Understanding Service Retention

Within and Across Cohorts Using Limited Information

David A. Schweidel*

Peter S. Fader

Eric T. Bradlow

* David A. Schweidel is Assistant Professor of Marketing at the University of Wisconsin, Madison. Peter S. Fader is the Frances and Pei-Yuan Chia Professor, Professor of Marketing, and Eric T. Bradlow is the K.P. Chao Professor, Professor of Marketing, Statistics, and Education, and Academic Director of the Wharton Small Business Development Center, The Wharton School of the University of Pennsylvania. The authors express their appreciation to the anonymous firm that graciously provided the data used here. Please address all correspondence on this manuscript to David Schweidel, dschweidel@bus.wisc.edu, 4172 Grainger Hall, 975 University Avenue, Madison, WI 53706.

Understanding Service Retention

Within and Across Cohorts Using Limited Information

Abstract

Service churn and retention rates remain central as constructs in marketing activities such as valuation of service subscribers and resource allocation. While extant approaches have been proposed to relate service churn to external factors such as reported satisfaction, marketing mix activities, and the like, managers often face situations in which the only information available is the duration for which subscribers have had service. In such cases, can they forecast service churn and understand the contributing factors, which may allow for subsequent intervention?

The authors propose a framework to examine factors that may underlie service retention in a contractual setting. Specifically, they utilize a model of retention that takes into account: (1) duration dependence, (2) promotional effects, (3) subscriber heterogeneity, (4) cross-cohort effects, and (5) calendar-time effects (e.g., seasonality). The framework is then applied to subscription databases of seven services offered by a telecommunications provider, mirroring the format commonly used to forecast future service churn (and to make managerial decisions).

Across all seven services, the inclusion of promotional effects always improves the forecast accuracy of retention behavior, while including cross-cohort effects does not significantly improve it. In five of the services, customer heterogeneity, calendar-time effects, and duration dependence also contribute to improved forecasts. These results are then used to understand how the expected value of a subscription differs across model specifications. They find considerable variation across model specifications, indicating that model misspecification can affect resource allocation decisions and other marketing efforts that are important to a firm.

Introduction

Retention remains a key construct for contractual service providers, as it is essential for determining the value of existing and future subscriptions, as well as for making resource allocation decisions. Modeling approaches have been proposed to understand churn patterns based on numerous factors, such as perceptions of quality (e.g., Zeithaml, Berry and Parasuraman 1996), customer satisfaction (e.g., Bolton 1998), and the firm's marketing activities (e.g., Lewis 2005; Bolton, Kannan and Bramlett 2000). Such research has furthered our overall understanding of the antecedents of service retention and the consequences of activities based on these factors. In practice, however, managers often face situations in which they have little or no external information available to them beyond the number of subscribers to a particular service. In these (realistic) situations, how accurately can one model (and forecast) service retention patterns?

Consider, for example, the case of Shaun, a hypothetical analyst working for a large telecommunications firm, who only has standard billing information available for a particular service offered by a cable TV provider (e.g., high-speed Internet access). He knows the number of subscribers at any given time, and the duration for which each of them has had service. He also knows about past promotional programs that the firm has run. However, the firm has not conducted surveys of its subscribers, so he does not have access to any information about customer characteristics or attitudinal measures. While recognizing that he cannot target individual subscribers with such limited information, can he understand the effects of different factors on retention and predict the number of subscribers who will discard service each month?

From his reading of the literature on retention forecasting, he is aware of several factors that he can (and should) incorporate into his analysis. Past research supports his initial observation that service churn decreases as a subscriber's tenure increases, i.e., *negative duration dependence* (e.g., Reichheld 1996; Hughes 2006).

In his investigation, he also notices that promotional offers (or, more generally, *marketing activities*) appear to affect the observed service churn (e.g., Lewis 2004). Some subscribers may leave after a promotional period ends, not because of increased dissatisfaction with the service, but simply because of the resulting change in their price plan. Isolating subscribers' responsiveness to these offers can also enable the company to determine the best marketing strategy to help maximize the value of subscriptions (e.g., Lewis 2005).

He is also aware of the importance of accounting for unobserved differences across subscribers (i.e., *subscriber heterogeneity*) when modeling duration data (e.g., Morrison and Schmittlein 1980). Another concern he has is about differences that may exist between subscribers who started service last month versus those who started last year. If the forecasting model has "staying power," he can use it to predict the service retention behavior of future subscribers (e.g., Neslin et al 2006). However, he hesitates to apply the same forecasting model to the latest "cohort" of subscribers, as there may be systematic differences in its behavior in comparison to older cohorts (i.e., *cross-cohort effects*). Another factor of interest that he's identified from past literature is *seasonality* in retention patterns (e.g., Danaher 2002; Radas and Shugan 1998).

