NET Institute*

www.NETinst.org

Working Paper #05-28

October 2005

Cell Phone Demand and Consumer Learning – An Empirical Analysis

Martin Gaynor, Yunfeng Shi, Rahul Telang, and William Vogt
H. John Heinz III School of Public Policy and Management
Carnegie Mellon University

* The Networks, Electronic Commerce, and Telecommunications ("NET") Institute, http://www.NETinst.org, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, "virtual networks" comprised of computers that share the same technical standard or operating system, and on network issues in general.

Cell Phone Demand and Consumer Learning – An Empirical Analysis

Martin Gaynor

Yunfeng Shi *

Rahul Telang

William Vogt

H. John Heinz III School of Public Policy and Management
Carnegie Mellon University
Pittsburgh, PA, 15213

* Corresponding author: yungfens@andrew.cmu.edu; 412-268-8892

The financial support from the NET institute (http://www.NETinst.org) is gratefully acknowledged.

Abstract:

A structural model is used in this paper to analyze the demand and learning behavior in cell phone market. We assume that the cell phone consumption can be divided into a high-value part and a low-value part. The consumers are assumed to be uncertain about the exogenous shock of the need for high-value usage and also their preferences over the low-value usage. Meanwhile, we assume that the consumers' knowledge improves over time. As a result, the match between their plan choice and consumption pattern becomes better. Such a learning behavior is supported by the data set. Bayesian updating is used to represent the learning. The estimates of the parameters are obtained and compared to the benchmarks from previous research.

1. Introduction

The major purpose of this paper is to empirically analyze the demand and consumers' learning behavior in the cell phone market, using a panel data set from a large provider in Asia. Cell phone consumption is special for two important reasons. First, it is a two-stage decision problem with uncertainty. In most cases, the consumer has to choose a fee schedule and then decides how many minutes to call. Since the final payment is based on the number of minutes used and the chosen plan, a rational consumer needs to predict the usage when choosing a plan. However, the prediction is rarely accurate and the chosen plan might not be optimal ex-post. This is similar to the health (and other) insurance market. On the other hand, the consumer may not be completely clear about her own preference (utility function). This is particularly true if cell phone is a relatively new product in the market. In this paper, we develop a structural model that accounts for the sequential decision and both sources of uncertainty.

Second, under different situations, the consumer may value the cell phone usage differently. Emergency calls could be of much higher value than regular chatting. Therefore, at the same consumption level, the marginal utility derived from each additional minute could differ due to the different purpose of calling. Our model explicitly accounts for this difference by separating the total usage into two parts, the high-value part and the low-value (normal-value) part.

We assume the following scenario. The need for high-value usage is random and exogenously driven and the consumer uses exactly the number of high-value minutes that she needs. The consumer's preference over the low-value minutes is indexed by two key parameters in the utility function, one of which is assumed to be uncertain. At the first stage, the consumer chooses a plan based on her prediction of the need for high-value usage and her prior belief of the preference over low-value usage. After the plan is chosen, she receives a signal on her preference and updates her belief. The posterior belief of the preference is used in the utility maximization, which leads to the demand for low-value usage. Then, the need for high-value usage is realized for that period and the consumer updates her prediction of the next period high-value usage. This procedure repeats in each period, beginning with the updated prediction and belief from the previous period. The dynamic learning behavior is captured by the two updating structures, both following the Bayes's rule. The time dimension of the panel data has enabled us to examine the learning empirically.

Our approach builds upon previous research mainly from two lines, the literature of telecommunication demand modeling (Train, et al 1987, Kling and Ploeg 1991, Park, et al 1991, and Sung and Lee 2002), and that of consumer learning (McFadden and Train 1996, Miravete 2002a, Ackerberg 2003, and Clay, et al 2004). To our knowledge, this paper is the first empirical analysis of cell phone demand with consumer learning. Two recent articles are closely related to our research. Telang (2004) uses the same data set to estimate a demand model, but that paper does not consider the consumers' learning behavior and the utility function is assumed to be deterministic. Miravete (2003) models consumers' learning in the traditional land line phone service, using a data set from the 1986 Kentucky experiment. The current paper contributes primarily to the understanding of the consumer side in the emerging cell phone market. We observe a significant level of learning behavior in the data, which motivates our structural model.

In Telang (2004), the author reports a much higher level of price elasticity for the cell phone users, as compared to the results from the previous research on land line phone services. We ask the question whether the demand for cell phone is still highly elastic under the learning model and compare our results to the benchmarks from the previous literature. More generally, we hope this paper would add new empirical knowledge to the telecommunication demand and consumer learning. Besides, the separation of high-value and low value usage in the utility function seems to be a new treatment in the literature. The data are collected from a large cell phone service provider in Asia. The company certainly has some market power, but it is far from being a monopoly in the market. Due to the lack of information on the major competitors in that region, we do not consider strategic pricing behavior of the firm and thus ignore the question of optimal tariff design and its possible interaction with the consumers. Although our model focuses on the consumer's learning about preferences and exogenous shock, it can be extended to accommodate learning about product quality.

The next section discusses the data and the motivation for our model. Section 3 explains the details of the demand model, followed by the empirical specification and estimation in section 4. The final section discusses the results and concludes the paper.

2. Data and Preliminary Analysis

The data set used in this study comes from a large wireless service provider in Asia. The original data include about 10,000 subscribers over 12 months. However, due to missing observations, the initial test period (the first two months), and the promotional plan offered to some special groups of consumers, the final sample we use includes 6625 subscribers over 10 months (from March to December). The company offers five regular cell phone plans, with a fixed fee and a certain number of free minutes associated with each one. A same per minute price is charged for all usage above the free minutes. The table below, adapted from Telang (2004), describes all those plans.

