Reprinted with permission from the *Journal of Marketing Research,* published by the American Marketing Association, Raghuram Iyengar, Asim Ansari, and Sunil Gupta, Vol. XLIV (November 2007), 529–544

RAGHURAM IYENGAR, ASIM ANSARI, and SUNIL GUPTA*

In many services (e.g., the wireless service industry), consumers choose a service plan according to their expected consumption. In such situations, consumers experience two forms of uncertainty. First, they may be uncertain about the quality of their service provider and can learn about it after repeated use of the service. Second, they may be uncertain about their own usage of minutes and learn about it after observing their actual consumption. The authors propose a model to capture this dual learning process while accounting for the nonlinearity of the pricing scheme used in wireless services. The results show that both quality learning and quantity learning are important. The authors conduct several policy experiments to capture the effects of consumer learning, pricing, and service quality on customer lifetime value (CLV). They find that consumer learning can result in a win–win situation for both consumers and firm; consumers leave less minutes on the table, and the firm experiences an increase in overall CLV. For example, the authors find that there is a 35% increase (approximately $75) in overall CLV with consumer learning than without. The key driver of this result is the change in the retention rate with and without learning.

A Model of Consumer Learning for Service Quality and Usage

In many service industries, consumers choose a service plan according to their expected consumption. However, choice of plans and consumption of services differ from the traditionally analyzed choice and quantity decisions of grocery products in two distinct ways. We focus on wireless services to illustrate these differences. First, unlike the situation in products, the quality of a service is difficult to assess because of the large variability inherent in service delivery. For example, the quality of Coke is fairly consistent over time, but the quality of a wireless service may vary depending on customer contact. Therefore, a customer can learn about the true quality of a service provider only after repeated usage. Second, services are perishable. If a consumer buys more Coke than he or she wants to consume in a period, it can be easily stored for future usage. However, except for a few wireless service providers (e.g., Cingular Wireless), if a consumer does not use his or her free minutes for a month, they cannot be carried over to the next month. In other words, consumers observe their usage and, over time, learn about their own consumption. In turn, this can lead them to change their service plan in the future. We model this dual learning process of quality and quantity within a Bayesian learning framework.

Modeling quality uncertainty reduction as a Bayesian learning process has a long history, beginning with studies by Stoneman (1981), Meyer and Sathi (1985), and Roberts and Urban (1988). In marketing, Erdem and Keane (1996) were among the first to model formally consumer uncertainty and learning about grocery product attributes as a Bayesian updating process. More recently, similar Bayesian learning models have been used for other scenarios (Ching 2002; Narayanan, Manchanda, and Chintagunta 2005).

Some researchers have also focused on usage uncertainty. Nunes (2000) shows that as the two consumer decisions of choice of service and consumption of service are separated, consumers are uncertain about how much they might eventually consume. Similarly, Lemon, White, and

*Raghuram Iyengar is Assistant Professor of Marketing, The Wharton School, University of Pennsylvania (e-mail: riyengar@wharton.upenn.edu). Asim Ansari is Professor of Marketing, Columbia Business School, Columbia University (e-mail: maa48@columbia.edu). Sunil Gupta is Edward W. Carter Professor of Business Administration, Harvard Business School, Harvard University (e-mail: sgupta@hbs.edu). This article is based on the first author's doctoral dissertation at Columbia University. The two coauthors are his dissertation advisors. The authors thank the other dissertation committee members (Kamel Jedidi, Yuxin Chen, and K. Sudhir) for their extensive comments and suggestions. The first author thanks Z. John Zhang for insightful discussions. The authors also thank Carl Mela and the Teradata Center at Duke University for providing the data. Finally, the authors thank the Marketing Science Institute for providing financial support for this project.

To read and contribute to reader and author dialogue on JMR, *visit http://www.marketingpower.com/jmrblog.*

ISSN: 0022-2437 (print), 1547-7193 (electronic)

Winer (2002) analyze the effect of consumers' future use of an interactive television entertainment service on whether they will continue the service. Narayanan, Chintagunta, and Miravete (2005) analyze data from an experiment conducted by South Central Bell. They developed a model for plan choice and consumption that incorporates consumers' usage uncertainty. In their experiment, people had a choice between a flat-rate pricing scheme and a two-part tariff.

Service contexts also have another unique aspect—namely, the presence of nonlinear pricing schemes. Within the wireless industry, pricing schemes are typically characterized by an access fee, included free minutes, and a per-minute marginal price for any consumption in excess of the free minutes. Such pricing schemes are termed "increasing block" because the applicable marginal price increases with consumption. Although we focus on the wireless industry, increasing-block schemes are also used for other services. For example, in the electricity and water supply industry, the total charges payable by consumers are based on their consumption in the billing period, and the applied per-unit rates typically increase with increasing consumption[1] (Herriges and King 1994; Maddock, Castano, and Vella 1992; Reiss and White 2005). These kinds of pricing schemes create a simultaneity between price and consumption; the applicable marginal price depends on consumption, and vice versa (Iyengar 2006).

In summary, there are three aspects we address herein: consumer uncertainty about and learning of a service provider's quality, consumer uncertainty about and learning of their own consumption pattern, and nonlinear pricing that creates a simultaneity between price and quantity. Whereas previous research has considered some of these issues, this article addresses all three aspects together. For example, Erdem and Keane (1996) consider quality learning in grocery settings but do not incorporate either usage uncertainty or nonlinear pricing schemes. Erdem, Imai, and Keane (2003) model both quality and quantity learning but do not incorporate nonlinear pricing schemes. Lambrecht, Seim, and Skiera (2005) develop a model for Internet usage under increasing-block tariffs that incorporates usage uncertainty but does not include quality uncertainty. Similarly, Iyengar (2006) analyzes a pricing scheme similar to that herein but does not incorporate quality or usage learning. Finally, Narayanan, Chintagunta, and Miravete (2005) incorporate quantity learning and analyze a two-part tariff pricing scheme. In a two-part tariff scheme, there is no simultaneity of marginal price with consumption. The pricing scheme we use (and our modeling framework) is more general and can incorporate a two-part tariff scheme, flat-fee pricing, and a per-minute pricing as special cases. In addition, our modeling framework can be readily extended to other, more complex increasing-block tariff settings. Other key differences exist between Narayanan, Chintagunta, and Miravete's approach and our specification. We model customer defection (churn), whereas Narayanan, Chintagunta, and Miravete do not. Modeling churn allows us to perform policy simulations that speak toward how consumer learning is beneficial not only for consumers but also for the firm. We derive this and other such substantive implications by using predictions of customer defection from our model and then calculating the overall customer value. In addition, whereas we allow for heterogeneity in all individual-level parameters, Narayanan, Chintagunta, and Miravete assume that certain coefficients (e.g., the price coefficient) are the same for all people in the population.

We propose a model within a Bayesian learning framework. We apply the model to customer-level monthly billing data from a single wireless service provider, and we use hierarchial Bayesian methods to estimate the model. In this data set, we observe consumers' choice of plans, their consumption of minutes, and their decision to leave the service provider.[2] We obtain several notable results. Specifically, we find that past underutilization of free minutes either increases current consumption or influences customers to downgrade their service plans and, in some cases, even leave the service provider. In contrast, past overutilization either decreases current consumption or influences customers to upgrade their service plans. We also find that both quality learning and quantity learning are important aspects of the data. In addition, in our application, consumers learn about service quality rapidly. Indeed, more than 90% of quality learning occurs within the first five service encounters. This suggests that firms need to manage the first few service encounters strategically.

We then use policy experiments to investigate the effects of consumer learning. We find that consumer learning can result in a win–win situation for both consumers and the firm; consumers leave less minutes on the table, and the firm experiences an increase in overall customer lifetime value (CLV). In particular, we estimate that there is approximately a 35% increase in CLV (approximately $75) in the presence of consumer learning. The key driver of this difference is the change in the retention rate with and without consumer learning. We also perform simulations that relate service quality to CLV. We determine that, on average, a 1% increase in mean service quality leads to approximately a $2 increase in CLV. Because the service provider has 21 million customers, this increase in mean quality results in an overall long-term increase in profit of approximately $42 million. Finally, policy experiments related to pricing show that changes in access fees have a significantly greater influence on the CLV of "light users." The primary contributor to this result is the change in retention rate of light users following a change in the access fee. The retention rate of heavier users is less affected by such price variations.

We organize the rest of the article as follows: We begin with an outline of the dual consumer learning process and describe nonlinear pricing schemes. Next, we develop a model that incorporates nonlinear pricing and consumer learning about service quality and consumption quantity. We then describe the data. Thereafter, we develop three competing models. We then discuss estimation results and

[1]If the pricing scheme includes a single marginal price, the scheme is the two-part tariff (Danaher 2002).

[2]Because these data are from a single service provider, we cannot distinguish between customers who defect to another provider and those who leave the wireless service category altogether.

report the results of policy experiments. We conclude with our contributions and key results.

CONSUMER LEARNING AND NONLINEAR PRICING STRUCTURES

Dual Learning

We assume that consumers make a decision at the beginning of each period; they either choose a plan (i.e., remain on the current plan or change plans) or leave the company. At the beginning of each period, before choosing a plan, consumers have prior beliefs about the service quality and their consumption quantity over that period. They use these beliefs to assess the expected utility from each service plan and then choose the plan that yields the maximum expected utility. They defect if the expected utility from defection is higher than the expected utility from any of the plans.

After choosing a plan, consumers receive two signals. First, they receive a noisy signal of the service provider's quality (quality signal). This signal might come from their service encounter with a customer service representative. It is noisy because just a few encounters with the provider do not inform consumers about its true quality. Second, because the two decisions of plan choice and consumption are temporally separated, consumers observe their actual usage at the end of the month (usage signal). They use these two signals to update their uncertainty. The quality signal is used to learn about the service quality of the provider. Similarly, consumers use the usage signal to learn about the distribution of their consumption quantity. Subsequently, they use these updated quality and quantity beliefs to choose plans in the subsequent month. Figure 1 describes this process in a flowchart. Subsequently, we formalize this conceptual framework in a Bayesian learning model.

Nonlinear Pricing Structures

Nonlinear pricing structures are characterized by a fixed fee and a set of marginal prices. The fixed fee is the amount paid to access the service. If the pricing scheme includes multiple marginal prices, the price for consuming an additional unit of service depends on the total consumption. There are two main forms of nonlinear pricing schemes: increasing-block schemes and decreasing-block schemes. In an increasing-block structure, the marginal prices increase with consumption, and in a decreasing-block scheme, the marginal prices decrease with consumption.

