Consumer Choice of Service Plan: Switching Cost and Usage Uncertainty

Ping Xiao
Tat Chan
Chakravarthi Narasimhan

May, 2007

---

Ping Xiao is a doctoral student in Marketing, Tat Chan is an Assistant Professor of Marketing, and Chakravarthi Narasimhan is the Philip L. Siteman Professor of Marketing, all at Washington University in St. Louis. We are grateful to Matthew Gentzkow, Anja Lambrecht and seminar participants at the 4th QME Conference for many helpful comments. Chan and Narasimhan acknowledge the generous financial support from The Boeing Center on Technology, Information and Manufacturing. An early version of this paper was circulated under the title “Service Bundles under Three-Part Tariffs.”

Corresponding author: Chakravarthi Narasimhan: narasimhan@wustl.edu, Ph. (314) 935-6313
Consumer Choice of Service Plan: Switching Cost and Usage Uncertainty

Abstract

A strategy used by firms in industries such as wireless is to bundle several services into a service plan and administer a nonlinear price with different marginal prices for different usage levels. Usually firms offer more than one service plan. While product bundling and nonlinear pricing have been extensively studied separately in the economics and marketing literature there are no theoretical or empirical studies that consider them simultaneously. Consumers may be uncertain about their tastes for different services, their preferences may be correlated, and they may face psychological and out of pocket costs to switch among plans. A firm needs to take these into account in designing its pricing strategy. To explore these issues we develop a structural model of consumers' choice of plans and usage decisions, and estimate the model using a dataset from a wireless service provider. We find that consumers derive utility from both voice call and text message and ignoring the fact that consumers simultaneously determine consumption levels of multiple services may lead to biased estimation results. Moreover, we find a somewhat counter-intuitive result that consumer preferences for voice call and text message are positively correlated. We show that incorporating switching costs, consumer uncertainty and learning are important in explaining the switching patterns observed in the data. We also conduct several “what-if” scenarios to illustrate how the service penetration process may be different under various behavioral assumptions.

Keywords: Nonlinear Pricing, Wireless Service, Three-Part Tariffs, Switching Cost, Learning, Discrete/Continuous Model
1. INTRODUCTION

A service plan in industries such as wireless typically consists of two important components: (i) it is a bundle of multiple services and, (ii) the marginal price for usage often varies with the level of usage, i.e. the pricing scheme is nonlinear. The pricing structure under a commonly used three-part tariff with bundling includes an access fee, free usages (allowances) for multiple services, and a marginal price for additional usage above the free usage for each service. For example, under a service plan of two services 1 and 2, a consumer pays an access fee of $A$ per month that entails him to free usages $F_1$ and $F_2$. Beyond these free usages he will have to pay marginal prices of $p_1$ and $p_2$. If the consumer’s usage of service 1 ($q_1$) is larger than $F_1$ but his usage of service 2 is less than $F_2$, he has to pay an amount of $A + p_1(q_1 - F_1)$. If the consumer’s usages of both service 1 and 2 ($q_1$ and $q_2$, respectively) are both higher than free usage level, he has to pay for an amount of $A + p_1(q_1 - F_1) + p_2(q_2 - F_2)$. Since the access fee $A$ involves the usage of both services the consumer has to consider his expected usage of the two services when he makes service plan choice decision.

Prior research has examined how a monopolist could use bundling as a second-degree price discrimination device to extract greater consumer surplus. Adams and Yellen (1976) show that bundling can increase a monopolist’s profit when consumer preferences for the two products are negatively correlated and marginal costs are not too high. Subsequent studies (for example, see Schmalensee (1984) and McAfee, McMillan, and Whinston (1989) etc.) find that the negative correlation is not a necessary condition. Crawford (2005) tests the price discrimination hypothesis using data from the cable television industry, and finds that bundling can lead to six percent increase in firms’ profits and more than five percent reduction in consumer surplus.

Common practice of three-part tariff pricing in industries has also generated many studies of this practice. Jensen (2006) finds that in a differentiated product duopoly two-part tariff may not be an equilibrium strategy and shows that three-part tariff is an equilibrium pricing strategy. Grubb (2006) shows that when consumers
are over-confident in predicting their future usage, three-part tariff is an equilibrium. A few empirical studies have investigated the impacts of three-part tariff when firms offer a single product or service. For instance, Iyengar (2005) shows that the access fee affects consumer churn more than the marginal price. Lambrecht, Seim and Skiera (2006) find that demand uncertainty drives consumers' choice among three-part tariffs. They also find that access fee has a larger impact on choice of a plan than the marginal price and free usage.

We are not aware of any theoretical or empirical studies of three-part tariffs in a world where firms offer bundles of products or services. Since bundling and three-part tariff pricing have become more and more popular in many industries it is important for firms to understand the impact on consumer choice when these strategies are combined. Unfortunately, many results from previous literature cannot be applied to this situation. For instance, majority of the prior literature on product bundling assumes discrete demand, where consumers can only choose either zero or one unit of each product. This assumption comes at a cost of precluding the study of three-part tariffs that applies only when consumers choose multiple units of products or services (Mitchell and Vogelsang (1991)). Similarly, previous research on three-part tariffs only focuses on single product or service case. When firms offer multiple plans (e.g., voice call and SMS text message in our data) consumers’ choice of a plan then would depend on their expected level of usages of these services. Taking account of the potential correlation in preferences among the services and hence the impact of changing the pricing structure of one service on the other is important for a firm's pricing strategy to effectively discriminate among consumers. To summarize, we study consumer choices under bundling and three part tariffs. Towards this objective, we develop a structural demand model to explain consumers’ choice of a bundle and usage decisions of the bundled services. The model explicitly allows for continuous demand for each of the services and correlation in preferences between the services. Hence, this paper makes a contribution in filling out the gap in the consumer choice literature by considering the simultaneous impacts of bundling and three-part tariffs, a combination that is a popular strategy in many industries.
We use a panel dataset provided by a wireless company in a Chinese city. Consumers are not allowed to choose voice call and text message services separately; instead, they have to choose from a menu of multiple service plans offered by the company. A unique feature in our data is that a new service plan was introduced in the middle of the sample period. The new service plan and an existing service plan offered identical services but were under different pricing schemes. During the observational period, many joined in as new subscribers, some switched between the two service plans, and some switched out of the services provided by the company. In order to understand these observed dynamic switching patterns, we incorporate three important components in our model:

First, consistent with previous research, we explicitly account for the time lag between service plan choice decision and usage decision. At the time of choosing a service plan consumers form expectations about their usage levels of voice call and text message in the coming month. However, the actual usage can be different from these expectations. It is important to understand how the usage uncertainty affects a consumer’s choice of service plan (for example see Train, Ben-Akiva and Atherton 1989, Narayanan, Chintagunta and Miravete 2006 and Lambrecht, Seim and Skiera 2006).

Second, we model the switching cost consumers incur when they switch. It is widely recognized in the literature that consumer switching cost is one of firms' important strategic elements to retain consumers (Porter 1980, 1985; Klemperer 1995, Lieberman and Montegomery 1988, Kolter 1997 etc.). Moshkin and Shachar (2002) empirically find that switching cost can account for stickiness in plan choice. Although there is no out of pocket switching cost charged by the firm in our data, implicit switching cost may exist. For example, Iyengar (2005) finds that consumers may not choose the best service plan or switch to the best plan due to inertia. Epling (2002) shows that inherent switching cost is an important factor in subscribers’ switching decision to another service provider. In terms of switching among different service plans offered by the same firm, consumers may incur search and transaction costs (e.g. the time and effort to arrange for the switching). For firms the existence of switching cost has implications for customer relationship
management. For example, one way to increase switching cost is through loyalty programs (Kotler 1997). Regulators often make inference about market power based on the existence of significant switching costs.

Third, our model allows for the fact that consumers may be uncertain about their true preferences for voice call and text message. A learning model is embedded in our choice model where consumers will gradually learn about their preferences. Many previous empirical studies have examined consumer learning (for examples see Erdem and Keane (1996), Ackerberg (2002), Crawford and Shum (2002), Coscelli and Shum (2004), and Narayanan, Chintagunta and Miravete (2006)). Ignoring such consumer learning we find that our choice model cannot explain the penetration process in our data after the new service plan was introduced. We also find from our data that some consumers stay on the same plan for the entire period even when they would have been better off switching to the new plan, while others switch. Hence, we believe that it is important to incorporate both switching costs and consumer learning in explaining these differential switching behaviors.

Thus, while others have studied learning and switching cost on its own, we consider them simultaneously in consumers’ choices.

To conclude, our research objectives are the following:

- What are the preferences of voice call and text message, and what is the correlation between the two? Answers to these questions have important implications for the firm in its segmentation and pricing strategy.
- How do access fee, free usage and marginal price under service bundling affect consumers choice of service plan and usage decisions? Answer to this question has implications for a firm’s design of the pricing and bundling strategies.
- How do switching costs and consumer learning impact the decisions of switching among services plans? Answers to these questions help to understand the major factors driving the penetration process when a new product is introduced.
From our estimation we find a somewhat surprising result that preferences for voice calls and text messages are positively correlated, i.e., consumers who make more calls are also more likely to send out more text messages. The implication for customer segmentation is that a firm is more likely to find customers with high or low preferences for both of the services. Results show that access fee affects the choice probabilities and firm revenue more than marginal prices and free usage. Moreover, change in marginal price or free usage of one service will impact the aggregate usage level of another service, demonstrating a complementary relationship between services that are bundled together in a service plan. We also find that

- both voice call and text message have important weights in the consumer utility function; hence, ignoring either of them (e.g., by focusing only on voice calls) may lead to biased results in the choice model;
- switching costs are important in explaining why some consumers choose not switch to the new service plan though they would be benefited had they switched;
- consumer learning explains the gradual growth in market share of the new service plan.

An out-of-sample validation exercise confirms that our proposed model outperforms competing models in terms of predictability. We further conduct some “what-ifs” experiments to illustrate how the product penetration process will change under different assumptions of switching costs and consumer uncertainty about their preferences. These results help managers understand the underlying factors which determine the dynamic patterns of consumer service plan choices.