Plan Name	Code	Monthly Fee	Free Minutes	Overtime Charge
ORANGE2000	Plan 1	2000	2117	3 per minute
ORANGE1500	Plan 2	1500	1217	3 per minute
ORANGE1100	Plan 3	1100	817	3 per minute
ORANGE800	Plan 4	800	517	3 per minute
PIONEER	Plan 5	350	233	3 per minute

All the consumers in our sample had stayed with the company for the whole period and their plan choices and voice usage. The data record their plan choices and the number of minutes used for each period. Besides, some of the consumer characteristics are also observed. The table below summarizes the data month by month.

Variable	Description	Mean	Standard Deviation
age	Age of the subscriber	33.1	7.04
gd	1=female; 0=male	0.54	0.498
rs	1=owned; 0=other	0.27	0.441
mr	1=married; 0=other	0.28	0.447
Plan1	Plan chosen in the Month	5.00	0
Plan2		5.0003	0.05
Plan3		4.96	0.29
Plan4		4.93	0.37
Plan5		4.92	0.40
Plan6		4.91	0.42
Plan7		4.91	0.42
Plan8		4.91	0.45
Plan9		4.90	0.46
plan10		4.90	0.47
Voice1	Minutes used in the month	416.03	335.96
Voice2		373.28	275.28
Voice3		257.88	217.59
Voice4		267.72	308.20

Voice5		249.63	248.98
Voice6		273.77	245.56
Voice7		276.51	240.73
Voice8		269.21	240.16
Voice9		272.67	233.80
Voice10		276.02	232.49

Two trends are noticeable from the table. First, the average number of minutes used declined sharply during the first three months. Second, the plan choices diversified over the time, although the majority of the consumers had chosen the basic plan all the time. (The basic plan was the only plan offered in the beginning, and therefore, all consumers chose the same plan during the first month. Starting from the second month, all plans were on the menu.) It seems that the consumers were making adjustment of both their plan choices and their consumption behavior over time. While many of them reduced the usage to match the basic plan, some of them switched to higher plans to match their consumption.

In the table below, we find the "ideal plan" for each consumer conditional on the voice usage of that month and then calculate the distance between the real payment and the optimal payment.

Deviation from Optimal Payment	**Mean**	**Standard Deviation**
Month 1	176.7174	389.129
Month 2	97.45377	285.3432
Month 3	26.33237	162.9723
Month 4	34.99516	219.36
Month 5	35.23877	195.15
Month 6	34.93534	179.4459
Month 7	35.49614	176.1575
Month 8	30.43516	164.8827
Month 9	33.42444	174.5696
Month 10	33.48186	168.6575

It is clear that the deviation from the optimal payment shrinks quickly over the first three months. After that, it stays relatively stable. The possible intuition could be twofold.

First, for those consumers who did deviate from the optimal payment, they tend to "learn" when the deviation is large. Second, those consumers can only correct their "mistakes" up to certain extent. There might be some uncontrollable random factors that prevent them from being optimal ex post.

Next, we look at the proportion of consumers who actually did choose the optimal plan in the month.

Optimal Plan Choice 1=optimal; 0=non-optimal	**Mean**
Month 1	0.72
Month 2	0.77
Month 3	0.94
Month 4	0.93
Month 5	0.93
Month 6	0.92
Month 7	0.92
Month 8	0.92
Month 9	0.92
Month 10	0.91

Obviously, the number of consumers choosing the best plan increased rapidly during the first two months. This is consistent with the previous table. Both tables have suggested that the consumers were learning to improve the match between their plan choices and cell phone usage. Moreover, the learning behavior can be separated into two types, plan switching and usage adjustment. This motivates the demand model that we will discuss in the next section.

3. The Demand Model and Learning

3.1 Utility, Demand and Plan Choice - the Basic Model:

Following the literature of telecommunication demand analysis, we model the consumer behavior as a two-stage sequential decision problem. During the first stage, a consumer chooses a calling plan based on her prediction of the number of minutes she will call. Then, after the uncertainty is resolved, the consumer maximizes her utility by deciding how many minutes to use.

Calling plans are indexed by *k*, and are described by $\left(T_k, g_k, p_k\right)$, where T_k, g_k and p_k are the fixed monthly fee, the free minutes allowed and the per-minute rate above the free minutes for plan *k*. Consumer's preference is represented by the utility function $U\left(M,C\right)=\frac{1}{b}(\theta M-\frac{1}{2}M^2)+C$, where M is the total number of cell phone minutes consumed and C is the composite good representing everything else with unit price. Notice that the utility function is quasi-linear in C and quadratic in M.

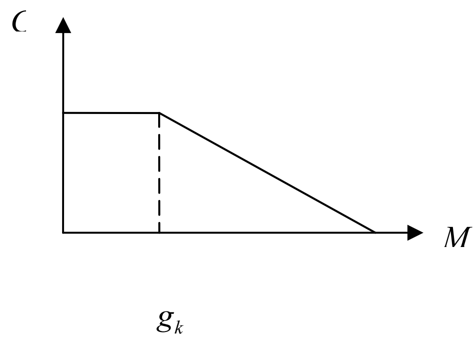
We deal with this discrete-continuous choice problem recursively, a typical approach in the related literature. Conditional on the choice of a calling plan *k* in the first stage, a consumer's utility maximization problem of the second stage is the following:

$$\underset{M,C}{Max}\ \ U\left(M,C\right)$$

Subject to the budget constraint: $C+p(M-g_k)=\omega-T_k$

Where ω is consumer's income and $p=\begin{cases} p_k & if\ M>g_k \\ 0 & if\ M\le g_k \end{cases}$.

