

NET Institute*

www.NETinst.org

Working Paper #04-10

October 2004

An Empirical Analysis of Cellular Voice and Data services

Rahul Telang

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.

An Empirical Analysis of Cellular Voice and Data services ¹

Rahul Telang

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

rtelang@andrew.cmu.edu

¹ Financial support from the NET Institute (<http://www.NETinst.org>) is gratefully acknowledged. Author would like to thank Atip Asvanund, Ramayya Krishnan, William Vogt and participants of the Austin Mobility Conference for their valuable feedback.

An Empirical Analysis of Cellular Voice and Data services

Abstract

Cellular telephony and associated data services has been a major social phenomena for well over a decade now. It has changed the way – in some countries more than others – in which people communicate. In many countries in Northern Europe and Asia, its penetration rates are very high and in others less so but in all cases it has engendered change at multiple levels – socially as noted and in terms of market structure and competition with the established Incumbent Local Exchange and Inter Exchange service providers. However, there has been little work published in the academic literature on user consumption of cellular voice and data services. This has been due to the unavailability of longitudinal data at the individual user level on their consumption of voice and data services. We have such data from a large cellular service provider in Asia. Demand for voice and data services is influenced by the tariffs or “service plans” offered by firms. In our analysis we empirically estimate the drivers for cellular services how demographic and plan characteristics affect the user choices. We first provide a theoretical model and then provide insight into consumption patterns over a one year period of cellular voice and data services and relate it to service plan design.

1. Introduction

In the last few years, the demand for cell phone voice as well data services have grown exponentially. It is making considerable inroads into voice services of incumbents and is generating options for new personal use data services such as SMS which did not exist before. For many users, Cell phone has become their only option for communication, replacing the use of land lines completely. As the reliability and ubiquity of wireless services have grown, the number of new land line connections have been steadily dropping all over the world. Moreover low capital expenditure, easiness of installation and generally fewer regulatory restrictions have made cell-phones accessible to even users in low income and developing countries. Not surprisingly, in African continent the penetration of cell-phones far exceeds that of land line phones (ITU report 2001).

Similar trends are visible in wireless data services where SMS (short messaging services) have been tremendously popular in European and African continent changing the way users use communication channel. With the emergence of devices such as PDA's being combined with cell phones and platforms such as J2ME (java2 2 mobile edition) and .Net mobile edition signal the emergence of richer platforms for data and voice services than in the past being available over cellular networks.

However, while these are interesting technology trends, there is little that is understood and published in the academic literature on individual level consumption of voice and data services for the cellular markets. Telecommunication demand modeling literature has a rich history (See Taylor 1994) for details. Generally, telecommunication pricing and consumption has been an

intriguing area of research because traditional phone services have unique characteristics ((for example, high fixed costs but almost zero marginal costs) which pose interesting econometric challenges. More importantly, telecommunication traditionally has been highly regulated industry and there has been lot of work on impact of regulatory changes on user consumption and social welfare. Generally researchers have made distinction between Access and Usage of telephones and have modeled them explicitly (Perl 1987; Taylor and Kridel 1990) while calculating the price elasticity for telephone demand. Researchers have also focused on the usage based (measured) and flat rate pricing and studied how users choose one over the other and that how their demand changes when they choose flat rate as opposed to usage based plans. Based on some structure then researchers have calculated the price elasticity (Kling and Van der Ploeg 1990; Kridel 1990; Miravette 2002). In doing so, they have either used aggregate data, or individual data or sometimes survey data. Park, Wetzel and Mitchell (1983) study the effect of mandatory switch from flat rate to measured service on local calling. Martin-Filhos and Mayo (1993) study the impact of EAS (extended area service) on both local and long distance demand and subsequently on social welfare. Again, local calls incur zero marginal costs while long distance calls incur positive marginal cost. Train, McFadden and Ben-Akiva (1987) take discrete choice modeling approach to model demand for local phone service. They characterize users as choosing a particular service option and a particular calling portfolio where portfolio is defined as calls to a particular number and average duration of calls at each time of day to each distance zone. Then they specify the nested logit model and estimate price elasticity. Miravette (2002) models the user uncertainty of number of minutes demanded in the next month when choosing a plan today.

The studies on cellular markets though are still only emerging. Recently, Miravette (2003) estimated a structural model for tariffs in US cellular market but their focus was different and data was not at the individual level. Sung and Lee (2002) study the substitution between mobile and fixed telephones in Korea.

