A Dynamic Structural Model for Heterogenous Mobile Data Consumption and Promotion Design*

Lizhen Xu
Scheller College of Business, Georgia Institute of Technology
lizhen.xu@scheller.gatech.edu

Jason A. Duan
McCombs School of Business, The University of Texas at Austin
duanj@mccombs.utexas.edu

Yu (Jeffrey) Hu
Scheller College of Business, Georgia Institute of Technology
jeffrey.hu@scheller.gatech.edu

Yuan Cheng
School of Economics and Management, Tsinghua University, China
chengyn@sem.tsinghua.edu.cn

Yan Zhu
School of Economics and Management, Tsinghua University, China
zhuyan@sem.tsinghua.edu.cn

*We thank the seminar participants at University of Florida, The University of Texas at Dallas, George Washington University, Johns Hopkins University, Nanjing University, Beijing University of Posts and Telecommunications, Workshop on Information Systems and Economics (2014), and China Summer Workshop on Information Management (2015) for their helpful feedback.

A Dynamic Structural Model for Heterogeneous Mobile Data Consumption and Promotion Design

Abstract

This paper studies how individual users who subscribe to limited monthly mobile data plans dynamically consume data on a daily basis. Employing a unique data set on users' daily mobile data usage from a large mobile carrier, we develop a dynamic structural model in which an individual user maximizes her data usage utility by considering the intertemporal trade-off between current and future consumption. The model contains individual-specific utility parameters to fully capture the heterogeneity in consumption behavior across mobile data users. Theoretical properties derived from the model enable us to discover distinct usage patterns for both forward-looking and myopic users. We then empirically test for forward-looking consumption patterns in order to control for possible consumer myopia. For each of the users with forward-looking consumption patterns, we estimate her structural utility parameters such as price sensitivity. We find simulated individual consumption based on our model estimates can correctly capture the dynamic trends observed in the actual individual consumption. We also conduct counterfactual simulations to investigate how month-end promotions and discounts can stimulate individual mobile data consumption and increase carrier profit.

Keywords: mobile data; individual heterogeneity; price sensitivity; forward-looking consumption pattern; myopic consumption pattern; month-end promotion; dynamic structural model

1 Introduction

Consumers nowadays increasingly access digital content via the mobile broadband on their smartphones. Worldwide there are more than 2.3 billion active mobile broadband subscriptions, translating to a mobile broadband penetration ratio of roughly 32% of the entire world population (ITU, 2014). As accessing mobile data becomes an essential part of everyday life for more and more people, understanding how users consume their mobile data is thus a critically important research question to both the academics and the industry.

As major mobile carriers around the globe gradually shift their emphasis from traditional voice and message services to mobile data services, they start to offer carefully-designed limited data plans to consumers (e.g., Segan, 2014). Such limited data plans typically consist of a pre-specified data quota in a billing cycle for a fixed fee and additional costs for extra usage exceeding the quota. As various activities ranging from web browsing to multimedia streaming can gradually use up users' allotted quota, users may need to carefully manage their data usage over time within a billing cycle to avoid excessive surcharges while achieving the highest utility from their quota. Therefore, when determining their daily data usage amount, users face the trade-off between the immediate utility from present consumption and the expected utility gain from saving for the future days. This is a dynamic optimization problem of allocating the consumption of a pre-specified quota over a finite period of time subject to usage need uncertainty, where the solution leads to interesting consumption patterns over time that is worth exploring. Furthermore, such dynamic usage decisions could differ significantly across users, as some are more price sensitive or have greater usage needs than others. Thanks to the era of big data, detailed usage data at individual user level are increasingly available, making it imperative and intriguing to examine the dynamic consumption behavior of individual mobile data users in this ideal context.

Understanding individual users' dynamic behavior in consuming mobile data provides valuable managerial implications for mobile carriers. The users who dynamically optimize their usage will typically reduce their data usage when the remaining quota diminishes. As a result, when it approaches the end of a billing cycle (usually a month), total mobile data usage volume tapers off, leaving excess bandwidth capacity in the carrier's mobile network. It thus opens great potential for innovative design of month-end promotions to boost usage and increase revenue by targeting the users who would actively respond. Naturally, it calls for a well constructed model of individual users' dynamic behavior in consuming mobile data.

Therefore, we aim our study at the dynamics of mobile data consumption, which has been left largely unexamined in the existing literature. First of all, we seek to properly model how mobile data users dynamically optimize their usage every day within a billing cycle given the fixed plan quota and to explore any notable usage pattern over time. Second, we aim to perform individual-specific analysis to fully capture the heterogeneity among users and identify which users are more price sensitive than others. Moreover, we intend to design and evaluate profitable promotion strategies for mobile carriers when consumers' quotas are mostly expended near the end of a billing cycle. Properly addressing these issues fills an important gap meaningful for both researchers and practitioners.

Employing a unique dataset on the daily mobile data usage over a nine-month period for a large number of subscribers from a leading mobile carrier in China, we are able to construct and estimate a dynamic structural model for daily data usage under limited data plans. We model an individual user's daily data usage in a framework of maximizing the sum of present and expected future utility facing day-to-day random utility shocks. Hence, our dynamic structural model captures the intertemporal substitution between current and future consumption. Taking advantage of the richness of our individual-level and multiple-month daily usage data, we fully capture the heterogeneity among users by specifying the model at the individual level so that the model parameters are user specific.

To correctly characterize mobile data users' dynamic usage at the individual level, an important factor to control for, yet often neglected in the literature, is that not all users may be *forward-looking*. A forward-looking user takes into account future utility when optimizing the present usage. In contrast, a *myopic* user focuses on the immediate utility only without considering the future. In dynamic models, such differentiation would be reflected in the values of the *discount factor*, which quantifies the present value of future utility. Unfortunately, it is well known, in both theory and application, that the discount factor is deeply confounded with other utility parameters and cannot be identified by observational data in general. A few recent studies attempt to identify a common discount factor for all individuals with the help of special structures in observational data, which can still be difficult in practice. In the general cases when such data are not available, the common approach is to set a fixed value as the common discount factor for all individuals. In either case, estimating individual-specific discount factors along with all other utility parameters based on observational data only has not been possible in the existing literature, as we discuss in more detail in Section 2.

In light of the above considerations, we develop an innovative approach to control for the fact that some users are forward-looking whereas others are myopic, while estimating other utility parameters individual specifically. We first derive and prove formal analytical results from our dynamic model, which uncover distinct usage patterns for forward-looking and myopic users. We next develop a reduced-form empirical test for forward-looking behavior of individual users by matching their observed usage patterns and the theoretical results. We then apply our dynamic model to the users with forward-looking patterns consistent with the proposed model. As we show, forward-looking users' daily usage is positively correlated with the remaining data quota, whereas myopic users' usage is uncorrelated with the remaining quota. Our test finds that about 40% of all users in our data exhibit forward-looking usage patterns, whereas the rest demonstrate myopic patterns. Our analytical results establish theoretical understanding of temporal patterns for the class of dynamic problem as we study, which has been absent in the literature. Our empirical findings provide evidence for individual forward-looking behavior and also suggest the necessity of controlling for possible consumer myopia in studying dynamic behavior at the individual level.

In estimating the dynamic model, we adopt a two-stage estimation approach similar to Bajari et al. (2007). The estimation method is computationally efficient and can be easily parallelized, which facilitates speedy estimation for hundreds of users in our sample and opens feasibility for a larger-scale implementation by mobile carriers. Our dynamic model is able to identify a user's price sensitivity from her intertemporal substitution of data consumption even if she has never exceeded her plan quota, whereas a static model cannot. By examining the relationship between estimated structural parameters and observed consumer demographics, we find older users generally have lower price sensitivity and less daily variation in usage. To validate the estimation results, we develop simulation method by solving the dynamic programming problem given the estimated parameter values. Simulated usage well capture the dynamic trends observed in the actual data, which convincingly validates our model development.

We further demonstrate how our model and results can help mobile carriers increase their revenue by properly devising promotions near month end. We conduct counterfactual simulations to investigate the design of two types of month-end promotions: discounting the price rate for extra usage beyond the plan quota, and offering unlimited-use passes with a fixed fee in the last several days of a month. As these promotions target forward-looking users who tend to respond actively, an accurate evaluation of their effectiveness is predicated on a deep grasp of the dynamic decision process. This is because users start to rationally adjust their usage right upon receiving the promotion information and strategically accept or decline the promotional offer by comparing the expected future utility under different options. We find that an optimal discount of 40% to 50% in the overage charge can increase the carrier's revenue by about 27%. Selling unlimited-use passes at a proper price level and at the right time (e.g., a ¥10 pass for the last 4 days of a month) can increase the expected revenue by as much as 20%.

In addition to the rich implications in the substantive domain, our paper also makes several contributions on the methodological front. First, we develop a dynamic structural model for daily usage of mobile data. The model is shown to well capture the dynamic trends observed in the actual usage data. Second, we theoretically establish the dynamic patterns for the consumption of a pre-specified quota over a finite period of time with random utility shocks. Theoretically uncovering the dynamic patterns and formally proving these results are not obvious given the complex nature of dynamic problems, which has not been documented in the literature. Third, we successfully implement individual-specific estimation for hundreds of users, which is particularly meaningful when significant behavioral variation exists and individual heterogeneity is important to address. Our estimation approach is fully parallel and computationally efficient, which enables large-scale implementation and precise targeting. Fourth, we propose an innovative approach of first deriving analytical properties of a dynamic structural model and then constructing a test of forward-looking behavior based on matching the observed usage patterns and the analytical properties. It allows us to control for possible myopic consumption behavior and avoid the strong assumption that all consumers share the same discount factor. Our proposed test shares the same notion with previous studies that test for collective forward-looking behavior in purchasing durable goods based on aggregate data (e.g., Chevalier and Goolsbee, 2009). In contrast, our paper is the first to conduct a formal test for forward-looking behavior at the individual consumer level and to provide evidence for both forward-looking and myopic behavior individually. Fifth, the analytical results derived from our dynamic model establish the monotone properties of the value function. In order to ensure the preservation of such monotonicity in numerically solving the value function, we apply the monotone piecewise cubic Hermite interpolation (Fritsch and Carlson, 1980) to approximate the value function instead of other commonly used approximating methods (Judd, 1998). It serves as a necessary and useful alternative especially when monotonicity needs to be preserved.

The rest of the paper is organized as follows. In the next section, we briefly review prior literature on mobile services and dynamic structural models and discuss how our paper differs from them. We then construct our model and derive formal analytical results from it in Section 3. We provide details of our data and conduct reduced-form analysis to test for forward-looking behavior in Section 4. In Section 5, we discuss the estimation and simulation methods for our model, followed by the results in Sections 6. We conduct counterfactual simulations of various promotion strategies to provide policy implications in Section 7. Section 8 concludes the paper.

2 Literature Review

The substantive context of our study is related to the existing literature on traditional wireless phone services and the burgeoning literature on consumer usage of mobile data and smartphone apps. Our modeling and estimation approach relates to the growing literature on dynamic structural models applied to various business and economics problems. We briefly review the relevant literature in both domains.

Previous studies have examined different aspects of traditional wireless phone services including voice call service and short message service (SMS). For example, Iyengar et al. (2007) study consumer learning over months in choosing their voice call plans. Kim et al. (2010) develop a static structural model to study the substitution between voice and SMS demands. Yao et al. (2012) estimate consumers' weekly discount rate in using voice call service. Recently, increasing research interest has been attracted to consumer behavior related to mobile data and smartphone apps. For example, Ghose and Han (2011) investigate the correlation between uploading and downloading for users using the mobile Internet. Niculescu and Whang (2012) examine the co-diffusion process of the adoption of wireless voice and mobile data services in Japan. Xu et al. (2014) explore the complementary effect between the introduction of a mobile app and website visits for news media. In contrast, our study differs from these papers in the fundamentals, such as research context (i.e., mobile data consumption), research methodology (i.e., dynamic structural model), and research focus (i.e., dynamic patterns of daily usage and month-end promotion design).