Notice that the budget constraint is non-linear, due to the free minutes offered under each plan. However, it is convex and piecewise linear, with the form below:



Substitute the budget constraint into the objective function, the utility maximization problem becomes:

$$\underset{m}{Max} \quad \frac{1}{b}(\theta M - \frac{1}{2}M^2) + \omega - p(M - g_k) - T_k$$

$$p = \begin{cases} p_k & if\ M > g_k \\ 0 & if\ M \le g_k \end{cases}$$

Depending on the value of θ, the first order conditions are:

$$\frac{\partial U}{\partial M} = \frac{1}{b}(\theta - M) - p_k = 0 \quad if \begin{cases} M > g_k \\ p = p_k \end{cases}$$

$$\frac{\partial U}{\partial M} = \frac{1}{b}(\theta - M) = 0 \qquad if \begin{cases} M \le g_k \\ p = 0 \end{cases},$$

And they imply the following demand equations:

$$M = \theta - bp_k \quad \text{if } \theta > g_k + bp_k \tag{3.1a}$$
$$M = \theta \quad \text{if } \theta \le g_k \tag{3.1b}$$
$$M = g_k \quad \text{if } g_k < \theta \le g_k + bp_k \tag{3.1c}$$

The above equations are intuitive. After a plan has been chosen and the monthly fee has been paid, if the satiation point is high enough (3.1a), the optimal consumption will be above the free minutes and up to the point where the marginal utility is equal to the marginal (per-minute) price. If the satiation point is lower than the number of free minutes (3.1b), the consumption should reach the satiation point and stop exactly there. If the satiation point is higher than the number of free minutes, but the marginal utility after the free minutes is lower than the marginal price (3.1c), the consumer should use up all the free minute but no more.

Back to the first stage, the consumer need to choose a plan based on the demand equations derived above, which lead to the following indirect utility function:

$$W(g_k, T_k, p_k) = \begin{cases} \frac{1}{2b}\theta^2 - \theta p_k + \frac{1}{2}bp_k + g_k p_k - T_k + \omega & \text{if } \theta > g_k + bp_k \quad (3.2a) \\ \frac{1}{2b}\theta^2 - T_k + \omega & \text{if } \theta \le g_k \quad (3.2b) \\ \frac{1}{b}\theta g_k - \frac{1}{2b}g_k^2 - T_k + \omega & \text{if } g_k < \theta \le g_k + bp_k \quad (3.2c) \end{cases}$$

Suppose the consumer has perfect knowledge of θ and b, then she would simply choose a plan with the combination of $g_k, T_k,$ and p_k that gives her the highest indirect utility at the first stage. Clearly, assuming b is constant and taking $g_k, T_k,$ and p_k as given, the choice of plan depends on the value of θ. Here, the ordering of g_k, T_k, and p_k across the plans is critical. Ideally, we would like the ranking of the plans to be monotone in θ. This means for each consumer, the fixed fee and the number of minutes of her preferred plan will increase as her satiation point gets higher. Such a ranking would be guaranteed by the following three conditions. The per-minute price $p_k \equiv p$ is constant for all plans; for any two plans *k* and *j*, it is always the case that $T_k > T_j$ if and only if $g_k > g_j$; and finally, the ratio T_k / g_k (the per-minute price of free minutes) decreases in T_k and is always below p_k. The five regular plans in our data satisfy those conditions and the above indirect utility function implies a set of threshold values:

$$\theta_0, \theta_1, \theta_2 ... \theta_K$$

Those values guarantee that the consumer will choose plan *k* if and only if $\theta_{k-1} \le \theta \le \theta_k$. Telang (2004) has details on the derivation of the specific values for our data.

Since θ is not observed by the econometrician, the above threshold values, together with the distributional assumption on θ, would allow us to estimate an ordered probit (or logit) model with maximum likelihood, as in Telang (2004). In the following section, we take a different approach by assuming that the consumer is uncertain about her satiation point when choosing the plan.

3.2 Two-part Consumption - the Extended Model:

The basic model discussed above implicitly assumes that the consumer treats all the voice minutes in the same manner when maximizing utility. It means at any given consumption level, the marginal utility of each additional minute, whatever the purpose of calling, is the same. This is inconsistent with what actually happens in real life. Most subscribers use cell phone for many different purposes, and the marginal utility of each additional minute could potentially depend on its function. The consumer may put very high value on certain types of usage (e.g. family emergency, change of schedule or canceling appointment) and relatively low value on others. Moreover, in most cases, it could be hard to predict in advance how many high-value minutes a consumer will need in the next period. To formalize this idea, we extend the basic model by decomposing the total number of used minutes into two parts, the high-value minutes (M_h) and the low-value minutes (M_l). We assume the consumer has the following utility function:

$$U\left(M_h, M_l, C\right) = VM_h I_{(M_h \leq \theta_h)} + \frac{1}{b}[\theta_l(M_l + M_h) - \frac{1}{2}(M_l + M_h)^2] + C \qquad (3.3)$$

Where $I_{(M_h \leq \theta_h)}$ is an indicator function and θ_h is a parameter reflecting the number of high-value minutes the consumer needs. Notice that the above utility function has a couple of features. First, if the value of "*V*" is high enough, then upon observing θ_h, the consumer will always choose M_h equal to θ_h, but not more than θ_h, regardless of the plan choice and the consumption of M_l. Intuitively, this means that the consumer will use exactly the number of high-value minutes she needs. Second, at the optimum, an increase in M_h will generally reduce M_l, since the marginal utility of the low-value minutes is decreasing. This is particularly true if the total number of minutes has reached the satiation point in the second term of the utility function.