Consider a service that has a two-tier increasing-block pricing structure characterized by a fixed fee and two marginal prices. In Figure 2, which graphically depicts this scheme, F represents the access price for the service. The applicable marginal price changes when the consumption exceeds the kink point (A). The marginal price, p_1, for consuming an additional unit before the kink is less than the marginal price, p_2, for consuming after the kink.

Consumers, however, do not make the decision to use a service in isolation from their other consumption decisions. At any point in time, they have several consumption opportunities, and they allocate their income among these opportunities. This trade-off across goods can be appropriately represented using a budget set representation. Figure 3 depicts a budget set that corresponds to an increasing two-tier pricing scheme. The vertical axis in the figure corresponds to the consumption of the outside good (z), and the horizontal axis corresponds to the consumption of units of the service (x).

Figure 1

OUTLINE OF THE LEARNING PROCESS

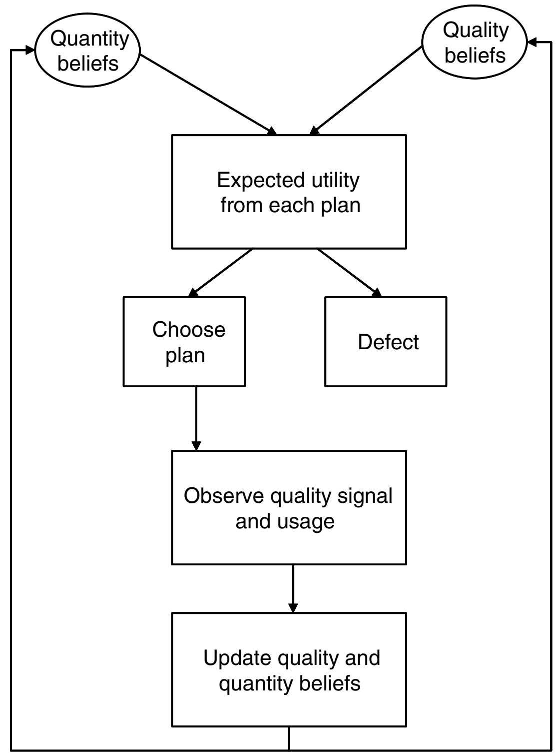
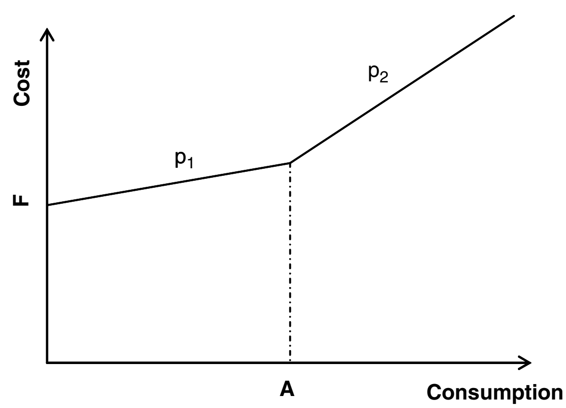


Figure 2

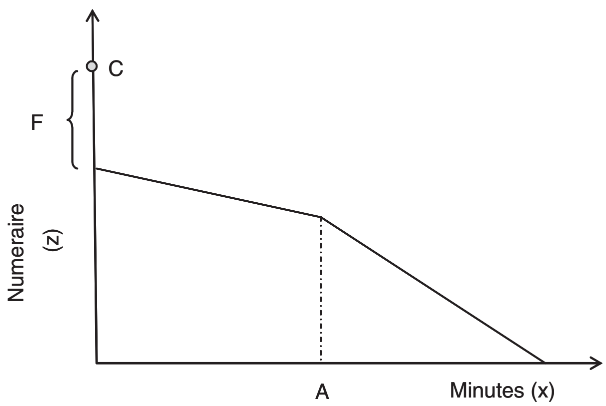
A TWO-TIER INCREASING-BLOCK PRICING SCHEME



Notes: F refers to the access fee, A is the kink point (free minutes), and p_1 and p_2 are the marginal prices before and after the kink point, respectively.

Figure 3 shows that the two-tier increasing-block pricing structure of the service results in a piecewise linear budget set with a kink point (A). A consumer who subscribes to the service faces a convex budget set, and his or her income (I) is lowered by the sum of the access fee (F) and the variable

Figure 3
A BUDGET SET REPRESENTATION OF A TWO-TIER INCREASING-BLOCK PRICING SCHEME



Notes: F refers to the access fee, and A is the kink point (free minutes). If the consumer does not subscribe to a plan, the total income I is used for consuming the outside good. Point C represents this situation.

charges for any consumed service. However, if he or she does not subscribe to the service, the entire income is used for consuming the outside good. This is represented by the point C on the vertical axis. If the marginal price of the outside good is normalized to 1 (numeraire), the following equations represent the piecewise budget set:

$$p_1x + z \leq I - F, \text{ if } x > 0 \text{ and } x \leq A, \text{ and} \tag{1}$$

$$p_2(x - A) + z \leq I - F - p_1A, \text{ if } x > A. \tag{2}$$

A restricted form of such a two-tier increasing-block pricing scheme is widely used in the wireless communications industry, where p_1 is 0. Therefore, consuming an additional minute before the kink point is costless.

This pricing scheme and the kinked nature of the budget set that ensues raise interesting econometric issues for any demand analysis. In the current setting, there is a simultaneity between the applicable marginal price and consumption; that is, the level of consumption determines the applicable marginal price, but at the same time, the pricing scheme influences the level of consumption.[3] Next, we describe a model that includes the dual learning process, captures consumers' choice and consumption decisions when they face nonlinear pricing schemes, and addresses the simultaneity between marginal price and consumption.

MODEL

We assume that consumers make a discrete choice decision at the beginning of each period; they either choose a single plan (i.e., remain with the current plan or change plans) or terminate the service. At the beginning of each period, before making a decision, consumers have prior beliefs about the service quality and their consumption quantity over that period. They use these beliefs to assess the expected utility from each service plan. They defect if the expected utility from defection is higher than the expected utility from any of the plans. Otherwise, they choose the plan that yields the maximum expected utility. They subsequently consume under that chosen plan. We begin by specifying the direct utility for each plan. Next, we describe how expected utility is computed and specify the belief distributions associated with service quality and consumption. We then show how the belief distributions evolve because of consumer learning. Finally, the model discussion concludes with the specification of an estimable model.

Utility Function

Let U_{ijt} be the direct utility function for a consumer i for consuming x_{ijt} minutes under plan j and a quantity z_{ijt} of the numeraire commodity during period t. We specify U_{ijt} as

$$U_{ijt}(x_{ijt}, z_{ijt}) = \beta_{ij} + S^e_{it} + \alpha_{i1}x_{ijt} + \alpha_{i2}z_{ijt} + \alpha_{i3}x^2_{ijt} + (\mathbf{r}'_{ijt}\boldsymbol{\gamma}_i)x_{ijt} + \mathbf{s}'_{ijt}\boldsymbol{\zeta}_i + \varepsilon_{ijt}. \tag{3}$$

The term β_{ij} represents an individual- and plan-specific intercept, and S^e_{it} is the noisy signal about the service provider's quality that consumer i receives at time t. We assume that the service quality is invariant across plans. The parameter α_{i1} represents the main effect of consumption of minutes, and α_{i2} represents the effect of consuming a unit of the numeraire. The term α_{i3} captures the effect of differential marginal impact of consuming an additional minute.[4] The vectors $\mathbf{r}_{ijt}$ and $\mathbf{s}_{ijt}$ contain consumer-, plan-, and time-specific covariates. The covariates in $\mathbf{r}_{ijt}$ affect the direct utility through their interaction with the consumed quantity (x_{ijt}), whereas the covariates in $\mathbf{s}_{ijt}$ have only a main effect. These two vectors can share variables because a covariate (e.g., prior consumption) can have both a direct and an interactive effect. The parameter vectors $\boldsymbol{\gamma}_i$ and $\boldsymbol{\zeta}_i$ contain individual-specific sensitivities to the covariates in $\mathbf{r}_{ijt}$ and $\mathbf{s}_{ijt}$, respectively. Finally, the random errors that are unobservable to the researcher but are known to the consumer are contained in ε_{ijt}. For example, these errors can contain any plan-specific promotional activities that are unknown to the researcher but influence consumers to choose a certain plan. We assume that these errors are normally distributed.

We also use Equation 3 to specify the utility associated with defection (churn). This is the utility the consumer receives if he or she does not choose any of the available plans. Thus, the minutes consumed are 0 (because the consumer leaves the company and does not choose any service plan), and the entire income for consumer i, denoted by I_i,

[3]A similar form of simultaneity is also present in the relationship between nonlinear income taxes and hours of work (for a discussion of the issues and methods to address such simultaneity, see Hausman 1985; Moffitt 1990).

[4]A logarithmic specification can also capture diminishing utility. We chose a quadratic specification because it is easy to impose the Slutsky constraints at the individual level (the Appendix shows these constraints for our model). A logarithmic specification results in complex nonlinear Slutsky constraints on the individual-level parameters.

is spent on other goods. Then, we can insert a value of 0 for consumed minutes and I_i for the numeraire commodity into Equation 3 to obtain the following expression:

$$(4)\qquad U_{ict}(x_{ict} = 0,\ z_{ict} = I_i) = \beta_{ic} + \alpha_{i2}I_i + \mathbf{s}'_{ict}\boldsymbol{\zeta}_i + \varepsilon_{ict}.$$

In Equation 4, the churn option is denoted as c, and the utility associated with defection for consumer i at time t is denoted as U_{ict}. The vector $\mathbf{s}_{ijt}$ contains any churn-specific covariates, and the error ε_{ict} includes any unobserved influences specific to the exit decision that are known to the consumer but are unknown to the researcher. As an example, this error can contain any competitive promotional activities that are unknown to the researcher but influence consumers to leave the service provider. We assume that ε_{ict} is normally distributed as well.

Expected Utility and Prior Beliefs

Note that consumers are uncertain about two components in the direct utility function (Equation 3) when they make a plan choice decision at time t. First, the quality signal, S^e_{it}, is revealed to a consumer only after he or she chooses a plan. Second, because plan choice and consumption are temporally separated, the consumption, x_{ijt}, occurs subsequent to the choice of plan j and is uncertain as well.

At any point in time, consumers have belief distributions about these two unknowns. Before choosing a plan, consumers use these beliefs to assess the expected utility associated with the different plans. The expected utility that consumer i associates with plan j at time t can be written as

$$(5)\qquad EU_{ijt} = E^t_{qual}\left\{E^t_{usage}\left[U_{ijt}(x_{ijt},\ z_{ijt})\right]\right\}.$$

The two expectation operators, $E^t_{qual}[\cdot]$ and $E^t_{usage}[\cdot]$, denote the expectation with respect to the quality beliefs and quantity beliefs at time t, respectively.