Dynamic structural models and their estimation have long been applied to a large variety of economics and business problems involving forward-looking agents, for example, from bus engine replacement (Rust, 1987) to sales-force compensation (e.g., Chung et al., 2014; Misra and Nair, 2011), and from online grocery service (e.g., Goettler and Clay, 2011) to social media contribution (e.g., Huang et al., 2015; Tang et al., 2012). Enriching this increasingly vibrant area, our study contributes in several aspects. First, we study the daily consumption of mobile data, a new phenomenon with distinctive features in the underlying dynamic problem. Second, we formally establish the temporal patterns associated with the class of dynamic problem as we study based on rigorous theoretical analysis. Such theoretical results are absent in the existing literature of dynamic models and thus valuable for future studies in relevant contexts. Third, our study is among the few performing individual-specific analysis. Panel data with sufficient observations longitudinally are not often available, and the computational burden for dynamic models is typically high. Therefore, most existing studies of dynamic models usually specify and estimate the key structural parameters either uniformly across all individuals (e.g., Goettler and Clay, 2011; Tang et al., 2012) or in segments (e.g., Arcidiacono and Miller, 2011; Chung et al., 2014), foregoing individual heterogeneity to a certain extent. In our context of individual usage of mobile data, significant heterogeneity exists among users, and a small number of segments cannot capture the rich variation in multiple heterogeneous parameters. Therefore, we take advantage of the fine granularity of our data and develop computationally efficient estimation based on Hotz and Miller (1993) and Bajari et al. (2007) so as to specify and estimate our dynamic model individual specifically. In this respect, our estimation strategy is similar to Misra and Nair (2011), who perform individually separate estimation for 87 sales agents. Our successful implementation of speedy estimation for hundreds of users adds to the examples meaningful for future studies when significant individual heterogeneity needs to be taken full account of and/or large-scale parallel implementation is important.

Another unique aspect of our study is that we take account of the fact that some consumers are forward-looking whereas others are myopic. Previous studies of non-structural analysis using aggregate data have found empirical evidence that consumers *collectively* are forward-looking in purchasing durable goods (e.g., Busse et al., 2013; Chevalier and Goolsbee, 2009). On the other hand, common wisdom on consumer behavior believes some individuals may be myopic or insufficiently comprehend future utility (Chevalier and Goolsbee, 2009). It is possible that some consumers may not have a good command of their own states when planning their consumption (e.g., Frederick et al., 2002) or display impulsive consumption behavior (e.g., Hoch and Loewenstein, 1991). Therefore, theoretical studies often explicitly consider some consumers to be forward-looking whereas others myopic (e.g., Gabaix and Laibson, 2006; Pashardes, 1986).

Existing literature on dynamic structural analysis, however, largely neglects this issue. This is mostly because the discount factor, being the key structural parameter characterizing forward-looking behavior, is notoriously difficult to identify. Theoretically, it is shown that the discount factor in dynamic structural models generally cannot be separately identified from other structural parameters by observational data (Magnac and Thesmar, 2002; Manski, 1993). Theoretical results on strict identification conditions for discount factors have yet to be established. Magnac and Thesmar (2002, Proposition 4) theoretically derive an “exclusion restriction” condition in the discrete choice context, which might not be easily verifiable in practice (Yao et al., 2012, p.824). In applied studies with observational data, the common approach is to set a fixed value as the common discount factor for all individuals (e.g., Misra and Nair, 2011; Rust, 1987). Some studies attempt to identify the discount factor by imposing restrictions on the model structure. For example, Goettler and Clay (2011) assume homogeneity across all consumers with different pricing structures, which enables the identification of the common price sensitivity through cross-sectional variation and thus the common discount factor through intertemporal variation. A few recent studies manage to identify a common discount factor for all individuals utilizing special structures in their data. For example, Yao et al. (2012) use a data set spanning before and after a natural experiment where there is a shift from linear pricing to non-linear pricing; Chung et al. (2014) exploits particular sales-force compensation structures to find variables that do not affect current utility but affect future utility so as to form exclusion restrictions. Nevertheless, even under these situations, it is still possible that a common discount factor cannot be reliably estimated in practice. For example, Chung et al. (2014, p. 178) find in practice that the objective function is “relatively flat with respect to changes in the discount factor,” and hence they have to rely on grid search to find the most appropriate discount factor instead of the usual estimation methods. Therefore, it is not surprising that successful estimation of individual-specific discount factors along with all utility parameters using observational data only has not been possible according to the existing literature.[1]

In our study, when estimating individual-specific utility parameters such as price sensitivity and usage need, we also control for the fact that some users may be myopic. The approach we take shares the same spirit of previous non-structural studies that develop empirical tests for consumers' forward-looking behavior collectively in purchasing durable goods using aggregate data (e.g., Chevalier and Goolsbee, 2009). To the best of our knowledge, we are the first to conduct formal tests for consumer forward-looking behavior at the individual level and to provide empirical evidence for both forward-looking and myopic behavior of individual consumers.

[1]Some experimental studies, by carefully controlling the utility function, manage to estimate heterogeneous discount factors following a common distribution (e.g., Dube et al., 2014).

3 Model

In this section, we first lay out the model details; we then explore the analytical properties implied by the model. To take full account of the heterogeneity among users, we exploit the granularity of our individual-level daily usage data over multiple months and specify our model such that the model parameters are user specific and individually estimated.

3.1 Model Development

Consider an individual user i with monthly mobile data plan j, which adopts a typical nonlinear pricing scheme of two-part tariff. With a fixed monthly fee F_j (in the unit of Chinese yuan, CNY or ¥, as we use throughout this paper), the plan includes a monthly data quota Q_j (in the unit of megabyte, MB, as we use throughout this paper). The user can consume mobile data up to Q_j without incurring any extra cost. If the total data usage within a month exceeds Q_j, the user needs to pay a unit price p_j per MB of additional usage. Denote user i's daily mobile data usage as a_{it} ($a_{it} \geq 0$) for any given day t within a month, $t = 1, ..., T$, where T is the total number of days in the month.[2] At the beginning of day t, the remaining data quota user i has is thus $q_{it} = \max\left\{Q_j - \sum_{\tau=1}^{t-1} a_{i\tau}, 0\right\}$.

For any given day t, user i's per-period utility from consuming a_{it} amount of mobile data for that particular day is modeled as follows.

$$u\left(a_{it}\right) = \left(\mu_i + \xi_{it}\right) a_{it} - \frac{1}{2} a_{it}^2 - \eta_i p_j \max\left\{a_{it} - q_{it}, 0\right\} \tag{1}$$

The first two terms of the above utility function (i.e., $\left(\mu_i + \xi_{it}\right) a_{it} - \frac{1}{2} a_{it}^2$) captures the direct utility from mobile data consumption. We adopt the typical quadratic functional form, which is commonly used for utility functions in the literature related to mobile communication (e.g., Iyengar et al., 2007; Kim et al., 2010). The quadratic functional form properly models the decreasing marginal utility as the consumption volume increases, where the linear term can be viewed as reflecting the utility gain from the data consumption and the quadratic term as the disutility associated with the time and effort spent as well as the opportunity cost. In order to account for the variation in an individual

[2] We explicitly account for different numbers of days in different months (e.g., 28 days for February and 31 days for March).

user's mobile data usage need from day to day, in addition to the mean level μ_i, we incorporate a daily random shock, ξ_{it}, into the utility function (1). The private shock ξ_{it} is observed only to each individual user but not to the researchers. We assume ξ_{it} follows a normal distribution with individual-specific variance, $N\left(0, \sigma_i^2\right)$ and is independently and identically distributed across time. As we discuss in Section 4, a normal distribution of the daily random utility shocks closely approximates the distribution of daily usage observed in the data. Notice that the daily mobile data usage a_{it} is bounded below at zero, and we allow both μ_i and ξ_{it} to take either positive or negative values. Therefore, a negative value of $\mu_i + \xi_{it}$ indicates little need for consuming mobile data on that particular day and will generally lead to a zero usage volume, whereas a positive and large value of $\mu_i + \xi_{it}$ indicates the opposite and can lead to a high usage volume on that particular day.

The last term in the per-period utility function (1) (i.e., $\eta_i p_j \max\left\{a_{it} - q_{it}, 0\right\}$) accounts for the cost of extra usage exceeding the plan quota. Here, η_i captures the individual-specific price sensitivity, measuring the utility cost per unit of monetary spending. Notice that the coefficients in the utility function (1) is only identifiable up to a constant scale. Therefore, for identification purpose, we normalize the coefficient before the second term (i.e., a_{it}^2) to a fixed constant $\frac{1}{2}$. It is also worth noting that the focal mobile carrier deducts overage charges daily from users' account balances, so we include the associated utility cost in the per-period utility function rather than till the end of a month.

When determining the amount of mobile data to consume each day, rational users who are *forward-looking* not only consider the utility for the present day but also take into account the expected utility for the future days in the rest of the month. To properly model such a decision process, we assume that at the beginning of each day t, each individual user observes her private daily shock ξ_{it}, and *then* decide her data usage for that day, a_{it}. A user optimizes her usage in any day by weighing her present utility (given the observed shock for that day) against her expected future utility (by taking expectation over all the future utility shocks). The optimization problem for each user is thus to maximize the sum of the present utility and the expected future utility, which can be formulated as

$$u\left(a_{it};q_{it},\xi_{it}\right)+E_{\{\xi_{i\tau}\}}\left[\sum_{\tau=t+1}^{T}\beta^{\tau-t}u\left(a_{i\tau};q_{i\tau},\xi_{i\tau}\right)\right], \tag{2}$$

where $\beta\in(0,1)$ is the discount factor for future utility as in the usual dynamic models. Notice that the expectation in (2) is taken with respect to all the future shocks $\{\xi_{i\tau}\}$ for $\tau=t+1,...,T$, which in turn determine the optimal future consumption path $\{a_{i\tau}\}$ and the associated utilities $\{u\left(a_{i\tau}\right)\}$.

The problem described above can be formulated as a dynamic programming problem with T periods (i.e., days). At the beginning of each period t, a user i decides her optimal usage based on three state variables: the remaining unused data quota from her monthly plan, q_{it}, the number of days left in the month, d_t, and the private utility shock for the current period, ξ_{it}. Among them, q_{it} and d_t are evolving states transitioning from period to period deterministically given the previous usage, that is,

$$q_{i,t+1}=\max\left\{q_{it}-a_{it},0\right\}\quad\text{and}\quad d_{t+1}=d_t-1. \tag{3}$$

Note that here we assume each user knows her remaining data quota each day. This is a reasonable assumption considering the wide availability of built-in tools in most smartphones and various third-party apps that can easily track the cumulative mobile data usage with a single finger tap.

The value function, $V\left(q_{it},d_t,\xi_{it}\right)$, can be defined recursively backwards by the Bellman equation,

$$V\left(q_{it},d_t,\xi_{it}\right)\;=\;\max_{a_{it}\geq 0}u\left(a_{it};q_{it},\xi_{it}\right)+\beta E_{\xi_{i,t+1}}V\left(q_{i,t+1},d_{t+1},\xi_{i,t+1}\right). \tag{4}$$

Starting backwards from the last period (i.e., the last day in a month so that $t=T$ and $d_t=1$), where $V\left(q_{it},d_t=1,\xi_{it}\right)=\max_{a_{it}\geq 0}u\left(a_{it};q_{it},\xi_{it}\right)$, (4) determines the maximized sum of the present utility and the expected future utility as a function of the state variables in any period. The solution to (4) yields the policy function, $a_i^*\left(q_{it,}d_{it},\xi_{it}\right)$, which determines the optimal mobile data usage in any day as a function of the state variables q_{it}, d_{it} and ξ_{it}. Because the policy function directly links to the daily usage observed in our data, it hence plays a central role in our estimation strategies.

The structural model parameters to be estimated for each individual user are therefore μ_i, σ_i^2, and η_i. Note that we assume users have full information about their structural parameter values and do not model users' initial learning processes, because for this study, we examine the general population of mobile data users rather than the newly acquired users. As we discuss in detail in Section 4, the vast majority of the users in our data set are experienced users who have been with the carrier for a long time. Therefore, we focus on experienced users and abstract away from the learning problem in the scope of this study. It is also worth mentioning that our study focuses on the consumption of mobile data through the mobile network of the focal carrier. When users have access to alternative data connections (e.g., WiFi), it diminishes their need of using mobile data, which will be reflected in the changing utility shocks in our model.

3.2 Analytical Properties

We next derive important analytical properties from our proposed model. We are particularly interested in exploring any notable patterns in daily usage that will differentiate those users who are forward-looking from those who are not. Hence, the formal results established in this section serve two purposes. First, based on rigorous analysis, they offer theoretical understanding of the consumption patterns for different mobile data users. Second, the formal results provide a theoretical foundation for our initial data analysis in Section 4.2, where we develop a reduced-form test to provide evidence of forward-looking behavior and identify users applicable to our dynamic model.