Consider the following scenario: after choosing a plan, the consumer has to choose M_l first and then observe θ_h and finally choose M_h. With the above structure, the consumer's utility maximization now becomes a three-stage decision problem. Recursively, we start from the final stage. After observing θ_h, the consumer decides M_h, which will always be equal to θ_h, due to our assumption on *V*. For almost all consumers, the phone bill usually comprises only a small proportion of the income; hence it is reasonable to assume that the total number of minutes $(M_l + M_h)$ is always strictly inside the budget constraint. At the second stage, if the consumer knows her θ_h, then the problem becomes the following:

$$\underset{(M_l, C)}{Max}: U\left(M_l, C \mid \theta_h\right) = V\theta_h + \frac{1}{b}[\theta_l(M_l + \theta_h) - \frac{1}{2}(M_l + \theta_h)^2] + C$$

Subject to the budget constraint: $C + p(M_l + \theta_h - g_k) = \omega - T_k$

Where ω is consumer's income and $p = \begin{cases} p_k & \text{if } M_l + \theta_h > g_k \\ 0 & \text{if } M_l + \theta_h \le g_k \end{cases}$.

Substitute the budget constraint into the objective function, the utility maximization problem becomes:

$$\underset{(M_l)}{Max}: U\left(M_l\right) = V\theta_h + \frac{1}{b}[\theta_l(M_l + \theta_h) - \frac{1}{2}(M_l + \theta_h)^2] + \omega - p(M_l + \theta_h - g_k) - T_k$$

Where $p = \begin{cases} p_k & \text{if } M_l + \theta_h > g_k \\ 0 & \text{if } M_l + \theta_h \le g_k \end{cases}$

The first order conditions are:

$$\frac{dU}{dM_l} = \frac{1}{b}(\theta_l - \theta_h - M_l) - p_k = 0 \qquad if \begin{cases} M_l + \theta_h > g_k \\ p = p_k \end{cases}$$

$$\frac{dU}{dM_l} = \frac{1}{b}(\theta_l - \theta_h - M_l) = 0 \qquad if \begin{cases} M_l + \theta_h \le g_k \\ p = 0 \end{cases},$$

And they lead to the following demand equations:

$$
\begin{aligned}
M_l &= \theta_l - \theta_h - bp_k && \text{if } \theta_l > g_k + bp_k \\
M_l &= \theta_l - \theta_h && \text{if } \theta_l \le g_k \\
M_l &= g_k - \theta_h && \text{if } g_k < \theta_l \le g_k + bp_k
\end{aligned}
$$

Back to the first stage, the consumer chooses a plan based on the demand equations derived above, which lead to the following indirect utility function:

$$
W(g_k, T_k, p_k) = \begin{cases} V\theta_h + \dfrac{1}{2b}\theta_l^2 - \theta_l p_k + \dfrac{1}{2}bp_k + g_k p_k - T_k + \omega & \text{if } \theta_l > g_k + bp_k \\ V\theta_h + \dfrac{1}{2b}\theta_l^2 - T_k + \omega & \text{if } \theta_l \le g_k \\ V\theta_h + \dfrac{1}{b}\theta_l g_k - \dfrac{1}{2b}g_k^2 - T_k + \omega & \text{if } g_k < \theta_l \le g_k + bp_k \end{cases}
$$

Suppose the consumer knows θ_l and θ_h, then she would simply choose a plan with the combination of g_k, T_k, and p_k that gives her the highest indirect utility at the first stage. Based on above equations, the choice of plan obviously depends on the value of θ_l. Similar to the basic model in the previous section, the five regular plans in our data and the above indirect utility function implies a set of threshold values:

$$
\theta_{l0}, \theta_{l1}, \theta_{l2} ... \theta_{lK} \tag{3.4}
$$

Those values guarantee that the consumer will choose plan *k* if and only if $\theta_{l,k-1} \le \theta_l \le \theta_{lk}$.

3.3 Uncertainty and Learning about Preference:

Up to now, our model assumes that the consumer knows her preference with certainty (i.e. all the parameters in the utility function). The data suggest otherwise, however. We observe that over time, particularly during the first two months, many consumers have substantially shortened the distance between their actual payment and the ideal payment (or consumption, depending on perspective) in theory. It seems that those consumers are uncertain about their preferences but their knowledge improves over time. Notice that although θ_h is a parameter in the utility function, it can be interpreted as an exogenous shock, since it represents the number of high-value minutes a consumer needs during that period. With that in mind, we consider the following learning behavior: the consumer is uncertain about both θ_l and θ_h. At the first stage, she chooses a plan based on her initial beliefs of those two parameters. Then she observes a signal on θ_l and updates her belief of it. The consumption of low-value minutes is based on the updated belief of θ_l and the initial belief of θ_h. Then, at the third stage, θ_h is realized and the consumer chooses M_h equal to the realized θ_h and updates her belief of it. For each period, the consumer repeats this procedure, using the updated beliefs from the last period as the initial beliefs of the current period.