To further complicate his work, he has been asked to analyze separately the churn patterns for multiple different services. While he has explained that his ability to do so with only subscriber counts is limited, he is interested in knowing whether or not the same set of factors affect retention of each service. If so, he can apply the same forecasting model to each of the company's services; if not, though, he needs a flexible model that will let him understand which factors affect retention of which services. He decides to develop a general framework that he can easily apply to subscriber data from separate datasets to address this empirical question and understand the drivers of service retention. This is the main objective of this research.

As a starting point, consider the retention of Service A by four cohorts of subscribers that first signed up between 2/2002 and 5/2002.[i] These service retention curves are presented in Figure 1a, and the empirical hazard rates is presented in Figure 1b. Thus, Figure 1a shows the proportion of initial subscribers from each cohort who still have service after *t* months, and Figure 1b shows the proportion of remaining subscribers at month *t-1* who discontinue service at month *t*.

[Insert Figure 1a and 1b]

Based on these curves, a pattern emerges in service retention. Specifically, subscribers seem less likely to discard service within the first three months, then suddenly increase in their likelihood to discard service, before ultimately slowing down in their likelihood of discarding service in later months. In this research, we explore the aforementioned five factors, which may provide explanations for the observed service retention and hazard rate patterns both within and across cohorts (and for different services). These include: (1) duration dependence, (2) time-varying marketing activity (e.g., promotional effects), (3) subscriber heterogeneity, (4) cross-cohort effects, and (5) calendar-time effects.[ii] Since each of these factors may influence the hazard rate, we utilize a standard paradigm that allows for the incorporation of these factors: the proportional hazard framework (e.g., Seetharaman and Chintagunta 2003). We now describe these factors briefly (a more formal technical description will be provided in the Model Development section).

Duration dependence allows the service churn rate for a subscriber to vary based on the length of time that he has had service. For example, a decreasing churn rate as a function of time since acquisition might explain the observed decrease in overall service churn during the later months; however, this factor alone would not explain the increased service churn in the intermediate months in Figure 1a (and the increase in the hazard rate in Figure 1b). Hence, it is unlikely to be the sole contributing factor, but instead may be one of many.

Promotional effects, associated with a well-defined period of short-term marketing activity, can be seen very clearly from the first portion of each cohort's retention curve. Throughout our dataset, the service provider always offered a special three-month introductory period for new customers. Many service subscribers kept the service for the duration of the promotional period, and then discarded it soon after the promotion ended, as reflected by the sudden increase in the empirical hazard rate in Figure 1b.

While promotional activity might explain the initially low service churn and hazard rate, followed by the sudden increase after the promotional period, it is not enough to account for the decreased churn in the later months. This is where *subscriber heterogeneity*, the idea that service churn rates vary across subscribers, may come in. Subscribers with higher churn rates may drop service relatively quickly, leaving the firm with a smaller set of more stable subscribers in the long run and therefore a curve that reflects greater retention over time. By simultaneously considering both duration dependence and heterogeneity, we can disentangle two competing effects: differences in service churn rates across subscribers and differences associated with tenure. While both may lead to related retention patterns, they imply very different stories and, consequently, different estimates of managerially relevant metrics, such as the value of a subscriber's subscription to a service, which we explore later in the paper.

Next, we note that Figures 1a and 1b also show some variability across the four cohorts shown, which may be attributable to systematic differences across cohorts. While no obvious trend emerges graphically from Figures 1a and 1b, we need to allow for the possibility of such a *cross-cohort effect* when we examine a broader set of cohorts.

Lastly, we consider *calendar-time effects* on service retention rate. In contrast to duration dependence, calendar-time effects focus on differences linked to the time of year rather than to the duration of a subscriber's service. Month-specific differences or seasonality can potentially influence the likelihood that subscribers will churn, regardless of their tenure, causing service churn to systematically vary for cohorts of all ages.

In summary, our objective is to build a general predictive model that can accommodate and sort out these competing explanations. By developing a predictive model, Shaun can not only forecast the future retention behavior of existing subscribers (in aggregate), but also predict the retention patterns of future subscribers for which he currently has no information. The model itself is not revolutionary; it closely resembles other duration models that have attempted to explain a varying set of drivers (e.g., Fader, Hardie, and Zeithammer 2003; Vanhuele et al. 1995). However, much of the extant research on service retention and churn heavily relies on external information at the level of the individual subscriber (including customer characteristics and attitudinal measures) to predict future subscriber behavior. In this research, we put forth a general modeling framework that can easily incorporate such information if it is available, but can forecast service retention of existing and future subscribers, even when it is not available to the service provider, such as our hypothetical analyst Shaun.