The rest of the paper is organized as follows. In section 2 we discuss the bundling and pricing practices in the wireless service industry and the data. Section 3 presents our structural demand model. In section 4 we first discuss the details of the model estimation and then the identification issues. In section 5 we present the estimation results and discuss other experimental outcomes. Section 6 presents our conclusion and discussions on future research.
2. DATA

In the context of the Chinese provider, a wireless service plan offers local, long distance including international calls, text, GPRS\(^2\), and other features such as call waiting, caller ID, three-way calling, call forwarding, and coloring ring back tones. Some sub-bundles also exist in each of these service components. For example, a voice call ("voice" hereon) is a bundle of on-net and off-net calls.\(^3\) Text messages ("text" hereon) are also bundles of on-net and off-net text messages. To simplify the analysis, we only focus on the two most popular services, local calls and text messages, in this study. Consumers are required to pay a monthly access fee with fixed quantities of free usages (allowances) for voice and text, followed by a positive marginal price once the usage exceeds the free allowance.

Our dataset contains a sample of 2,357 consumers from a Chinese city. The dataset consists of monthly service plan choices and usage levels for local voice and text from July 2003 to September 2005. Numerous service plans with different pricing schemes are offered by the firm. In this study we focus on the two most popular wireless service plans: the "voice centric" and "data centric" plan. Table 1 provides details of the two plans. The voice centric plan charges different access fees to users based on the availability of roaming\(^4\) (if a consumer on the plan uses roaming service in a particular month, he will pay ¥30\(^5\) as monthly access fee, otherwise he pays only ¥15 as monthly access fee), while the data centric plan allows roaming at a single access fee (¥20). The voice centric plan allows a higher free voice usage but no free text usage, while the data centric plan allows 300 free text a month. Beyond the allowances, marginal prices for both services are similar, except that the data centric plan charges a lower price for off-net outgoing calls.

---

\(^2\) GPRS (General Packet Radio Service) is an emerging technology standard for high speed data transmission over GSM networks. (www.wirelessadvisor.com)

\(^3\) An on-net local call is a call that originates and terminates in the network provided by the same company, and an off-net call is a call that originates and terminates in networks provided by different firms. Since costs may be different between on- and off-net calls, a service provider with a larger network of subscribers usually has a competitive advantage in the market.

\(^4\) We do not observe the marginal price charged for roaming in our data (prices may vary a lot depending where to call); hence, we only model the usage decisions of local phone calls in this paper. We use the average access fees of roaming and non-roaming weighted by their relative market share as the access fee of the voice centric plan.

\(^5\) Chinese dollar ¥1 is approximately equal to $0.125 in US dollar.
Off-net incoming and outgoing marginal prices are higher than on-net prices for both voice and text because the firm is charged an access fee or termination fee by local land-line phone firms or other wireless firms. Since we only have data on the total minutes of voice and number of text, we use the weighted average\(^6\) of on- and off-net prices as the marginal prices for voice and text.

Two major wireless service providers co-exist in our city. The focal firm in our data (Firm A) is the market leader with greater than 60% share during the sample period (see Figure 1). The other firm (Firm B) has a market share between 20 and 30 percent. The third firm (Firm C) is the land-line provider in the city. Traditionally this land-line provider is not a competitor in the wireless market; however, it introduced a wireless fixed line service in March 2004. By subscribing to this service, consumers can carry their phones like a cell phone. This new service proved to be very successful, with the firm’s share of the wireless market increasing to about 12 percent by the end of the sample period. In response to this new threat, firm A introduced in August 2004 a new data centric plan targeting consumers with high demand for text. This soon became a popular plan and firm A’s market share has rebound since then. As shown in Figure 2, the new plan gained new users quickly after its introduction. By the end of the sample period the number of users in this plan almost equaled the existing voice centric plan.\(^7\) To explain the growth of the data-centric plan we need to model the choices of not only the existing users but also new consumers. Therefore, unlike most of the previous studies (e.g., Lambrecht, Seim and Skiera (2006) do not consider new consumers during the sample period, and Narayanan, Chintagunta and Miravete (2006) only consider switchers between plans), it is important for us to model the join-in decisions of those users who originally chose the outside option.\(^8\)

---

\(^6\) We assume a balanced calling pattern, i.e., when a consumer makes a telephone call, the receiver of the call can be any other consumer with probability independent of his current subscribed network.

\(^7\) Because of the growing acceptance of cell phone, total market demand in the city grew over time during the sample period. As a result, number of users for both plans and for the three firms in the city also grew overtime.

\(^8\) The outside option includes the choice of not using any wireless services and the choice of using other services from competing firms. Since we do not observe these outside choices from data, they are not separately identified in our model. More details will be provided in the model section.
To understand the difference in usage among consumers, we break down the sample into (i) those who stay on the voice centric plan and do not switch, (ii) those who join the new data centric plan and then stay, (iii) those who switch from the voice centric plan to data centric plan, (iv) those who join the data centric plan and then switch to the voice centric plan, and (v) those who drop out. Table 2 reports some summary statistics of the usage patterns for both voice and text for the whole sample as well as for these five segments. While the average voice usage is not significantly different across groups of users the average text usage for those who stay on the voice centric plan is significantly lower than the other groups. The level of voice usage of switchers from data centric to voice centric plan and drop-outs are also lower than those who stay with or switch from voice centric to data centric plan. Nevertheless, we note that usage levels are endogenous depending on the amount of free usages and marginal prices. Usage differences may reflect the difference in either inherent consumer preferences or prices or both.

To investigate whether consumers change their usage patterns and whether they gain from switching, we examine the switchers. We compute the expenditure under the voice centric and data centric plans based on the observed quantity. There are 131 users that switch from the voice to the data centric plan. Based on the level of usage after they switch, 78 percent gain from switching with an average saving of ¥16.2, and 22 percent are worse off from switching with an average loss of ¥3.8. The average cost saving is ¥11.7 among all users. This provides an approximate upper bound on the savings of switching from voice centric to data centric plan. Based on the level of usage before they switch, 60 percent of switchers gain with an average saving of ¥10.5, and 40 percent lose from switching with an average loss of ¥3.85. The average cost saving is ¥4.69. This can be viewed as a lower bound on the savings of switching to the data centric plan. Figure 3 illustrates the patterns of average gain and loss of switchers from the voice centric to the data centric plan. Similar patterns of gain and loss are found among those switching from the data to the voice centric plan. The results imply that consumers are on average making optimal decisions. Moreover, consumers will
change their usage patterns once they switch to another plan with different prices; hence, we observe the difference between upper and lower bounds.

We want to highlight two issues related to the switching patterns. First looking at Figure 2 we see that the market share of the new data centric plan increases gradually in the 12 months after it was introduced. This is consistent with consumers learning their preferences gradually hence switching over time. To accommodate this we propose a consumer learning model that could explain this dynamic pattern of switching. Second, though we find that most switches seem to be optimal, there are still a significant proportion of consumers that choose to stay with either voice centric or data centric plan but could have benefited if they had switched. For example, among the 965 consumers who choose to stay with the voice centric plan, 25 percent of them would have benefited had they switched to the data centric plan, with an average saving of ¥5.04. Similarly, among those 728 consumers who join and stay with the data centric plan until the end of sample period, about 15 percent of them would have been better off had they switched to the voice centric plan, with a saving of ¥4.79. Moreover, there are some switchers incurring losses by switching but choose to stay with the new plan. Why don’t these consumers switch back to the original plan? A consumer learning story may not be sufficient to explain the fact that so many users choose to stay with the existing plan after 12 months of the introduction of the new plan. An alternative story that the firm may put in different marketing efforts for the two service plans during the sample period cannot explain why a significant proportion of consumers from both plans do not switch to another less costly plan. We hypothesize that consumers incur switching costs to switch among plans or to drop out to better account for this behavior.

We randomly select about 25% consumers from the 2,357 consumers for our model estimation. This selected sample consists of 6,774 monthly observations with a 55-45 split between the voice centric and data centric plans. We compare the usage and switching patterns of this selected sample and found them to be very similar to that of the whole sample.
3. THE MODEL

In this section we develop a structural model to explain consumers’ choice of service plan and usage under three-part tariff pricing. Consumers maximize expected utility and we specify a direct utility function over voice and text usage. Allowing for consumer preference heterogeneity we derive consumer usage decisions conditional on the choice of a service plan. Based on these we can derive the indirect utility function of service plan choice and probability of choosing the different service plans when consumers form expectations about their usages. We explain in detail how switching costs and consumer learning are incorporated in our model. Finally we discuss identification issues in model estimation.

3.1 The Utility Function

We model a consumer’s decisions using an integrated framework: consumer first chooses from among many service plans including an outside option and then, conditional on the choice of a plan, makes the usage decisions for voice and text. While the service plan choice is a discrete decision, usage decisions are continuous. This set-up is analogous to the discrete/continuous demand model (for example see Hanemann (1984) and Dubin and McFadden (1984)). For recent empirical work see Hendel (1999), Kim, Allenby and Rossi (2002), Dube (2004) and Chan (2005).) with the difference that we also account for the time lag between the plan choice and usage decisions (e.g., see Iyengar (2005) and Narayanan, Chintagunta and Miravete (2006)), where in each stage consumers’ information sets may be different.

We assume that consumers’ utility is derived from using both voice and text. Consumer \( i, i=1, \ldots, N \), chooses a service plan from the available service plan options at time \( t \). The existing consumer has three options: stay in the same plan as time \( t-1 \), switch to another service plan from the same firm and drop out of the existing service plans (i.e., choose the outside option). On the other hand the new consumer has two options: stay out or sign in for one of the service plans as a new user. If she chooses a service plan, indexed by \( j = 1, \ldots, J \), from the focal firm at time \( t \), she will then choose the number of voice minutes \( x_{it}^V \), the number of text \( x_{it}^D \), and quantity of the outside good \( x_{it}^0 \) which is the consumption of products and
services other than the wireless services. To consume a wireless service bundle \( \{x^V, x^D\} \) from service plan \( j \), the consumer pays an access fee \( A_j \), enjoys a free usage for voice \( F^V_j \) and for text \( F^D_j \), and then pays a marginal price for voice \( p^V_j \) if \( x^V > F^V_j \), and for text \( p^D_j \) if \( x^D > F^D_j \). For model estimation we normalize the price of the outside good to 1. We allow for the case that the consumption of either voice or text may be zero in some periods, i.e., corner solutions for usages could exist.