Following the common approach in the consumer learning literature, as shown in Clay, Goettler and Wolff (2004), we assume the learning is represented by Bayesian updating. The uncertainty of θ_l is characterized by a normal distribution with mean μ_l, which is unknown and variance σ_l^2, which is known. At the beginning of the first period, the consumer has a prior distribution on μ_l, which is also normal with mean μ_{l0} and variance σ_{l0}^2. Similarly, we assume that θ_h is also normal, with unknown mean μ_h and known variance σ_h^2. The consumer has a normal prior on μ_h, with mean μ_{h0} and variance σ_{h0}^2.

With the normal prior, after observing the signals s_t (t=1, 2…T), the posterior distribution of μ_l at time "t" is also normal, with mean μ_{lt} and variance σ_{lt}^2. Similarly, the realizations of θ_h, r_t (t=1, 2…T-1), lead to the posterior distribution of μ_h, with mean μ_{ht} and variance σ_{ht}^2. To summarize:

$\theta_l \sim N(\mu_l, \sigma_l^2)$: The distribution of θ_l, with unknown μ_l and known σ_l^2;

$\mu_l \sim N(\mu_{l0}, \sigma_{l0}^2)$: The prior distribution of μ_l at t=0;

$s_1, s_2 ... s_T$: The signals observed by the consumer at each period;

$f(\mu_l \mid s_1, s_2 ... s_t) \sim N(\mu_{lt}, \sigma_{lt}^2)$: The posterior distribution of μ_l at each period;

Where $$\mu_{lt} = \frac{\sigma_{l0}^2 \sum_{i=1}^{t} s_i + \mu_{l0}\sigma_l^2}{\sigma_l^2 + t\sigma_{l0}^2} \tag{3.5a}$$

and $$\sigma_{lt}^2 = \frac{\sigma_{l0}^2 \sigma_l^2}{\sigma_l^2 + t\sigma_{l0}^2} \tag{3.5b}$$

(t=1, 2…T)

$\theta_h \sim N(\mu_h, \sigma_h^2)$: The distribution of θ_h, with unknown μ_h and known σ_h^2;

$\mu_h \sim N(\mu_{h0}, \sigma_{h0}^2)$: The prior distribution of μ_h at t=0;

$r_1, r_2 ... r_{T-1}$: The realizations of θ_h observed by the consumer at the end of each period;

$f(\mu_h \mid r_1, r_2 ... r_{T-1}) \sim N(\mu_{ht}, \sigma_{ht}^2)$: The posterior distribution of μ_h at the end of each period;

Where $$\mu_{ht} = \frac{\sigma_{h0}^2 \sum_{i=1}^{t} r_i + \mu_{h0}\sigma_h^2}{\sigma_h^2 + t\sigma_{h0}^2} \tag{3.6a}$$

and $$\sigma_{ht}^2 = \frac{\sigma_{h0}^2 \sigma_h^2}{\sigma_h^2 + t\sigma_{h0}^2} \tag{3.6b}$$

(t=1, 2…T, and the learning of θ_h is one period lagged.)

Notice there is a slight difference between our assumptions on the learning of the two parameters. We interpret θ_h as an exogenous shock and the consumer learns about its distribution. The consumption of high-value minutes depends on the ***realization*** of this random variable in each period. On the other hand, θ_l is considered the true satiation point, but the consumer never knows it for sure. Therefore, although the learning is also about its distribution, she uses the ***posterior belief*** to choose the consumption of low-value minutes. When choosing the plan in the first stage, the prior distributions of both parameters are used.

With the above learning structure, for each period "t" the consumer chooses $M_h = r_t$, the realization of θ_h at time "t". Back to the second stage, the consumer maximizes the expected utility by choosing M_l: (strictly speaking, it is the expectation of the expected utility.) Assuming θ_l and θ_h are independent and using the current beliefs at time "t", the problem becomes the following:

$$\underset{(M_{lt})}{Max}: EU\left(M_{lt}\right) = E\left\{V\theta_h + \frac{1}{b}[\theta_l(M_{lt}+\theta_h) - \frac{1}{2}(M_{lt}+\theta_h)^2] + \omega - p(M_{lt}+\theta_h - g_k) - T_k\right\}$$

$$= V\mu_{h,t-1} + \frac{1}{b}[(\mu_{lt}M_{lt} + \mu_{lt}\mu_{h,t-1}) - \frac{1}{2}(M_{lt}^2 + 2M_l\mu_{h,t-1} + \mu_{h,t-1}^2 + \sigma_{h,t-1}^2)] + \omega - p(M_{lt} + \mu_{h,t-1} - g_k) - T_k$$

Where the "expected price" $p = \begin{cases} p_k & if\ M_{lt} + \mu_{h,t-1} > g_k \\ 0 & if\ M_{lt} + \mu_{h,t-1} \le g_k \end{cases}$

The first order conditions are:

$$\frac{dU}{dM_l} = \frac{1}{b}(\mu_{lt} - \mu_{h,t-1} - M_{lt}) - p_k = 0 \qquad if \begin{cases} M_{lt} + \mu_{h,t-1} > g_k \\ p = p_k \end{cases}$$

$$\frac{dU}{dM_l} = \frac{1}{b}(\mu_{lt} - \mu_{h,t-1} - M_{lt}) = 0 \qquad if \begin{cases} M_{lt} + \mu_{h,t-1} \le g_k \\ p = 0 \end{cases},$$

