	(0.6574)	(0.7658)	(1.0229)	(0.8669)
`Treat2` $\times$ `Test`	1.4265**	0.0681	1.9186**	1.1356	2.3803**
	(0.4091)	(0.6499)	(0.7836)	(1.0773)	(0.8895)
`Test`	-7.6452***	-6.3499***	-8.2233***	-8.8396***	-8.8396***
	(0.2945)	(0.4852)	(0.5456)	(0.7545)	(0.6449)
Observations	220,702	71,643	72,189	28,727	48,143

Notes:

***$p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Standard errors are shown in parentheses.

`Treat1`: Indicator of price promotion treatment

`Treat2`: Indicator of free content promotion treatment

`Test`: Indicator of treatment period

8. Conclusion

The main goal of the present study was to model the changes of mobile users' engagement in an app platform, the results of which provide us with a plethora of managerial implications for both marketers and researchers. We based our research on an analysis of individual mobile users' continuous reading records on a mobile reading app. We proposed and estimated a structural framework that combines two individual-level behavioral models: the single-agent dynamic discrete choice model and the HMM. Such combination allows us to detect individuals' time-varying engagement evolution in consideration of their forward-looking behaviors. Our estimation suggests that mobile users can be classified into four engagement stages. Then, we examined the value of detecting mobile user engagement stages by designing the optimal identical and personalized pricing strategies. We also examined the effects of personalized targeting strategies from the causal perspective by designing and conducting a randomized field experiment. Our proposed method can be applied to other digital infrastructures and platforms in similar marketing settings to help platforms to identify their users' behavior and adjust their targeting strategies and improve their business models accordingly. In addition to the digital platform market, our method can also provide guidelines for other durable products, the demand for which comes from consumers' forward-looking behaviors.

Our paper has a few limitations that nonetheless provide interesting opportunities for future research. First, our data lack information such as exact reading time. Future studies could take into account a more precise measurement for the period of reading time if this information is available. Second, due to the current design of app setting, we did not consider the product selection problem in mobile users' decision-making process. This can be treated in future work in order to understand how users might search and choose the next content/book to read. Third, there is an engagement hierarchy in terms of book genres, books, and chapters (books within a genre, chapters within a book), and engagement evolution and utility preference vary with multiple factors, including time-of-day, weather, and others. For example, people in metropolitan cities would spend more time on public transportation, which might allow them to be more engaged in the app during their commuter time. Further study respecting these issues might be more interesting and practical to mobile market managers. Our model, with the necessary adjustments, offers the potential to add in such factors.

Acknowledgments

References

Abhishek, Vibhanshu, Peter Fader, Kartik Hosanagar. 2012. Media exposure through the funnel: A model of multi-stage attribution. *Working Paper* .

Aguirregabiria, Victor, Pedro Mira. 2007. Sequential estimation of dynamic discrete games. *Econometrica* **75**(1) 1–53.

An, Yonghong, Yingyao Hu, Johns Hopkins, Matt Shum. 2013. Identifiability and inference of hidden markov models. Tech. rep., Citeseer.

Andrews, Michelle, Xueming Luo, Zheng Fang, Anindya Ghose. 2015. Mobile ad effectiveness: Hyper-contextual targeting with crowdedness. *Marketing Science* .

Ascarza, Eva, Bruce GS Hardie. 2013. A joint model of usage and churn in contractual settings. *Marketing Science* **32**(4) 570–590.

Bakos, Yannis, Erik Brynjolfsson. 1999. Bundling information goods: Pricing, profits, and efficiency. *Management science* **45**(12) 1613–1630.

Brodie, Roderick J, Linda D Hollebeek, Biljana Juric, Ana Ilic. 2011. Customer engagement: conceptual domain, fundamental propositions, and implications for research. *Journal of Service Research* 1094670511411703.

Choudhary, Vidyanand. 2010. Use of pricing schemes for differentiating information goods. *Information Systems Research* **21**(1) 78–92.