We assume that a consumer’s utility is additively separable in voice and text.\(^9\) We define the utility when consumer \( i \) chooses a service plan \( j \) as:

\[
U^i_j(x^V_{it}, x^D_{it}, x^D_{it}) = \delta_j + x^0_{it} + \left[ \omega^V_i \beta^V_i x^V_{it} - \beta^V_i \left( \frac{x^V_{it}}{2} \right) \right] + \left[ \omega^D_i \beta^D_i x^D_{it} - \beta^D_i \left( \frac{x^D_{it}}{2} \right) \right] + \epsilon_{ijt} (1)
\]

where \( U^i_j(\cdot) \) is the consumer’s direct utility function, and \( \delta_j \) is a plan-specific preference intercept representing benefits other than voice and text offered by the plan that are not in our data. For simplicity we assume that \( \delta_j \) is homogeneous among consumers. The marginal utility derived from the outside good \( x^0_{it} \) is normalized to 1. The first square bracketed term is utility from voice and the second square bracketed term is utility from text. Note that this additive separable structure implies that there is neither complementarity nor substitutability between voice and text at the individual level, and the utilities from using voice and text are independent of the utility from consuming the outside good. Parameters to be estimated in (1) include \( \delta_j, \theta^V_i, \theta^D_i, \beta^V_i \) and \( \beta^D_i \), where the last two parameters are restricted to be positive. The last component in (1) \( \epsilon_{ijt} \) is the individual-, plan- and

\(^9\) We note that such additive separability assumption may restrict the substitution pattern among usages of the two services. A more flexible specification is to allow an interaction term between the voice and text usages in the utility function. However, since there is no price variation in either plan during our sample period, it is not possible to identify this from our data. Furthermore, as will be discussed below, we allow a consumer’s preferences for voices and texts to be correlated. Therefore our model is able to generate a flexible substitution pattern between the two service plans at the aggregate level.
period-specific \textit{i.i.d.} random shock that follows a double exponential distribution which affects the consumer’s choice of service plan but unobserved by researchers. The quadratic sub-utility functions with positive $\beta_i^V$ and $\beta_i^D$ imply that consumers are risk averse and there is a satiation point in usage.\footnote{This quadratic utility specification is consistent with previous literature such as Wilson (1992), Miravete (2002), Iyengar (2005) and Economides, Seim and Viard (2005) etc.}

3.2 Preference Heterogeneity

We assume that a consumer's preference parameter for using voice and text, $\theta_{it}^L, L = \{V, D\}$, has two components. First there is an individual-specific and time-invariant preference component $\theta_i^L$, and then an individual- and time-specific \textit{i.i.d.} unobserved preference shock $\xi_{it}^L$. That is,

$$\theta_{it}^L = \theta_i^L + \xi_{it}^L, L = \{V, D\}$$

Let $\theta_i = \begin{pmatrix} \theta_i^V \\ \theta_i^D \end{pmatrix}$. We assume that $\theta_i \sim \text{normal}\left(\begin{pmatrix} \bar{\theta}_i^V \\ \bar{\theta}_i^D \end{pmatrix}, \Sigma_\theta \right)$, and $\Sigma_\theta = \begin{bmatrix} \sigma_{V}^2 & \sigma_{VD} \\ \sigma_{VD} & \sigma_{D}^2 \end{bmatrix}$.

Consumer preferences for voice and texts are correlated when $\sigma_{VD} \neq 0$. We also assume the time-varying preference shocks $\xi_{it} = \begin{pmatrix} \xi_{it}^V \\ \xi_{it}^D \end{pmatrix} \sim \text{i.i.d. normal}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_\xi \right)$,

where $\Sigma_\xi = \begin{bmatrix} \sigma_{V}^2 & \sigma_{VD} \\ \sigma_{VD} & \sigma_{D}^2 \end{bmatrix}$. This allows for the correlation in time-varying preference shocks for using voice and text.

Finally, as we will show below, $1/\beta_i^V$ and $1/\beta_i^D$ are marginal prices effect on voice and text usage, respectively. Because there is no price variation within each service plan in our data, consumer heterogeneity of usage responsiveness cannot be identified. We impose the homogeneity assumption that $\beta_i^V = \beta^V$ and $\beta_i^D = \beta^D$ for all consumers.

3.3 Usage Decisions

We start with modeling a consumer’s usage decisions for voice and text, conditional on choosing a service plan $j$. We assume that at this stage the
consumer’s preferences for voice and text $\theta^V_n$ and $\theta^D_n$ in (2) are realized. This is in contrast to the consumer’s decision at the service plan choice stage when consumers have expectations about their preferences. However, the consumer cannot infer the time-invariant preference component $\theta^t_i$ from the time-specific preference shock $\xi_{it}$. That is, she only knows the sum of preferences in period $t$ and therefore does not know how this changes over time. We assume that the consumer maximizes her utility subject to a budget constraint. The direct utility maximization problem is specified as

$$
\max_{x^0_n, x^V_n, x^D_n} U^i_j\left(x^0_n, x^V_n, x^D_n \mid d_n = j\right)
$$

s.t.

$$
x^0_n + \left[p^V_j \cdot (x^V_n - F^V_j)\right] \left\{x^V_n \geq F^V_j\right\} + \left[p^D_j \cdot (x^D_n - F^D_j)\right] \left\{x^D_n \geq F^D_j\right\} + A_j \leq Y_i
$$

where $U^i_j(\cdot)$ is consumer $i$'s direct utility in (1) conditional on choosing plan $j$, $d_n$ is the discrete service plan choice at time $t$, and $\left[x^0_n, x^V_n, x^D_n\right]$ are the endogenous usage decisions for the outside good, voice, and text, respectively. $F^V_j$ and $F^D_j$ are the corresponding free usages for voice and text, and $A_j$ is the access fee, $p^V_j$ and $p^D_j$ are marginal prices for voice and text, and $Y_i$ represents the income of the consumer. Finally $\{\cdot\}$ in the second line in (3) is an indicator function which equals 1 if the logical expression inside is true, and 0 otherwise. The consumer will be charged the marginal prices only if her usage exceeds the number of free minutes or the number of free text in the plan. If she chooses the voice centric plan in our data, there is no free usage for text in the service plan, therefore $F^D_j = 0$. Solving the maximization problem in (3) we have the following optimal usage levels for voice ($x^V_{it}^{*}$) and text ($x^D_{it}^{*}$):

\[\text{(3)}\]

\[\text{Following convention we assume there always exists an interior solution for the outside good consumption, i.e., } x^0_{it}^{*} > 0.\]
for $L=\{V, D\}$. The condition in (4) allows for the existence of corner solutions, i.e.,

$$x_u^L = 0.$$ Moreover, when $0 < \theta_u^L \leq F_j^L$, $x_u^L = \theta_u^L$; and when $\theta_u^L > F_j^L + \frac{1}{\beta_u^L} \cdot p_j^L$,

$$x_u^L = \theta_u^L - \frac{1}{\beta_u^L} \cdot p_j^L.$$ There is a sticky point in (4) when $F_j^L < \theta_u^L \leq F_j^L + \frac{1}{\beta_u^L} \cdot p_j^L$,

under which condition the usage amount will be at the free usage level when $\theta_u^L$ changes.

### 3.4 Service Plan Choices

Based on the direct utility function in (1) and the objective function in (3), we can plug in the optimal usages in (4) to derive an indirect utility function of choosing service plan $j$ as the follows:

$$V_{j,u}(A_j; F_j^V, F_j^D, p_j^V, p_j^D; Y)$$

$$= \delta_j + [Y_i - A_j] + \left(\frac{\theta_j^V}{2}\right)^2 \beta^\nu \cdot \{\theta_j^V > 0\} + \left(\frac{\theta_j^D}{2}\right)^2 \beta^\rho \cdot \{\theta_j^D > 0\}$$

$$+ \left[\theta_j^V \beta^\nu F_j^V - \left(\frac{F_j^V}{2}\right)^2 \beta^\nu \theta_j^V + \left(\frac{\theta_j^V}{2}\right)^2 \beta^\nu \right] \cdot \{F_j^V < \theta_j^V \leq F_j^V + \frac{1}{\beta_j^V} \cdot p_j^V\}$$

$$+ \left[\theta_j^D \beta^\rho F_j^D - \left(\frac{F_j^D}{2}\right)^2 \beta^\rho \theta_j^D + \left(\frac{\theta_j^D}{2}\right)^2 \beta^\rho \right] \cdot \{F_j^D < \theta_j^D \leq F_j^D + \frac{1}{\beta_j^D} \cdot p_j^D\}$$

$$+ \left[-\theta_j^V p_j^V + \frac{1}{2} \beta_j^V \left(p_j^V \right)^2 + p_j^V F_j^V\right] \cdot \{\theta_j^V > F_j^V + \frac{1}{\beta_j^V} \cdot p_j^V\}$$

$$+ \left[-\theta_j^D p_j^D + \frac{1}{2} \beta_j^D \left(p_j^D \right)^2 + p_j^D F_j^D\right] \cdot \{\theta_j^D > F_j^D + \frac{1}{\beta_j^D} \cdot p_j^D\}$$