And they imply the following demand equations:

$$M_{lt} = \mu_{lt} - \mu_{h,t-1} - bp_k \quad \text{if } \mu_{lt} > g_k + bp_k \tag{3.7a}$$
$$M_{lt} = \mu_{lt} - \mu_{h,t-1} \quad \text{if } \mu_{lt} \le g_k \tag{3.7b}$$
$$M_{lt} = g_k - \mu_{h,t-1} \quad \text{if } g_k < \mu_{lt} \le g_k + bp_k \tag{3.7c}$$

It is important to notice that at each period "t", while the consumer knows μ_{lt} and observes r_t after she chooses M_{lt}, none of those values are observed by the econometrician separately. For each period, what we observe in the data is M_t^* the total consumption ex post. Since the consumption of high-value minutes M_{ht} is always equal to the current realization r_t, the equation $M_t^* = M_{lt} + r_t$ must hold all periods. Therefore, the above demand equations (3.7a, 3.7b and 3.7c) imply the following:

$$M_t^* = \mu_{lt} - \mu_{h,t-1} + r_t - bp_k \quad \text{if } \mu_{lt} > g_k + bp_k \tag{3.8a}$$
$$M_t^* = \mu_{lt} - \mu_{h,t-1} + r_t \quad \text{if } \mu_{lt} \le g_k \tag{3.8b}$$
$$M_t^* = g_k - \mu_{h,t-1} + r_t \quad \text{if } g_k < \mu_{lt} \le g_k + bp_k \tag{3.8c}$$

To complete the model, we consider the consumer's choice of the plans in the first stage. With unknown θ_l and θ_h, at the beginning of each period, the expected indirect utility under plan ***k*** is: $E_t W(g_k, T_k, p_k)$

$$
=\begin{cases} V\mu_{h,t-1}+\dfrac{1}{2b}(\mu_{l,t-1}^2+\sigma_{l,t-1}^2)+p_k(g_k+\dfrac{1}{2}b-\mu_{l,t-1})-T_k+\omega & \text{if } \mu_{l,t-1}>g_k+bp_k \\ V\mu_{h,t-1}+\dfrac{1}{2b}(\mu_{l,t-1}^2+\sigma_{l,t-1}^2)-T_k+\omega & \text{if } \mu_{l,t-1}\le g_k \\ V\mu_{h,t-1}+\dfrac{1}{b}\mu_{l,t-1}g_k-\dfrac{1}{2b}g_k^2-T_k+\omega & \text{if } g_k<\mu_{l,t-1}\le g_k+bp_k \end{cases}
$$

At the first stage, the consumer chooses a plan that gives her the highest indirect utility at the first stage based on above equations, which lead to the same set of threshold values we derived previously $\theta_{l0},\theta_{l1},\theta_{l2}...\theta_{lK}$ (3.4). Those values guarantee that the consumer chooses plan k if and only if $\theta_{l,k-1}\le\mu_{l,t-1}\le\theta_{lk}$.

This concludes the theoretical part of the paper. The essential results of the demand and learning model are the threshold values (3.4), the equations (3.5a), (3.5b), (3.6a), (3.6b), (3.8a), (3.8b) and (3.8c), which we take to the data in the next section.

4. Estimation

4.1 Empirical Specification:

First, notice that the preference signals of θ_l, $s_1,s_2...s_T$ and the realizations of θ_h, $r_1,r_2...r_{T-1}$ are deterministic for each consumer at each period (they are assumed to be observed by the consumer without error). However, those values are not observed by us and are stochastic from our perspective. We assume both sets are i.i.d. random variables with normal distribution across the population:

$$s_t \sim i.i.d.\ N(\mu_s,\sigma_s^2) \qquad \text{(t=1, 2…T)} \qquad (4.1a)$$

$$r_t \sim i.i.d.\ N(\mu_r,\sigma_r^2) \qquad \text{(t=1, 2…T-1)} \qquad (4.1b)$$

The same situation applies to the two prior means μ_{l0} and μ_{h0}, as the cell phone user knows them while we do not. We make the simplifying assumption that μ_{l0} is constant across all users. Since the data show the same plan choice during the first month for all users, this assumption might not be too restrictive. Meanwhile, μ_{h0} is also assumed to be normally distributed across the population:

$$\mu_{h0} \sim N(\mu_p, \sigma_p^2) \tag{4.1c}$$

We make another simplifying assumption that all the variance terms involved in the Bayesian updating (σ_h^2, σ_{h0}^2, σ_l^2 and σ_{l0}^2) are constant for all users. Our goal is to estimate the structural parameter *b*, which measures the slope of the demand curve, and all the parameters associated with the learning behavior.

The above specification (4.1a), (4.1b), and (4.1c), combined with (3.5a) and (3.6a) from the previous section, implies that the current beliefs (the posterior means) of each period, μ_{lt} and μ_{ht} are also random variables, both with normal distribution across the population:

$$\mu_{lt} \sim N(\upsilon_{lt}, \tau_{lt}^2) \text{ and } \mu_{ht} \sim N(\upsilon_{ht}, \tau_{ht}^2)$$