We incorporate the direct utility in Equation 5 and rewrite Equations 3 and 5 as

$$(6)\qquad EU_{ijt} = E^t_{qual}\left[S^e_{it}\right] + E^t_{usage}\left[g(x_{ijt},\ z_{ijt})\right] + \beta_{ij} + \mathbf{s}'_{ijt}\boldsymbol{\zeta}_i + \varepsilon_{ijt},\ \text{and}$$

$$g(x_{ijt},\ z_{ijt}) = \alpha_{i1}x_{ijt} + \alpha_{i2}z_{ijt} + \alpha_{i3}x^2_{ijt} + (\mathbf{r}'_{ijt}\boldsymbol{\gamma}_i)x_{ijt}.$$

Here, we assume that quality and quantity beliefs are independent. Thus, we can separate the components that involve quality uncertainty, quantity uncertainty, and no uncertainty. The first term, $E^t_{qual}[S^e_{it}]$, represents the expectation of the quality signal that a consumer i will receive after choosing a plan at time t. The term, $E^t_{usage}[g(x_{ijt},\ z_{ijt})]$, captures the effect of quantity uncertainty on the overall expected utility. Finally, the parameters β_{ij}, $\boldsymbol{\zeta}_i$, and $\boldsymbol{\gamma}_i$ and the error ε_{ijt} are known to the consumer. Next, we specify the quality and quantity beliefs.

Quality beliefs. We assume that the experienced quality signal, S^e_{it}, comes from a normal distribution, $N(\mu,\ \delta^2)$, where μ is the true quality of the service provider and δ^2 represents the variance in quality. Thus, S^e_{it} is a noisy signal about the underlying true service quality μ. Consumers know that the experienced quality comes from a normal distribution but are unaware of the true quality of the provider (Erdem and Keane 1996). Thus, they have a belief distribution over μ given by

$$(7)\qquad f^{qual}_{it}(\mu) = N(m_{it},\ \sigma^2_{it}),$$

where m_{it} is the mean of this belief distribution for consumer i at time t and σ^2_{it} is the variance that captures his or her uncertainty about μ. We further assume that these quality beliefs are independent from beliefs about consumption. Given our assumptions about the quality signal and the quality beliefs, the expectation for the quality signal at any time is given by

$$(8)\qquad E^t_{qual}[S^e_{it}] = E^t_{qual}[\mu] = m_{it}.$$

Subsequently, we describe how the parameters m_{it} and σ^2_{it} of the belief distribution are updated in every period.

Quantity beliefs. The quantity x_{ijt} that can be consumed under a plan j at time t is random from the viewpoint of consumer i because many factors that are not under his or her control can affect it. We model this quantity in terms of a systematic component, $E(x_{ijt})$ and a random component $\eta_{ijt} \sim N(0,\ \tau^2_{ij})$. Thus, for each plan j, we have an individual-specific and time-varying distribution from which consumption quantities are realized; that is,

$$(9)\qquad x_{ijt} \sim N\left[E(x_{ijt}),\ \tau^2_{ij}\right].$$

The systematic component is known to the consumer and is obtained from maximizing the direct utility in Equation 3 subject to the nonlinear pricing constraints imposed by plan j. Specifically, the optimization process for obtaining the expected quantity $E(x_{ijt})$ can be written as

$$(10)\qquad \max_x U_{ijt}\left[x,\ z(x)\right]$$

subject to

$$\text{Constraint I: } p_{1j}x + z = I_i - F_j,\ \text{if } 0 < x \le A_j,$$

$$\text{Constraint II: } p_{2j}(x - A_j) + z = I_i - F_j - p_{1j}A_j,\ \text{if } A_j < x < B,$$

where F_j is the access fee; A_j is the kink point; p_{1j} and p_{2j} are the two marginal prices for consumption below and above the kink point, respectively; and I_i is the income of consumer i. The two constraints characterize the piecewise linear budget set, and B is an upper bound on consumption and often corresponds to the physical limit on the total consumption that can occur in a period. The Appendix details the conditions for a unique optimal and also outlines how the optimal can be computed. Note that the computation of $E(x_{ijt})$ is conditioned on the other individual-specific parameters and covariates in the direct utility function. Because these individual-specific parameters and covariates are known to the consumer, the plan-specific expected consumption can be determined.

We further assume that the consumer does not know the variance τ^2_{ij} for the random component and learns about the variance from observing his or her actual usage over time. Let $f^{\tau}_{ijt}(\tau^2_{ij})$ denote the belief of consumer i at time t about the usage variance under a plan j. We assume the following inverse gamma distribution for this belief:

$$(11) \qquad f^{\tau}_{ijt}(\tau^2_{ij}) \sim IG(a_{ijt}, b_{ijt}),$$

where a_{ijt} and b_{ijt} are consumer-, plan-, and time-specific parameters.

The overall beliefs about usage x_{ijt}, denoted by $f^{usage}_{ijt}(x)$, can be obtained by combining the uncertainty of usage conditioned on the variance τ^2_{ij} with the uncertainty in the variance itself. We can represent this as

$$(12) \qquad x_{ijt} \sim f^{usage}_{ijt}(x) = \int_0^{\infty} N\Big[E(x_{ijt}),\ \nu\Big] f^{\tau}_{ijt}(\nu)d\nu.$$

Thus, $f^{usage}_{ijt}(x)$ is a t-distribution with a mean of $E(x_{ijt})$, a variance of $(a_{ijt}\, b_{ijt})^{-1}$, and degrees of freedom of $2a_{ijt}$. This distribution implies that a consumer's beliefs about his or her usage in a given month for a plan j are anchored on $E(x_{ijt})$, which comes from the previously described rational optimization process. Recall that this optimization process incorporates the budget constraints imposed by the nonlinear pricing scheme of a plan.

Using the quantity belief distribution, $f^{usage}_{ijt}(x)$, for plan j and its budget constraints, we can compute the component, $E^t_{usage}[\cdot]$, of the overall expected utility (Equation 6). The budget constraints for the plan impose a relationship between the consumed minutes (x_{ijt}) and the numeraire (z_{ijt}), as we show in Equation 10. For example, if Constraint I holds, then $z_{ijt} = I_i - F_j - p_{1j}x_{ijt}$. Similarly, if Constraint II holds, then $z_{ijt} = I_i - F_j - p_{1j}A_j - p_{2j}(x_{ijt} - A_j)$. In other words, we can rewrite $g(x_{ijt}, z_{ijt})$ in Equation 6 as a function of x_{ijt} only. Let $g(x_{ijt}, z_{ijt})$ be denoted by $h_1(x_{ijt})$ if $x_{ijt} \le A_j$ and by $h_2(x_{ijt})$ if $x_{ijt} > A_j$. The quantity expectation is as follows:

$$(13) \qquad E^t_{usage}[g(x_{ijt}, z_{ijt})] = \int_0^{Aj} h_1(x) f^{usage}_{ijt}(x)dx + \int_{Aj}^{B} h_2(x) f^{usage}_{ijt}(x)dx.$$

Overall expected utility. After substituting the two components $E^t_{qual}[\cdot]$ (Equation 8) and $E^t_{usage}[\cdot]$ (Equation 13) into the overall expected utility (Equation 6), we obtain the following:

$$(14) \qquad EU_{ijt} = m_{it} + \int_0^{Aj} h_1(x) f^{usage}_{ijt}(x)dx + \int_{Aj}^{B} h_2(x) f^{usage}_{ijt}(x)dx + \mathbf{s}'_{ijt}\boldsymbol{\zeta}_i + \beta_{ij} + \varepsilon_{ijt}.$$

Using the expected utilities for the plans, consumers choose a plan and subsequently consume minutes or terminate the service.

Consumer Learning

Updating of quality beliefs. The prior beliefs that a consumer i has about the true quality, μ, at time t is given by Equation 7. The experience signal, S^e_{it}, obtained after the consumer chooses a plan is normally distributed with mean μ and variance δ^2. We can then use Bayes' theorem to specify the posterior distribution for quality beliefs. Thus,

$$(15) \qquad f^{qual}_{it+1}(\mu) = N(m_{it+1}, \sigma^2_{it+1}),$$

$$\sigma^{-2}_{it+1} = \sigma^{-2}_{it} + \sigma^{-2}, \text{ and}$$

$$m_{it+1} = \sigma^2_{it+1}\Big[\sigma^{-2}_{it} m_{it} + \delta^{-2} S^e_{it}\Big].$$

The posterior belief distribution, $f^{qual}_{it+1}(\mu)$, then becomes the prior belief distribution for time t + 1.

Updating of quantity beliefs. We assume that consumers are uncertain about the consumption variance, τ^2_{ij}. For a consumer i at time t and plan j, the prior belief about the usage variance is represented by $f^{\tau}_{ijt}(\tau^2_{ij})$ (Equation 11). Suppose that a consumer chooses plan j*; then, the actual quantity x^{actual}_{it} consumed in that month is given by

$$(16) \qquad x^{actual}_{it} = E(x_{ij^*t}) + \eta_{ij^*t}.$$

On observing the actual quantity, the consumer knows the overall deviation (η_{ij^*t}) that has occurred. The consumer then updates his or her uncertainty associated with the usage variance for the chosen plan j*. Thus,

$$(17) \qquad f^{\tau}_{ij^*t+1}(\tau^2_{ij^*}) = IG(a_{ij^*t+1}, b_{ij^*t+1}),$$

$$a_{ij^*t+1} = a_{ij^*t} + \frac{1}{2}, \text{ and}$$

$$b^{-1}_{ij^*t+1} = b^{-1}_{ij^*t} + \left[\frac{(\eta_{ij^*t})^2}{2}\right].$$

This updating stems from the conjugacy of the inverse gamma prior with a Gaussian likelihood for the actual quantity. Note that if a plan is not chosen, the belief distribution for the consumption variance associated with that plan remains unchanged.

Note that in our model, usage uncertainty is distinct from usage variability. Usage variability is captured by the variance (τ^2_{ij}) of the usage shock (η_{ijt}). Different usage shocks over different periods result in the usage being variable over time (Equation 9). We assume that consumers do not know both η_{ijt} and τ^2_{ij} and that this lack of knowledge leads to usage uncertainty. We model this usage uncertainty by assuming that consumers have a belief distribution about the variance (τ^2_{ij}), which is updated on the basis of the observed quantity (see Equations 11 and 17). In turn, this belief distribution on variance induces another belief distribution over the quantity that could be consumed under a plan (Equation 12).

In summary, consumers know the choice errors (ε_{ijt}) before choosing a plan. However, the consumer is uncertain about the service quality and quantity that he or she will consume in the following month. The consumer uses the two belief distributions to compute the expected utility for each plan. The quality signal (S^e_{it}) and the demand shocks (η_{ijt}) are revealed to the consumer after the plan choice. These are then used to update the quality and quantity belief distributions, respectively.