$$+ \epsilon_{yi}$$
Interpretations of equation (5) are as follows: If both \( \theta^v_i < 0 \) and \( \theta^d_i < 0 \), the indirect utility function will be \( V_{j,it} = \delta_j + \left[ Y_i - A_j \right] + \epsilon_{ijt} \); if both \( 0 < \theta^v_i \leq F^v_j \) and \( 0 < \theta^d_i \leq F^d_j \), we will have \( V_{j,it} = \delta_j + \left[ Y_i - A_j \right] + \frac{\left( \theta^v_i \right)^2 \beta^v}{2} + \frac{\left( \theta^d_i \right)^2 \beta^d}{2} + \epsilon_{ijt} \);
if both \( F^v_j < \theta^v_i \leq F^v_j + \frac{1}{\beta^v} \cdot p^v_j \), \( F^d_j < \theta^d_i \leq F^d_j + \frac{1}{\beta^d} \cdot p^d_j \), we will have
\[
V_{j,it} = \delta_j + \left[ Y_i - A_j \right] + \left[ \frac{\left( \theta^v_i \right)^2 \beta^v}{2} - \theta^v_i p^v_j + \frac{1}{2 \beta^v} \left( p^v_j \right)^2 + p^v_j F^v_j \right] + \epsilon_{ijt};
\]
finally if both \( \theta^v_i > F^v_j + \frac{1}{\beta^v} \cdot p^v_j \) and \( \theta^d_i > F^d_j + \frac{1}{\beta^d} \cdot p^d_j \), we will have
\[
V_{j,it} = \delta_j + \left[ Y_i - A_j \right] + \left[ \frac{\left( \theta^v_i \right)^2 \beta^v}{2} - \theta^v_i p^v_j + \frac{1}{2 \beta^v} \left( p^v_j \right)^2 + p^v_j F^v_j \right] + \epsilon_{ijt}.
\]

The indirect utility function of choosing service plan \( j \) depends on the consumer preferences for using both voice and text.

If the consumer chooses the outside option, the indirect utility function is
\[
V_{0,it} = Y_i + \epsilon_{0it} \quad (6)
\]
We assume that \( \epsilon_{0it} \) is independently and identically distributed with double exponential distribution.

The consumer makes the service plan choice at the beginning of each period \( t \). As opposed to the usage decisions in (4), she does not exactly know the value of \( \theta^l_i \), where \( L=V \) or \( D \) (see equation (2)), which consists of the time-varying idiosyncratic demand shock \( \varepsilon^l_i \) of which the consumer only knows the distribution, and the time-invariant usage preference \( \theta^l_i \) that she may only learn over time (we will further discuss the learning model later). The consumer has to form an expectation for \( V_{j,it} \) conditional on her information set \( \Omega_i \) which consists of her
past usage history, i.e., \( E[V_{j,t} \mid \Omega_{jt}] \). The consumer will choose the option with the highest expected indirect utility.

The new data centric plan was introduced in the middle of the sample period. Before the new plan is available, consumers’ consideration set consists of the existing voice centric plan and the outside option. After the new plan is available, consumers’ consideration set now consists of the voice centric plan, the data centric plan and the outside option.\(^\text{12}\)

### 3.5 Switching Costs and Learning

To explain the dynamic switching patterns observed among consumers, we allow for the existence of switching costs and consumer learning of their own preferences. First, we assume that consumer \( i \) incurs switching cost, \( SC_i^1 \), when she switches to a different service plan provided by the same firm and by assuming that the switching cost is distributed as \( N(SC^1, \sigma_{sc}^2) \) we allow this cost to vary across consumers. To identify the parameters of the model, we do not allow this switching cost to be time-varying. However, one may interpret the fluctuation in switching cost in each period as part of the time-specific i.i.d. shock \( \epsilon_{jt} \) (see equation (1) and (5)). If the consumer decides to drop out as a user of the firm, or join in from outside, another type of switching cost would incur which may be different from switching between service plans within a firm. In our estimation we find that allowing for heterogeneity of this cost leads to difficulty in identification, perhaps due to the lack of variation in data among consumers within these categories. Therefore we restrict this cost to be homogeneous across consumers as \( SC^2 \).

Suppose the consumer is a user of the voice centric plan (plan 1) in period \( t-1 \). Based on the above discussion, she will again choose the same plan in period \( t \), instead of the data centric plan (plan 2) or the outside option, if the following is true:

\[
E[V_{1,jt} \mid \Omega_{jt}] \geq \max \left\{ E[V_{2,jt} \mid \Omega_{jt}] - SC_i^1, V_{0,jt} - SC^2 \right\},
\]

\(^{12}\) We ignore in our model that some consumers may not be aware of the new service plan. Though it may take time for consumers to learn the existence of the service plan, such a learning process is difficult to be identified from our data, especially that it cannot be distinguished from the process of consumers learning own usage preferences, which is modeled in the paper.
Based on the double exponential distribution assumption for plan choice shocks \( (e_{i1}, e_{i2}, e_{i0}) \) the choice probability for service plan 1 in period \( t \) is

\[
\text{Pr}_i(d_{it} = 1) = \text{Pr}\left\{ (e_{i1}, e_{i2}, e_{i0}) : E\tilde{V}_{1,it} + e_{i1} \geq \max( E\tilde{V}_{2,it} - SC^i_i + e_{i2}, V_{i0} - SC^2 + e_{i0}) \right\} \\
= \frac{\exp\left( E\tilde{V}_{1,it} \right)}{\exp(V_{i0} - SC^2) + \exp\left( E\tilde{V}_{1,it} \right) + \exp\left( E\tilde{V}_{2,it} - SC^i_i \right)}
\]

(7)

where \( d_{it} \) is the discrete service plan choice of consumer \( i \) in period \( t \), \( E\tilde{V}_{j,it} \) is the deterministic part in \( E[V_{j,it} | \Omega_{it}] \) without the idiosyncratic component \( \epsilon_{j,it} \), and \( V_{i0} \) is similarly defined.

Similarly, the consumer will switch to the data centric plan if the following condition is true:

\[
E[V_{2,it} | \Omega_{it}] - SC^i_i \geq \max\left\{ E[V_{1,it} | \Omega_{it}], V_{0,it} - SC^2 \right\}
\]

And she will choose to drop out if

\[
V_{0,it} - SC^2 \geq \max\left\{ E[V_{1,it} | \Omega_{it}], E[V_{2,it} | \Omega_{it}] - SC^i_i \right\}
\]

with choice probabilities similar to equation (7).

If the consumer is not a user of the firm in period \( t-1 \), by joining either service plan it will incur a switching cost \( SC^2 \). Hence, she will only choose the voice centric or the data centric plan if either

\[
E[V_{1,it} | \Omega_{it}] - SC^2 \geq \max\left\{ E[V_{2,it} | \Omega_{it}] - SC^2, V_{0,it} \right\}
\]

or

\[
E[V_{2,it} | \Omega_{it}] - SC^2 \geq \max\left\{ E[V_{1,it} | \Omega_{it}] - SC^2, V_{0,it} \right\}
\]

is true, respectively. Based on that, the probability of joining either plan can be derived similarly as equation (7). We will discuss how we compute \( E\tilde{V}_{j,it} \) in detail later.
Turning to consumer learning and consistent with previous literature on learning, we assume that consumers may not know their own time-invariant preference types \( \{\theta_i^v, \theta_i^d\} \) hence form expectations based on past experiences. This type of learning could be one source in explaining why in our data consumers only switched to the new data centric plan several periods after the plan had been introduced (and some did not switch at all) even when their savings are large had they switched earlier. At the end of each period, consumer \( i \) observes her usages \( x_{it}^v \) and \( x_{it}^d \). Suppose these usages are positive and not exactly equal to the free usages \( F_j^v \) and \( F_j^d \), the optimal conditions of usages imply that

\[
\begin{align*}
\theta_{it}^v &= \theta_{it}^v + \xi_{it}^v = x_{it}^v + \frac{1}{\beta^v} \cdot p_j^v \cdot \left\{ \theta_{it} > F_j^v + \frac{1}{\beta^v} \cdot p_j^v \right\} \\
\theta_{it}^d &= \theta_{it}^d + \xi_{it}^d = x_{it}^d + \frac{1}{\beta^d} \cdot p_j^d \cdot \left\{ \theta_{it} > F_j^d + \frac{1}{\beta^d} \cdot p_j^d \right\}
\end{align*}
\]

(8)

Though there is a one-to-one mapping between usage \( x_{it}^l \) and the time-varying preference \( \theta_{it}^l \) (we assume that the price sensitivity coefficient \( b^l = 1/\beta^l \) is known to the consumer) so that she can use the observed usages to infer preferences, she cannot separate \( \theta_{it}^l \) and \( \xi_{it}^l \) from the sum \( \theta_{it}^l \). As prior knowledge, we assume that the distribution function of \( (\xi_{it}^v, \xi_{it}^d), N(0, \Sigma_\xi) \), is known by all consumers. Regarding \( \theta_{it}^l \), we assume that in the first period \( t=1 \) consumers have the same prior beliefs distributed as \( \begin{pmatrix} \theta_{i1}^v \\ \theta_{i1}^d \end{pmatrix} \sim N\left(\begin{pmatrix} \theta_{0}^v \\ \theta_{0}^d \end{pmatrix}, \Sigma_{\theta_0}\right) \), where \( \begin{pmatrix} \theta_{0}^v \\ \theta_{0}^d \end{pmatrix} \) represents the prior preference means, and \( \Sigma_{\theta_0} = \begin{bmatrix} \sigma_{0}^{v2} & 0 \\ 0 & \sigma_{0}^{d2} \end{bmatrix} \) is the prior variance-covariance matrix which measures the consumer uncertainty.\(^\text{13}\) At the end of every period, after observing their usages, consumers update the beliefs of their true preferences using the Bayesian rule (Degroot (1970)): Assuming after \( t \) periods the