We know of no study where individual level data on cellular voice and data services has been analyzed. We have collected detailed usage (voice and data services) data on 10,000 customers for over a period of one year from January 2002 to December 2002. We also have their demographic information. Therefore, we investigate how demographic characteristics affect the demand for voice and data services and how stable is this demand over time. Therefore first research question we focus on is to empirically estimate the drivers of cellular voice and data services.

As we noted earlier, telecommunication data poses econometric challenges in estimating price elasticity because of fixed and marginal pricing. In our data too, we find that users select a plan first (by paying a fixed price) and then utilizes their voice minutes depending on the plan chosen. If they exceed the minutes then they pay the marginal price of each minute. But the choice of the plan first and consumption of minutes later is endogenous. Using a structural model, we simultaneously model user's choice of the plan and number of minutes consumed. Then, we calculate the price elasticity of cellular voice services and compare it with the traditional telephone elasticity. This is the second goal of this paper.

Finally, we are interested in the cross price elasticity of cellular data and voice services. More precisely, we want to examine how consumption of data services affects the consumption of voice services. Are they substitutes or complements? This is the third goal of this paper.

Using a structural model, we estimate the price elasticity of users subscribed to the lowest value plan (95% users in our sample subscribe to that plan) to be 1.12. This is in contrast to some of landline estimates where typically elasticity is reported to be quite small and in the range of 0.1-0.2 for local phones and 0.4-0.5 for long distance phone service. Such high elasticity, among other things, indicates that slight change in price can induce large changes in user demand.

We also find that SMS use complements the demand for voice services. Higher use of SMS indicates higher use of voice services as well. Therefore, even though SMS is considered a low price substitute for voice minutes, in our data we find that opposite is true. On the other hand, WAP service use was not found to be significant. We also find that men use cellphone more often than women and older users (30-50 years) tend to use cellphones more than younger ones. All else equal, singles use it more often than married users.

The rest of the paper is organized as follows. In the next section, we outline the basic technology trends. We outline our theoretical and estimable model in section 3. In section 4, we provide details on our data. In section 5, we provide the results of our estimation and finally we conclude our paper in section 6.

2. Basic Technology Trends in Cellular voice and Data Services

There currently are three major types of cellular standards in deployment – GSM, CDMA and TDMA². The three standards are based on CDMA and TDMA multiplexing technologies and used for voice communication. CDMA uses code division multiplexing, while TDMA uses time division for sharing air wave spectrums. GSM is a variant of the traditional TDMA in that it uses CLIP for voice digitization rather than the traditional ADPCM. While CDMA and TDMA are the dominant standards in the United States, GSM is dominant in European and other markets. Nevertheless, GSM has been growing rapidly in the United States recently.

For data services, following two technologies are used -

SMS: Short message service is a store and forward service of data between handsets. It is implemented as a data network overlay on cellular networks. It was initially designed as a two-way text paging service, but recently more interactive applications – email, stock quotes, and others – have been implemented using SMS as a platform. SMS was initially developed for GSM-based networks and benefit from GSM-specific features. In particular, it can interact with the profile editing feature of GSM SIM cards to provide personalization for GSM-based wireless e-commerce and other advanced applications.

WAP: Wireless application protocol is introduced as an air interface-neutral protocol (i.e., independent of GSM, CDMA or TDMA) that enables applications such as access to WAP web sites, email and other online services. It is a competing technology with SMS. WAP typically

² GSM – Global System for Mobile Communication, CDMA – Code Division Multiple Access, TDMA - Time Division Multiple Access

works in conjunction with WML, and now XHTML, to provide interactive web-like services that can be accessed from a cellular handset. While it is argued that WAP will supplant SMS as an e-commerce protocol, in many markets, content provision is still based on SMS because of its historical backgrounds and links to GSM dominance.

As we describe in the following section, in our data we have detailed consumption information on voice, SMS and WAP services.

3. Data

The data comes to us from a large cellular service provider in Asia. The firm is the 3rd largest firm in Thailand having a customer base of more than 2 millions. The firm offers wireless voice and wireless data services to its customers. The data collection began on January 2002 and continued for one year. We have detailed data on about 10,000 customers on their use of voice (minutes of use per month) for each month in the whole year. For wireless data we have information on their number of SMS exchanged and WAP use. WAP use typically means accessing email, stock quote or any other information on the cell phone. We have information on how many minutes of WAP service were used by each user in each month.