We start with a benchmark case in which users are *myopic*, who determine their daily mobile data usage only based on the present utility (as in (1)), without considering the future utility (as in (2)). Myopic behavior can be viewed as degeneration of the dynamic model with $\beta = 0$. In such a case, the dynamic programming problem reduces to a static optimization problem, which can be fully solved in a closed form, leading to the following result.

Proposition 1. *For myopic users, given any utility shock ξ_{it} in any given day before the day when the data plan quota is fully expended (i.e., $q_{it} > a_i^*(q_{it}, d_t, \xi_{it})$), their daily mobile data usage $a_i^*(q_{it}, d_t, \xi_{it})$ is independent of the remaining data plan quota q_{it}.*

Proof. The proof is detailed in the Appendix. □

For myopic users, in maximizing the per-period utility (1) in a given day, as long as there is enough plan quota remaining (i.e., q_{it} is large enough), they simply maximize the direct utility from data consumption (i.e., the first two terms in (1) as the last term equals zero when $q_{it} > a_{it}^*$). In other words, they just use as much as they need, which maximizes the direct consumption utility given the realized utility shock for that day, and does not depend on q_{it} at all.

Next, we turn to our dynamic model and consider users who are *forward-looking.* They optimize their daily usage to maximize the sum of the present utility and the expected future utility as in (2). As a result, their optimal usage typically changes dynamically over time, depending on the state in each period. In particular, we are interested in whether and how the policy function $a_i^*(q_{it}, d_t, \xi_{it})$ depends on the remaining plan quota q_{it}, so as to contrast with the result for myopic users.

In order to prove the result formally and generally, we first prove two lemmas with regard to some important properties of the expected value function $\bar{V}(q_{it}, d_t)$, which is the value function $V(q_{it}, d_t, \xi_{it})$ taken expectation over the random utility shock ξ_{it},

$$\bar{V}(q_{it}, d_t) = E_{\xi_{it}} V(q_{it}, d_t, \xi_{it}) . \tag{5}$$

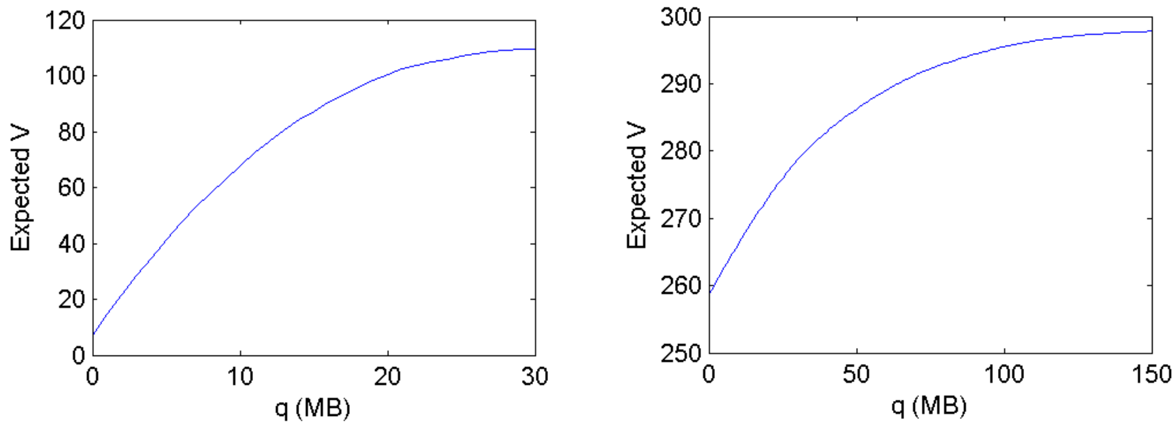
The two lemmas and their proofs are presented in the Appendix due to their technical nature. Briefly, Lemma 1 shows that the expected value function in the last period (i.e., when the number of remaining days $d_t = 1$) is *strictly concave* in the remaining quota q_{it}; Lemma 2 further shows that if the next-period expected value function $\bar{V}(\cdot, d-1)$ is strictly concave, then the current-period expected value function $\bar{V}(\cdot, d)$ is also strictly concave. By backward induction utilizing both lemmas, we can conclude that $\bar{V}(q_{it}, d_t)$ is strictly concave in the remaining quota q_{it} in any period. To illustrate this result, in Figure 1, we present two examples of the expected value function numerically calculated under different sets of parameter values. The algorithm we developed to numerically solve the value function will be discussed in detail in Section 5.2.

The strict concavity of the expected value function leads to the following properties of the policy function for forward-looking users.

Proposition 2. *For forward-looking users, given any utility shock ξ_{it} in any given day (except the last day of a month) before the day when the data plan quota is fully expended (i.e., $q_{it} > a_i^*(q_{it}, d_t, \xi_{it})$), the optimal daily usage $a_i^*(q_{it}, d_t, \xi_{it})$ is strictly increasing in q_{it} if $0 < a_i^* < q_{it}$.*

Proof. The proof is detailed in the Appendix. □

To understand the reasoning leading to Proposition 2, recall the optimization problem facing



(a) $\mu_i = 8, \sigma_i = 3, \eta_i = 8; Q_j = 30, p_j = 1; d_t = 3$ (b) $\mu_i = 0.5, \sigma_i = 10, \eta_i = 1; Q_j = 150, p_j = 1; d_t = 25$

Figure 1: Illustration of the Expected Value Function $\bar{V}\left(q_{it}, d_t\right)$

forward-looking users as in (4), which can be written more simply as

$$a_i^*\left(q_{it}, d_t, \xi_{it}\right) = \arg\max_{a_{it}\geq 0} u\left(a_{it}\right) + \beta\bar{V}\left(\max\left\{q_{it} - a_{it}, 0\right\}, d_{i,t+1}\right). \tag{6}$$

From (6), it is clear that users face the trade-off between their immediate utility from current consumption, $u\left(a_{it}\right)$, and their future utility given the quota remaining after the current-period consumption, $\bar{V}\left(\max\left\{q_{it} - a_{it}, 0\right\}, d_{i,t+1}\right)$. Users will increase their usage amount if the marginal benefit from current consumption exceeds the marginal benefit from saving for the future, and decrease their usage amount if the opposite. If the remaining quota at the beginning of the current period q_{it} increases, with the same amount of data usage a_{it}, the quota left for the next period $\left(\max\left\{q_{it} - a_{it}, 0\right\}\right)$ also increases in general. Because $\bar{V}\left(\cdot, d_{i,t+1}\right)$ is strictly concave as we have shown, the marginal utility gain from saving for the future will then decrease, which gives users the incentive to increase their consumption in the current period. As a result, the optimal usage a_{it}^* strictly increases with q_{it} in general. Exceptions arise after the plan quota is used up or in the last day of the month. In these cases, there is simply nothing to save for the future, so the dynamic programming problem reduces to the same static optimization problem facing myopic users as is discussed in Proposition 1.

The distinct usage patterns that distinguish forward-looking users from myopic ones are testable given our observed data. We can use each individual user's daily usage before the plan quota is used up or the last day of a month to test for any positive relationship with the remaining quota so as to discover evidence of forward-looking consumption behavior, as we discuss in detail in Section 4.2. It is worth clarifying that the dependence of the policy function on the remaining quota shown in the two propositions hold *given* the random utility shock ξ_{it} and a particular day d_t. The said relationship with the remaining quota may not hold across different days. Moreover, even when d_t and q_{it} are given, the optimal usage a_{it}^* is still a random variable depending on the realization of the unobservable utility shock ξ_{it}. Therefore, what we should test for is the average trend between the daily usage and the remaining quota in terms of expectation rather than each realization *per se*.

4 Data

4.1 Data Overview

We obtain the data for this study from one of the leading telecommunication companies that provide mobile services in China.[3] This carrier has nationwide mobile network coverage and has hundreds of millions of subscribers. We collaborate with the company's subsidiary in a certain province and obtain the mobile data usage records of its subscribers in the capital city of that province. The data furnished to us span a nine-month period from January to September, 2013, and include all subscribers in the focal city who had been using the company's mobile data service with monthly data plans throughout the nine months.

A mobile data user with the focal carrier usually has a main service package, which typically consists of a combination of voice call minutes, text messages, and other value-added mobile services (e.g., free music download, real-time stock market information), marketed under different brand series with different cost structures. On top of the main service package, a user may also have a mobile data plan, which adopts a two-part tariff price structure. After paying an upfront monthly fee for a limited data quota (e.g., ¥50.00 for 500 MB), a user can consume mobile data up to the plan quota without any additional cost within a month.[4] If the total data usage within a month

[3] We are unable to reveal the identity of the company because of a nondisclosure agreement.

[4] Some main service packages come with an allowance of free data as well. In these cases, a user can consume up to the combined total of the quotas included in both the main package and the data plan before incurring any extra cost. Note that while we include the data quota contained in the main service package in our analysis, we take users' main service package as exogenously given and do not model users' choice of main service package. This is simply to avoid the labyrinthine main service offerings which total up to more than 100 different packages.

exceeds the plan quota, overage charges will be levied as a fixed unit price rate per megabyte of additional usage. The billing cycle is based on calendar months and hence the same for all users. At the beginning of each month, unused data quota from the previous month is automatically cleared, and the accumulated data usage is reset to zero. For our study period, there are seven different data plans available, with plan quota ranging from 30 MB to 5,120 MB and monthly fee from ¥5.00 to ¥200.00. The extra rate, however, is the same across all these data plans,[5] which simplifies our estimation approach as we discuss in Section 5.

For this study, we select from the original data a random sample of 1,000 users who have not changed their data plan within the nine-month period. We focus on these users because of practical consideration. Note first that our model can be easily extended to incorporate users who changed their data plan. We can simply allow the mean and variance of the daily usage to differ before and after the plan change, because it reflects their rational choices based on varying consumption needs. Nevertheless, the increased number of model parameters results in fewer data points applicable to the identification of each parameter. Given that we have only nine months of data, for the purpose of better identification and more efficient estimation, we choose to focus on the users who have not changed their data plan within the study period. We also exclude from our sample those newly joined users. Because we focus on the general mobile data users instead of new users for this study, we do not model users' initial learning process. For this reason, we only include in our sample those experienced users who have been with the focal carrier for at least one year, which account for over 85% of all users in our original data.

For each individual user in our sample, we include the following information for the analysis: (i) the monthly plan information, including the chosen main service package and data plan each month, the plan quota, and the monthly fees and overage price rates; (ii) the daily mobile data usage for each day throughout the entire study period, resulting in 273 data points for each user; (iii) the user profile information, including age, gender, and customer history (i.e., how long, measured in months, the user has been with this carrier at the beginning of the study period). Table 1 provides the summary statistics of the user profile information.

Figure 2 shows the histograms of the daily mobile data usage for four representative users. As

[5]The rate is ¥1.00 per extra megabyte regularly, whereas some users' main service packages offer discounts in the extra rate (e.g., ¥0.29 / MB). Therefore, depending on their main service packages, different users may have different effective extra rates. In any case, for a given user, the extra rate stays the same across different data plans.

Table 1: Summary Statistics of the User Profile Information

	Mean	St. Dev.	Max	Min
Age	32.8	9.87	67	16
Male	0.697	0.460	1	0
Customer History	58.7	35.4	192	12

we can see, there exists variance in a user's mobile data consumption from day to day, and the distribution of the daily usage can be well approximated by a normal distribution truncated at zero. Notice that with all negative values replaced by zero, there is a point mass at zero accounting for the mass of the negative values of a normal distribution. The data patterns justify our model setting with regard to the distribution of the random utility shock ξ_{it}. It is also worth noticing that while some users incur positive usage almost every day, there are others with zero usage for many days. When reflected in our model, the latter are likely to be those users with negative μ_i's so that they have no need for mobile data usage often (i.e., when $\mu_i + \xi_{it}$ is negative) and incur positive usage sometimes (i.e., when ξ_{it} is high so $\mu_i + \xi_{it}$ becomes positive).

4.2 Model-Free Analysis

In this section, we develop a reduced-form empirical test based on the theoretical results in Section 3.2 and apply it to our data sample. The purpose of this test is two-fold. First, we seek to provide empirical evidence prior to fitting our structural model that a certain proportion of users are indeed forward-looking and dynamically optimize their usage over time. Second, we want to apply our dynamic model only to those users who exhibit dynamic usage patterns consistent with the model.