\(^\text{13}\) For the simplicity of model estimation we assume that consumers believe that their preferences for voices and texts are independent. It is generalizable to the case when preferences in prior beliefs are not independent.
consumer’s belief of her preference type for using service $L$ is $N(\theta_{i,t}^L, \Sigma_{\theta,i,j})$, where the subscript "i" denotes consumers’ beliefs of own preference types after period t. Suppose in period $t$ usages of both voice and text are above the free usage levels, then her belief at time $t + 1$ are distributed as $N(\theta_{i,t+1}^L, \Sigma_{\theta,i,j})$ where $[\theta_{i,t+1}^L, \theta_{i,t+1}^D]'$ and $\Sigma_{\theta,i,j}$ follow the specifications below:

$$
\begin{pmatrix}
\theta_{i,t+1}^V \\
\theta_{i,t+1}^D
\end{pmatrix} = \Sigma_{\theta,i,j}^{-1} \begin{pmatrix}
\begin{pmatrix} x_{it}^V + \frac{1}{\beta^V} \cdot p_{j}^V \\
 x_{it}^D + \frac{1}{\beta^D} \cdot p_{j}^D
\end{pmatrix} + \begin{pmatrix}
\theta_{i,t}^V \\
\theta_{i,t}^D
\end{pmatrix}
\end{pmatrix}
\begin{pmatrix}
\Sigma_{\theta,i,j}^{-1}
\end{pmatrix}^{-1}
$$

(9)

If usages are below free usages, the terms $\frac{1}{\beta^V} \cdot p_{j}^V$ and $\frac{1}{\beta^D} \cdot p_{j}^D$ will not be in the above equation. Zero usage, though infrequent, does exist for either voice or text in the data. In this case the one-to-one relationship between usages and preferences does not exist. We assume that, for example, when $x_{it}^V = 0$, the consumer can only infer that $\theta_{i,t}^V \leq 0$ at time $t$. The updated belief for $\theta_{i,t}^V$ will follow the prior belief distribution at time $t$ conditional on the truncated usage, i.e.,

$$
N(\theta_{i,t}^V, \sigma_{i,t}^{V,2} | \theta_{i,t}^V + \xi_{it}^V \leq 0)
$$

(10)

where $\theta_{i,t}^V$ and $\sigma_{i,t}^{V,2}$ are the mean and variance in prior beliefs for $\theta_{i,t}^V$ at time $t$, respectively. There is no closed-form expression for these posterior beliefs, so in model estimation we simulate this conditional distribution to obtain the estimates for $\theta_{i,t+1}^V$ and $\sigma_{i,t+1}^{V,2}$. We do similar simulations for $\theta_{i,t+1}^D$ and $\sigma_{i,t+1}^{D,2}$ when $x_{it}^D = 0$.

### 4 MODEL ESTIMATION

There is another complication in the updating process when the consumer’s usages are exactly equal to $F_j^V$ or $F_j^D$, in which case $\theta_{i}^V$ and $\theta_{i}^D$ are defined within a range. This never happens in our data; hence, we can ignore such updating process in our model estimation.
4.1 Estimation Procedure

We first discuss the estimation procedure for the usage decisions, conditional on service plan choice. The optimal usages of voice and text in equation (4) imply a Tobit-type regression: if the observed usages \( x_{it}^L, L = V \) or \( D \), is positive, we can derive that \( \theta_i^L + \xi_{it}^L = x_{it}^L + \frac{1}{\beta_i^V} \cdot p_{ij}^L \cdot \{ x_{it}^L > F_{ij}^L \} \). If \( x_{it}^L = 0 \), we can infer that \( \theta_i^L + \xi_{it}^L \leq 0 \).\(^{15}\) Conditional on the service plan choice \( d_{it} \), we can write down the probability function of the observed usages \((x_{it}^V, x_{it}^D)\) as

\[
\Pr(x_{it}^V, x_{it}^D | d_{it}) = \int \Pr(x_{it}^V, x_{it}^D | d_{it}, \theta_{it}^V, \theta_{it}^D) dF(\theta_{it}^V, \theta_{it}^D),
\]

where \( F(\cdot) \) is the joint distribution function of \((\theta_{it}^V, \theta_{it}^D)\). Note that \( \theta_{it}^L \) here is the true preference of consumer \( i \) and not the prior belief because the true value has been realized in the usage decision stage. Since \( \theta_{it}^L = \theta_i^L + \xi_{it}^L \) and \( \theta_i^L \) is time-invariant for each consumer \( i \), \( \theta_{it}^L \) will be correlated overtime. Further, we have to account for the correlation between preferences for voice and text in \( \theta_i \) and \( \xi_{it} \). Direct evaluation of the likelihood of the consumer’s whole usage history of both voice and text will be difficult in model estimation. Instead, we simulate the unconditional probability function of usages by using a frequency simulator for \[ \left( \theta_{i}^V, \theta_{i}^D \right) \]: Let \( s \) represents a draw in the simulation. We draw, for each consumer \( i \), \[ \left( \theta_{i}^V,s, \theta_{i}^D,s \right) \] from the assumed population distribution \( N\left( \bar{\theta}_V, \Sigma_v \right) \) for \( ns \) times and fixed these simulated draws over time for each consumer. Because of our \( i.i.d. \) assumption for \( \xi_{it} \), the likelihood of the whole observed history of usages for consumer \( i \) can be evaluated for each period \( t, t=1,\ldots, T \), as the follows:

\(^{15}\) Again we ignore the case that the consumer’s usages are exactly equal to \( F_{ij}^V \) or \( F_{ij}^D \), which we do not observe in our data.
\[
\Pr_i \left( x_{i1}^y, x_{i1}^d, \ldots, x_{iT}^y, x_{iT}^d, d_i, d_{iT} \right) = \prod_{t=1}^{T} \left[ \Pr \left( x_{it}^y, x_{it}^d, | \theta_{it}^y, \theta_{it}^d, d_i, d_{iT} \right) \right] \\
= \frac{1}{n_s} \sum_{s=1}^{ns} \left[ \Pr(x_{i1}^y, x_{i1}^d | d_i, \theta_{i1}^y, \theta_{i1}^d) \ldots \Pr(x_{iT}^y, x_{iT}^d | d_{iT}, \theta_{iT}^y, \theta_{iT}^d) \right]
\]

where in the first line \( \tilde{\Pr}_i \) represents the simulated version of \( \Pr_i \), and \( \Pr(x_{it}^y, x_{it}^d | d_i, \theta_{it}^y, \theta_{it}^d) \) in the second equality is evaluated conditional on the simulated time-invariant preferences and the distribution assumption \( \xi \sim \text{normal}(0, \Sigma) \).

Let us turn to the likelihood function of service plan choices in each period. Equation (7) provides an example of the choice probability that a consumer chooses service plan \( l \) at time \( t \). All other probability functions are similarly defined. The difficulty in actual model estimation comes from evaluating \( E \tilde{V}_{j,t} \), which is the deterministic part in \( E[V_{j,t} | \Omega_t] \) without the idiosyncratic component \( \epsilon_{jt} \) (also see equation (5)). As opposed to evaluating the probability function for usages above, \( E \tilde{V}_{j,t} \) is now a function of the consumer’s beliefs of \( \left( \theta_i^y, \theta_i^d \right) \) and not her true preferences. Let \( (\tilde{\theta}_{ij}^y, \tilde{\theta}_{ij}^d) \) be a pair of random variables that represent the consumer’s beliefs of her true preference types at time \( t \), with the distribution function \( F(\tilde{\theta}_{ij}^y, \tilde{\theta}_{ij}^d) = N \left( (\theta_i^y, \theta_i^d) | \theta_{ij}, \Sigma_{\theta_{ij}} \right) \), where \( [\theta_i^y, \theta_i^d] \) and \( \Sigma_{\theta_{ij}} \) are the updated means and variances of her preferences, respectively. \( E \tilde{V}_{j,t} \) can be written as the follows:

\[
E \tilde{V}_{j,t} = \delta_j + \left[ Y_t - A_j \right]
\]

\[
\left\{ \left( \tilde{\theta}_{ij}^y + \xi_i^y \right) \beta^y + \left( \tilde{\theta}_{ij}^d + \xi_i^d \right) \beta^d \right\} \left\{ \tilde{\theta}_{ij}^y + \xi_i^y > 0 \right\} + \frac{1}{2} \beta^y \left( \tilde{\theta}_{ij}^y + \xi_i^y \right)^2 + \frac{1}{2} \beta^d \left( \tilde{\theta}_{ij}^d + \xi_i^d \right)^2 \right\} dF(\tilde{\theta}_{ij}^y, \tilde{\theta}_{ij}^d) dF(\xi_i^y, \xi_i^d)
\]

\[
\left\{ \left( \tilde{\theta}_{ij}^y + \xi_i^y \right) p_j^y + \frac{1}{2} \beta^y \left( p_j^y \right)^2 + p_j^y F_j \right\} \left\{ \tilde{\theta}_{ij}^y + \xi_i^y > F_j + \frac{1}{\beta^y} \right\} \left\{ \tilde{\theta}_{ij}^y + \xi_i^y > F_j + \frac{1}{\beta^y} \right\} dF(\tilde{\theta}_{ij}^y, \tilde{\theta}_{ij}^d) dF(\xi_i^y, \xi_i^d)
\]

\[
\left\{ \left( \tilde{\theta}_{ij}^d + \xi_i^d \right) p_j^d + \frac{1}{2} \beta^d \left( p_j^d \right)^2 + p_j^d F_j \right\} \left\{ \tilde{\theta}_{ij}^d + \xi_i^d > F_j + \frac{1}{\beta^d} \right\} \left\{ \tilde{\theta}_{ij}^d + \xi_i^d > F_j + \frac{1}{\beta^d} \right\} dF(\tilde{\theta}_{ij}^y, \tilde{\theta}_{ij}^d) dF(\xi_i^y, \xi_i^d)
\]