We have detailed demographic information about the users (subscribers). We know the *gender, age, occupation, residence type and city, the method for bill payment, income, and marital status* of each user in our data set. We also have information on the pricing plans that were being offered each month and which plan the users selected. We have detailed information about each plan like, fixed price to enroll in the plan, number of free available minutes, price per minute if

the allocated number of minutes were exceeded etc. Finally, we also have information on the type of handset the users were using.

The following tables provide some descriptive statistics on the average characteristics and consumption pattern.

Table 1: Average Usage Statistics

Monthly Usage per Individual	Average	Minimum	Maximum
Voice (Minutes)	167.99	0	11845.31
SMS (Messages)	7.10	0	3106
WAP (Minutes)	0.05	0	360.8

In Table 2, provide details on the demographic characteristics of the users in our sample. About 58% are females and about 38% users are below age 30. We also have information of residence type (own, rental etc), the city (metro or small city), and language used in billing. Since they do not seem to be of any significance in our estimation, we do not report them. We also had some information on the income. But the data was not too reliable and hence we do not use it.

Table 2: Demographic Characteristics

Age	Gender	Payment Type	Marital Status	Employment
38% under 30	58% Female	98% Cash	68% Single	Business Employee 80%
55% in 30-50	42% Male	1% Credit Card	30% Married	Self Employee 5%
7% over 50		1% Debit Card	1% Divorced	Government 3% Student 3%

Finally, we also have detailed information on the plans offered. During the year, the firm offered 6 major plans with the following characteristics.

Table 3: Plan characteristics

	Fixed Fee	Free minutes	Overtime charge
Pioneer	350	400	3 per minutes
Option800	800	567	3 per minute
Option1100	1100	967	3 per minute
Option1500	1500	1267	3 per minute
Option2000	2000	2167	3 per minute

Another plan Option500 (with the fixed fee of 500 and free minutes of 317) was also offered but it is dominated by Pioneer plan and hence we ignore it (Few users had selected that plan and we drop them as well). In the first two months of the data (January and February 2002), not all plans were being offered and they were changing frequently. Therefore, we drop all observations for the first two months. Many users in our sample had also selected themselves in a “family” plan for which the individual minutes were not available and we drop those observations as well. In our final sample, we had 6257 users selecting one of the five plans of table 3, over the 10 month period.

3. Model

The economic theory of consumer behavior assumes that the decision maker maximizes her utility of consumption of goods; in our case it is the mobile services. We start with a utility

function. But as we noted, our analysis is complicated by the fact that user demand critically depends on the given tariff structure she chooses. Similarly, the choice of tariff depends on user demand and utility. Therefore, our model needs to accommodate this simultaneity.

Since the major goal of this paper is to estimate key elasticity parameters, we take the tariffs as given.³ We assume that the utility function of the user is given by

$$U = \frac{1}{b} \left[\theta q - \frac{1}{2} q^2 \right] \quad (1)$$

where θ is customer type, q is the number of minutes used by user and b is the parameter to be estimated. θ captures the user heterogeneity and all else equal, a high θ customer demands more minutes. The marginal utility is given by $\frac{1}{b}[\theta - q]$, which much be equal to the marginal price

for utility maximization. Therefore, we have $\frac{1}{b}[\theta - q] = p$. Therefore the demand equation is given by⁴

$$q = \theta - bp \quad (2)$$

Where b is the price elasticity of demand. When $p = 0$, demand is given by simply θ , which is the satiation point for the user with type θ . In other words, the user never demands more than θ .

We assume that θ is distributed truncated normal with mean μ and variance σ , i.e.,

$\theta \sim TN[\mu, \sigma^2]$. The truncation is from below at zero. This ensures that θ is never negative as we would expect. The user's choice of tariff depends on the tariff that gives her the highest surplus.

³ Miravette (2003) has first derived the optimal tariffs and then fit the lower envelop of non-linear tariffs to observed tariffs.

⁴ We chose a linear demand curve for simplicity in estimation. Other demand curves are also possible.