Thus far, we have developed the model from the consumers' perspective. However, there are a few aspects of the model that are unobservable to the researcher. Next, we consider the restricted information set of the researcher and show how the model can be estimated.

Estimable Model

In the model, both consumers and the researcher observe monthly usage, x_{it}^{actual}. However, the quality signal, S_{it}^{e}, that consumer i experiences after choosing a plan is unknown to the researcher. For example, only the consumer knows the nature of the service contact after choosing the plan. This asymmetry in the information set must be accounted for when estimating the model.

Recall that from a consumer's perspective, the evolution of the mean of the quality belief from time t to time t + 1 occurs deterministically as $m_{it+1} = \sigma_{it+1}^{2}[\sigma_{it}^{-2}m_{it} + \delta^{-2}S_{it}^{e}]$. However, because the researcher does not observe S_{it}^{e}, this evolution is stochastic from his or her perspective. This stochastic evolution can be specified as follows: We assumed that S_{it}^{e} comes from a normal distribution with mean μ and variance δ^2. Thus, let $S_{it}^{e} = \mu + \kappa_{it}$, where $\kappa_{it} \sim N(0, \delta^2)$. We can then rewrite the equation relating m_{it+1} to m_{it} as

$$(18) \qquad m_{it+1} = \sigma_{it+1}^{2}\left[\sigma_{it}^{-2}m_{it} + \delta^{-2}(\mu + \kappa_{it})\right].$$

Let w_{1it} be $\sigma_{it+1}^{2}/\sigma_{it}^{2}$ and w_{2it} be $\sigma_{it+1}^{2}/\delta^2$; then, we can express the stochastic transition equation as

$$(19) \qquad m_{it+1} = w_{1it}m_{it} + w_{2it}\mu + \chi_{it}, \text{ and}$$
$$\chi_{it} = w_{2it}\kappa_{it} \sim N(0,\ w_{2it}^{2}\delta^2).$$

We can now use this specification of the transition equation to write the complete model:

$$(20) \qquad EU_{ijt} = m_{it} + \int_{0}^{Aj} h_1(x)f_{ijt}^{x}(x)dx + \int_{Aj}^{B} h_2(x)f_{ijt}^{x}(x)dx$$
$$+ s_{ijt}'\zeta_i + \beta_{ij} + \varepsilon_{ijt},$$
$$E(x_{ijt}) = \arg\max_{x} U_{ijt}\left[x,\ z(x)\right]$$
$$\text{subject to}$$
$$\text{Constraint I: } p_{1j}x + z = I_i - F_j, \quad \text{if } 0 < x \le A_j,$$
$$\text{Constraint II: } p_{2j}(x - A_j) + z = I_i - F_j - p_{1j}A_j,$$
$$\text{if } A_j < x < B,$$
$$\varepsilon_{ijt} \sim N(0,\ \lambda_j^2),$$
$$m_{it+1} = w_{1it}m_{it} + w_{2it}\mu + \chi_{it},$$
$$\chi_{it} \sim N(0,\ w_{2it}^{2}\delta^2),$$
$$x_{it}^{actual} = E(x_{ij^*t}) + \eta_{ij^*t}, \text{ and}$$
$$\eta_{ij^*t} \sim N(0,\ \tau_{ij^*}^{2}).$$

Here, j* in the quantity equation indexes the chosen plan, and x_{it}^{actual} is the consumption under this plan. We use the nonlinear pricing constraints of a plan to determine the expected consumption $[E(x_{ijt})]$ under the plan. (Further details appear in the Appendix.) We can interpret this model as an example of a general state–space model (Harvey 1989; West and Harrison 1997), in which the expected utility equation acts as the observation equation and the equation describing the stochastic evolution of the mean of the quality beliefs is the transition equation.

Model identification. In the preceding model, there are two intercept terms associated with the utility of each plan. One is a plan-specific intercept, β_{ij}, and the other is a common intercept, m_{it}. In addition, we have an intercept that is associated with the utility of defection, β_{ic}. This specification is different from that in a standard multinomial choice model, in which there are only alternative-specific intercepts with no common intercept. Therefore, for model identification, setting the intercept of only one of the alternatives to zero is not sufficient. Here, we need to fix the intercepts of two of the alternatives to zero. Thus, we set the intercept in the utility of churn, β_{ic}, and the intercept in the utility of any one of the plans, β_{ij}, to zero. There are four plans in the data set (which we discuss subsequently), and we fix the intercept of Plan 4 to zero. In addition, to set the scale of the utilities, we set the variance of the churn utility to one. The variation in a customer's consumed quantity across the different months and the choice of different plans helps in the identification of the individual-specific parameters, α_{i1}, α_{i2}, and α_{i3}. In addition, the variation in the consumed minutes aids in the identification of the parameters of usage variation (τ_{ij}^{2}) as well as in the updating of usage uncertainty.

Until now, we focused on one consumer and showed how the model can be specified. Next, we incorporate heterogeneity across consumers.

Heterogeneity specification. The model contains several individual-level coefficients. Let the vector $\boldsymbol{\omega}_i$ contain the coefficients (α_{i1}, ψ_{i2}, ψ_{i3}, $\boldsymbol{\gamma}_i$, $\boldsymbol{\zeta}_i$, $\{\boldsymbol{\beta}_i\}$), where ψ_{i2} and ψ_{i3} are individual-specific parameters such that $\exp(\psi_{i2}) = \alpha_{i2}$ and $-\exp(\psi_{i3}) = \alpha_{i3}$ (see the Appendix). We specify the heterogeneity across consumers by assuming that $\boldsymbol{\omega}_i$ is normally distributed in the population. Thus,

$$(21) \qquad \boldsymbol{\omega}_i \sim N(\boldsymbol{\psi}_{\omega},\ \boldsymbol{\Lambda}_{\omega}).$$

Here, the vector $\boldsymbol{\psi}_{\omega}$ contains population-level coefficients, and $\boldsymbol{\Lambda}_{\omega}$ is the population covariance matrix.

Apart from $\boldsymbol{\omega}_i$, the variance τ_{ij}^{2} of the usage error, η_{ijt}, is also individual specific. For each plan j, a heterogeneity distribution over τ_{ij}^{2} can be specified as an inverse gamma distribution. Thus,

$$(22) \qquad \tau_{ij}^{2} \sim IG(c_j,\ d_j).$$

Note that this population distribution is distinct from the belief distribution of consumers about their consumption variance. In other words, this heterogeneity distribution is across consumers as opposed to the previously specified belief distribution within a consumer.

We adopt a Bayesian framework for simulation-based inference. The Web Appendix (see http://www.marketingpower.com/content84060.php) contains the details of the priors for the unknowns and the full conditionals.

DATA

The data set contains information of subscriber-level monthly billing records and promotions sent by a wireless service provider. The billing data are from September 2001 to May 2003.[5] New customers joined between August 2001 and December 2001. There is no left truncation in the data (i.e., we have billing data for customers from their first month onward).

The monthly billing data provide information on the customer's current service plan and his or her monthly consumption of minutes. Typically, the included minutes are of two types: included off-peak/weekend minutes and included peak minutes (also called the plan allowance). Usually, off-peak/weekend minutes are free. Peak minutes are charged according to a two-tier pricing scheme similar to that shown in Figure 2, except that the marginal price p_1 of consuming before the kink point is zero. For a given plan, this kink point represents the plan allowance. All the plans in our data set have such a two-tier pricing scheme for the consumption of peak minutes.

In this article, we model the usage of peak minutes. The analysis is restricted to four types of service plans that differ in their included peak number of minutes—namely, 200, 300, 350, and 500 minutes. These service plans are available with many different features. For example, a 200-peak-minutes plan can have short messaging system capability or not. In the data set, we determined that there is negligible use of features such as short messaging system, long distance, and roaming. Therefore, we ignore these differences and characterize a plan by three features: the included peak minutes, the access fee, and the marginal price for consumption exceeding the included minutes. Table 1 describes these characteristics for the four types of plans. The plans are numbered in an ascending order of included peak minutes. We refer to a plan switch from a lower plan to a higher plan as an "upgrade" and a switch from a higher plan to a lower plan as a "downgrade." Together, these four types of plans account for more than 70% of the chosen plans in the data.

For the analysis, we consider subscribers who switch only among the four types of service plans and have no contractual relationship with the company. We also have one other restriction primarily to increase the signal from plan switching. As we mentioned, each plan in the data set is described by several attributes, such as long distance, roaming, and data transfer availability, that are either present or absent. We choose consumers who changed at least one option within a plan or the plan itself within their entire tenure period. The first selection criterion of considering customers who switch only among the four plans differs from that used in prior scanner panel research, in which brands not focal to the analysis are typically aggregated into an "other" brand. In our context, because each service plan has a nonlinear pricing scheme that results in a specific nonlinear budget set, it is not clear how such an "other" service plan can be specified. The second criterion

[5]The usage data for one month were missing. For each customer, we replaced the consumption in the missing month by the average of his or her usage in a month before and after.

Table 1
DESCRIPTIVES OF DATA

Variables	*Plan 1*	*Plan 2*	*Plan 3*	*Plan 4*
Number of free minutes	200	300	350	500
Access price ($)	30	35	40	50
Marginal price ($/minute)	.40	.40	.40	.40
Share (%)	47.36	9.92	32.1	10.62
Consumption (minutes/month)				
M	147.72	190.35	215.14	318.9
Variance	107.84	109.93	148.36	199.93
Underage (minutes/month)	73	118	150	200
Overage (minutes/month)	20	10	15	19
Past Underage Revenue ($)				
M	8.6	14.54	18.73	29.48
Variance	10.14	11.8	13.18	14.75
Past Overage Revenue ($)				
M	19.5	8.95	5.76	1.42
Variance	38.26	26.53	21.47	10.53
Cumulative Underage Revenue ($)				
M	93.44	152.05	193.59	298.94
Variance	96.75	130.44	156.06	212.87
Cumulative Overage Revenue ($)				
M	168.76	72.36	44.37	9.7
Variance	287.89	161.91	116.7	38.61
Promotion Dummy				
M	.15			
Variance	.36			
Minutes Consumed/Month				
M	191.78			
Variance	144.28			

of selecting customers with no contractual relationship ensures that customers are free to change their service plans at the beginning of every month.