According to the discussion in Section 3.2, we examine each individual user's daily mobile data usage over the nine months for the days before the plan quota is used up (excluding the last day in each month). For users who are forward-looking and dynamically optimize their usage, their daily usage given the number of remaining days, which is a random variable depending on the unobservable demand shock, should demonstrate positive correlation with the remaining quota at the beginning of each day. In contrast, for myopic users, their daily usage should be uncorrelated with the remaining quota. To test for such usage patterns, we apply the following reduced-form

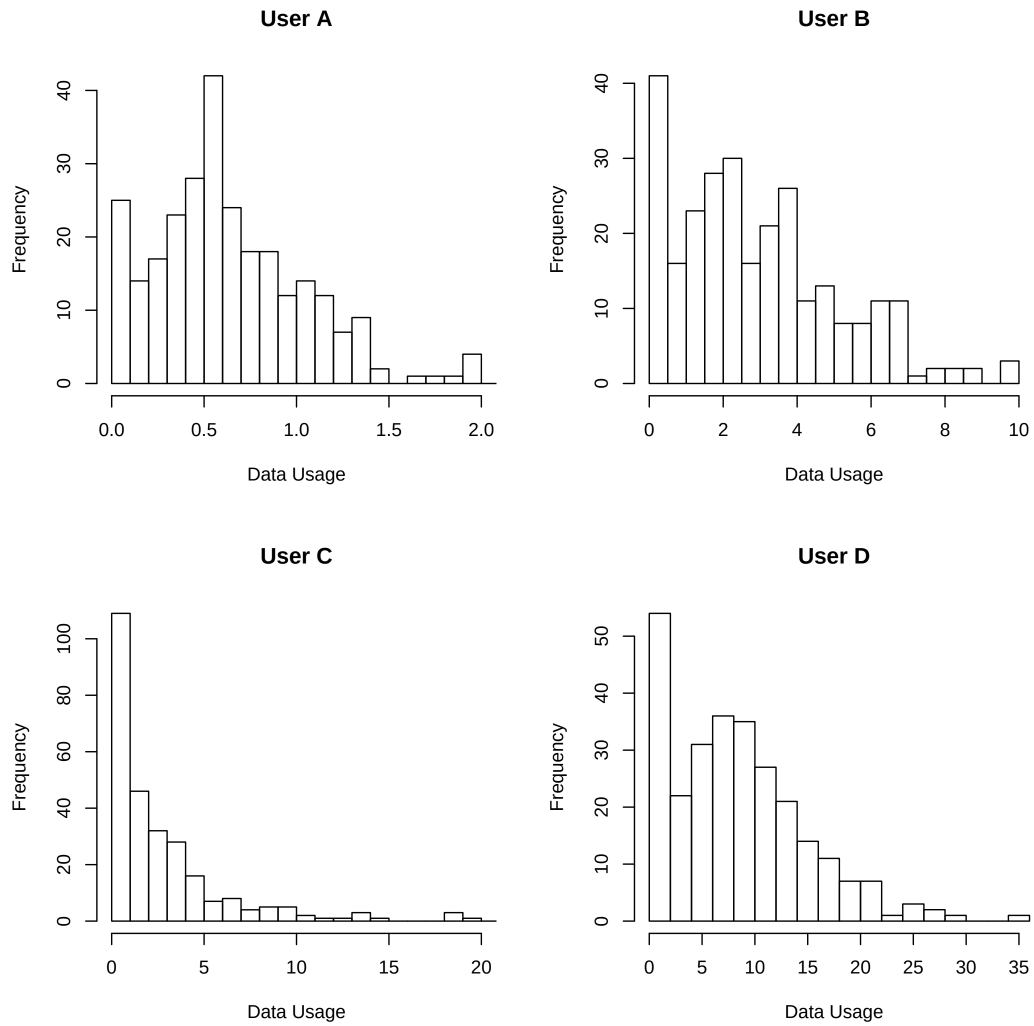


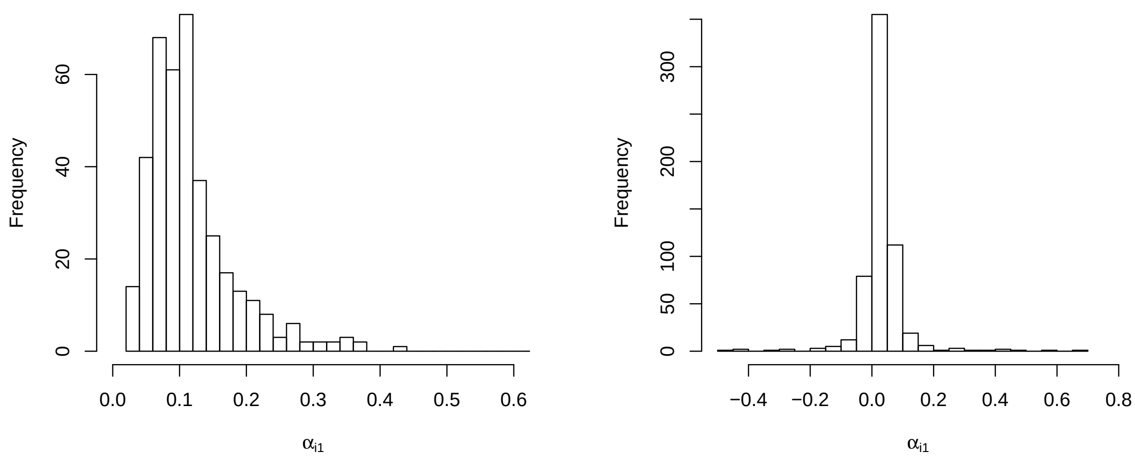
Figure 2: Histogram of Daily Data Usage

empirical model

$$
\begin{aligned}
a_{it} &= \begin{cases} \tilde{a}_{it} & \text{if } \tilde{a}_{it} > 0 \\ 0 & \text{otherwise} \end{cases} \\
\tilde{a}_{it} &= \alpha_{i0m} + \alpha_{i1} q_{it} + \alpha_{i2} d_t + \varepsilon_{it},\ \varepsilon_{it} \sim N\left(0, \omega_i^2\right).
\end{aligned} \tag{7}
$$

This is a typical Type I Tobit model, where the dependent variable a_{it}, which is the observed usage for user i on day t of a month, reflects a latent variable $\tilde{a}_{it}$ such that a_{it} equals $\tilde{a}_{it}$ if it is positive and equals zero otherwise. The latent variable $\tilde{a}_{it}$ has a regression structure regarding the remaining quota q_{it}, controlling for the number of remaining days d_t. The error term ε_{it} follows a normal distribution, which is consistent with the actual distribution of daily usage observed in our data as is discussed in Section 4.1 (Figure 2). We allow the coefficients (α_i's) and the variance of the error term (ω_i^2) to be individually different for different users. We also account for the monthly fixed effects by allowing the intercept α_{i0m} to differ across different months. The key parameter of interest here is α_{i1}, which is expected to be significantly greater than zero for forward-looking users, and insignificantly different from zero for myopic users.

We use maximum likelihood estimation to estimate the model and obtain the p-values for coefficient estimates. We apply the model for each individual user within our data sample one by one, and hence need to run 1,000 separate estimations. Among all 1,000 users in our sample, we found 392 users with a positive estimate of α_{i1} that is significant at 1% level. Among the rest 608 users, 597 users have α_{i1} estimates insignificantly different from zero. There are only 11 users (i.e., 1.1%) with significantly negative α_{i1}, which can be considered as singular outliers. Figure 3 depicts the histogram of the estimated α_{i1}'s.



(a) Users with Forward-Looking Usage Patterns

(b) Users with Myopic Usage Patterns

Figure 3: Histogram of the Test Results (Estimated α_{i1})

To ensure the reliability of the test results, we perform further robustness checks. Notice that the model parameters in (7) are individually specific, meaning that the model already captures various types of individual heterogeneity. To further take account of possible time-varying effects, such as weekend or national holidays, we add additional dummies into the regression specification of $\tilde{a}_{it}$ in (7) as control variables. We found that both weekends and national holidays do not have significant effects on daily usage. Furthermore, after controlling for the weekend and holiday effects, the test results (in terms of the significance of α_{i1}) remain unchanged for 98% of all users. For the very few whose test results change, the cause is merely a slight change in the p-value of the α_{i1} estimate around the significance level, whereas the sign of α_{i1} remains the same and the value is hardly changed.

The reduced-form empirical test in this section provides formal evidence that there is a considerable proportion of users (about 40%) who are forward-looking and dynamically plan their mobile data usage in a way consistent with our model results. Meanwhile, the findings also suggest the necessity of controlling for the possibility of some consumers being myopic when studying dynamic behavior at the individual level.

5 Estimation Strategies

5.1 Dynamic Model Estimation

To estimate the dynamic model, we adopt the two-stage estimation strategy proposed by Bajari, Benkard, and Levin (2007, BBL henceforth). We choose this estimation approach for its advantage in computational efficiency. The general idea of the two-stage estimation strategy is as follows. In the first stage, we empirically estimate the policy function by fitting a distribution of the daily usage from the observed data. In the second stage, we estimate the structural parameters by ensuring that the policy function estimated in the first stage is indeed optimal. As shown by BBL, such a two-stage estimation yields consistent estimates of the structural model parameters.

For each individual user, the structural parameters to be estimated are $\{\mu_i, \sigma_i, \eta_i\}$. As discussed in Section 2, the discount factor cannot be identified by general observational data like ours. Therefore, we follow the common approach in the literature of dynamic modeling to set $\beta = 0.9$ for forward-looking users.[6] To simplify notations, we will suppress the subscripts i and/or t henceforth in this section whenever not causing confusion.

[6] We also tried alternative values of β (e.g., 0.95, 0.99) and found that the estimation results are insensitive to the value of β.

5.1.1 First-Stage Estimation

In the first stage, we empirically estimate the policy function $a^*(q_t, d_t, \xi_t)$ actually adopted by users in the observed data. Recall that without observing the private utility shock ξ_t, given the two observed states (q_t and d_t) only, the daily usage $a^*(q_t, d_t, \xi_t)$ is observed as a random variable. Therefore, to estimate the policy function, we first empirically estimate the conditional distribution of the daily usage a_t given the two observed state variables, defined by its cumulative distribution function (cdf) $F(a_t|q_t, d_t)$. We then back out the policy function $a^*(q_t, d_t, \xi_t)$ based on the estimated $F(a_t|q_t, d_t)$ and the distribution of the private shock ξ_t.

Following BBL, we specify a flexible functional form to approximate the conditional distribution $F(a_t|q_t, d_t)$, as follows.

$$\begin{aligned} a_t &= \begin{cases} \breve{a}_t & \text{if } \breve{a}_t > 0 \\ 0 & \text{otherwise} \end{cases} \\ \breve{a}_t &\sim N\left(\theta(q_t, d_t), \tau^2(q_t, d_t)\right) \end{aligned} \tag{8}$$

We let daily usage a_t be the censored observation of a latent variable $\breve{a}_t$ so that a_t equals $\breve{a}_t$ if it is positive and equals zero otherwise. The latent variable $\breve{a}_t$ follow a normal distribution whose mean and variance are flexible functions of (q_t, d_t). We specify $\theta(q_t, d_t)$ and $\log \tau^2(q_t, d_t)$ in flexible parametric forms as polynomial functions of q_t and d_t. Specifically, we let $\theta(q_t, d_t) = [1, q_t, q_t^2, d_t, d_t^2, q_t d_t] \cdot \mathbf{\Theta}$ and $\tau^2(q_t, d_t) = \exp([1, q_t, d_t] \cdot \mathbf{\Gamma})$, where $\mathbf{\Theta}$ and $\mathbf{\Gamma}$ are a series of coefficients to be estimated. We use the maximum likelihood estimation to estimate the coefficients $\mathbf{\Theta}$ and $\mathbf{\Gamma}$ based on the observed data $\{a_t, q_t, d_t\}$ for all days throughout the entire observation period (excluding the days with no quota left and the last day of a month) for each individual user.

After obtaining the estimated coefficients, we determine $F(a_t|q_t, d_t)$ according to (8) and then derive the policy function given $F(a_t|q_t, d_t)$. We utilize the fact that $a^*(q_t, d_t, \xi_t)$ is increasing in ξ_t, so $F(a^*(q_t, d_t, \xi_t)|q_t, d_t) = G(\xi_t)$,[7] where $G(\cdot)$ is the cdf of ξ_t. As a result, the policy function can be written as

$$a^*(q_t, d_t, \xi_t) = F^{-1}(G(\xi_t)|q_t, d_t). \tag{9}$$

[7]Note that $F(a_t|q_t, d_t)$ has a mass point at $a_t = 0$. Therefore, $\forall \xi_t$ such that $G(\xi_t) \leq F(0|q_t, d_t)$, we define $F^{-1}(G(\xi_t)|q_t, d_t) = 0$ in (9).