In our data, we observe five different plans a user can choose from. By definition, a high θ user consumes more minutes than a low θ user. Therefore, highest θ user will choose the largest tariff offered in our data (Option2000). User surplus from choosing the option2000 (see data section for details on various plans) is

$$U_{2000} = \frac{1}{b} \left[\theta q - \frac{1}{2} q^2 \right] - P_1 - p_1(q - 2167)$$

where P_1 is the fixed price of the plan and p_1 is the marginal price, q is the number of minutes used and $(q - 2000)$ incorporates the fact that marginal price is applicable only after 2000 free minutes have been exhausted. Hence, $P_1 + p_1(q - 2000)$ combined together is the cost of the plan. Similarly, surplus from choosing plan Option1500 can be written as

$$U_{1500} = \frac{1}{b} \left[\theta q - \frac{1}{2} q^2 \right] - P_2 - p_2(q - 1267)$$

Similarly, we can write the surplus from other plans as well.

Now, we want to find an indifferent user of type θ_1 who is indifferent between Option2000 and Option1500. Note that if a user prefers Option2000 over Option1500 then Option2000 will dominate other plans as well due to the natural ordering of these plans. First, note that for such a θ_1 to exist, it must be that such a user will not use any “extra” minutes (that is more than 2167) in Option2000. In case of Option1500, the user may or may not decide to either use “extra” minutes (i.e. more than 1267 minutes) depending on the marginal price. Second, for a given $\theta_1 < 2167$, user demand is given by (2) when the user uses “extra minutes”, otherwise demand is simply $q = \theta$. In short, for any $1267 < \theta_1 < 2167$, if the user uses plan Option2000 then he will be always

satiated but not when using plan Option1500. Substituting the appropriate value of q^5 in U_{2000} and U_{1500} and comparing them leads to

$$\frac{1}{b} \left[\theta^2 - \frac{1}{2} \theta^2 \right] - P_1 = \frac{1}{b} \left[\theta(\theta - bp_2) - \frac{1}{2} (\theta - bp_2)^2 \right] - P_2 - p_2(\theta - bp_2 - 1267) \quad (3)$$

Simplifying and solving for θ give us an inequality such that all users with

$$\theta_1^* \geq \frac{bp_2}{2} + \frac{P_1 - P_2 + 1267p_2}{p_2}$$

The constraint here is that such a θ_1 should satisfy the constraint that $\theta_1 - bp_2 - 1267 > 0$; i.e. when the user is consuming the extra minutes in plan Option1500. This happens only when b is not very high. Otherwise, user will not consume marginal minutes in both plans and hence the surplus comparison will be

$$\frac{1}{b} \left[\theta^2 - \frac{1}{2} \theta^2 \right] - P_1 = \frac{1}{b} \left[\theta 1267 - \frac{1}{2} 1267^2 \right] - P_2$$

Note that now such a user is satiated when using Plan Option2000 but not when using Option1500. But he still does not find it beneficial to use marginal minutes in Plan Option1500 because of high b . Solving above yields

$$\theta_1^* \geq 1267 + \sqrt{2b(P_1 - P_2)b}$$

The restriction here is that $\theta_1 < 2167$. When b is very high then this restriction is violated⁶. In such case, when $\theta_1 > 2167$, the surplus comparison should be

⁵ Note that the value of q is not yet realized. The user only knows its θ , and based on that it infers how many minutes it will use in both plans.

⁶ In general, estimated b has to be really high for this to hold. In our data, only for the lower two plans, Option800 and Pioneer we find the case when $\theta_4 > 567$.

$$\frac{1}{b} \left[\theta 2167 - \frac{1}{2} 2167^2 \right] - P_1 = \frac{1}{b} \left[\theta 1267 - \frac{1}{2} 1267^2 \right] - P_2$$

Note that now the user is unsatiated in both plans but does not use marginal minutes in both of them. Solving this yields,

$$\theta_1^* \geq 1717 + \frac{b}{900}(P_1 - P_2)$$

Following the similar process we can compare the surplus for all subsequent plans to find different indifferent θ ($\theta_1, \theta_2, \theta_3, \theta_4$).

Since $\theta \sim TN[\mu, \sigma^2]$, for a user of type θ , we can find the probability of him choosing Option2000, Option1500 and so on from the inequality outlined in Appendix A. This is the first part of the estimation where a user knowing his θ , anticipates the minutes he will use in various plans and choose one of the plans.