There were 3010 customers who satisfied the selection criteria. We randomly chose 300 customers to form the calibration data set. The calibration data set has 5281 observations, each representing monthly bill information. On average, we have 17 months of data for each customer. Of these 17 months, on average, customers experienced 14 months (82%) of underage and 3 months (18%) of overage, which is consumption in excess of free minutes. Table 1 also shows the market shares of the four types of plans, the average monthly consumption of minutes, its variance, and the underage and overage in the plans. The table shows that consumption variance is different across plans. This is consistent with our model, in which the consumption variance, τ^2_{ij}, is plan specific. The average underage and overage results show that consumers experience more underage than overage.

On average, a customer changed his or her plan once during his or her tenure, and approximately 17% of customers never changed plans. In addition, of the total number of plan changes, 61% were upgrades (i.e., a customer moved from a plan with lower number of minutes to one with higher number of minutes), and the rest (39%) were downgrades. The churn rate is approximately 27%; 82 of the 300 people left the company during the data period. Of these, 46% left after being on Plan 1 as their last chosen plan, and approximately 50% were on Plan 3 and Plan 4 before leaving the company (the remaining were on Plan 2).

Variables

We define the following variables that describe promotional activities and past consumption dynamics:

- Promotion (Prom): The data set contains information on several promotions, such as free roadside assistance, a Valentine's Day promotion, free accessories, and dot-com back-to-school credit. Because there is considerable variability in the types of promotions, we abstract away from these differences and capture the overall impact of a promotional event by using a dummy variable that records whether any promotion was sent to a customer in a given month.
- State dependence (State_Dep): Several researchers in marketing and economics (Heckman 1981; Seetharaman, Ainslie, and Chintagunta 1999) have established that previous choices significantly affect current decisions. In our context, subscribers have no contractual relationship with the service provider, and yet they might show a tendency to retain the current plan. To capture this effect, for each plan, we create a variable that takes a value of 1 if plan j was chosen in period t – 1 and 0 if otherwise.
- Past quantity (Past_Qty): The data set has a record of consumers' consumption decisions. At each time t, we use the lagged quantity consumed to capture these consumption effects.

In each month, customers choose only one service plan and then either underconsume or overconsume relative to the plan's allowance. As we mentioned previously, the unconsumed minutes in a month are called underage, and consumption above the included minutes is called overage. Although overage and underage are known for the chosen plan, they are estimated for the unchosen plans as follows: At time t, we use the actual consumption for the consumer at time t – 1 to calculate the underage or overage if he or she had that consumption under plan j. Note that for plan j at a time t, only one of these (i.e., underage or overage) is positive, and the other is zero. We use these two measures to create the following variables:

- Past underage revenue (Past_Under_Rev): This variable represents the counterfactual underage revenue the consumer leaves on the table. We compute this by multiplying the underage associated with plan j with the cost per minute of plan j.
- Past overage revenue (Past_Over_Rev): This variable represents the overage revenue and is computed analogously to the corresponding underage variable.
- Cumulative underage revenue (Cum_Under_Rev): This variable is the sum of past underage revenues from month 1 to month t – 1. We include this variable to characterize the overall penalty associated with a plan j if the consumer had chosen that plan from his or her first period and retained that plan.
- Cumulative overage revenue (Cum_Over_Rev): This variable is the sum of past overage revenues from month 1 to month t – 1.

Recall that for consumer i at time t, the plan-specific covariates for plan j enter the direct utility (Equation 3) either as an interaction through $\mathbf{r}_{ijt}$ or as a main effect through $\mathbf{s}_{ijt}$. All the variables we described are included in the vector $\mathbf{r}_{ijt}$. We specify the vector $\mathbf{s}_{ijt}$ to contain only the variables associated with plan-specific underage and overage—namely, past underage revenue, past overage revenue, cumulative underage revenue, and cumulative overage revenue. We do not include the promotion variable and the past quantity in $\mathbf{s}_{ijt}$ for model parsimony. Each inclusion leads to four additional parameters because these variables are the same across plans, and identification would then require a different parameter for each plan. In addition, $\mathbf{s}_{ict}$, which contains covariates that affect the utility of defection, is a zero vector. Table 1 contains the descriptive statistics for the variables.

NULL MODELS

Null Model 1: No Learn (No Quality or Quantity Learning)

In this model, consumers are assumed to know the true service quality as well as their consumption distribution. Thus, there is no consumer learning. We specify the model as follows:

$$
\begin{aligned}
(23)\qquad EU_{ijt} &= \beta_{ij} + m_i + \alpha_{i1}E(x_{ijt}) + \alpha_{i2}(z_{ijt}) + \alpha_{i3}\left[E(x_{ijt})\right]^2 \\
&\quad + (\mathbf{r}'_{ijt}\boldsymbol{\gamma}_i)E(x_{ijt}) + \mathbf{s}'_{ijt}\boldsymbol{\zeta}_i + \varepsilon_{ijt}, \\
z_{ijt} &= I_i - F_j - p_{1j}E(x_{ijt}), \\
&\quad \text{if } E(x_{ijt}) \geq 0 \text{ and } E(x_{ijt}) \leq A_j, \\
z_{ijt} &= I_i - F_j - (p_{2j} - p_{1j})A_j - p_{2j}E(x_{ijt}), \\
&\quad \text{if } E(x_{ijt}) > A_j \text{ and } E(x_{ijt}) < B, \\
\varepsilon_{ijt} &\sim N(0, \lambda^2_j), \\
x^{actual}_{it} &= E(x_{ij^*t}) + \eta_{ij^*t}, \text{ and} \\
\eta_{ij^*t} &\sim N(0, \tau^2_{ij^*}).
\end{aligned}
$$

Here, j* in the quantity equation indexes the chosen plan, x_{it}^{actual} is the quantity consumed under this plan, and $E(x_{ijt})$ is the unique utility-maximizing quantity (also called the expected quantity). Note that the expression for the expected utility changes depending on the location of the expected quantity along the budget set. In addition, β_{ij} is a consumer- and plan-specific intercept, and m_i is a consumer-specific intercept that is common to all plans. The latter variable is analogous to the service quality variable in the full model. In this null model, we assume that consumers know these variables when choosing a plan.

For model identification, we set the intercept of the utility for churn and for Plan 4 to be zero. In addition, variance of the utility of churn is set to one. These restrictions are the same as those for the full model.

Null Model 2: Quant Learn (Only Quantity Learning)

Here, we assume that consumers are certain about the service quality but are uncertain about their consumption quantity and its variance. Thus, we include only the quantity learning component of the full model. A comparison of the fit of this model with that of Null Model 1 can indicate the importance of the quantity learning component. We express this model mathematically as follows:

$$\begin{aligned}(24)\qquad EU_{ijt} &= \beta_{ij} + m_i + \int_0^{Aj} h_1(x) f_{ijt}^x(x)dx + \int_{Aj}^{B} h_2(x) f_{ijt}^x(x)dx \\ &\quad + \mathbf{s}_{ijt}'\boldsymbol{\zeta}_i + \varepsilon_{ijt}, \\ \eta_{ijt} &\sim N(0,\ \tau_j^2), \\ x_{it}^{actual} &= E(x_{ij^*t}) + \eta_{ij^*t},\ \text{and} \\ \eta_{ij^*t} &\sim N(0,\ \tau_{ij^*}^2).\end{aligned}$$

Null Model 3: Qual Learn (Only Quality Learning)

In this model, we assume that consumers are certain about their consumption distribution but are uncertain about the quality of the service provider. Therefore, this model allows only for quality learning but not quantity learning. We express this model as follows:

$$\begin{aligned}(25)\qquad EU_{ijt} &= m_{it} + \alpha_{i1}E(x_{ijt}) + \alpha_{i2}(z_{ijt}) + \alpha_{i3}\left[E(x_{ijt})\right]^2 \\ &\quad + (\mathbf{r}_{ijt}'\boldsymbol{\gamma}_i)E(x_{ijt}) + \mathbf{s}_{ijt}'\boldsymbol{\zeta}_i + \varepsilon_{ijt}, \\ z_{ijt} &= I_i - F_j - p_{1j}E(x_{ijt}),\ \text{if } E(x_{ijt}) \geq 0 \text{ and } E(x_{ijt}) \leq A_j, \\ z_{ijt} &= I_i - F_j - (p_{2j} - p_{1j})A_j - p_{2j}E(x_{ijt}), \\ &\quad \text{if } E(x_{ijt}) > A_j \text{ and } E(x_{ijt}) < B, \\ \varepsilon_{ijt} &\sim N(0,\ \lambda_j^2), \\ m_{it+1} &= w_{1it}m_{it} + w_{2it}\mu + \chi_{it}, \\ \chi_{it} &\sim N(0,\ w_{2it}^2\delta^2), \\ x_{it}^{actual} &= E(x_{ij^*t}) + \eta_{ij^*t},\ \text{and} \\ \eta_{ij^*t} &\sim N(0,\ \tau_{ij^*}^2).\end{aligned}$$

RESULTS

We estimated the full model and the three competing null models using Markov chain Monte Carlo (MCMC) methods. For each model, we obtained parameter draws based on 100,000 iterations after a burn-in period of 30,000 iterations of the MCMC.

Model Comparison

For model comparison, we used these draws to calculate the log-marginal likelihoods for each model. Low absolute values denote a better model. The log-marginal likelihoods are –7918.06 for Null Model 1 (No Learn), –7736.27 for Null Model 2 (Quant Learn), –7464.07 for Null Model 3 (Qual Learn), and –7432.52 for the full model. Thus, according to Kass and Raftery's (1995) criterion, the full model is best supported by the data. A comparison of the log-marginal likelihoods shows that both quantity and quality learning are important elements of the data. It also appears that quality learning is more important than quantity learning in our application.

Parameter Estimates

The parameter estimates for the full model are given in Table 2 and Table 3. The tables present the estimates for the population means of the parameters. (In Table 2, the numbers in parentheses are the 95% posterior intervals around the mean, and the significant posterior means appear in bold.) The subscript r refers to the covariates within $\mathbf{r}_{ijt}$, which affect the utility through their interaction with quantity, and the subscript s refers to the covariates within $\mathbf{s}_{ijt}$, which have only a main effect.

The parameters β_1, β_2, and β_3 are the plan-specific intercepts for Plan 1, Plan 2, and Plan 3, respectively. Note that we set the intercept for Plan 4 and for churn to 0. The population-level estimate of α_1, which can be interpreted as the intercept in the demand equation, is positive, as we expected. The population coefficients ψ_2 and ψ_3 can be transformed to yield mean estimates of the income effect (α_2) and the concavity parameter of the utility function (α_3). Recall that for a consumer i, α_{i2} was constrained as $\exp(\psi_{i2})$, and α_{i3} was constrained as $-\exp(\psi_{i3})$. Thus, the approximate population income coefficient is $e^{-7.26} = .001$, and the approximate quadratic component is $-e^{61} = -1.84$.