Equation (9) determines the policy function estimated from the observed data as the outcome of the first-stage estimation.

5.1.2 Second-Stage Estimation

In the second stage, we estimate the structural parameters such that the policy function estimated in the first stage and the observed data plan choice are both optimal under these parameter values.

We first define the expected utility given a usage strategy $a(q_t, d_t, \xi_t)$ at any state (q_0, d_0) *before* the realization of random shock ξ_0 as

$$\bar{U}(q_0, d_0; a) = E_{\{\xi_t\}} \sum_{t=0}^{d_0-1} \beta^t u\left(a\left(q_t, d_t, \xi_t\right)\right). \tag{10}$$

Under the true model parameters, the policy function estimated in the first stage, $a^*(q_t, d_t, \xi_t)$, should be optimal such that for any state, the expected utility given the policy function $a^*(q_t, d_t, \xi_t)$ is no less than that given any alternative $a^{'}(q_t, d_t, \xi_t)$. Following BBL, we construct the alternative policy functions by adding arbitrary perturbations to the observed policy function, that is,

$$a^{'}(q_t, d_t, \xi_t; e) = \max\left\{a^*(q_t, d_t, \xi_t) + e, 0\right\}. \tag{11}$$

Note that the perturbation e can be positive or negative, whereas $a^{'}(q_t, d_t, \xi_t)$ is bounded below at zero because the data usage cannot be negative. Thus, the optimality of $a^*(q_t, d_t, \xi_t)$ implies that the following inequality holds for any set of (q, d, e),

$$g(q, d, e) = \bar{U}(q, d; a^*) - \bar{U}\left(q, d; a^{'}(e)\right) \geq 0. \tag{12}$$

For estimation, we generate $n_I = 500$ different sets of states and policy function perturbation, $\{q_k, d_k, e_k\}_{k=1}^{n_I}$, and evaluate (12) at all of these 500 different sets of values.

In addition to $a^*(q_t, d_t, \xi_t)$ being optimal, under the true model parameters, users' data plan choices should also be optimal. Given the chosen data plan j, let F_j be its monthly fee and Q_j be the total plan quota accordingly. A user's rational selection of data plan j implies that the expected utility at the beginning of a month under this particular plan is no less than that under any other data plan $j^{'}$. Therefore, the following inequalities should also hold,

$$g^{'}\left(Q_j, Q_{j^{'}}, T; F_j, F_{j^{'}}\right) = \bar{U}\left(Q_j, T; a^{*}\right) - \bar{U}\left(Q_{j^{'}}, T; a^{*}\right) - \eta\left(F_j - F_{j^{'}}\right) \geq 0, \quad \forall j^{'} \neq j \tag{13}$$

Note that $\bar{U}\left(Q_j, T; a^{*}\right)$ is evaluated at the beginning of a month, and T is the total number of days in a month, which can be 31, 30, or 28. In formulating (13), we exploit the fact that the price rate for overage remains unchanged across different data plans for any given user, as discussed in Section 4.1. For this reason, even if a user were to switch to an alternative data plan, her policy function would remain the same as the one observed in the data. Therefore, we can use the same $a^{*}\left(q_t, d_t, \xi_t\right)$ estimated in the first stage to compute $\bar{U}\left(Q_{j^{'}}, T; a^{*}\right)$ for $\forall j^{'} \neq j$.

We obtain the estimators of the structural parameters $\{\mu, \sigma, \eta\}$ by minimizing the overall violation of the two sets of inequalities in (12) and (13), that is, by minimizing the following objective function

$$\sum_{k=1}^{n_I}\left(\min\left\{g\left(q_k, d_k, e_k\right), 0\right\}\right)^2 + \sum_{j^{'}}\left(\min\left\{g^{'}\left(Q_j, Q_{j^{'}}, T; F_j, F_{j^{'}}\right), 0\right\}\right)^2. \tag{14}$$

In evaluating (14), we apply Monte Carlo methods to compute the expected utility $\bar{U}\left(q, d; a\right)$, which cannot be derived in a closed form. For each expected utility, we simulate $n_S = 10,000$ random shock paths of $\{\xi_t\}_{t=0}^{d-1}$, calculate the usage path $\{a_t\}$ according to the given policy function and generate the state path $\{q_t, d_t\}$. We then calculate the sum of discounted utility in each simulation round and average over all n_S simulation paths to approximate the expected utility $\bar{U}\left(q, d; a\right)$,[8] based on which we can compute the objective function and obtain the structural parameter estimates.

5.2 Simulation Method

In this section, we briefly describe the method we develop to simulate each individual user's daily usage based on the estimated parameter values. We use this method to validate our estimation results by comparing the simulated usage with the actual usage, as we report in Section 6. It will also be used in the policy experiments in Section 7, where we explore various promotion designs.

Given the estimated parameter values, to simulate users' daily usage according to the dynamic

[8] We exploit the fact that the utility function (1) is linear in all model parameters (we can rewrite ξ_t as $\sigma_i \xi_{0t}$ where $\xi_{0t} \sim N\left(0, 1\right)$). Therefore, in computing each $\bar{U}\left(q, d; a\right)$, we only need to simulate the n_S paths once independent of the parameter values, which significantly expedites the estimation.

model, we first numerically solve the dynamic programming problem. In particular, we need to calculate the expected value function $\bar{V}(q, d)$ (defined by (5)) at any state (q, d). We use backward recursion to solve $\bar{V}(q, d)$. We start from the last period (i.e., the number of remaining days $d = 1$), where we can analytically derive the optimal data usage, and hence the value function $V(q, d = 1, \xi)$, in a closed form. We then compute $\bar{V}(q, d = 1)$ by simulating from the distribution of ξ and taking the average.

Note that we need to solve $\bar{V}(q, d)$ as a function of q continuously and obtain the value of $\bar{V}(q, d)$ at any possible q given d. To do so in a computationally feasible way, we first compute $\bar{V}(q, d)$ at multiple different values of q, and then interpolate the function values at other points. Recall that we have shown $\bar{V}(q, d)$ is increasing in q (Lemmas 1 and 2 in the Appendix). Therefore, the interpolation algorithm needs to preserve the monotonicity of the original function. For this reason, we choose to use the monotone piecewise cubic Hermite interpolation (Fritsch and Carlson, 1980), which is a variant of spline interpolation that preserves monotonicity of the original data points being interpolated while ensuring smoothness of the interpolated function. Thus, our method introduces a useful alternative to the dynamic structural model literature, which is different from the commonly used interpolation methods (e.g., Erdem and Keane, 1996; Judd, 1998) and especially relevant in the context when monotonicity is clearly established.

Knowing $\bar{V}(q, d)$ at any possible q given d, we can recursively derive $V(q, d + 1, \xi)$ by solving the optimization problem in (4). Repeating the steps described above, we can compute $\bar{V}(q, d + 1)$ at multiple points of q, and then interpolate the function values at other points. By backward recursion, we numerically solve the expected value functions day by day till the first day of the month. Figure 1 presents two examples of the expected value functions solved under different sets of parameter values.

After solving all expected value functions, the optimal data consumption $a^*(q, d, \xi)$ can be obtained by solving the maximization problem in (6) given any simulated utility shock ξ.

6 Results

We follow the two-stage estimation strategies described in Section 5.1 and estimate our dynamic model individual by individual for the 392 users who exhibit forward-looking usage patterns consistent with the model based on the test results in Section 4.2. We write the estimation programs in MATLAB and run them on a Windows workstation computer with multi-core processors, which facilitates parallel computing and significantly reduces the computing time despite the hundreds of separate estimations. We are able to obtain the estimates of the structural model parameters for each of the 392 users individually. Figure 4 shows the histogram of the estimates of $\{\mu_i, \sigma_i, \xi_i\}$ across all these users.

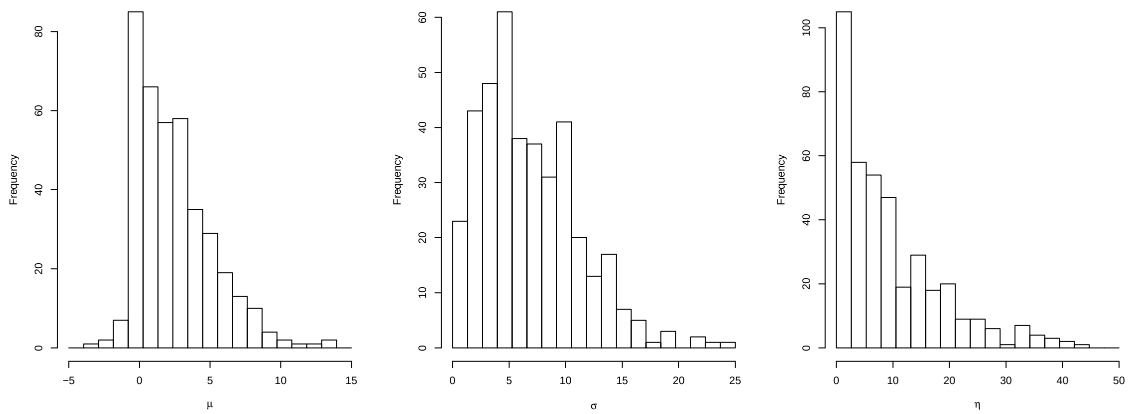


Figure 4: Dynamic Model Estimation Results

As Figure 4 shows, significant variation exists in the estimates of all three structural parameters across users, reflecting considerable heterogeneity in users' dynamic behavior of mobile data consumption. Note that though the mean value of the daily data consumption need, μ_i, is positive for most users, a small portion of users have negative μ_i's, indicating that they oftentimes do not need to use mobile data, which is consistent with the observations in our data discussed in Section 4.1. The parameter of particular interest is η_i, which measures each individual user's price sensitivity in mobile data consumption. As we can see, while the less price-sensitive users have η_i close to zero, some users' price sensitivity level could be considerably high. Such variation reaffirms the necessity of fully capturing individual heterogeneity and avoiding the bias of assuming a common parameter value or just a small number of segment values for all users. It is also worth noting that our dynamic model is able to identify a user's price sensitivity through intertemporal substitution reflected in daily usage, even if she has never exceeded her plan quota and thus has never actually incurred any overage charge. In contrast, a static model would be unable to identify the price sensitivity for the users who have never exceeded their plan quota. In this sense, capable of capturing the subtle dynamics in user consumption behavior and properly estimating individual-level structural parameters like the price sensitivity, our modeling approach is especially valuable for examining the effectiveness of various promotion designs, as we explore in Section 7.

To validate that the estimates of model parameters properly reflect the true dynamics in the actual data, we simulate each user's daily usage paths given the estimated structural parameter values by following the approach described in Section 5.2. For each user, we first simulate the daily utility shocks for a 30-day period and compute the optimal usage for each day given these shocks. We simulate such a 30-day usage path for 100 times, and calculate the average usage for each day. We then contrast the simulated data with the actual data (i.e., the average usage for each day of a month over the nine-month observation period). Figure 5 depicts the simulation results for six representative users.

In Figure 5, the solid curves represent the average daily usage over 100 rounds of simulation outputs, whereas the dashed curves represent the average daily usage over the nine months from the actual data. The vertical bars in grey indicate the variation of daily usage in the actual data (i.e., two standard deviations representing the 95% interval). As we can see, the simulated average usage falls into the intervals for almost all days and closely approximate the actual average daily usage, indicating good fit of the estimation in general. Especially, it is noteworthy that the simulated usage paths capture the dynamic trends observed in the actual data very well. Recall that by weighing immediate utility against future utility, forward-looking users restrain their data usage when the remaining quota is less because the marginal benefit of saving for the future becomes higher. Such intertemporal substitution results in a smooth and gradual decrease in daily usage over time towards the end of the month, as we observe from the actual usage in Figure 5. The simulated usage well captures such intertemporal trends, which validates our proposed model for the dynamic consumption of mobile data.