Choudhary, Vidyanand, Anindya Ghose, Tridas Mukhopadhyay, Uday Rajan. 2005. Personalized pricing and quality differentiation. *Management Science* **51**(7) 1120–1130.

Claussen, Jörg, Tobias Kretschmer, Philip Mayrhofer. 2013. The effects of rewarding user engagement: the case of facebook apps. *Information Systems Research* **24**(1) 186–200.

ComScore. 2016. The 2015 u.s. mobile app report .

Connault, Benjamin. 2014. Hidden rust models. *Working Paper* .

Faugère, Christophe, Giri Kumar Tayi. 2007. Designing free software samples: a game theoretic approach. *Information Technology and Management* **8**(4) 263–278.

Ghose, Anindya, Avi Goldfarb, Sang Pil Han. 2012. How is the mobile internet different? search costs and local activities. *Information Systems Research* **24**(3) 613–631.

Ghose, Anindya, Sang Pil Han. 2014. Estimating demand for mobile applications in the new economy. *Management Science* **60**(6) 1470–1488.

Goettler, Ronald L, Karen Clay. 2011. Tariff choice with consumer learning and switching costs. *Journal of Marketing research* **48**(4) 633–652.

Han, Sang Pil, Jinpyo Hong, Jae Yun Moon, Wonseok Oh. 2015a. The economics of all-you-can-read pricing: Tariff choice, contract renewal, and switching for e-book purchases. *Working Paper* .

Han, Sang Pil, Sungho Park, Wonseok Oh. 2015b. Mobile app analytics: A multiple discrete-continuous choice framework. *Management Information Systems Quarterly (MISQ)* .

Hartigan, John A, Manchek A Wong. 1979. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **28**(1) 100–108.

Hollebeek, Linda D, Mark S Glynn, Roderick J Brodie. 2014. Consumer brand engagement in social media: Conceptualization, scale development and validation. *Journal of interactive marketing* **28**(2) 149–165.

Hotz, V Joseph, Robert A Miller, Seth Sanders, Jeffrey Smith. 1994. A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies* **61**(2) 265–289.

Huang, Ke-Wei, Arun Sundararajan. 2011. Pricing digital goods: Discontinuous costs and shared infrastructure. *Information Systems Research* **22**(4) 721–738.

Huang, Yan, Param Vir Singh, Anindya Ghose. 2015. A structural model of employee behavioral dynamics in enterprise social media. *Management Science* **61**(12) 2825–2844.

Kim, Young Hoon, Dan J Kim, Kathy Wachter. 2013. A study of mobile user engagement (moen): Engagement motivations, perceived value, satisfaction, and continued engagement intention. *Decision Support Systems* **56** 361–370.

Kumar, V, S Sriram, Anita Luo, Pradeep K Chintagunta. 2011. Assessing the effect of marketing investments in a business marketing context. *Marketing Science* **30**(5) 924–940.

Lambrecht, Anja, Bernd Skiera. 2006. Paying too much and being happy about it: Existence, causes, and consequences of tariff-choice biases. *Journal of marketing Research* **43**(2) 212–223.

Laxman, Srivatsan, Vikram Tankasali, Ryen W White. 2008. Stream prediction using a generative model based on frequent episodes in event sequences. *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 453–461.

Lee, Clarence, Vineet Kumar, Sunil Gupta. 2013. Designing freemium: a model of consumer usage, upgrade, and referral dynamics. *MimeoLeung, Michael (2013), Two-step estimation of network formation models with incomplete information* .

Li, Hui. 2015. Intertemporal price discrimination with complementary products: E-books and e-readers. *Working Paper* .

Liu, Charles Z, Yoris A Au, Hoon Seok Choi. 2012. An empirical study of the freemium strategy for mobile apps: Evidence from the google play market. *ICIS* .

Localytics. 2016. An even better recipe for perfect push notifications .