Where $\upsilon_{lt} = \dfrac{t\sigma_{l0}^2\mu_s + \mu_{l0}\sigma_l^2}{\sigma_l^2 + t\sigma_{l0}^2}$, and $\tau_{lt}^2 = \left(\dfrac{\sigma_{l0}^2}{\sigma_l^2 + t\sigma_{l0}^2}\right)^2 t\sigma_s^2$ (4.2a)

$$\upsilon_{ht} = \frac{t\sigma_{h0}^2\mu_r + \mu_p\sigma_h^2}{\sigma_h^2 + t\sigma_{h0}^2}, \text{ and } \tau_{ht}^2 = \left(\frac{\sigma_{h0}^2}{\sigma_h^2 + t\sigma_{h0}^2}\right)^2 t\sigma_r^2 + \left(\frac{\sigma_h^2}{\sigma_h^2 + t\sigma_{h0}^2}\right)^2 \sigma_p^2 \tag{4.2b}$$

Similarly, following the equations (3.8a), (3.8b) and (3.8c), the total consumption ex post M_t^* is also normally distributed, with mean and variance depending on the value of μ_{lt}. Let Φ denote the standard normal CDF and ϕ the corresponding PDF:

$$f(M_t^* \mid \mu_{lt} > g_k + bp_k) = \frac{f(\mu_{lt} - \mu_{h,t-1} + r_t - bp_k)}{P(\mu_{lt} > g_k + bp_k)} = \frac{\phi\left(\dfrac{M_t^* - \mu_{m1}}{\sigma_{m1}}\right)}{1 - \Phi\left(\dfrac{g_k + bp_k - \upsilon_{lt}}{\tau_{lt}}\right)} \tag{4.3a}$$

(Where $\mu_{m1} = \upsilon_{lt} - \upsilon_{h,t-1} + \mu_r - bp_k$, and $\sigma_{m1}^2 = \tau_{lt}^2 + \tau_{ht}^2 + \sigma_r^2$)

$$f(M_t^* \mid \mu_{lt} \le g_k) = \frac{f(\mu_{lt} - \mu_{h,t-1} + r_t)}{P(\mu_{lt} \le g_k)} = \frac{\phi\left(\dfrac{M_t^* - \mu_{m2}}{\sigma_{m2}}\right)}{\Phi\left(\dfrac{g_k - \upsilon_{lt}}{\tau_{lt}}\right)} \tag{4.3b}$$

(Where $\mu_{m2} = \upsilon_{lt} - \upsilon_{h,t-1} + \mu_r$, and $\sigma_{m2}^2 = \tau_{lt}^2 + \tau_{ht}^2 + \sigma_r^2$)

$$f(M_t^* \mid g_k < \mu_{lt} \le g_k + bp_k) = \frac{f(g_k - \mu_{h,t-1} + r_t)}{P(g_k < \mu_{lt} \le g_k + bp_k)} = \frac{\phi\left(\dfrac{M_t^* - \mu_{m3}}{\sigma_{m3}}\right)}{\Phi\left(\dfrac{g_k + bp_k - \upsilon_{lt}}{\tau_{lt}}\right) - \Phi\left(\dfrac{g_k - \upsilon_{lt}}{\tau_{lt}}\right)} \tag{4.3c}$$

(Where $\mu_{m3} = g_k - \upsilon_{h,t-1} + \mu_r$, and $\sigma_{m3}^2 = \tau_{ht}^2 + \sigma_r^2$)

4.2 The Maximum Likelihood Estimation:

In the data, we observe the plan choice and the total consumption for each consumer during each period. Assuming the decisions of the consumers are mutually independent, the joint probability of the observed plan choices and total consumptions for *N* consumers during *T* periods is simply the product of the individual probabilities. This leads us to the following likelihood function:

$$\boldsymbol{L}\ (\gamma \mid \text{plans, consumptions}) = \prod_{t=1}^{T} \prod_{n=1}^{N}\ [P_{nt}(plan = k \mid \gamma_1) \cdot p_n(M_t^* \mid \gamma_2, plan = k)] \tag{4.4}$$

Where γ is vector of the parameters we want to estimate; γ_1 and γ_2 are the two subsets of γ that affect the plan choice probability and the consumption probability respectively. Subscript t indexes time periods and subscript n indexes consumers. Based on the results of the previous section, for consumer n at time t, the probability of plan choice is the following:

$$P_{nt}(plan = k \mid \gamma_1) = P(\theta_{l,k-1} \le \mu_{l,t-1} \le \theta_{lk}) = \Phi\left(\frac{\theta_{lk} - \upsilon_{l,t-1}}{\tau_{l,t-1}}\right) - \Phi\left(\frac{\theta_{l,k-1} - \upsilon_{l,t-1}}{\tau_{l,t-1}}\right) \qquad (4.5)$$

Where Φ denotes the standard normal CDF; $\upsilon_{l,t-1}$ and $\tau_{l,t-1}$ correspond with (4.2a). The probability (density) of the observed total consumption for consumer n at time t, conditional on the chosen plan, is the following: $p_n(M_t^* \mid \gamma_2, plan = k)$

$$= f(M_t^* \mid \mu_{lt} > g_k + bp_k) + f(M_t^* \mid \mu_{lt} \le g_k) + f(M_t^* \mid g_k < \mu_{lt} \le g_k + bp_k) \qquad (4.6)$$

The right-hand side corresponds with (4.3a), (4.3b) and (4.3c) derived above. The parameters are estimated by maximizing the log-likelihood function based on (4.4).