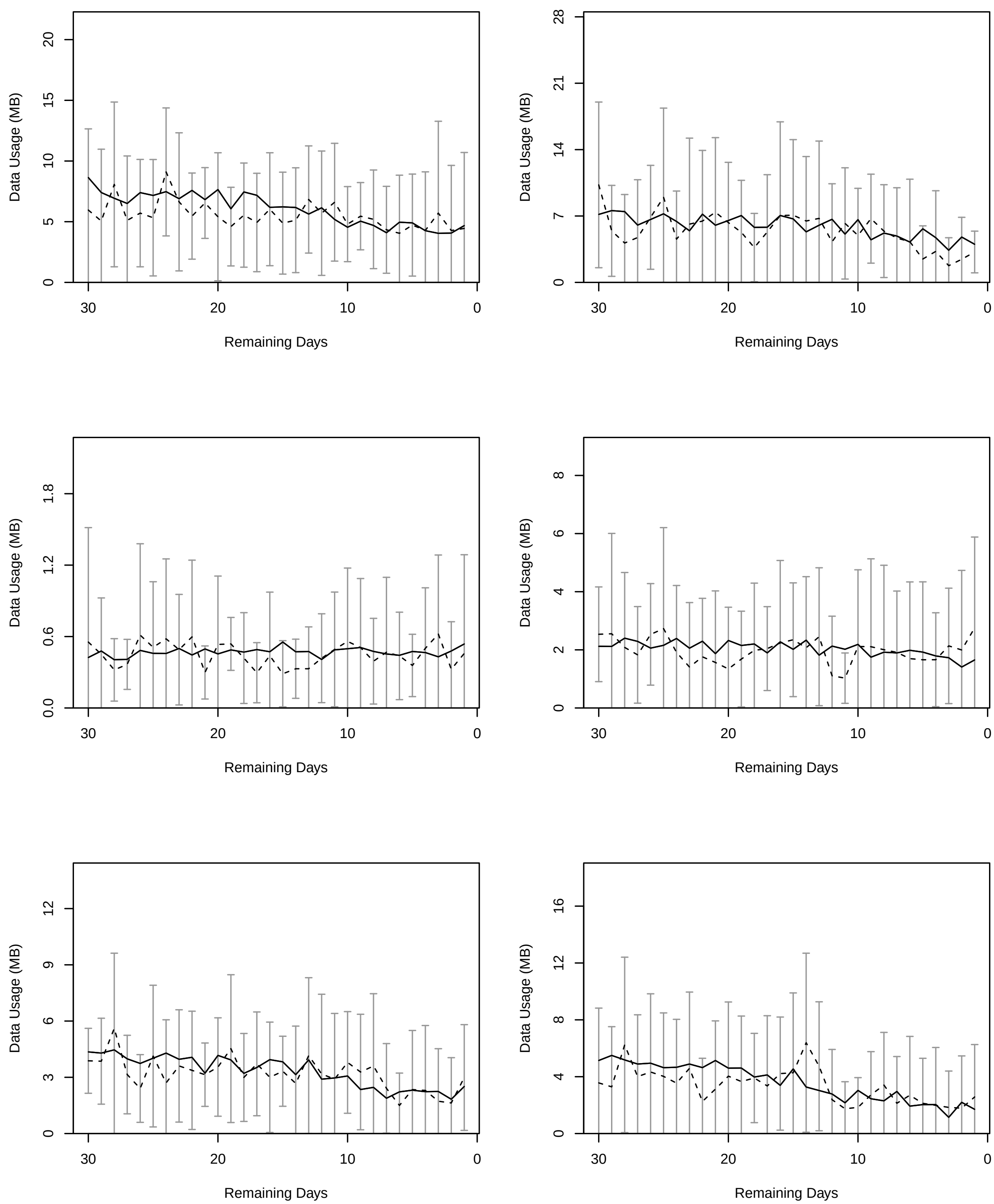


Figure 5: Simulation Output vs. Actual Data

Table 2: SUR Analysis for Estimated Model Parameters and Demographics

Demographic Var.	Estimated μ		Estimated $\log\sigma$		Estimated $\log\eta$	
	Estimate	St. Dev.	Estimate	St. Dev.	Estimate	St. Dev.
Intercept	3.33**	0.48	2.18**	0.17	2.39**	0.30
Age	-0.016	0.015	-0.013*	0.0053	-0.027**	0.0096
Gender	0.21	0.28	0.024	0.10	-0.20	0.18
Customer History	-0.082	0.043	-0.017	0.015	0.019	0.027

Note: **: p-value $<$ 0.01; *: p-value $<$ 0.05

We are also interested in how the estimated structural parameters relate to users' demographic information. Thus, we regress the estimated $\{\mu_i, \sigma_i, \eta_i\}$ on user demographics, including age, gender, and customer history (converted into the unit of 10 months). To account for the possible correlation among the three estimates for each individual, we apply the seemingly unrelated regression (SUR) so that a system of equations is estimated with the estimated $\{\mu_i, \sigma_i, \eta_i\}$ as the vector dependent variable. Table 2 shows the results. Somewhat consistent with the general notion that demographic information is not most informative in predicting user behavior, we find that neither gender nor customer history have any significant effects on the estimated mean usage level μ_i, the standard deviation of the daily random shock σ_i, or the price sensitivity η_i. On the other hand, age is found to have significantly negative effects on η_i and σ_i, which indicates older users generally have lower price sensitivity and less daily variation in usage. The results can be explained by the fact that income increases with age in general, so the price sensitivity is lower for older users accordingly. Meanwhile, older users are likely to have more stable usage habit, resulting in lower usage variance in general.

7 Policy Implications

Capturing the dynamics of mobile data consumption, we are able to estimate individual-specific structural parameters such as price sensitivity, which are invariant to policy changes. Therefore, it allows us to predict each individual user's rational response in adjusting to new pricing policies that have not been implemented yet. Thus, our structural model enables mobile carriers to evaluate the effectiveness of various promotion strategies before carrying them out.

The focal carrier is particularly interested in exploring promotions near the end of a month, when users are likely to have expended their quotas and the overall data usage typically drops.

Effective month-end promotions can help the carrier make full use of the excess bandwidth capacity, improve customer satisfaction, and increase revenue. Because forward-looking users greatly restrain their usage as their quotas are close to being expended, they tend to actively respond to month-end promotions. Therefore, these users are the natural target group of interest. In this section, we conduct counterfactual experiments to examine two types of month-end promotions targeting forward-looking users, as an illustration of how our model can be applied to deliver valuable managerial implications.

7.1 Discounts in Overage Charges

One of the most common strategies to promote usage or sales is offering discounts in price. Therefore, we first investigate how the focal carrier could increase revenue by offering discounts in the price rates for overage when users are about to deplete their quotas. Note that a dynamic model is essential in order to obtain reliable revenue implications. In addition to the obvious bias of static models in modeling forward-looking users' dynamic behavior, there are more reasons for the necessity of a dynamic model. First, as discussed in Section 6, a dynamic model can identify individual user's price sensitivity through intertemporal variation in usage. In contrast, a static model would be unable to identify the price sensitivity levels for users who have never exceeded their quotas during the observation period, which would lead to biased evaluation of the promotion effects. Second, the carrier usually sends out promotion offers before users completely expend their plan quota. Upon being notified of the discounted rate, users start to adjust their usage immediately, even if they still have some quota left and do not incur overage charges right away. Such effects can be captured only by a dynamic model properly incorporating users' forward-looking decision processes.

To explore the optimal promotional rate, we conduct the following counterfactual simulation study. We consider the carrier notifies users about a discount, as a percentage δ ($0 < \delta < 1$), of the price rate for overage when users are about to deplete their plan quota. Under such a promotion strategy, we simulate each user's usage individual by individual for all 392 users estimated using the dynamic model. For each user, we simulate their usage during a 30-day period for 100 rounds. In each round, the user starts using mobile data given the price structure of her chosen data plan. We simulate a user's usage path by first simulating the daily utility shocks and then computing the optimal usage for each day according to (6). Notice that we need to solve the expected value functions $\bar{V}(q,d)$ for $d = 1, ..., 30$ beforehand based on the structural parameter estimates, following the method described in Section 5.2. Once the remaining quota falls below a certain threshold (i.e., 10% of each user's total monthly quota as we set for this analysis), the user is notified of the discounted extra rate and starts adjusting her usage given this new rate. She re-solves the dynamic programming problem in (4) by replacing the original price rate p_j with $p_j^{'} = \delta p_j$ in (1). For this reason, we need to re-calculate the expected value functions $\bar{V}(q,d)$ by backward recursion using the new discounted rate $p_j^{'}$ and find the new optimal usage for the rest days till the end of the month. We can then calculate the total usage simulated over the 30-day period and the total extra charge (beyond the fixed plan fees) incurred. Averaging the extra charges simulated over all 100 rounds yields the expected (extra) revenue for each individual user. Finally, we can obtain the expected average (extra) revenue under a certain promotion strategy by averaging over all 392 users. Table 3 shows the results with different discount levels.

Table 3: Results of Counterfactual Analysis 1: Discounts in Overage Charges

Discount (δ)	Average Revenue (CNY)	Revenue Increase (%)
100%	5.66	-
90%	6.12	8.1
80%	6.56	15.9
70%	6.93	22.4
60%	7.18	26.9
50%	7.21	27.4
40%	6.92	22.3

As we can see from Table 3, the expected average revenue without discount (i.e., $\delta = 100\%$) is ¥5.66. As discount deepens, the revenue increases, indicating that some users, especially those with high price sensitivity, do respond actively to price reduction by increasing their usage significantly. Facing the usual trade-off between profit margin and sales volume, the revenue follows a typical inverted-U curve, reaching its peak around $\delta = 50\%$ and starting to decrease as price further goes down. As the counterfactual analysis shows, a month-end promotion offering 40% to 50% off the original overage charges could increase the expected average revenue by as much as 27%.

7.2 Month-End Unlimited-Use Passes

The mobile carrier is also interested in another type of month-end promotion, which offers unlimited-use passes that grant users unlimited mobile data usage during the last several days of a month with a fixed fee. Notice that besides the reasons discussed in Section 7.1, assessing the effectiveness of this promotion strategy is even more critically dependent on a dynamic model, because we need to evaluate whether and how likely each user will accept the promotion offer. Users make their acceptance/rejection decisions by comparing their expected future utility in both cases, which can be properly evaluated only by modeling users' forward-looking decision processes.

We conduct the second counterfactual simulation study as follows. We consider the mobile carrier sends out a promotion on the nth day to the end of a month. The promotion offers users the option to purchase a month-end unlimited-use pass such that after paying a fixed fee P, users can freely use mobile data till the end of the month without any additional charge, regardless of the remaining quota from their original data plan. If a user chooses to decline the offer, she continues with her original data plan and pays for any extra usage according to the plan rate. Following the same approach as in Section 7.1, we simulate each user's behavior under this promotion strategy individual by individual for all 392 users estimated using our dynamic model. For each user, we simulate a 30-day usage path for 100 rounds. In each round, we simulate a user's daily usage for the first $(30-n)$ days according to the original dynamic programming problem given her chosen data plan and the structural parameter estimates. At the beginning of the $(30-n+1)$th day, the user is notified of the promotional offer and decides whether to accept it or not. If she chooses to accept the offer, the expected utility can be derived as

$$\bar{V}^{Accept}(n,P)=\sum_{\tau=0}^{n-1}\beta^{\tau}E_{\xi_{it}}\frac{1}{2}\left(\max\left\{\mu_i+\xi_{it},0\right\}\right)^2-\eta_i P. \tag{15}$$

The first term of (15) represents the sum of discounted expected utility from unconstrained data usage over the last n days of a month. When a user can freely consume mobile data without monetary cost, the optimal usage every day is simply $\max\left\{\mu_i+\xi_{it},0\right\}$, which maximizes the utility function in (1) without the last term. The second term of (15) represents the utility cost associated with the fee of the pass. Recall that the expected utility from continuing with the original plan when there are $d=n$ days left and q_{it} quota remaining is $\bar{V}\left(q_{it},n\right)$ by (5). Therefore, a user accepts the promotional offer if and only if $\bar{V}^{Accept}(n,P) > \bar{V}(q_{it},n)$. Notice that while $\bar{V}^{Accept}(n,P)$ is independent of the remaining quota q_{it}, $\bar{V}(q_{it},n)$ does depend on q_{it}. As a result, whether a user will accept the promotional offer or not depends on her usage history in the earlier part of the month. Therefore, we can determine a user's acceptance decision and continue to simulate her usage for the last n days, which yields the total surcharge (beyond the monthly plan fee) a user pays in each simulation round. We can then compute the average expected revenue and the average probability of accepting the offer by averaging over all simulation rounds and all users. Table 4 shows the results with different values of n and P.