23
The above expression, unfortunately, does not have a closed-form expression. To compute this function we again have to rely on the simulation method: we draw 
\((\tilde{\theta}^V_{i,t}, \tilde{\theta}^D_{i,t})\) from the prior belief distribution \(N\left(\left[\theta^V_{i,t}, \theta^D_{i,t}\right], \Sigma_{\theta_{i,t}}\right)\), and \((\tilde{\xi}^V_{it}, \tilde{\xi}^D_{it})\) from the assumed distribution \(N(0, \Sigma_s)\), where “s” in the superscript denotes a simulated draw, for \(ns\) times. Each pair of \((\tilde{\theta}^V_{i,t}, \tilde{\theta}^D_{i,t})\) and \((\tilde{\xi}^V_{it}, \tilde{\xi}^D_{it})\) are plugged into the above equation, then we take the average to form the simulated \(E\tilde{V}^*_j\). Let 
\(\Pr\left(d = j \mid x^V_{i_{t-1}}, x^D_{i_{t-1}}\right)\) be the probability function of consumer \(i\) choosing service plan \(j\) in period \(t\) conditional on past usages \((x^V_{i_{t-1}}, x^D_{i_{t-1}})\) before period \(t\) (which are used to form the prior belief distribution \(N\left(\left[\theta^V_{i,t}, \theta^D_{i,t}\right], \Sigma_{\theta_{i,t}}\right)\)). We plug the simulated \(E\tilde{V}^*_j\) into the multinomial logit function analogous to equation (7) for every service plan \(j\) to become the corresponding simulated probability function 
\(\tilde{\Pr}_i\left(d = j \mid x^V_{i_{t-1}}, x^D_{i_{t-1}}\right)\). This procedure is repeated for every period: in period \(t\) with prior belief distribution \(N\left(\left[\theta^V_{i,t}, \theta^D_{i,t}\right], \Sigma_{\theta_{i,t}}\right)\) the simulated service plan probability function 
\(\tilde{\Pr}_i\left(d = j \mid x^V_{i_{t-1}}, x^D_{i_{t-1}}\right)\) is evaluated. Then the beliefs will be updated to 
\(N\left(\left[\theta^V_{i,t+1}, \theta^D_{i,t+1}\right], \Sigma_{\theta_{i,t+1}}\right)\) after usages \((x^V_{i_{t+1}}, x^D_{i_{t+1}})\) are revealed. We will draw 
\((\tilde{\theta}^V_{i,t+1}, \tilde{\theta}^D_{i,t+1})\) based on this updated distribution to evaluate the simulated service plan probability function 
\(\tilde{\Pr}_i\left(d = j \mid x^V_{i_{t+1}}, x^D_{i_{t+1}}\right)\) for period \(t+1\).

We jointly estimate the likelihoods of both usage and service plan choice decisions. Let \(\Theta\) be the vector of parameters in the model. Our estimator \(\hat{\Theta}\) maximizes the following joint likelihood:
\[ \prod_{i=1}^{N} \{ \Pr_i (x_i^v, x_i^d, d_i; \ldots; x_{iT}^v, x_{iT}^d, d_{iT} | \Theta) \} \]
\[ \approx \prod_{i=1}^{N} \left( \frac{1}{N} \sum_{n=1}^{N} \Pr_i (d_i | x_{i,t-1}, x_{i,t-1}^v; \Theta) \cdot \Pr(x_{i,t-1}, x_{i,t}^v | d_i, \theta_{i,t}, \theta_{i,t}^{D^v}; \Theta) \ldots \right) \]

A few more details about the estimation model as follows: For simplicity we assume that at time \( t=1 \) all consumers have the same prior beliefs \( \left( \tilde{\Theta}^v, \tilde{\Theta}^d \right) \sim N \left( \left( \Theta_0^v, \Theta_0^d \right), \Sigma_\theta_0 \right) \), where \( \Sigma_\theta_0 = \begin{bmatrix} \sigma_0^{v^2} & 0 \\ 0 & \sigma_0^{D^2} \end{bmatrix} \). Since these parameters relate to consumers’ beliefs that are unobserved, they can only be inferred from the service plan choices in particular in the early periods before beliefs are updated. Their exact values are difficult to identify from the data (for related discussion see Chan and Hamilton (2006)). Hence, we restrict them to be equal to \( \tilde{\Theta}^v \) and \( \tilde{\Theta}^d \), the population mean preferences for using voice and text, respectively. We also impose the restriction that \( \sigma_0^{v^2} \) and \( \sigma_0^{D^2} \) are equal to the variances of preferences among the population. Though our model will be mis-specified if such assumptions are invalid, we believe that our estimation results will not be much affected since with consumer learning the impact of these prior beliefs on the consumer choice quickly declines. Chan and Hamilton (2006) also use similar assumptions and find that their estimation results are robust to different assumptions of prior beliefs. Since under this set-up the learning model is restrictive and does not increase the number of parameters in the model, any improvement in terms of fit will clearly indicate the advantage in incorporating consumer learning in explaining the consumers’ service plan choices.

A unique feature in our data is that many new consumers join (especially after the data centric service plan was introduced) and existing consumers drop out in different time periods. Past histories of these consumers before they join or usages after they drop out are not observed from data. In order to evaluate the probability of choosing the outside option we have to make assumptions about these customers’ beliefs of own preferences. In the model estimation, we assume that
these consumers maintain the same prior beliefs \( N\left( \begin{pmatrix} \theta_0^V \\ \theta_0^D \end{pmatrix}, \Sigma_{\theta_0} \right) \) as in period 1 until they join in. Then they will update their beliefs using the observed usages. After they drop out, we assume that they will maintain their updated beliefs right before they leave. Choices of the outside option before consumers join in or after drop out are included in the likelihood function in (11). An interpretation which rationalizes this assumption of constant beliefs is that by choosing the outside option consumers do not use any wireless service hence they do not know what their usages would be had they subscribed to the service plans. We note that this assumption may be restrictive; however, we are not able to infer how beliefs are updated since usages of wireless services during these periods are unobserved.

We estimate three types of model: Model 1 is a base model that does not allow for switching costs or consumer learning. Model 2 allows switching costs but no consumer learning. Model 3 allows for both switching costs and consumer learning. Comparing these three models helps to shed light on the robustness of our estimation results. It is also useful in understanding how much better, by adding switching costs and consumer learning, helps explain the observed patterns of switching.

4.2 Model Identification

Using the data on voice and text usages helps us identify many of the parameters in the utility function. In particular, the average usages over time across consumers identify the mean and variances preferences, \( \bar{\theta}^V, \bar{\theta}^D, \sigma^2_v \) and \( \sigma^2_D \), for voice and text. Moreover, the correlation of observed usages of voice and text across consumers will identify the covariance parameter, \( \sigma_{VD} \) (see discussion in Section 3.2). The fluctuation of usages for same consumer in different periods helps to identify the variances of time-varying usage shocks, \( \sigma^2_v \) and \( \sigma^2_D \). The correlation of the usage fluctuations for voice and text will identify the covariance parameter \( \sigma_{VD} \).

26
Two parameters in the utility function in (1) that are difficult to estimate are the price coefficients $b^V = 1/\beta^V$ and $b^D = 1/\beta^D$, since there is no price change for either service plan in the data. However, two features in the data help to identify these parameters. First, about 10 percent of consumers switched either from voice centric to data centric plan or from data centric to voice centric plan. Given that marginal prices are different in the two service plans, changes in usages among these switchers help to identify the price coefficients among the whole population, as these coefficients are restricted to be homogeneous in our model. Second, usages of choosers of both plans may be above the amount of free usages, at which level marginal prices will affect the usage decisions. For example, there is no free usage of text for the voice centric plan while that for the data centric plan is 300; hence, the difference in text usage between choosers of both plans will, conditional on the preference parameters in the utility function, identify the price coefficient $b^D$.

Since there are no additional parameters in the learning model, the identifiability of our proposed model (Model 3) is virtually the same as the switching cost only model (Model 2). Conditional on the prior beliefs of true preferences (which is derived from the prior belief in period 1 and observed past usages), if a consumer switches from one service plan to the other later than another consumer, though the expected benefit of switching is the same, it implies that her switching cost of changing plans $SC^1$ is higher. Finally, since we do not have the usage data of those consumers before they joined in as users of the firm, identification of the switching cost $SC^2$ would solely rely on the observations of those consumers who drop out from the firm - conditional on the beliefs, if some consumers choose to stay with the firm though their expected indirect utility generated from either of the service plans is much lower than the outside option value, it implies that $SC^2$ is large.

5 RESULTS

5.1 Estimation Results
We report the results from the estimation in Table 3. All three models suggest that, while $\bar{\theta}^v$, the mean preference for using voice (voice demand intercept), is higher than $\bar{\theta}^D$, the mean preference for using text (text demand intercept), there is a larger heterogeneity in text preference, i.e., $\sigma_D$ (standard deviation of text demand intercept). The demand slopes for using both voice and text, $b^v = 1/\beta^v$ and $b^D = 1/\beta^D$, are significantly negative, suggesting that consumers do respond to marginal price changes in voice and text. To examine the importance of incorporating preference for text in the utility function, we simulate the utility of using both services based on the estimates of Model 3. After that, we compute the preference weight of using text, relative to that of using voice. Two situations are evaluated: 1. every consumer uses the voice centric plan; and 2. everyone uses the data centric plan. The resulting preference weights have a mean of 0.7 and 0.8 under above situations, respectively. That is, the utility of using text messages accounts for at least 40 percent of the total utility. Furthermore, the utility of using text service is higher than using voice service for about 10% of consumers.