4.3 The Effect of Observed Consumer Heterogeneity:

To incorporate the observed consumer characteristics, we assume those covariates shift the prior mean of low-value minutes. Thus, for a consumer with characteristics X_i , his or her prior distribution of low value minute becomes the following: $\theta_l \sim N(\mu_l, \sigma_l^2)$ and $\mu_l \sim N(\mu_{l0} + X_i\beta, \sigma_{l0}^2)$. Based on the structure of our model, the consumer characteristics could in principle affect several different aspects of demand and learning. Besides the prior mean of low value minutes, they may shift the prior mean of high-value minutes or prior variances (and thus, the posterior variances). It is not hard to imagine that the consumers are heterogeneous in the effectiveness (speed) of learning. However, to identify those effects separately would be hard. The assumption we make is reasonable, since the low value usage should be the part that the consumers have the best knowledge, and therefore, the part that might be best explained by the observed characteristics.

The covariates that we include in the model are age, gender, residence type and marital status. The data actually contains more variables. Some of them have very little variation, such as payment type and account type. The other two variables that we do not use are Short Message Service and Web Access usage. It would be hard to treat them as exogenous, since both of them are decision variables from the perspective of the consumers. On the other hand, the model would have been much complicated if we include them as endogenous variables. In our data, the usages of SMS and WAP are very low compared to the voice minutes. Hence their exclusion from the model should not substantially affect the results.

5. Results and Discussions

The point estimates and the standard deviations of the parameter are reported in the table below.

Parameters	**Point Estimates**	**Standard Deviations**
b	58.9 (**)	8.7
μ_{l0}	345.6	22.3
σ_{l0}^2	129.7	16.4
σ_l^2	201.4	28.3
μ_p	805.8	35.7
σ_p^2	289.4	19.4
σ_{h0}^2	207.9	78.4
σ_h^2	225.5	36.5
μ_s	1791.9	122.6
σ_s^2	8.1	1.7
μ_r	865.9	24.8
σ_r^2	132.1	9.5
Age	0.74(**)	0.02
Gender	-0.28	0.16
Residence type	0.65	0.34
Marital status	-0.69(**)	0.05

The demand slope estimated from the model (58.9) is lower than the previous research with the same data set (99.9), as reported in Telang (2004). However, the interpretation of this key structural parameter has to be careful. It potentially affects both the plan choice and the consumption level. Moreover, the latter is affected only when the (predicted) satiation point is high enough conditional on the chosen plan. Although it in general measure the consumers' responsiveness towards price change, to extract the demand elasticity from it is not straightforward. However, based on its scale, the cell phone users do seem to be more responsive to price change than the traditional land line phone users.

Similarly, the interpretation of the coefficients of the consumer characteristics should not be in the usual way. They do not affect the demand directly, but rather, affect the prior mean. Over time, their influence reduces as the posteriors depend more and more on the signals and realizations. Moreover, considering the scale of the observed usage, all the four covariates we use have coefficients close to zero. Although two of them are statistically significant, they are still practically insignificant in the sense that the change in the usage they could cause is almost none.

The other parameters in the table measure the learning behavior. The shrinking posterior variances of both high-value and low-value usage can be obtained from the estimates.

We conclude the paper by discussing several issues in this research. First, as mentioned above, the strategic pricing behavior of the firm is completely ignored. Second, our data only include those consumers whole stay with this provider for the whole period. We do not have adequate information on those who signed up later or cancelled the service earlier. Service cancellation would be particularly interesting to explore, since the consumers who were already in the basic plan might find out even that plan was not worth the price based on their preferences. This could further illustrate the learning behavior. Third, as also discussed in Telang (2004), we do not have reliable income data, and this restricts our choice of the utility function. In case the above information becomes available in the future, it will be very interesting to extend our study to address those issues.

References

Ackerberg, D. A. (2003): "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination," *International Economic Review*, 44, 1007-1040.

Clay, K., R. Goettler and E. Wolff (2004): "Consumer Learning about Experience Goods: Evidence from and Online Grocer," Working Paper, Carnegie Mellon University.

Hausman, J. A. (1985): "The Econometrics of Nonlinear Budget Sets," *Econometrica*, 53, 1255-1282.

Kling, J. P. and V. D. Ploeg (1991): "Estimating Local Call Elasticities with a Model of Stochastic Class of Service and Usage Choice", in *Telecommunications Demand Modeling*, ed. by A. de Fonteny, M. H. Shugard and D. S. Sibley, New York: North-Holland, 1990.

Martins-Filho, C. and J. W. Mayo (1993): "Demand and Pricing of Telecommunications Services: Evidences and Welfare Implications," *Rand Journal of Economics*, 24, 439-454.

McFadden, D. L and K. E. Train (1996): "Consumers' Evaluation of New Products: Learning from Self and Others," *Journal of Political Economy*, 104, 683-703.

Miravete, E. J. (2002a): "Choosing the Wrong Calling Plan? Ignorance and Learning," *American Economic Reivew*, 93, 297-310.

Miravete, E. J. (2002b): "Estimating Demand for Local Telephone Service with Asymmetric Information and Optimal Calling Plans," *Review of Economic Studies*, 69, 943-971.

Mitchell. B. M. and I. Vogelsang (1991), "Telecommunication Pricing: Theory and Practice", New York: Cambridge University Press, 1991.

Park, R. E., B. M. Wetzel and B. M. Mitchell (1991), "Price Elasticities for Local Telephone Calls," *Econometrica*, 51, 1699-1730.

Sung, N. and Y. Lee (2002), "Substitution Between Mobile and Fixed Telephones in Korea," *Review of Industrial Organization*, 20(4), 367.

Train, K. E., D. L. McFadden and M. Ben-Akiva (1987): "The Demand for Local Telephone Services: A Fully Discrete Model of Residential Calling Patterns and Service Choices," *Rand Journal of Economics*, 18, 109-123.