Luo, Xueming, Michelle Andrews, Zheng Fang, Chee Wei Phang. 2013. Mobile targeting. *Management Science* **60**(7) 1738–1756.

MacDonald, Iain L, Walter Zucchini. 1997. *Hidden Markov and other models for discrete-valued time series*, vol. 110. CRC Press.

Magnac, Thierry, David Thesmar. 2002. Identifying dynamic discrete decision processes. *Econometrica* **70**(2) 801–816.

Miller, Robert A. 1984. Job matching and occupational choice. *The Journal of Political Economy* 1086–1120.

Montgomery, Alan L, Shibo Li, Kannan Srinivasan, John C Liechty. 2004. Modeling online browsing and path analysis using clickstream data. *Marketing Science* **23**(4) 579–595.

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

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

Punera, Kunal, Srujana Merugu. 2010. The anatomy of a click: modeling user behavior on web information systems. *Proceedings of the 19th ACM international conference on Information and knowledge management*. ACM, 989–998.

Rust, John. 1987. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society* 999–1033.

Shiller, Ben, et al. 2016. Personalized price discrimination using big data. Tech. rep.

Vigna, Stefano Della, Ulrike Malmendier. 2006. Paying not to go to the gym. *The American Economic Review* **96**(3) 694–719.

Vivek, Shiri D, Sharon E Beatty, Robert M Morgan. 2012. Customer engagement: Exploring customer relationships beyond purchase. *Journal of Marketing Theory and Practice* **20**(2) 122–146.

Wagner, Thomas M, Alexander Benlian, Thomas Hess. 2013. The advertising effect of free–do free basic versions promote premium versions within the freemium business model of music services? *System Sciences (HICSS), 2013 46th Hawaii International Conference on*. IEEE, 2928–2937.

Xu, Kaiquan, Jason Chan, Anindya Ghose, Sang Pil Han. 2016. Battle of the channels: The impact of tablets on digital commerce. *Management Science* .

Online Appendix

Appendix A: Identification and Estimation

The structural parameters in our model are $\{\alpha, \boldsymbol{\omega}, \boldsymbol{\delta}, \boldsymbol{h}, \beta\}$. The three elements in bold are parameter vectors of engagement-specific coefficients: for example, $\boldsymbol{\omega} = \{\omega_1, ..., \omega_{nE}\}$. Due to the identification issue of discount factor, we assume $\beta = 0.99$. Thus, the model primitives that need to be estimated are: $\Pi = \{\alpha, \boldsymbol{\omega}, \boldsymbol{\delta}, \boldsymbol{h}\}$.

A.1. Identification

Our model combines both HMM model and single-agent dynamic discrete choice model. We will first discuss the identification issue from the theoretical perspective using several existing theorems in the literature, and then discuss empirically the identification issue from our data set.

A.1.1. Theoretical Identification An et al. (2013) proves that to identify the transition probability of unobserved states and conditional probability of outcome given states, we need four assumptions: (1) The time series process of unobserved states and outcome should be strictly stationary and ergodic; (2) The probability matrix $S_{Y_t,Y_{t-1}}$ has full rank; (3) The conditional probability should be different with different value of unobserved states; (4) There exists a functional $F(\cdot)$ such that $F(f_{Y_t|S_t})$ is monotonic in S_t (Y_t is outcome variable and is the unobserved state). We argue that our model satisfies the above four assumptions:

(1) We allow mobile users to move upward and downward in the engagement funnel, which satisfies the ergodicity assumption; To meet the stationarity assumption, we need to assume the initial probability of state $\boldsymbol{\Phi_i}$ is stationary. To achieve that, we follow Netzer et al. (2008) and Huang et al. (2015), by calculate the stationary $\boldsymbol{\Phi_i}$ from the equation: $\boldsymbol{\Phi_i} = \boldsymbol{\Phi_i}\tilde{\boldsymbol{Q}}_{\boldsymbol{i}}$, where $\tilde{\boldsymbol{Q}}_{\boldsymbol{i}}$ is the transition matrix calculated at the mean value of covariates for individual i.