Table 4: Results of Counterfactual Analysis 2: Month-End Unlimited-Use Passes

Last n Days	Fee P (CNY)	Acceptance Rate	Average Revenue (CNY)	Revenue Increase (%)
3	5.00	0.3964	6.27	10.8
3	10.00	0.1245	6.32	11.7
3	15.00	0.0176	5.82	2.8
4	5.00	0.4438	6.01	6.2
4	10.00	0.2428	6.81	20.3
4	15.00	0.0547	6.12	8.1
5	5.00	0.4764	5.68	0.35
5	10.00	0.2851	6.78	19.8
5	15.00	0.1235	6.61	16.8

As Table 4 shows, the average acceptance rate of the promotional offer increases as the fee P reduces or the number of days n increases. For the same n, the average revenue follows an inverted-U curve as the fee increases, which can be easily explained by the usual trade-off between price and demand. In comparison with the average revenue without any promotion (i.e., ¥5.66)[9], selling the unlimited-use pass at a price too low (e.g., a 5-day pass for ¥5.00) can result in high acceptance rate but may not help increase the revenue at all. Instead, a properly priced promotion (e.g., a 4-day pass for ¥10.00) can increase the expected revenue by as much as 20%.

There are a few caveats worth mentioning with regard to interpreting the results of the counterfactual analyses in this section. First, the analyses here assume users do not anticipate any

[9] In order to make the results comparable across different scenarios, we use the same set of random draws of utility shocks for all promotion scenarios in both Table 3 and Table 4, so as to eliminate any differences caused by simulation variation.

promotion before being notified. If the company were to offer the same promotion routinely, users could form expectations about the promotion offers, which would change their dynamic usage behavior from the very beginning and hence change the promotion effects. To keep the promotion unexpected from users' perspective, the mobile carrier can offer promotions sporadically, randomly select a subset of all users each time, and constantly switch among different classes of promotion designs. Second, the revenue analyses here do not take costs into account. In general, when making use of the available bandwidth near month end, the carrier's marginal cost of providing mobile data communication is negligible. Nevertheless, if the costs of mobile data provision do need to be accounted for, the company can easily incorporate the actual cost information into the analysis and re-evaluate the effects of various promotion designs on profit.

8 Conclusion

With the proliferation of modern mobile devices, the expansion of fast-speed data networks, and mobile carriers shifting emphasis to limited data plans, our paper addresses the important yet under-examined issue of the dynamic consumption behavior of mobile data users. Properly characterizing the heterogeneous behavior of different individual users requires well developed modeling and estimation approaches. Gaining insights into users' dynamic consumption behavior can help mobile carriers design profitable promotions near month end targeting the users who are forward-looking. Our study thus provides meaningful implications for both researchers and practitioners.

We enrich the literature on dynamic consumer behavior by systematically examining the dynamic patterns in the consumption of a pre-specified quota over a finite period of time with uncertain utility shocks. With both theoretical results and empirical evidence, we uncover distinct patterns for forward-looking and myopic behavior, which are easily testable with observational data. In this sense, our study not only deepens the theoretical understanding of dynamic consumer behavior but also provides a useful instrument for future research exploring consumer dynamics in relevant contexts.

We also highlight the necessity and feasibility of fully capturing individual heterogeneity in modeling dynamic consumer behavior. We demonstrate that well-structured data and properly-optimized estimation can achieve sufficient computational efficiency for separate estimation of individual dynamic structural models in a speedy fashion. Our results show significant heterogeneity among users in their latent characteristics such as price sensitivity and data usage needs. They can all be individually identified based on the intertemporal substitution observed in the daily usage data, as we show. An individual estimation of these underlying structural parameters enables companies to perform more precise targeting for various marketing objectives beyond the usual demographic targeting. Moreover, our study also provides empirical evidence for considerable individual differentiation in terms of forward-looking and myopic behavior. Thus, we demonstrate the necessity of controlling for possible consumer myopia in modeling individual dynamic behavior. We propose an innovative approach for such a purpose, which offers a valuable alternative in the face of the inherent difficulty in identifying individual discount factors.

In addition, we illustrate the great potential of dynamic promotion in accordance with consumers' dynamic behavior, and our modeling and estimation approach furnishes a helpful tool for this purpose. As we show, forward-looking users restrain their usage as their data plan quota gets depleted towards the end of month. Our counterfactual analysis suggests that properly designed month-end promotions targeting forward-looking users can considerably increase revenue for mobile carriers. Such revenue opportunities are possible to exploit only with a well developed dynamic analysis. It empowers companies to employ various innovative promotions dynamically at the right time to the right group, complementing the traditional static promotion techniques.

For future research, opportunities abound in the flourishing area related to mobile data usage. For example, first, it would provide greater insights into consumer behavior if we could know what amount of data is used in which type of mobile applications. Second, as a large amount of mobile data is used for user-generated content and social media (e.g., Susarla et al., 2012), it would be interesting to know the mutual influence of data usage between consumers. Third, if we had data on consumer usage before and after a carrier changes its plan and fee structures, we could utilize the natural experiment to study how consumers adjust their behavior in both the amount of data usage and the choice of applications. Finally, a deeper understanding of consumer behavior in data consumption could help mobile carriers explore various new marketing strategies. These extensions could be developed based on our proposed modeling framework with the availability of new data, which are all promising directions for future research.

References

Arcidiacono, Peter, Robert A. Miller. 2011. Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica* **79**(6) 1823–1867.

Bajari, Patrick, C. Lanier Benkard, Jonathan Levin. 2007. Estimating dynamic models of imperfect competition. *Econometrica* **75**(5) 1331–1370.

Busse, Meghan R., Christopher R. Knittel, Florian Zettelmeyer. 2013. Are consumers myopic? evidence from new and used car purchases. *The American Economic Review* **103**(1) 220–256.

Chevalier, Judith, Austan Goolsbee. 2009. Are durable goods consumers forward-looking? evidence from college textbooks. *The Quarterly Journal of Economics* **124**(4) 1853–1884.

Chung, Doug J., Thomas Steenburgh, K. Sudhir. 2014. Do bonuses enhance sales productivity? a dynamic structural analysis of bonus-based compensation plans. *Marketing Science* **33**(2) 165–187.

Dube, Jean-Pierre, Gunter J. Hitsch, Pranav Jindal. 2014. The joint identification of utility and discount functions from stated choice data: An application to durable goods adoption. *Quantitative Marketing and Economics* **12**(4) 331–377.

Erdem, Tulin, Michael P. Keane. 1996. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science* **15**(1) 1–20.

Frederick, Shane, George Loewenstein, Ted O'Donoghue. 2002. Time discounting and time preference: A critical review. *Journal of Economic Literature* **40**(2) 351–401.

Fritsch, Fred N., Ralph E. Carlson. 1980. Monotone piecewise cubic interpolation. *SIAM Journal of Numerical Analysis* **17**(2) 238–246.

Gabaix, Xavier, David Laibson. 2006. Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics* **121**(2) 505–540.

Ghose, Anindya, Sang Pil Han. 2011. An empirical analysis of user content generation and usage behavior on the mobile internet. *Management Science* **57**(9) 1671–1691.

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

Hoch, Stephen J., George F. Loewenstein. 1991. Time-inconsistent preferences and consumer self-control. *Journal of Consumer Research* **17**(4) 492–507.

Hotz, V. Joseph, Robert A. Miller. 1993. Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies* **60**(3) 497–529.

Huang, Yan, Param Vir Singh, Anindya Ghose. 2015. A structual model of employee behavioral dynamics in enterprise social media. *Management Science* forthcoming.

ITU. 2014. Mobile-broadband penetration approaching 32 percent: Three billion Internet users by end of this year. Market research report, International Telecommunication Union (ITU).

Iyengar, Raghuram, Asim Ansari, Sunil Gupta. 2007. A model of consumer learning for service quality and usage. *Journal of Marketing Research* **44**(4) 529–544.

Judd, Kenneth. 1998. *Numerical Methods in Economics*. MIT Press, Cambridge MA.

Kim, Youngsoo, Rahul Telang, William B. Vogt, Ramayya Krishnan. 2010. An empirical analysis of mobile voice service and SMS: A structural model. *Management Science* **56**(2) 234–252.

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

Manski, Charles F. 1993. Dynamic choice in social settings: Learning from the experiences of others. *Journal of Econometrics* **58**(1-2) 121–136.

Misra, Sanjog, Harikesh S. Nair. 2011. A structural model of sales-force compensation dynamics: Estimation and field implementation. *Quantitative Marketing and Economics* **9**(3) 211–257.

Niculescu, Marius F., Seungjin Whang. 2012. Codiffusion of wireless voice and data services: An empirical analysis of the Japanese mobile telecommunications market. *Information Systems Research* **23**(1) 260–279.

Pashardes, Panos. 1986. Myopic and forward looking behavior in a dynamic demand system. *International Economic Review* **27**(2) 387–397.

Rust, John. 1987. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica* **55**(5) 999–1033.

Segan, Sascha. 2014. The death of unlimited mobile data. *PC Magazine Online* March 10.

Susarla, Anjana, Jeong-Ha Oh, Yong Tan. 2012. Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research* **23**(1) 23–41.

Tang, Qian, Bin Gu, Andrew B. Whinston. 2012. Content contribution for revenue sharing and reputation in social media: A dynamic structural model. *Journal of Management Information Systems* **29**(2) 41–75.

Xu, Jiao, Chris Forman, Jun B. Kim, Koert van Ittersum. 2014. News media channels: Complements or substitutes? Evidence from mobile phone usage. *Journal of Marketing* **78**(4) 97–112.

Yao, Song, Carl F. Mela, Jeongwen Chiang, Yuxin Chen. 2012. Determining consumers' discount rates with field studies. *Journal of Marketing Research* **49**(6) 822–841.

A Appendix

Proofs of Propositions in Section 3.2

To simplify notations, without causing confusion, below we suppress subscripts i and t. Also notice that throughout all the proofs below, wherever *strict* monotonicity or concavity applies, we explicitly stress it; without explicit stress of strictness, we mean *weakly* increasing/decreasing (or concave/convex).

Proof of Proposition 1. Myopic users determine their daily usage by maximizing the per-period utility only, which is defined in (1). The optimal daily usage a^* can thus be derived as

$$a^* = \begin{cases} \mu + \xi - \eta p \, (> q) & \text{if } 0 \leq q < \mu + \xi - \eta p \\ q & \text{if } \mu + \xi - \eta p \leq q \leq \mu + \xi \\ \max\{\mu + \xi, 0\} \, (< q) & \text{if } q > \mu + \xi \end{cases} \tag{16}$$

In any day before the day when the data plan quota is fully expended, $a^* = \max\{\mu+\xi, 0\} (< q)$, which is obviously independent of q. Q.E.D.

In order to prove Proposition 2, we first prove two important lemmas with regard to the properties of the expected value function $\bar{V}(q,d)$ as defined in (5).

Lemma 1. *For the last period, the expected value function $\bar{V}(q, d=1)$ is continuous, increasing, differentiable, and strictly concave in the remaining data plan quota q.*

Proof. Recall that in the last period,

$$V(q, d=1, \xi) = \max_{a \geq 0} (\mu+\xi) a - \frac{1}{2} a^2 - \eta p \max\{a-q, 0\} \tag{17}$$

We can thus explicitly solve the value function as

$$V(q, d=1, \xi) = \begin{cases} \frac{1}{2}(\mu+\xi-\eta p)^2 + \eta p q & \text{if } 0 \leq q < \mu+\xi-\eta p \\ (\mu+\xi) q - \frac{1}{2} q^2 & \text{if } \mu+\xi-\eta p \leq q \leq \mu+\xi \\ \frac{1}{2}(\max\{\mu+\xi, 0\})^2 & \text{if } q > \mu+\xi, \end{cases} \tag{18}$$

where $q \geq 0$. It is easy to show that $V(q, d=1, \xi)$ is continuous, increasing, differentiable, and concave in q given any ξ. The continuity can be easily verified by checking the function value at each endpoint. (Notice that because $q \geq 0$, if $\mu+\xi-\eta p < \mu+\xi < 0$, only the third segment applies and (18) reduces to a constant so that $V(q, d=1, \xi) \equiv 0$ for $\forall q \geq 0$; if $\mu+\xi-\eta p < 0 < \mu+\xi$, (18) reduces to two segments.) The monotonicity is immediate because the piecewise function is continuous and piecewise increasing in q. $V(q, d=1, \xi)$ is differentiable because the left and right derivatives are equal at each endpoint: $\frac{\partial}{\partial q} V_{q \to (\mu+\xi-\eta p)^+}(q, d=1, \xi) = \frac{\partial}{\partial q} V_{q \to (\mu+\xi-\eta p)^-}(q, d=1, \xi) = \eta p$, and $\frac{\partial}{\partial q} V_{q \to (\mu+\xi)^+}(q, d=1, \xi) = \frac{\partial}{\partial q} V_{q \to (\mu+\xi)^-}(q, d=1, \xi) = 0$. $V(q, d=1, \xi)$ is concave in q because it is differentiable and piecewise concave in q.