The state dependence variable has a positive effect. Thus, even in the absence of any contractual agreement, there exists a strong inertia effect, which can be attributed to hassle costs of making plan changes. For the demand equation, a positive value of state dependence suggests that as consumers stay longer in a plan, they consume more minutes. The coefficient of past quantity is positive. For the utility equation, it suggests that as consumers use more minutes under a plan in the previous month, the utility of that plan increases. In the quantity equation, a positive coefficient indicates that previous consumed quantity positively influences current consumption. Not surprisingly, promotion has a significant, positive effect on consumer decisions.

We now discuss the parameters in the second column of Table 2. These are variables that capture the effects of past consumption dynamics. We begin with the effects of underage. In the quantity equation, the coefficient associated with immediate past underage is positive, whereas the coefficient for cumulative past underage is negative. This suggests that

Table 2
FULL MODEL: UTILITY COEFFICIENT ESTIMATES

Parameter	*M*	*Parameter*	*M*
Plan 1 intercept (β_1)	**3.06** (2.83, 3.28)	Past underage revenue$_r$ (Past_Under_Rev$_r$)	**1.20** (.68, 1.70)
Plan 2 intercept (β_2)	**2.09** (1.90, 2.31)	Cumulative underage revenue$_r$ (Cum_Under_Rev$_r$)	**–.14** (–.20, –.06)
Plan 3 intercept (β_3)	**2.18** (1.98, 2.39)	Past underage revenue$_s$ (Past_Under_Rev$_s$)	**–.41** (–.56, –.25)
Linear effect of minutes consumption (α_1)	**1.57** (1.38, 1.75)	Cumulative underage revenue$_s$ (Cum_Under_Rev$_s$)	**–.75** (–.93, –.59)
Linear effect of outside good (ψ_2)	**–7.26** (–7.64, –6.86)	Past overage revenue$_r$ (Past_Over_Rev$_r$)	**.17** (–.07, .39)
Quadratic effect of minutes consumption (ψ_3)	**.61** (.55, .66)	Cumulative overage revenue$_r$ (Cum_Over_Rev$_r$)	**–.46** (–.55, –.36)
State dependence$_r$ (State_Dep$_r$)	**2.95** (2.79, 3.12)	Past overage revenue$_s$ (Past_Over_Rev$_s$)	**–.89** (–1.03, –.74)
Past quantity$_r$ (Past_Qty$_r$)	**1.41** (1.33, 1.49)	Cumulative overage revenue$_s$ (Cum_Over_Rev$_s$)	**–1.39** (–1.59, –1.20)
Promotion$_r$ (Prom$_r$)	**.58** (.45, .73)		

Notes: The subscripts "r" and "s" correspond to the covariates in vectors $\mathbf{r}_{ijt}$ and $\mathbf{s}_{ijt}$, respectively. The covariates in vector $\mathbf{r}_{ijt}$ affect the direct utility through their interaction with consumption, and the covariates in vector $\mathbf{s}_{ijt}$ have a main effect.

Table 3
FULL MODEL: LEARNING, VARIANCE, AND HETEROGENEITY PARAMETERS

Parameter	*M*	*95% Posterior Interval*
Mean of service quality (μ)	.81	(.46, 1.34)
Variance of service quality (δ^2)	.30	(.19, .54)
Variance of Plan Utility		
Plan 1	1.84	(1.27, 2.30)
Plan 2	2.04	(1.64, 2.53)
Plan 3	1.99	(1.59, 2.44)
Plan 4	2.38	(1.81, 3.07)
Heterogeneity Distribution for Usage Variance		
Plan 1: c_1	1.12	(.89, 1.38)
Plan 1: d_1	3.57	(2.61, 4.79)
Plan 2: c_2	2.69	(1.71, 4.77)
Plan 2: d_2	.92	(.42, 1.56)
Plan 3: c_3	1.03	(.79, 1.33)
Plan 3: d_3	2.69	(1.80, 4.13)
Plan 4: c_4	1.90	(1.21, 2.95)
Plan 4: d_4	.45	(.26, .76)

the immediate impact of past underage results in consumers' desiring to consume more minutes. However, if underage persists, accumulated underage causes consumers to downgrade their plan (utility of plan is decreased) and lower their usage (negative effect on demand). This is consistent with the notion that consumers regret leaving unused minutes on the table and strive to reduce the possibility of reexperiencing this regret; they respond by increasing their consumption, and they switch to plans with a lower plan allowance. The negative main effects of immediate and cumulative underage are consistent with this explanation as well.

The effects of overage conform to intuition. The cumulative overage variable, both as part of $\mathbf{r}_{ijt}$ and $\mathbf{s}_{ijt}$, has a negative coefficient. For the expected consumption, this result suggests that the higher the accumulated penalty from past overage, the greater is the reduction in the expected quantity. For the choice equation, the negative coefficient suggests that in addition to the negative effect through the expected quantity, accumulated overage lowers the utility of plans. The effect of immediate past overage is negative (as part of $\mathbf{s}_{ijt}$). Again, this suggests that immediate past overage lowers the utility of plans. It is not clear why the past overage in the interaction term is not significant.

Table 3 contains the estimates of the service quality (μ) and the variance of the service (δ^2). The variances of the utilities for plans (τ^2_{ij}) also appear in Table 3. Recall that the variance for the utility of defection is set to one for identification purposes. The plan-specific parameters capturing the customer heterogeneity in the variance of usage (Equation 22) for the four plans (c_j, d_j) also appear in the table.

The results from the table show that the estimate of the service quality is positive, which is not surprising. The variance of the service quality (δ^2) is reasonably small and suggests that each service encounter provides a strong signal for the overall service quality.

The model also provides individual-level and plan-specific estimates for the variance of the random shocks that influence usage. We can calculate the average of these estimates across the 300 consumers. These averages are 1.31, .76, 2.50, and 3.00 for Plan 1, Plan 2, Plan 3, and Plan 4, respectively. As we expected, the variance of usage under plans with more free minutes (Plan 3 and Plan 4) is higher than that under the low-free-minutes plans (Plan 1 and Plan 2).

Quantity and Quality Learning

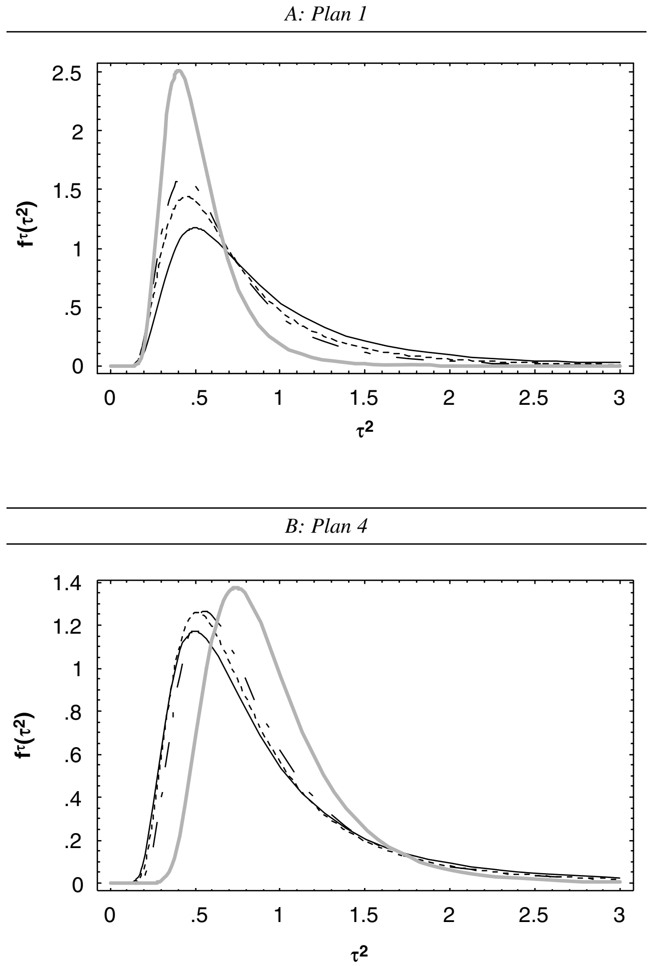
Figure 4 graphically shows quantity learning. Here, we show learning about the variance of usage under Plan 1 and Plan 4, respectively, for two randomly chosen consumers. We draw this figure using consumers' estimated quantity belief parameters based on our model. The top panel of Figure 4 shows the evolution of a consumer's beliefs about the usage variance under Plan 1 with successive choices of Plan 1. Recall that these are beliefs for a consumer that change as he or she gathers information about usage. There is a shift of this belief distribution to make small values of variance more likely. This accounts for low usage under Plan 1. Furthermore, the beliefs become more precise (shrinkage of the distribution) as more information about usage under Plan 1 is accrued. The bottom panel of Figure 4 illustrates the evolution of a single consumer's beliefs about the variance of usage under Plan 4. There is a shift of the belief distribution to make large values of variance more likely. This accounts for the high usage patterns under Plan 4. Figure 5 plots the change in the coefficient of variation with the number of observations for a randomly chosen consumer. The coefficient of variation of a distribution captures how the standard deviation changes relative to the mean. The figure suggests that as more observations are gathered, the standard deviation of consumers' quantity beliefs shrinks with respect to its mean.

Figure 6 illustrates quality learning. For a randomly chosen consumer, this figure shows the evolution of the mean of his or her beliefs for service quality. We find that the mean of the consumer's beliefs for quality rapidly converges to the estimated value. This is consistent with our estimation of the variance in service quality (δ^2) as being small, and thus each service encounter provides a strong signal for the overall service quality. Figure 6 shows that the mean is approximately .72 after the first five service encounters, which is approximately 90% of the estimated value of .81 (see Table 3). Thus, for this randomly chosen consumer, 90% of his or her learning about service quality occurs in the first five periods. We estimated this percentage across all consumers and found that it is approximately 92%. Thus, across all consumers in our data set, more than 90% of the learning about service quality occurs in the first five periods. This suggests that the first few service encounters are critical for a firm because consumers rapidly form beliefs about quality from these encounters.

The results from the three null models are mostly consistent with those from the full model. The parameter estimates for the null models are available on request.

Figure 4

EVOLUTION OF THE BELIEFS FOR THE VARIANCE OF THE USAGE ERROR



Notes: Here, $f_\tau(\tau^2)$ is the belief about the variance. The top panel corresponds to Plan 1, and the bottom panel corresponds to Plan 4. In each panel, the solid line represents the prior beliefs, the dashed line shows the beliefs after a single choice, the dash–dot line shows the beliefs after two choices, and the solid gray line shows the beliefs after nine choices.

POLICY EXPERIMENTS

Our model allows us to investigate several aspects associated with consumer learning, service quality, and pricing. We begin this section by exploring the consequences of consumer learning for both consumers and the firm. To study these effects, we use the CLV framework.