In the second stage, once the choice of plan has been made, actual number of minutes used is realized. The demand for the user, conditioned on him choosing Plan Option2000 is

$$\begin{aligned} q &= \theta - bp_1 + \varepsilon && \text{if } \theta > 2167 + bp_1 \\ q &= 2167 + \varepsilon && \text{if } 2167 \leq \theta < 2167 + bp_1 \\ q &= \theta + \varepsilon && \text{if } \theta_1^* < \theta \leq 2167 \end{aligned} \quad (4)$$

where ε is the error term which is normal with mean 0 and standard deviation σ_ε^2 . Since the user first chooses the plan based on the expectation number of minutes he plans to consume, this error term reflects the random shock (or the uncertainty) the user may experience in the given month. The higher the variance of ε , more will be difference between the actual minutes and estimated minutes. The error term may also reflect some measurement error.

In equation (4) above, the first term indicates the fact that the marginal price kicks in only for those users who have high enough θ . The second term indicates that for some range θ , the users restrict their demand and stop using the marginal minutes. The final term indicates that for when θ is low enough then users are satiated without the marginal minutes. The lower bound θ_1^* indicates that the plan Option2000 is chosen only when $\theta > \theta_1^*$. The distribution of q depends on both θ and ε . We can write the similar demand functions for different plans.

Now we can write the joint distribution of a user i choosing the plan given θ and then choosing the number of minutes given the plan and θ . Therefore the likelihood function for a user is

$$LL = \sum_{j=1}^5 \text{Pr ob}(\text{plan}_j/\theta) \cdot f(q_i/(\text{plan}_j, \theta)) \quad (5)$$

where probability of choosing a plan given θ is outlined previously. The second term is $f(q)$, probability of observing q , conditioned on a plan being chosen needs elaboration. In the following, we will provide details on calculating $f(q)$, for plan Option2000, the rest follows the same.

Let $f(\theta)$ be the conditional normal density when plan2000 is chosen with associated distribution function $F(\theta)$ and $\phi(\cdot)$ and $\Phi(\cdot)$ be the associated standard normal density and distribution respectively. Recall also that ε is distributed normal $\varepsilon \sim N(0, \sigma_\varepsilon)$ Therefore, the probability of observing q_i given that user has chosen plan Option2000 is

$$f(q_i) = \int_{\theta_i^*}^{2167} \frac{1}{\sigma_e} \phi\left(\frac{q_i - \theta}{\sigma_e}\right) f(\theta) d\theta + \int_{2167}^{2167+bp_2} \frac{1}{\sigma_e} \phi\left(\frac{q_i - 2167}{\sigma_e}\right) f(\theta) d\theta + \int_{2167+bp_2}^{\infty} \frac{1}{\sigma_e} \phi\left(\frac{q_i - \theta - bp_2}{\sigma_e}\right) f(\theta) d\theta \quad (6)$$

While the integrals look formidable, they can be simplified considerably such that evaluation of $f(q_i)$ requires no more than evaluating $\phi(\cdot)$ and $\Phi(\cdot)$ which can be performed by any computer program. We provide these details in Appendix A.

We calculate $f(q_i)$ for each plan similarly and substitute the values in (5). Since we have n users in our sample with 10 observations for each user for $N = 10*n$, the final likelihood function looks like

$$LL = \prod_{i=1}^N \sum_{j=1}^5 \text{Prob}(\text{plan}_j / \theta) \bullet f(q_i / (\text{plan}_j, \theta)) \quad (7)$$

We maximize this function to estimate our parameters. Recall that σ_e is the standard deviation of the ε and we estimate the structural parameters b , μ , σ , and σ_e .

4.1 Incorporating Covariates

To incorporate other covariates (like use of SMS, or WAP) and demographic variables to investigate how they affect the voice minutes consumption, we let some of the variation in μ be explained by covariate. Therefore, we let

$$\tilde{\mu} = \mu \exp(\xi X) \quad (8)$$

where X is the vector of covariates and ζ the parameter vector to be estimated. Hence, we use new value $\tilde{\mu}$ in the likelihood equation (4) to estimate ζ as well as μ, σ^2 , and b . Note that the way we have incorporated the covariates, they shift the mean of θ . One of the hypothesis we are interested in testing is how the use of SMS and WAP affects the use of voice minutes. Are they complements such that more use of SMS and WAP leads to more use of voice?, Or are they substitute where higher SMS and WAP also signal high voice minutes? We also examine the impact of demographic variables on demand. In particular we want to test whether younger users use more voice minutes compared to older user or how do males and females differ in their use of cell phone?