(2) Assumption 2 requires that $P(Y|S=j)$ is not a linear function of $P(Y|S=k), \forall k \neq j$. This can be easily satisfied from our model setting.

(3) In our model, different values of unobserved state are corresponding with different coefficients in the utility function, which will lead to different value functions and probability of decision choices.

(4) One easy way to interpret assumption 4 is that given Y, different values of unobserved state have different conditional probability, so that the conditional probabilities can be correctly ordered. Similar to (2), our model setting meets this requirement.

According to the identification theorem discussed in Magnac and Thesmar (2002), given the distribution of error term (assumed as Type I Extreme Value distribution), the transition matrix (identified from the above assumption), discount factor (fixed to 0.99) and utility of outside option normalized to 0, the utility function can be identified for all states and all decision choices.

A.1.2. Empirical Identification As shown in Equation (1), ω_e is the engagement-specific coefficient of reading utility. The repeated reading activities of the same user or reading activities from similar users at the same engagement stage, conditional on the same price (i.e., same free content after subscription, same pay-per-content unit price before subscription, or same bundling price upon subscription decision), can help us identify the reading utility coefficient at each engagement stage. More specifically, conditional on the same engagement stage and price, if we observe that users choose to read the content rather than switch to outside options, we can infer the reading utility coefficient ω_e is high.

Conditional on the reading utility, we can then identify the price coefficient, α. Note that due to the unique setting of our context, we only observe three levels of price in the data: free, pay-per-content price, and subscription price. Nevertheless, we observe users' repeated purchases (reading activities). We are able to identify the price elasticity from users' frequency of repeated reading activities after subscription. Specifically, a high magnitude of indicates higher reading frequency in post-subscription period, meaning that the users are more price sensitive and more reading activities can eliminate the waste of money on subscription. Moreover, if there is a subscription behavior, we observe the user's repeated reading activities before and after the subscription. Then we can also identify the price coefficient with the change in the frequency of repeated reading activities from the same user before and after subscription. For example, if we observe a significant increase in the frequency of repeated reading activities from the users after subscription, this may indicate that the users are relatively price sensitive (with high magnitude of α). In addition, the dynamics in the mobile users' forward-looking behavior can also help us identify the price elasticity. For example, if we observe high frequency of repeated pay-per-content reading activities from the users with no subscription, this may indicate the users are less price-elastic (with low magnitude of α).

A.2. Estimation Procedure

We follow the nested pseudo maximum likelihood procedure that was used in Aguirregabiria and Mira (2007) and Huang et al. (2015), to solve the hidden-state single-agent dynamic problem.

Step 1: Estimate $f_{\boldsymbol{X}}(\boldsymbol{X'}|S, d, \boldsymbol{\lambda_1})$.

In the observed reading experience variable vector $\boldsymbol{X}$, we consider three elements: (1) number of content units that have been read; (2) number of free content units that have been read; (3) average reading time per day in the previous reading experience. Those three elements, however, are all continuous. To simplify the estimation process, we discretize each of the three variables into two levels: high and low, based on the average value of each variable. As we discussed before, the transition probability of this observed variable vector is a function of previous state variables. We assume the function to be a logit function and use data to estimate the coefficients vector $\boldsymbol{\lambda_1}$ out of the iteration loop. Then we will get $\hat{f}_{\boldsymbol{X}}(\boldsymbol{X'}|S,d,\boldsymbol{\lambda_1})$.

Step 2: Conduct the joint likelihood function.