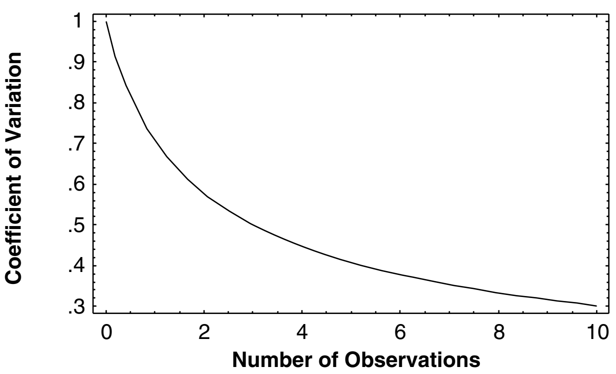
For all our policy experiments, we simulated the data for a maximum of 15 months (periods) for each customer. We assumed that all customers joined the company in the first month, and we set the promotion dummy for all periods across customers to zero. For each customer, using the MCMC draws, we generated 300 choice paths; a path for a customer represented his or her choices of service plans and the associated consumption for a maximum of 15 months. We then averaged across these paths to obtain our simulation outcomes.

Effects of Consumer Learning

Figure 4 shows the presence of quantity learning. Such consumer learning can happen in different ways. First, customers might actively track their usage over time. Second, several wireless providers (e.g., Verizon) now proactively send their customers information about their usage patterns. It is not clear a priori whether such proactive measures are beneficial from a company's perspective. Wireless providers generate revenue either when customers choose a plan that has more minutes (and consequently has higher access fees) than they require or when customers have an overage. If customers have a better knowledge of their usage, they can potentially give less revenue to the firm. We investigate the effects of customer learning from both a consumer's and a firm's perspective.

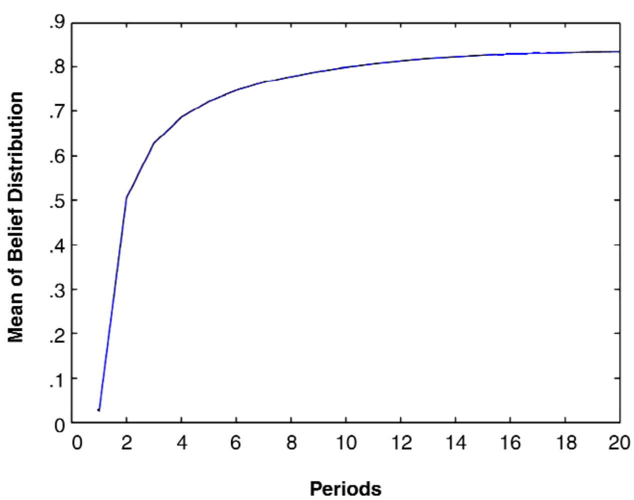
We begin with the consumers' viewpoint. We compare the average monthly underage and overage per consumer from a simulation using the full learning model with those from a policy experiment using Null Model 1. Table 4 shows these results.

Figure 5
COEFFICIENT OF VARIATION OF THE QUANTITY BELIEFS AS A FUNCTION OF NUMBER OF OBSERVED SIGNALS



Notes: The coefficient of variation of a distribution captures how the standard deviation changes relative to the mean.

Figure 6
EVOLUTION OF THE MEAN OF THE BELIEF DISTRIBUTION FOR SERVICE QUALITY



We find that with quantity learning, consumers have less underage than when there is no learning. The level of overage is relatively small, and there is little difference between the two models.

Next, we consider the firm's viewpoint. As we suggested previously, with low underage due to learning, it might appear that quantity learning is not beneficial for firms. However, quantity learning has another effect. Recall that in our learning model, consumers are uncertain about the usage variance and thus choose plans according to their expected utility. The quantity expectation accounts for consumer-specific beliefs about usage variance. Figure 4 shows that as these quantity beliefs evolve over time, there is a change in the mean and in the variance of its distribution, which in turn leads to certain values of the usage variance being more probable. This change in belief distribution translates into an increase in the expected utility of plans (i.e., a decrease in the overall churn rate). This increase in expected utility is coupled with an additional increase in the plan utility due to low underage. Within a CLV framework, this increase in customer retention leads to an increase in the overall CLV and should be beneficial for the firm.

To explore this hypothesis, we computed the average retention rate and the average revenue per user (ARPU) on the basis of the full model and the null model. We combined these two quantities to calculate the CLV. We used a simple expression that assumes a constant margin, m, and a constant retention rate, r. Then, CLV = mr/(1 + d – r), where d is the discount rate (Gupta, Lehmann, and Stuart 2004). The gross profit margin was 53% for this particular service provider. The annual discount rate was set at 10%. We found that the retention rate increased by 14% when we included consumer learning. This difference in retention rate translates into a difference in CLV; we computed that the overall CLV in the presence of consumer learning was 35% higher (approximately $75) than the CLV with no learning. We also compared the CLV from another model with that from our full model to evaluate the benefits from learning. This model was similar to our full model but with the learning parameters "turned off." Here, we assumed that consumers had priors about quality and quantity, but these priors were not updated (i.e., there was uncertainty about both quality and quantity, but there was no learning). Intuitively, we would expect that in such a scenario, the retention rate and the CLV should be much lower than that from a learning model, and indeed, we found that this was the case. The retention rate was approximately 30% lower than that which we obtained from our full learning model, and this translated into a 49% (approximately $145) decrease in CLV. This suggests that consumer learning can result in a win–win situation for both consumers and the firm. Thus, firms should proactively help their customers track their usage patterns because it lowers the defection rate.

Change in Service Quality

Companies can improve their service quality in several ways (e.g., make call centers more efficient, hire experienced employees). However, a firm will incur costs in doing so. The benefit of improving quality should be reflected in the defection rate; the higher the service quality, the lower defection should be. Because we do not have the data, we cannot ascribe service quality within our

Table 4
UNDERAGE AND OVERAGE FROM POLICY EXPERIMENTS

	Underage (Minutes)		*Overage (Minutes)*	
Plan	*Null Model*	*Full Model*	*Null Model*	*Full Model*
Plan 1	95	85	5	5
Plan 2	174	165	3	8
Plan 3	206	200	10	8
Plan 4	303	260	7	9

model to a particular managerial action. However, we can use the model to quantify the maximum dollar amount that a company should be willing to invest for improvements in quality. We investigate the ramifications of two types of changes in quality: permanent and transient (short-term).

Permanent change in mean quality. We introduce a mean quality change in the first period and keep the quality constant at this altered level throughout the entire 15 periods. Recall that the estimated mean service quality from our data (Table 3) was $\mu = .81$. To consider the impact of an increase in service quality, we augment the service quality in increments of 5%. Thus, we use levels from a minimum of a 5% increase to a maximum of a 35% increase. For each quality-change scenario, we computed the retention rate and the revenue per user on the basis of our model. We combined these two quantities to calculate CLV under the different scenarios. We compared the CLV from each of the quality change scenarios with the CLV of the baseline condition (i.e., no change in service quality).

We find that, on average, a 1% increase in quality leads to a \$2 increase in CLV. This service provider has 21 million customers. Therefore, a 1% increase in quality results in an overall long-term increase in profit of approximately \$42 million. This provides a measure of the maximum investment that should be made for improving quality.

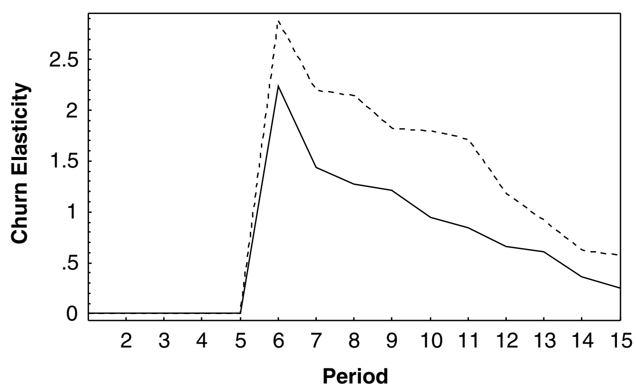
Temporary change in mean quality. Service providers can experience temporary fluctuations in service quality because of unforeseen random shocks, such as from attrition of personnel. Here, we estimate the effect of a temporary change in mean quality on customer defection. We shock the mean quality in the fifth month of a customer's relationship with the company, after which quality reverts to its original value for the remainder of the simulation. We chose the fifth month for this one-time shock because it was shown that more than 90% of the learning occurred in the first five periods. Thus, by their fifth month of the relationship with the company, customers would have learned about the baseline level of quality. We report the results using time-series plots for the percentage change in defection resulting from this one-time change in mean quality. For brevity, we show only the results for a 25% and a 35% temporary decrease in the mean service quality in Figure 7.

These results show that a temporary decrease in service quality not only increases the defection rate immediately (i.e, in the subsequent month) but also continues to have a lingering effect on churn. It takes approximately four to five periods for the churn elasticity to reduce in magnitude by approximately 50%. The results from the 35% decrease in service quality are similar. In addition, the lingering effects on churn reduce to almost zero by the end of the simulation period. The effect on CLV from such a temporary change is minimal. We find that the 25% temporary mean decrease in quality results in a decrease of approximately \$3 in the CLV of a customer, whereas the 35% temporary mean decrease leads to a loss of approximately \$6 in the long-term profit from a customer.

Change in the Pricing Scheme

A wireless service provider can alter the pricing schemes for the available plans to affect the revenue from the customer base as well as customer retention. In this section, we explore the impact of changing the access fee of the plans on the overall CLV.[6] For these policy experiments, we changed (both increased and decreased) the access fee of plans, one plan at a time, and computed the overall CLV. We then compared the CLV associated with each price change with the CLV of the baseline condition in which the prices were not changed. On the basis of this comparison, we computed the elasticity of CLV with respect to changes in access fee. Table 5 shows these elasticities for the full model.

Figure 7
CHURN ELASTICITY WITH A TEMPORARY CHANGE IN MEAN QUALITY



Notes: The dashed line represents 35% change in mean quality. The solid line represents 25% change in mean quality.

We find that, in general, a price decrease for a plan leads to a higher CLV than that from an equivalent price increase. Lifetime value embodies the trade-offs that are present in pricing decisions. A price increase for a plan results in higher ARPU but negatively affects retention. In contrast, a price decrease for a plan enhances retention but lowers the ARPU. The CLV results suggest that an increase in retention is more effective for increasing the CLV than an increase in the ARPU. We also find that the biggest effect on CLV is from changing the access fee of Plan 1.