5. Results and Discussion

We first estimate the model without demographics and other explanatory variables. The results of the structure parameters are presented in Table 1. The standard errors are reported in parenthesis.

Table 1: Elasticity Estimates

Parameter	Estimates
b	99.8 ^{**} (0.10)
μ	296.5 ^{**} (1.36)
σ	236.1 ^{**} (0.77)
σ_e	220.8 ^{**} (0.7)
<i>Log Likelihood</i>	475670

^{**} suggests significant at 1%

Low standard errors highlight the precision of our estimates because of the large sample we have. All the parameters are highly significant and in expected direction. The price elasticity parameter b is quite large and significant. With the linear demand function specification, it suggests that one unit increase in the marginal price of the call leads to reduction of about 100 minutes. The mean value of the satiation point θ is about 296 minutes and standard deviation of about 236. Finally, there is large variation in the number of minutes actually consumed which suggests the inability of the users to precisely estimate their expected usage.

Price Elasticity

One key focus of our paper is how users respond to usage price (which is fixed at $p = 3$ /minute) in our data. Since we use a linear demand curve, the elasticity calculation depends on quantity consumed as well. The elasticity equation is

$$\begin{aligned}\varepsilon &= \frac{\Delta q}{\Delta p} \bullet \frac{p}{q} \\ &= b \bullet \frac{p}{q}\end{aligned}$$

Clearly, elasticity depends on value of q . We calculate the elasticity at the mean value of minutes used in each plan. The estimated elasticity is presented in the table below.

Table 5: Price Elasticity of Demand

Quantity q	Elasticity
$q = 266$	1.12
$q = 642$	0.46
$q = 1010$	0.30
$q = 1556$	0.19
$q = 1994$	0.15

Since a large number of users have signed up for the lowest value plan (about 95%), clearly, the users in our sample exhibit very high level of price elasticity. As a point of comparison, for the landline local telephone use several authors report elasticity less than 0.1 (Mitchell et al 1983) or about 0.17 (Kling et al 1991). For the long distance phone use, authors have reported elasticity in the range of 0.4-0.5 (Taylor 1996; p 132). In contrast, for the cellphone use (at least in an Asian country), we report very high level of price elasticity.

Such high level of elasticity suggests that even small changes in the marginal price can induce large changes in user consumption. Given that the bandwidth of the cellphone service providers is more limited than that of land line phone companies, besides profit maximizing, pricing is also viable tool to optimally plan for the capacity and congestion. To see the importance of this, keeping all else constant, if the firm decreased its per minute price from 3 to 2, the users of the pioneer plan (lowest price plan) would have consumed about 30,000 more minutes per month. And, this is with the sample of only about 6300 users. With thousands of users as its customer

base, our results imply that small change in pricing could lead to potentially large impact on both profits and capacity utilization.

Estimates with Covariates

As noted before, we let the covariates shift the mean of θ . We incorporate SMS use, WAP use, age, gender and marital status of the users as potential covariates. We also tried other factors like payment method or employment but they did not seem very relevant. We use the log of SMS, WAP and age while gender and marital status were binary variables. The other structural parameters are same as in the previous case. The result of estimation is shown below.

Table 4: Estimates with Covariates

Parameter	Estimates
b	99.9 ^{**} (0.08)
μ	56.1 ^{**} (0.02)
σ	221.8 ^{**} (0.4)
σ_e	216.6 ^{**} (0.7)
SMS	0.24 ^{**} (0.01)
WAP	-0.15 (0.09)
Age	0.35 ^{**} (0.01)
Gender	0.10 ^{**} (0.01)
Marital Status	-0.44 ^{**} (0.01)
Log Likelihood	473250

The key parameter estimate for price elasticity ' b ' is virtually same as in the previous case. This reassures us about the robustness of our estimation model. The parameter for the mean μ is much lower but it is to be expected because the other covariates would shift it upwards (see equation 8). The standard deviation of the both θ and ε are similar to previous estimates which suggest that inclusion of covariates does not affect other parameters and hence our key structural estimates are very robust.