Let $D_i = \{d_{i1}, d_{i2}, \ldots, d_{iT_i}\}$ represent the choice sequence for individual i, $S_i = \{S_{i1}, S_{i2}, \ldots, S_{iT_i}\}$ represent the state sequence, and $\boldsymbol{\Phi_i}(1 \times n^{\{S\}})$ represent the initial state distribution. Due to the identification concern, to compute the initial state probability $\boldsymbol{\Phi_i}$, we need to solve the equation $\boldsymbol{\Phi_i} = \boldsymbol{\Phi_i}\boldsymbol{\tilde{Q}_i}$ (under the constraint $\sum^{n^{\{S\}}} \pi_{iS} = 1$) where $\boldsymbol{\tilde{Q}_i}$ is the transition matrix calculated at the mean value of covariates for individual i. For simplicity, we use O_i to represent the observed state sequences (i.e., $O_i = \{S_i \setminus e_i\}$). Then, the probability (i.e., individual likelihood function) of the observed outcome sequence D_i and observed state sequence O_i is:

$$l(D_i, O_i) = P(D_i, O_i|\boldsymbol{\lambda_2}) = \sum_{e_i} \big(P(D_i|O_i, e_i) \times P(e_i|O_i, \boldsymbol{\lambda_2}) \times P(O_i|\boldsymbol{\lambda_1})\big). \tag{A.1}$$

Step 3: Calculate the parameters in the transition matrix and utility function.

Let $\texttt{CCP}(n^{\{S\}} \times n^{\{D\}})$ be a matrix of conditional choice probabilities, where each element $\texttt{CCP}_{jk}$ is the probability of choosing decision k given the state j. Let Q_d represent the action-specific state transition matrix. Then the unconditional state transition matrix $Q = \sum_d \tilde{\texttt{CCP}(\cdot,d)} \odot Q_d$, where $\tilde{\texttt{CCP}(\cdot,d)}$ is a $n^{\{S\}}$-times copy of the d^{th} column of $\texttt{CCP}$ (i.e., $\texttt{CCP}(\cdot,d)$) and its dimension is $n^{\{S\}} \times n^{\{S\}}$. We use $\odot$ to denote an element-by-element multiplication. By following Aguirregabiria and Mira (2007), we can compute the state-specific value function V as:

$$V = (I - \beta Q)^{-1}[\sum_d \texttt{CCP}(\cdot,d) \odot (\boldsymbol{u_d}(\alpha,\boldsymbol{\omega}) + \boldsymbol{\epsilon_d})], \tag{A.2}$$

where I is an $n^{\{S\}} \times n^{\{S\}}$ identity matrix, $\boldsymbol{u_d}$ is an $n^{\{S\}} \times 1$ action-specific period utility absent the random shock, and $\boldsymbol{\epsilon_d}$ is an $n^{\{S\}} \times 1$ action-specific random shock vector with a closed form under the logit assumption: $\boldsymbol{\epsilon_d} = \text{Euler's constant} - log(\texttt{CCP}(\cdot,d))$.

The action-specific value function $V_d(n^{\{S\}} \times 1)$ is:

$$V_d = \boldsymbol{u_d}(\alpha,\boldsymbol{\omega}) + \beta \cdot Q_d \cdot V. \tag{A.3}$$

Then the conditional choice probability is:

$$\texttt{CCP}_{jk} = P(d = k|S = j) = \frac{\exp(V_d(S))}{\sum_{k\in\{0,1,2\}} \exp(V_k(S))}. \tag{A.4}$$

With the calculated CCP matrix, we can use Equation (A.1) to conduct the overall likelihood function $L = \sum_i l(D_i, O_i)$ and estimate $(\alpha, \boldsymbol{\omega}, \boldsymbol{\lambda_2})$ by maximizing L. With the estimated $(\hat{\alpha}, \hat{\boldsymbol{\omega}}, \hat{\boldsymbol{\lambda_2}})$, we can update $\texttt{CCP}, Q_d, Q$ and then redo the above procedure. We will stop the iteration until $\|Q^{(n+1)} - Q^{(n)}\| < \varepsilon_Q$, where n is the n^{th} iteration.