Because CLV is composed of ARPU and the retention rate, we can calculate the elasticity of these two components with respect to changes in the access fee. This can help us understand which component plays a greater role in the differences among the four plans. Table 5 also shows these results. We noted previously that the effect of changing the access price of Plan 1 has the highest effect on changing the overall CLV. We now observe that the primary contributor to this result is a change in retention rate of the light users on Plan 1.

CONCLUSIONS

In this article, we developed a model that incorporates consumer learning of both quality and quantity under nonlinear pricing schemes. Our model captures nonlinear pricing schemes in the form of budget constraints, and we for-

[6]We also calculated the effect of a change in marginal price on CLV. The effect was much smaller than that for the access fee. The results are available from the authors on request.

Table 5
ELASTICITY OF CLV, ARPU, AND RETENTION RATE WITH ACCESS FEE

	CLV		*ARPU*		*Retention Rate*	
Plan	*Fee Up*	*Fee Down*	*Fee Up*	*Fee Down*	*Fee Up*	*Fee Down*
Plan 1	–.93	1.42	.30	–.31	–.58	.75
Plan 2	–.15	.06	.20	–.22	–.16	.12
Plan 3	–.56	.25	.26	–.28	–.38	.24
Plan 4	.02	–.02	.07	–.08	–.02	.03

mally model consumer uncertainty and learning with Bayes' theorem. We estimated the model on customer-level billing data from a wireless service provider and employed hierarchical Bayesian methods for drawing inferences for the customer-level parameters.

The model provided several notable results on the effect of past consumption dynamics on current decisions. For example, past underutilization of free minutes either increased the current consumption or influenced customers to downgrade their service plans. The results also showed that both quality learning and quantity learning are important aspects of the model. In addition, we find that consumers in our data set learn about service quality quickly. More than 90% of quality learning occurred within the first five service encounters. Policy experiments conducted to investigate the effects of consumer learning and changes in service quality and pricing also gave several managerially relevant results. First, we found that consumer learning can be a win–win situation for both consumers and the firm; consumers leave fewer minutes on the table, and the firm increases overall CLV. In particular, we estimated that there was 35% increase in CLV (approximately \$75) in the presence of consumer learning. The key driver of this difference is the change in the retention rate with and without consumer learning. This suggests that the proactive measures that service providers, such as Verizon, undertake to let customers track their usage can be beneficial. Second, we found that a change in access fee influences the CLV of light users through a change in their retention rate.

The current analysis has limitations. In our data, consumers are not on contract and could change their plans at the beginning of each month. Within contractual situations, our model in its full generality cannot be directly applied, and thus some substantive conclusions may not be generalizable. We also lack information on the quality of the wireless service in different geographical regions. If such data were available, it would be worthwhile to study the implications for quality learning. We also assumed that consumers are myopic. Prior research has shown that people generate beliefs about the future and then make current decisions (Erdem and Keane 1996; Gonul and Srinivasan 1996). Such forward-looking behavior can also be incorporated.

APPENDIX: COMPUTATION OF EXPECTED CONSUMPTION

As we described in the text, consumers face a piecewise linear budget constraint. For a plan j, we denote its access fee by F_j and the kink point by A_j. The marginal price when the total consumption is less than A_j is p_{1j}, and when the total consumption is greater than A_j, the marginal price is p_{2j}. Finally, B is an upper bound on consumption.

The quantity x_{ijt} that can be consumed under a plan j at time t is random from the viewpoint of consumer i and is modeled using a systematic component $E(x_{ijt})$ and a random component $\eta_{ijt} \sim N(0, \tau_{ij}^2)$. Thus, for each plan j, we have an individual-specific and a time-varying distribution from which consumption quantities are realized; that is,

$$x_{ijt} \sim N\left[E(x_{ijt}),\ \tau_{ij}^2\right].$$

The systematic component is known to the consumer and is obtained from maximizing the direct utility in Equation 3, subject to the nonlinear pricing constraints imposed by plan j. Specifically, we can write the optimization process for obtaining the expected quantity $E(x_{ijt})$ as

$$\max_x U_{ijt}\left[x,\ z(x)\right]$$

subject to

$$\text{Constraint I: } p_{1j}x + z = I_i - F_j, \text{ if } 0 < x \le A_j,$$

$$\text{Constraint II: } p_{2j}(x - A_j) + z = I_i - F_j - p_{1j}A_j,$$

$$\text{if } A_j < x < B.$$

To ensure a unique solution to this maximization problem, the utility function should be quasi concave. This requires the Slutsky constraints: $\alpha_{i2} > 0$, and $\alpha_{i3} < 0$ in Equation 3. For estimation, we set $\alpha_{i2} = \exp(\psi_{i2})$ and $\alpha_{i3} = -\exp(\psi_{i3})$, where ψ_{i2} and ψ_{i3} are unconstrained individual-specific parameters.

For quasi-concave utility functions, the unique optimal solution x* can be at an interior point (between 0 and A_j or between A_j and B) or one of the end points—0, A_j, or B. We can find the two candidates for an interior optimal solution by maximizing the utility function subject to the two linear constraints. The first-order conditions yield the following two interior candidate optima:

$$x^{\text{candopt,I}} = \frac{\alpha_{i2}p_{1j} - \alpha_{i1} - r'_{ijt}\gamma_i}{2\alpha_{i3}}, \text{ and}$$

$$x^{\text{candopt,II}} = \frac{\alpha_{i2}p_{2j} - \alpha_{i1} - r'_{ijt}\gamma_i}{2\alpha_{i3}}.$$

In these equations, $x^{\text{candopt,I}}$ ($x^{\text{candopt,II}}$) refers to the candidate optimal consumption when the utility function is maximized with Constraint I (Constraint II).

Given the uniqueness of the solution, at most, one of the two candidates will be attainable (i.e., will lie in the consumption interval in which its applied constraint holds). Because Constraint I holds for any positive consumption less than A_j minutes, though $x^{candopt,I}$ can lie anywhere on the real line, it is attainable only if it lies between 0 and A_j minutes. Similarly, $x^{candopt,II}$ is attainable only if it lies between A_j minutes and B. However, it is possible that neither of two candidates for an interior solution is attainable. Then, one of end points (0, A_j, or B) might be chosen. These cases are mutually exclusive and, together with any possible interior solution, form an exhaustive solution set; that is, $x^* \in \{0, A_j, B, x^{candopt,I}, x^{candopt,II}\}$. We denote this overall attainable expected quantity for consumer i, plan j, and time t as $E(x_{ijt})$.

REFERENCES

Ching, Andrew (2002), “Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs After Patent Expiry,” working paper, Rotman School of Management, University of Toronto.

Danaher, Peter (2002), “Optimal Pricing of New Subscription Services: Analysis of a Market Experiment,” *Marketing Science*, 21 (2), 119–38.

Erdem, Tulin, Susumu Imai, and Michael P. Keane (2003), “A Model of Consumer Brand and Quantity Choice Dynamics Under Price Uncertainty,” *Quantitative Marketing and Economics*, 1 (1), 5–64.

——— and Michael P. Keane (1996), “Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets,” *Marketing Science*, 15 (1), 1–20.

Gonul, Fusun and Kannan Srinivasan (1996), “Estimating the Impact of Consumer Expectations of Coupons on Purchase Behavior: A Dynamic Structural Model,” *Marketing Science*, 15 (3), 262–79.

Gupta, Sunil, Donald R. Lehmann, and Jennifer A. Stuart (2004), “Valuing Customers,” *Journal of Marketing Research*, 40 (February), 7–18.

Harvey, A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge, UK: Cambridge University Press.

Hausman, Jerry (1985), “The Econometrics of Non-Linear Budget Sets,” *Econometrica*, 53 (6), 1255–82.

Heckman, James J. (1981), “Heterogeneity and State Dependence,” in *Structural Analysis of Discrete Panel Data with Econometric Applications*, C.F. Manski and D. McFadden, eds. Cambridge, MA: MIT Press, 179–95.

Herriges, Joseph and K.A. King (1994), “Residential Demand for Electricity Under Inverted Block Rates: Evidence from a Controlled Experiment,” *Journal of Business and Economics Statistics*, 12 (4), 419–30.

Iyengar, Raghuram (2006), “A Structural Demand Analysis for Wireless Services Under Nonlinear Pricing Schemes,” working paper, The Wharton School, University of Pennsylvania.

Kass, Robert E. and Adrian Raftery (1995), “Bayes Factors,” *Journal of the American Statistical Association*, 90 (430), 773–95.

Lambrecht, Anja, Katja Seim, and Bernd Skiera (2005), “Pricing Internet Access with Three Part Tariffs,” working paper, Graduate School of Business, Stanford University.

Lemon, Katherine N., Tiffany Barnett White, and Russell S. Winer (2002), “Dynamic Customer Relationship Management: Incorporating Future Considerations into the Service Retention Decision,” *Journal of Marketing*, 66 (January), 1–14.

Maddock, Rodney, Elkin Castano, and Frank Vella (1992), “Estimating Electricity Demand: The Cost of Linearizing the Budget Constraint,” *Review of Economics and Statistics*, 74 (2), 350–54.

Meyer, Robert J. and Arvind Sathi (1985), “A Multiattribute Model of Consumer Choice During Product Learning,” *Marketing Science*, 4 (1), 41–61.

Moffitt, Robert (1990), “The Econometrics of Kinked Budget Sets,” *Journal of Economic Perspectives*, 4 (2), 119–39.

Narayanan, Sridhar, Pradeep Chintagunta, and Eugenio J. Miravete (2005), “The Role of Self-Selection and Usage Uncertainty in the Demand for Local Telephone Service,” working paper, Graduate School of Business, University of Chicago.

———, Puneet Manchanda, and Pradeep Chintagunta (2005), “Temporal Differences in the Role of Marketing Communication in New Product Categories,” *Journal of Marketing Research*, 42 (August), 278–90.

Nunes, Joseph C. (2000), “A Cognitive Model of People’s Usage Estimations,” *Journal of Marketing Research*, 37 (November), 397–409.

Reiss, Peter C. and Matthew W. White (2005), “Household Electricity Demand, Revisited,” *Review of Economic Studies*, 72 (3), 853–83.

Roberts, John H. and Glen L. Urban (1988), “Modeling Multiattribute Utility, Risk and Belief Dynamics for New Consumer Durable Brand Choice,” *Management Science*, 34 (2), 167–85.

Seetharaman, P.B., Andrew K. Ainslie, and Pradeep Chintagunta (1999), “Investigating Household State Dependence Effects Across Categories,” *Journal of Marketing Research*, 36 (November), 488–500.

Stoneman, P. (1981), “Intra-Firm Diffusion, Bayesian Learning and Profitability,” *Economic Journal*, 91 (June), 375–88.

West, M. and P. Harrison (1997), *Bayesian Forecasting and Dynamic Models*, 2d ed. New York: Springer.