Estimate for SMS use is highly significant and positive. Therefore, higher use of SMS is correlated with higher use of voice minutes. In short, SMS and voice act like a complement. A user who uses more SMS is likely to also consume more voice minutes. Clearly, this presents an opportunity for the firm to offer its bundle of voice and SMS services with the knowledge that high SMS use would not cannibalize high voice minutes use. Generally SMS has less revenue potential for cellphone operators. Therefore, given the significant use of SMS (especially in Asian and European countries), this results should be reassuring for telephone operators that it does not cannibalize their voice revenues. It seems SMS and voice services are being used for different purpose by users. Estimate for WAP use is negative but it is not significant. We did not have many users using WAP services and hence its impact on voice consumption would be marginal at best (and statistically insignificant in our sample).

Estimate for *Age* is positive and significant. Older users tend to use more minutes. Since most of the users in our sample are less than 50 years of age, results suggest that users in 30-50 years age group are more likely to use higher minutes than users in less than 30 years age category.

Generally, high income users have been shown to consume more minutes. Since we do not have income data, this is probably consistent with the fact that since people in the 30-50 age group are

likely to have higher income, they also are using more voice minutes. Estimate for Gender is positive and significant. Since male is coded as 1 and female as 0, this suggests that males, on an average, use their cellphone more often than females. Finally, estimate for the marital status is negative and significant. This suggests that, all else equal, singles use their cellphone more than married users.

With the inclusion of covariates, the fit of the model has gone up as well as reflected in the lower value of log likelihood.

6. Conclusion

In this research, we provide an economic model of users' consumption of cellular voice and data services and estimate the price elasticity of voice minutes. Typically, cellphone contracts are nonlinear tariffs with a fixed fee and some free minutes and a marginal price if the users overshoot their free minutes. Such nonlinear pricing creates difficulty in estimating the elasticity because users first choose a plan which forces their consumption decision in the month. At the same time, choice of such a plan itself is driven by their demand expectation. In this paper, we present a structural model to estimate the price elasticity while accommodating such simultaneity.

We collected a unique data set of cellphone voice and data services consumption for about 6300 users over a year. We also collect their monthly use of cellphone minutes and the plan they sign for. In addition, we also have information on their SMS and WAP service use and their demographic characteristics. We find that price elasticity of users for the lowest plan is more than 1 (1.12). This is in contrast to some of landline estimates where typically elasticity is

reported to be in the range of 0.1-0.2 for local phones and 0.4-0.5 for long distance phone service. Such high elasticity, among other things, indicates that slight change in price can induce large changes in user demand.

We also find that SMS use complements the demand for voice services. Higher use of SMS indicates higher use of voice services as well. Therefore, even though SMS is considered a low price substitute for voice minutes, in our data we find that opposite is true. On the other hand, WAP service use was not found to be significant. We also find that men use cellphone more often than women and older users (30-50 years) tend to use cellphones more than younger ones. All else equal, singles use it more often than married users.

Ours is one of the first studies to examine the individual user's use of cellular voice and data services in a formal setting. While our study contributes to the literature in several interesting ways, it is not without limitations. Our demand function is linear, future work should explore nonlinear demand functions. We make specific distributional assumption on some of the parameters for analytical tractability; sensitivity analysis is needed to test the robustness of those assumptions. In future, collecting income data of users would be an important step as income plays a crucial role in users telecommunication use.

6. References

Burtless G. and Jerry Hausman (1978), "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment", *Journal of Political Economy*, 86(6), 1103-1130.

ITU (2002), "Reinventing Telecom World Communication Indicators", *World Telecommunication Development Report*.

Kling, J.P. and Van Der Ploeg S (1991), "Estimating Local Call Elasticities with a Model of Stochastic Class of Service and Usage Choice", in *Telecommunications Demand Modeling*, ed. by A. de Fontenay, M.H. Shugard, and D.S. Sibley, North Holland Publishing Co., Amsterdam.

Martins Filho C. and Mayo J. W. (1993), "Demand and Pricing of Telecommunication Services: Evidence and Welfare Implications", *The Rand Journal of Economics*, 24(3), 439-454.

Miravete E J and Roller L H (2003), "Competitive Nonlinear Pricing in Duopoly Equilibrium: The Early US Cellular Telephone Industry", *Working paper*.

Miravete E J (2002), "Estimating Demand for Local Telephone Service with Asymmetric Information and Optional Calling Plans", *The 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.

Perl L (1983), "Residential demand for telephone service" in *Telecommunications Demand in Theory and Practice*, by Lester Taylor, 83-96, Kluwer Academic Publishers, Dordrecht, Netherlands.