One interesting result is that the correlation between time-invariant voice and text preferences (correlation between voice demand intercept and text demand intercept) is positive (the estimated correlations are 0.029, 0.160 and 0.057 in three models), implying that consumers who have a higher preference for voice are also likely to have a higher preference for text. Based on our draws of time-invariant preferences $\theta_i^v$ and $\theta_i^D$, we divide consumers into four segments using the median criterion: 1. high voice and high text preferences, 2. high voice but low text preferences, 3. low voice but high text preferences and, 4. low voice and low text preferences. The segment shares are 26.6%, 23.4%, 23.4%, 26.6%, respectively. That is, the segments of high-high type (segment 1) and low-low (segment 4) type in both usages are larger than the high-low (segment 2) or low-high (segment 3) types. We will not be able to recover such consumer segmentation pattern if we only model the demand for single services. Another intriguing result is that the

---

16 The utility is normalized by the price coefficients, “voice demand slope” and “text demand slope” in Table 3.
time-varying demand shocks for voice and text (parameter $l_{12}$) are also positively correlated, implying that consumers with a high voice usage shock in a period are also likely to have a high text usage shock. The estimated covariance parameter is 0.374, 0.530 and 0.389 in three models (see footnote in the table). These results are consistent with the data. For example, we find that consumers who stay with or switch to the data centric plan have a higher usage for both voice and text, compared to those consumers who stay with or switch to the voice centric plan (see Table 2). If consumer preferences are negatively correlated, we should expect to see from data that consumers who choose the data centric plan (those with high preference for text) would use fewer minutes of voice. These findings should have an important implication for the firm’s marketing policies in terms of which consumer segments to target – instead of segmenting consumers into high voice preference vs. high text preference categories, it may be more appropriate to segment consumers into high vs. low users of both voice and text.

Models 2 and 3 show that on average consumers incur a mean switching cost, $SC^1$ (switching cost within brands), of about ¥6.0 and ¥6.3 with standard deviations, $\sigma_{\text{sc}}^2$ (standard deviation of switching cost within brands), of about 0.8 and 1.0, respectively, and ¥6.4 when they switch to and from the outside option, i.e., $SC^2$ (switching cost across providers). This is consistent with our intuition that the latter switching should be more costly, either monetarily or psychologically, than the former. Although overall the switching costs are small (less than US $1), it may not be unreasonable considering that the purchasing power of the Chinese consumers in our city is relatively low (the estimated switching cost is about 1.5% of the monthly household income in our city.)

Based on the estimation results we can examine the impacts of the three-part tariff structure on consumer service plan choice and voice and text usages. We compute the elasticities of (1) service plan choice probability, (2) voice usage, (3) text usage, and (4) firm’s revenue, with respect to changes in the three-part pricing scheme for the voice and data centric plans. Results are reported in Table 4.
Regarding the service plan choice elasticities (see the first panel in Table 4), access fee has a larger impact than either marginal prices or free usages under the voice and data centric plans. This is consistent with the finding in Iyengar (2005) and Lambrecht et al. (2006). Magnitudes of the elasticities under the voice centric plan are greater than those under the data centric plan. In particular, the impact of change in marginal price or free usage for text under the data centric plan is almost negligible, perhaps because the free usage level offered by the data centric plan is much larger than an average user will consume each period. Our choice elasticities are in general smaller than in Lambrecht et al. (2006) and Narayanan et al. (2006), perhaps because our dataset comes from a different market (China) and we allow for zero usage in our model.

Turning to the usage elasticities (second and third panel in Table 4), we find that the voice usage elasticity with respect to marginal voice price change is larger under the data centric plan, while the text usage elasticity with respect to marginal text price change is larger under the voice centric plan. Note that the voice centric plan does not allow for free text usage hence it is virtually a two-part tariff for using text. The difference in text usage elasticity under the two plans is consistent with previous findings that usage demand is more elastic to marginal prices under the two-part tariff (voice centric) than the three-part tariff (data centric) pricing scheme (Lambrecht et al. (2006)). Further, change in marginal price and free usage of voice (text) affect text (voice) usage in both plans. In the data centric plan changes in marginal price or free usage for text have a negligible impact on text usage, as compared with changes in marginal price or free usage for voice. If, as discussed above, the free usage level offered by the data centric plan is larger than an average user will consume, change in price or free usage will not impact text usage. In contrast, by lowering marginal price or increasing free usage for voice attract many new users to join the data centric plan (see the first panel in Table 4) hence the overall usage of text will increase. These results demonstrate an asymmetric complementary relationship between voice and text in the data centric plan – increase in free usage or decrease in marginal price for voice will increase the
aggregate text usage level, but increase in free usage or decrease in marginal price for text has negligible impact on voice usage.

Finally, similar to the choice probabilities results, access fee has a larger impact on revenue than either marginal prices or free usages under both voice and data centric plans. In particular, one percent increase in access fee will lead to 0.3 percent increase in revenue for the voice centric plan. This probably implies that the firm can increase its profit from the voice centric plan by raising the access fee. Overall the revenue elasticities under the voice centric plan are larger than that under the data centric plan.

5.2 Model Comparisons

We compare the explanatory and predictive power of the three models in this section. Accounting for switching costs improves the model fit significantly as the log likelihood decreases from -19,019 in Model 1 to -12,980 in Model 2. Allowing for consumer learning in Model 3 the likelihood further improves to -12,954. Since the numbers of parameters in both models are identical, our learning Model 3 dominates Model 2 using any criterion such as AIC or BIC. Beyond the fit statistics we also compare the predicted switching patterns from the models with data. According to the data there are 5.6% switchers who switch from voice-centric plan to data-centric plan. The predicted switching proportions are 13.4%, 2.8%, 4.1% according to Models 1, 2 and 3, respectively. This result clearly shows that Model 3 has a better explanatory power for the switching pattern in data.

To further investigate the predictive power among the models, we conduct an out-of-sample validation exercise. We randomly select another 241 consumers that are not included in our estimation. Based on the estimation results we compute the likelihoods of the service plan choice of each consumer in each period observed from the data under the three models. The results are -6718, -4747, and -4656 for models 1, 2 and 3, respectively, a pattern very similar to the fit statistics above. We further use another measure to compare model predictions. Let $s_{data,t}$ be the observed market share in each period $t$ of the two service plans and the outside option for the 241 consumers in the out-of-sample data, and $s_{pred,t}$ the market share predicted by a specific model. We compute a “loss” function defined as
\[ \text{Loss} = \sum_{t=1}^{T} (s_{\text{pred},t} - s_{\text{data},t})^2, \]

which is the sum of squared of the model mis-prediction. The loss function values are 0.77, 0.39 and 0.13 for models 1, 2, and 3, respectively, demonstrating that model 3 outperforms the other two in terms of predictive power.

Perhaps instead of simply comparing the fit with data, it is more important to compare how the different models predict the dynamic patterns of the consumer service plan choices after the new data centric plan was introduced. Figure 4 provides a comparison of the dynamic changes of the data centric plan, using the out-of-sample data, as predicted by the three models. It is interesting to see that, without allowing for consumer learning, models 2 and 3 predict a “jump” of market share for the data centric plan in the first period, and after that market share stabilizes (its fluctuations are only due to idiosyncratic demand shocks). In contrast, our proposed model 3 predicts a gradual product penetration process, which is much closer to the observed pattern in the data.

5.3 Some “What-If” Experiments

Given that switching cost and consumer learning are important in the consumer service plan choice, we carry out some “what-if” experiments in this section to predict the dynamic changes in market share of the two service plans, after the new data centric plan was introduced, under various assumptions of the switching costs and the degrees of consumer learning. The results will be useful to shed light on how the firm could manage the penetration of the existing and new plans.\(^\text{17}\)

We assume the following four scenarios: First, given the consumer learning process, (1) switching costs (both \(SC^1\) and \(SC^2\)) decrease by 50 percent; (2) zero switching costs (both \(SC^1\) and \(SC^2\)); (3) switching costs remain the same, but consumers have perfect information of own usage preferences; and (4) zero switching costs (both \(SC^1\) and \(SC^2\)) and consumers have perfect information of own usage preferences. We compare the predicted penetration using estimates from

\(^{17}\) For instance, by improving customer service for current or new customers when they need to change cellular plans can reduce the switching costs for consumers. By providing more information about its services through advertising campaign the firm may help to increase consumers’ learning about their own preferences for the services.
model 3 under scenarios (1) and (2) to better understand the impact of switching costs, then contrast these with results in scenario (3) to understand the impact of consumer learning on the dynamics of service plan choices. Finally, comparing the original model to scenario (4) illustrates the interaction effects of switching costs and consumer learning.

Figures 5a and 5b shows the implied penetration process for voice centric and data centric plans under different scenarios. When switching costs are reduced by 50 percent the market share of the data centric plan increases by about 16 percent and that of the voice centric plan decreases by about 25 percent in the last period. When there is no switching cost, the market share of the data centric plan increases a further 7 percent and that of the voice centric plan decreases a further 25 percent. The results show that switching costs have a long term impact on the market share of both service plans even after 15 months of the introduction of the data centric plan.

When consumers have perfect information about their own usage preferences and do not need to learn, we observe an immediate “jump” in market share for both data centric and voice centric plans in the first period, and the market share will remain roughly the same over future periods. It is also important to note that learning is very quick when switching cost is low. For example, when there are no switching costs the market share of the data centric plan stabilizes five periods after its introduction as in Figure 5a. Finally, the market share of the data centric plan increases from 45 percent in the original model to about 70 percent when there is full information and zero switching costs, while the market share of the voice centric plan decreases from about 42 percent to 12 percent. These imply that potentially there may be large room for the firm to affect the penetration of its service plans by providing usage information or cutting down switching costs for consumers.

6 CONCLUSION

Product bundling and three-part tariff pricing are both popular market strategies. In this study, we use data from the wireless industry to study the impacts
of these pricing strategies on consumer choice of service plan and usage decisions. We develop a structural demand model to study the underlying consumer preferences structure for product bundles under the three-part tariff pricing scheme. Our model accounts for the impacts of switching costs and consumer learning on the consumer service plan choice. We find the following important results:

- Consumer preference for voice is positively correlated with that for text. Further, both voice and text services are important in the utility function of most consumers.

- Access fee affects the choice probabilities and firm revenue more than marginal prices and free usage. Marginal price and free usage for text have negligible impacts on choice, usages and revenue under the data centric plan, perhaps because the free usage level offered by the plan has been much larger than an average user will consume in each period. There is a complementary relationship of the price structure between voice and text that are bundled together in the service plans.

- Switching cost and learning play important roles in explaining the choice patterns observed in the data; hence, our model has a higher predictive power than those models without accounting for these factors. The simulation exercises based on the demand estimates show changes in switching costs or the information about own usage preferences among consumers significantly affect the penetration of the two service plans offered by the firm. This has obvious implications for how the firm could manage the process.