Given that $V(q, d=1, \xi)$ is continuous, increasing, and differentiable in q for any ξ, it is immediate that the expected value function $\bar{V}(q,d) = E_\xi V(q, d, \xi)$, as an integral over all ξ, is also continuous, increasing, and differentiable in q.

To show that $\bar{V}(q,d)$ is strictly concave in q, note that because $V(q, d=1, \xi)$ is concave in q, by the definition of concavity, for any $q_1, q_2 > 0$ and $\lambda \in (0,1)$, we have

$$V\left((1-\lambda) q_1 + \lambda q_2, d=1, \xi\right) \geq (1-\lambda) V\left(q_1, d=1, \xi\right) + \lambda V\left(q_2, d=1, \xi\right) \tag{19}$$

for any ξ. Because $V(q, d=1, \xi)$ is *strictly* concave in the second segment in (18), *strict* inequality holds in (19) when $\xi \in (-\mu + q_1, -\mu + q_1 + \eta p) \cup (-\mu + q_2, -\mu + q_2 + \eta p)$. Recall that ξ is a random variable with a continuous support over the entire real field. Therefore, when taking expectation over ξ on both sides of (19), we have

$$E_\xi V\left((1-\lambda) q_1 + \lambda q_2, d=1, \xi\right) > (1-\lambda) E_\xi V\left(q_1, d=1, \xi\right) + \lambda E_\xi V\left(q_2, d=1, \xi\right), \tag{20}$$

which shows $\bar{V}(q, d=1)$ is *strictly* concave in q. Q.E.D. □

Lemma 2. *If the expected value function for the next period, $\bar{V}\left(q^{'}, d-1\right)$, is continuous, increasing, differentiable, and strictly concave in $q^{'}$, then the expected value function for the current period, $\bar{V}(q, d)$, is also continuous, increasing, differentiable, and strictly concave in q.*

Proof. We first show that given any ξ, the value function $V(q, d, \xi)$ is continuous, increasing, differentiable, and concave in q if $\bar{V}\left(q^{'}, d-1\right)$ is continuous, increasing, differentiable, and strictly concave in $q^{'}$. Substituting (1) and (3) into (4), we can rewrite the current-period value function as

$$V(q, d, \xi) = \max_{a \geq 0} (\mu + \xi) a - \frac{1}{2} a^2 - \eta p [a - q]^+ + \beta \bar{V}\left([q-a]^+, d-1\right), \tag{21}$$

where $[\cdot]$ stands for $\max\{\cdot, 0\}$. To simplify notation, we use $\bar{V}^{'}(q, d)$ to represent $\frac{\partial}{\partial q} \bar{V}(q, d)$ for the rest of this proof.

Let $\tilde{a}$ be the solution to the first order condition (with respect to a) when $a < q$, that is,

$$\mu + \xi - \tilde{a} - \beta \bar{V}^{'}(q - \tilde{a}, d-1) = 0 \tag{22}$$

Therefore, $\tilde{a} < q$ if and only if $\mu + \xi - q - \beta \bar{V}^{'}(0, d-1) < 0$. When $a > q$, the first order condition yields

$$\mu + \xi - a^* - \eta p = 0. \tag{23}$$

$a^* > q$ if and only if $\mu + \xi - q - \eta p > 0$. Notice that $\beta \bar{V}^{'}(0, d-1) < \eta p$ given $\beta < 1$. Therefore, we can summarize the optimal usage in the current period as

$$a^*(q, d, \xi) = \begin{cases} \mu + \xi - \eta p \, (> q) & \text{if } 0 \leq q < \mu + \xi - \eta p \\ q & \text{if } \mu + \xi - \eta p \leq q \leq \mu + \xi - \beta \bar{V}^{'}(0, d-1) \\ \max\{\tilde{a}, 0\} \, (< q) & \text{if } q > \mu + \xi - \beta \bar{V}^{'}(0, d-1) \end{cases} \tag{24}$$

Again, because $q \geq 0$, if $\mu + \xi - \beta \bar{V}^{'}(0, d-1) < 0$ or $\mu + \xi - \eta p < 0$, (24) reduces to one or two segments only. Accordingly, the current-period value function can be written as

$$V(q, d, \xi) = \begin{cases} \frac{1}{2}(\mu + \xi - \eta p)^2 + \eta p q + \beta \bar{V}(0, d-1) & \text{if } 0 \leq q < \mu + \xi - \eta p \\ (\mu + \xi) q - \frac{1}{2} q^2 + \beta \bar{V}(0, d-1) & \text{if } \mu + \xi - \eta p \leq q \leq \mu + \xi - \beta \bar{V}^{'}(0, d-1) \\ F(q; \xi) & \text{if } q > \mu + \xi - \beta \bar{V}^{'}(0, d-1), \end{cases} \tag{25}$$

where $F(q; \xi)$ is defined by substituting the optimal usage $a^* = \max\{\tilde{a}, 0\} \, (< q)$ into (21), that is,

$$F(q; \xi) = (\mu + \xi) a^* - \frac{1}{2} a^{*2} + \beta \bar{V}(q - a^*, d-1). \tag{26}$$

It is easy to show that $V(q, d, \xi)$ is continuous in q by verifying the continuity of function value at the endpoints: for example, when $q = \mu + \xi - \beta \bar{V}^{'}(0, d-1)$, $a^* = q$ so $F(q; \xi) = (\mu + \xi) q - \frac{1}{2} q^2 + \beta \bar{V}(0, d-1)$. To show that $V(q, d, \xi)$ is increasing in q, we just need to show $F(q; \xi)$ is increasing in q, because it is obviously true for the first two segments of (25). Taking derivative with respect to q on both sides of (26), by Envelope Theorem, we have

$$F^{'}(q; \xi) = \beta \bar{V}^{'}(q - a^*, d-1) \geq 0, \tag{27}$$

because $\bar{V}\left(q^{'}, d-1\right)$ is increasing in $q^{'}$. Therefore, $F(q; \xi)$ is increasing in q, and so is $V(q, d, \xi)$.

It is easy to show that $V(q, d, \xi)$ is differentiable in q, noticing that

$$\frac{\partial}{\partial q} V_{q \to \left(\mu + \xi - \beta \bar{V}^{'}(0, d-1)\right)^-}(q, d, \xi) = \beta \bar{V}^{'}(0, d-1) \tag{28}$$

$$\frac{\partial}{\partial q} V_{q \to \left(\mu+\xi-\beta \bar{V}'(0,d-1)\right)^{+}} (q,d,\xi) = F' (q;\xi) = \beta \bar{V}' (0, d-1), \tag{29}$$

where (29) holds by (27) and the fact that $a^* = q$ when $q = \mu + \xi - \beta \bar{V}' (0, d-1)$.

We next show that $V(q,d,\xi)$ is concave in q. It is obvious that $V(q,d,\xi)$ is (weakly) concave when $0 \le q < \mu + \xi - \eta p$ and strictly concave when $\mu + \xi - \eta p \le q \le \mu + \xi - \beta \bar{V}'(0, d-1)$. Given that $V(q,d,\xi)$ is differentiable in q, therefore, we only need to show that $F(q;\xi)$ is (strictly) concave in q for $q > \mu + \xi - \beta \bar{V}'(0, d-1)$.

We prove by the definition of concavity. Consider any $q_1, q_2 \ge \max\left\{\mu + \xi - \beta \bar{V}'(0, d-1), 0\right\}$, let $\hat{q}_1 = q_1 - a^*(q_1)$ and $\hat{q}_2 = q_2 - a^*(q_2)$. In other words, we use ˆ to represent the remaining quota at the beginning of the next period as a result of the optimal amount of usage in the current period. Note that $0 \le \hat{q}_1 \le q_1$ and $0 \le \hat{q}_2 \le q_2$. Denote $\bar{q} = \lambda q_1 + (1-\lambda) q_2$ and $\bar{\hat{q}} = \lambda \hat{q}_1 + (1-\lambda)\hat{q}_2$ for $\forall \lambda \in (0,1)$. Clearly, $0 \le \bar{\hat{q}} \le \bar{q}$. In addition, define $U\left(q, q'\right) = (\mu + \xi)\left(q - q'\right) - \frac{1}{2}\left(q - q'\right)^2$. It is easy to show that $U\left(q, q'\right)$ is concave in $\left(q, q'\right)$ because it is a quadratic function with a negative semidefinite Hessian matrix. Hence, $F(\bar{q};\xi)$ from (26) can be rewritten as

$$\begin{aligned} F(\bar{q};\xi) &= U\left(\bar{q}, \hat{\bar{q}}\right) + \beta \bar{V}\left(\hat{\bar{q}}, d-1\right) \\ &\ge U\left(\bar{q}, \bar{\hat{q}}\right) + \beta \bar{V}\left(\bar{\hat{q}}, d-1\right) \\ &\ge \lambda U(q_1, \hat{q}_1) + (1-\lambda) U(q_2, \hat{q}_2) + \beta \bar{V}\left(\bar{\hat{q}}, d-1\right) \\ &> \lambda U(q_1, \hat{q}_1) + (1-\lambda) U(q_2, \hat{q}_2) + \beta \lambda \bar{V}(\hat{q}_1, d-1) + \beta(1-\lambda)\bar{V}(\hat{q}_2, d-1) \\ &= \lambda F(q_1;\xi) + (1-\lambda) F(q_2;\xi) \end{aligned} \tag{30}$$

The first inequality in (30) holds because of the optimality of $\hat{\bar{q}}$; the second inequality holds because of the concavity of $U\left(q, q'\right)$ in $\left(q, q'\right)$; the third (strict) inequality holds because of the strict concavity of $\bar{V}\left(q', d-1\right)$ in q'. As a result, $F(q;\xi)$ is strictly concave in q for any $q \ge \max\left\{\mu + \xi - \beta \bar{V}'(0, d-1), 0\right\}$. Therefore, $V(q,d,\xi)$ is concave in q for any $q \ge 0$ and strictly concave if $q \ge \mu + \xi - \eta p$.

Given we have shown that $V(q,d,\xi)$ is continuous, increasing, differentiable, and concave in q for any ξ, and it is strictly concave in q when $\xi < -\mu + q + \eta p$, following the same logic as the last part of the proof of Lemma 1, we conclude that $\bar{V}(q,d) = E_\xi V(q,d,\xi)$ is continuous, increasing, differentiable, and strictly concave in q. Q.E.D. □

Proof of Proposition 2. Recall the optimal usage $a^*(q,d,\xi)$ derived in (24) for any $d \geq 2$. In any day before the day when the data plan quota is fully expended, $a^*(q,d,\xi) = \max\{\tilde{a},0\}\,(< q)$. We want to show that $\tilde{a}$, the solution to (22), is *strictly* increasing in q.

Recall that $\tilde{a}(q)$ solves the first order condition

$$\mu + \xi - \tilde{a}(q) - \beta \bar{V}'(q - \tilde{a}(q), d-1) = 0 \tag{31}$$

By Lemma 1 and Lemma 2, the expected value function $\bar{V}(\cdot,\cdot)$ is increasing and strictly concave in the remaining quota for any period. Therefore, for any $q' > q$,

$$\mu + \xi - \tilde{a}(q) - \beta \bar{V}'\left(q' - \tilde{a}(q), d-1\right) > 0 \tag{32}$$

because $\bar{V}'(q' - \tilde{a}(q), d-1) < \bar{V}'(q - \tilde{a}(q), d-1)$ given the strict concavity of $\bar{V}(\cdot, d-1)$. As a result, $\tilde{a}\left(q'\right) > \tilde{a}(q)$. Therefore, $\tilde{a}$ is strictly increasing in q, which implies $a^*(q,d,\xi) = \max\{\tilde{a},0\}$ is strictly increasing in q if $0 < a^* < q$. Q.E.D.