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

Taylor, L.D. and Kridel, D.J. (1990), "Residential Demand for Access to the Telephone

Network,” in *Telecommunications Demand Modeling*, ed. by A. de Fontenay, M.H. Shugard, and D.S. Sibley, North Holland Publishing Co., Amsterdam.

Taylor Lester (1994), *Telecommunications Demand in Theory and Practice*. Kluwer Academic Publishers, Dordrecht, Netherlands, 1994.

Train Kenneth E., M Ben Akiva, T Atherton (1989), “Consumption Patterns and Self Selecting Tariffs”, *The Review of Economics and Statistics*, 17(1), 62-73.

Appendix A

The first and third integrals in equation (6) are similar. The first integral has the form

$$\int_{\theta_1^*}^{2167} \frac{1}{\sigma_e} \phi\left(\frac{q_i - \theta}{\sigma_e}\right) f(\theta) d\theta \quad (\text{A1})$$

$f(\theta)$ is conditional normal density of θ given the plan Option2000 has been chosen. Let $\hat{\mu}$ and $\hat{\sigma}$ be the mean and standard deviation of $f(\theta)$ when $\theta > \theta_1^*$. Since ε is normal distributed, q_i is normally distributed as well with mean $\hat{\mu}$ and variance $\hat{\sigma}^2 + \sigma_e^2$. Now consider the joint distribution of q_i and θ which can be written as

$$f(\theta, q_i) = f(\theta/q_i) f(q_i)$$

Using the Bayes theorem the first term can be written as

$$f(\theta/q_i) = \frac{f(q_i/\theta) f(\theta)}{f(q_i)}$$

With some algebraic manipulations, it can be verified that $f(\theta/q_i)$ is distributed normally with

mean $\hat{\mu} + \frac{\hat{\sigma}^2 (q_i - \hat{\mu})}{\hat{\sigma}^2 + \sigma_e^2}$ and variance $\frac{\hat{\sigma}^2 \sigma_e^2}{\hat{\sigma}^2 + \sigma_e^2}$. Now, (A1) can be written as

$$\int_{\theta_i^*}^{2167} f(q_i) f(\theta/q_i) d\theta = \frac{1}{\sqrt{\hat{\sigma}^2 + \sigma_e^2}} \phi\left(\frac{q - \hat{\mu}}{\sqrt{\hat{\sigma}^2 + \sigma_e^2}}\right) \Phi\left[\frac{2167 - \hat{\mu} - \frac{\hat{\sigma}^2(q_i - \hat{\mu})}{\hat{\sigma}^2 + \sigma_e^2}}{\frac{\hat{\sigma}\sigma_e}{\sqrt{\hat{\sigma}^2 + \sigma_e^2}}}\right]$$

This integral is simply a combination of normal density function and normal distribution which can be estimated easily on a computer.

The second term in equation (6) is

$$\int_{2167}^{2167+bp_2} \frac{1}{\sigma_e} \phi\left(\frac{q_i - 2167}{\sigma_e}\right) f(\theta) d\theta \quad (\text{A2})$$

This can be simplified as

$$\frac{1}{\sigma_e} \phi\left(\frac{q_i - 2167}{\sigma_e}\right) \left[\Phi\left(\frac{2167 + bp_2 - \hat{\mu}}{\hat{\sigma}}\right) - \Phi\left(\frac{2167 - \hat{\mu}}{\hat{\sigma}}\right) \right]$$

The third term is analogous to the first one and can be written as

$$\int_{2167+bp_2}^{\infty} f(q_i) f(\theta/q_i) d\theta = \frac{1}{\sqrt{\hat{\sigma}^2 + \sigma_e^2}} \phi\left(\frac{q - \hat{\mu} - bp_2}{\sqrt{\hat{\sigma}^2 + \sigma_e^2}}\right) \left[1 - \Phi\left(\frac{2167 + bp_2 - \hat{\mu} - \frac{\hat{\sigma}^2(q_i - \hat{\mu})}{\hat{\sigma}^2 + \sigma_e^2}}{\frac{\hat{\sigma}\sigma_e}{\sqrt{\hat{\sigma}^2 + \sigma_e^2}}}\right) \right]$$

We can write the similar expression for all the plans with appropriate condition mean $\hat{\mu}$ and conditional variance $\hat{\sigma}^2$.