The main contribution of this paper is to fill a gap in the empirical literature by studying the impacts of three-part tariffs under product bundling on the consumer choice decisions. There are some limitations in our modeling approach and hence would be important for future research to address these issues. First, while we incorporate learning and switching cost in our model, we do not model the consumers as forward-looking. It may be important to study how a forward-looking consumer may behave differently from our model predictions. Second, as opposed to our model based on rational choice consumers may systematically over- or
under-estimate their usage needs. For example, DellaVigna and Malmendier (2005) found that consumers were over-optimistic in predicting their usage of gym facilities and hence would choose sub-optimal plans. Our learning model imposes restrictive assumptions on how consumers form prior beliefs regarding their usage preferences and how they update the beliefs. We notice from our data that at least some of the consumers do not make correct service plan choices even with sufficient time to learn their true usages. It may be important to use the consumer service plan choice and observed usages from data to infer the potential bias in some consumers’ beliefs regarding their usage expectations. Finally, constrained by the lack of price variations in our data, we model the consumer utility function as additively separable for consuming voice and text. It will be interesting to allow for interactions of the two types of consumption in the utility function hence may generate a richer substitution or complementarity pattern. This objective can be achieved if the price of either service plan changes during the sample period of data.
Figure 1: Evolution of Market Share in the Industry

Figure 2: Customer Growth over Time
Figure 3: Gain vs. Loss of Switching from Voice Centric to Data Centric Plan

![Gain vs. Loss of Switching from Voice Centric to Data Centric Plan](image)

Figure 4: Actual and Predicted Market Share of Data Centric Plan

![Actual and Predicted Market Share of Data Centric Plan](image)
Figure 5a: Dynamic Changes in the Market Share of Data Centric Plan

Figure 5b: Dynamic Changes in the Market Share of Voice Centric Plan
### Table 1: Details of Pricing Schemes

<table>
<thead>
<tr>
<th></th>
<th>Voice-centric</th>
<th>Data-centric</th>
</tr>
</thead>
<tbody>
<tr>
<td>access fee (¥/per month)</td>
<td>15 (no roaming), 30 (roaming)</td>
<td>20</td>
</tr>
<tr>
<td>on-net outgoing voice price (¥/per minute)</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>on-net incoming voice price (¥/per minute)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>off-net outgoing voice price (¥/per minute)</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>off-net incoming voice price (¥/per minute)</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>on-net outgoing text price (¥/per message)</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>on-net incoming text price (¥/per message)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>off-net outgoing text price (¥/per message)</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>off-net incoming text price (¥/per message)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>free voice minutes (value in ¥/per month)</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>free text messages (number)</td>
<td>0</td>
<td>300</td>
</tr>
</tbody>
</table>

### Table 2: Usage Patterns among Consumers

<table>
<thead>
<tr>
<th></th>
<th>Average minutes of voice (standard deviation)</th>
<th>Average number of text (standard deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole sample (100%)</td>
<td>178 (232.3)</td>
<td>114.2 (191.9)</td>
</tr>
<tr>
<td>Stay with voice centric plan (40.9%)</td>
<td>170.8 (218.1)</td>
<td>28.3 (81.6)</td>
</tr>
<tr>
<td>Stay with data centric plan (30.9%)</td>
<td>184.5 (252)</td>
<td>217.9 (210.1)</td>
</tr>
<tr>
<td>Switchers from voice- to data-centric plan (5.6%)</td>
<td>212 (202.9)</td>
<td>164.2 (231.2)</td>
</tr>
<tr>
<td>Switchers from data- to voice-centric plan (3.1%)</td>
<td>189.9 (218.2)</td>
<td>113.3 (201.4)</td>
</tr>
<tr>
<td>Drop outs (19.5%)</td>
<td>168.9 (245.9)</td>
<td>129.9 (256.7)</td>
</tr>
</tbody>
</table>
Table 3: Estimation Results

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voice centric plan intercept</strong></td>
<td>0.320</td>
<td>-1.900</td>
<td>-3.257</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.064)</td>
<td>(0.101)</td>
</tr>
<tr>
<td><strong>Data centric plan intercept</strong></td>
<td>-0.380</td>
<td>-0.156</td>
<td>-1.620</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.032)</td>
<td>(0.192)</td>
</tr>
<tr>
<td><strong>Voice demand intercept</strong></td>
<td>1.545</td>
<td>1.535</td>
<td>1.448</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Text demand intercept</strong></td>
<td>0.460</td>
<td>0.687</td>
<td>0.630</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.031)</td>
<td>(0.050)</td>
</tr>
<tr>
<td><strong>Standard deviation of voice demand intercept</strong></td>
<td>0.383</td>
<td>0.157</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.032)</td>
<td>(0.088)</td>
</tr>
<tr>
<td><strong>Standard deviation of text demand intercept</strong></td>
<td>0.818</td>
<td>0.960</td>
<td>0.878</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.031)</td>
<td>(0.402)</td>
</tr>
<tr>
<td><strong>Correlation between voice demand intercept and text demand intercept</strong></td>
<td>0.029</td>
<td>0.160</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.032)</td>
<td>(0.020)</td>
</tr>
<tr>
<td><strong>Voice demand slope</strong></td>
<td>-1.645</td>
<td>-1.574</td>
<td>-1.89</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.018)</td>
<td>(0.012)</td>
</tr>
<tr>
<td><strong>Text demand slope</strong></td>
<td>-1.127</td>
<td>-1</td>
<td>-1.52</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.023)</td>
<td>(0.036)</td>
</tr>
<tr>
<td><strong>$l_{11}$ in unobserved preference covariance matrix</strong></td>
<td>2.214</td>
<td>2.245</td>
<td>2.233</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.032)</td>
<td>(0.046)</td>
</tr>
<tr>
<td><strong>$l_{22}$ in unobserved preference covariance matrix</strong></td>
<td>1.768</td>
<td>1.839</td>
<td>1.765</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>$l_{12}$ in unobserved preference covariance matrix</strong></td>
<td>0.169</td>
<td>0.236</td>
<td>0.174</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.032)</td>
<td>(0.06)</td>
</tr>
<tr>
<td><strong>Switching cost within brands</strong></td>
<td>5.981</td>
<td>6.275</td>
<td>6.275</td>
</tr>
<tr>
<td></td>
<td>(0.158)</td>
<td>(0.168)</td>
<td></td>
</tr>
<tr>
<td><strong>Standard deviation of switching cost within brands</strong></td>
<td>0.794</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.199)</td>
<td>(0.088)</td>
<td></td>
</tr>
<tr>
<td><strong>Switching cost across providers</strong></td>
<td>6.368</td>
<td>6.442</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.117)</td>
<td>(0.121)</td>
<td></td>
</tr>
<tr>
<td><strong>Log-Likelihood</strong></td>
<td>-19019</td>
<td>-12980</td>
<td>-12954</td>
</tr>
</tbody>
</table>

Note: standard errors in parentheses
Note: the usage shock variance-covariance matrix is

$$\Sigma_x = l l'$$

where $l = \begin{bmatrix} l_{11} & l_{12} \\ 0 & l_{22} \end{bmatrix}$

$$\begin{bmatrix} 4.903 & 0.374 & 5.038 & 0.374 \\ 0.374 & 3.155 & 0.530 & 3.155 \\ 4.990 & 0.389 & 5.038 & 0.389 \end{bmatrix}$$
Table 4: Choice, Voice Usage, Text Usage and Revenue Elasticities

<table>
<thead>
<tr>
<th></th>
<th>Voice-centric</th>
<th>Data-centric</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(1) Choice Elasticities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Fee</td>
<td>-0.184</td>
<td>-0.102</td>
</tr>
<tr>
<td>Marginal Voice Price</td>
<td>-0.113</td>
<td>-0.087</td>
</tr>
<tr>
<td>Marginal Text Price</td>
<td>-0.046</td>
<td>-0.001</td>
</tr>
<tr>
<td>Free Voice Usage</td>
<td>0.128</td>
<td>0.061</td>
</tr>
<tr>
<td>Free Text Usage</td>
<td>0.120</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>(2) Voice Usage Elasticities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Fee</td>
<td>-0.020</td>
<td>-0.046</td>
</tr>
<tr>
<td>Marginal Voice Price</td>
<td>-0.129</td>
<td>-0.276</td>
</tr>
<tr>
<td>Marginal Text Price</td>
<td>-0.004</td>
<td>-0.001</td>
</tr>
<tr>
<td>Free Voice Usage</td>
<td>0.026</td>
<td>0.038</td>
</tr>
<tr>
<td>Free Text Usage</td>
<td>0.014</td>
<td>0.014</td>
</tr>
<tr>
<td><strong>(3) Text Usage Elasticities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Fee</td>
<td>-0.009</td>
<td>-0.041</td>
</tr>
<tr>
<td>Marginal Voice Price</td>
<td>-0.041</td>
<td>-0.033</td>
</tr>
<tr>
<td>Marginal Text Price</td>
<td>-0.035</td>
<td>-0.001</td>
</tr>
<tr>
<td>Free Voice Usage</td>
<td>0.006</td>
<td>0.024</td>
</tr>
<tr>
<td>Free Text Usage</td>
<td>0.021</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>(4) Revenue Elasticities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Fee</td>
<td>0.332</td>
<td>0.145</td>
</tr>
<tr>
<td>Marginal Voice Price</td>
<td>0.146</td>
<td>0.087</td>
</tr>
<tr>
<td>Marginal Text Price</td>
<td>0.087</td>
<td>0.004</td>
</tr>
<tr>
<td>Free Voice Usage</td>
<td>-0.088</td>
<td>-0.042</td>
</tr>
<tr>
<td>Free Text Usage</td>
<td>-0.071</td>
<td>-0.009</td>
</tr>
</tbody>
</table>
Reference


Wilson, R.B. (1992), Nonlinear Pricing, Oxford University Press.