We carefully test these five factors, first for one service (as initially explored in Figures 1a and 1b), and then for a wider variety of services from the same provider. Across seven different services, we find strong support for the inclusion of promotional activity in modeling service retention. In addition, we find support for the inclusion of a combination of subscriber heterogeneity, duration dependence, and calendar-time effects. Interestingly, different services require a different set of components, highlighting the need for a general framework that can systematically examine competing explanations of behavior.

In the following section, we provide a review of the literature that has examined service retention. We then develop our modeling framework. Next, we describe our empirical analysis, in which we test various model specifications in a factorial design to systematically understand the effects of each of the five factors on customer retention. We then present detailed results of our empirical analysis for Service A and review the findings of our analyses for six other services offered. We conclude with a discussion of the implications and limitations of this research, as well as directions for future work.

Related Literature on Service Retention

In this section, we offer a brief review of some of the external factors that have been linked to service retention decisions, and we discuss potential limitations of these approaches for the context examined in this research. We then explore existing modeling approaches that can be used to model service retention solely based on subscriber counts.

A considerable body of research has examined the link between satisfaction with a service and the duration for which a subscriber maintains that service. Rust and Zahorik (1993) present a framework that links measures of satisfaction with retention and market share. Their framework allows a service provider to identify the dimensions of service satisfaction that will have the greatest impact on service retention and subsequently determine if expenditures toward improving satisfaction are justified. Bolton (1998) finds that increased satisfaction increases the duration of service tenure. In addition, she finds that this effect is larger for subscribers who have had service for a longer duration. Her research also demonstrates the importance of satisfaction by assessing the increase in customer equity after service improvements are implemented. In addition to considering the effect of current satisfaction on subscribers' decisions to retain service, Lemon, White, and Winer (2002) also incorporate expected future usage.

In addition to satisfaction, the link between service quality and service retention has also received much attention in research. For example, Boulding et al (1993) present a framework linking expectations of service quality to behavioral intentions. The authors hypothesize and find support for different effects of "should" and "will" expectations, which are combined with the service quality that actually occurred to develop an overall perception of service quality that impact future behaviors and expectations of quality. Rust, Zahorik and Keiningham (1995) outline a methodology to estimate the return on investments in service quality, allowing providers to decide which improvement efforts should be undertaken. Zeithaml, Berry and Parasuraman (1996) find that service quality affects subscribers' behavioral intentions with regard to the service, which in turn affect their retention decisions. They then link these decisions to financial outcomes. Bolton, Lemon and Bramlett (2004) explore the impact of service renewal decisions in a B2B setting and find that subscribers are more likely to continue service after experiencing a high level of service quality.

Other external factors that have been examined in relation to service retention include: channel of acquisition (e.g., Reinartz, Thomas and Kumar 2005), commitment and the effect of loyalty programs (e.g., Verhoef 2003; Bolton, Kannan and Bramlett 2000), and payment equity (e.g., Bolton, Kannan and Bramlett 2000; Bolton and Lemon 1999). In addition to these external factors, Keaveney (1995) identifies triggers that cause individuals to switch service providers. The most common type of trigger was a core service failure, such as billing errors or service mistakes. Among the other types of triggers are service encounter failures (such as speaking with an uncaring or unknowledgeable representative) and inconvenience to the subscriber.

When available, such attitudinal measures and external data can provide additional insight into the drivers of service retention decisions. However, many service providers, like the hypothetical Shaun, do not have easy access to such information. Instead, they often have little more than the number of subscribers from a cohort at a given time, as this can be directly extracted from internal billing information.

Fader and Hardie (2007) present a discrete-time probability mixture model that can be applied to the duration of service subscriptions. While their modeling framework incorporates unobserved heterogeneity (as we do here), they assume a constant rate at which subscribers discontinue service. In addition, their framework does not accommodate cross-cohort effects or time-varying covariates identified earlier in this research (or any of the other forms of external information that a provider could collect), the impact of which would be of interest to managers.

It should be noted that extant research has also explored the link between defection in non-contractual (transactional) exchanges and customer value. In contrast to service retention, explored in this research, the decision to defect in a non-contractual relationship is unobserved and models that estimate customer value (e.g., Fader, Hardie and Lee 2005; Schmittlein, Morrison, and Colombo 1987) rely on purchase histories, such as the time and the number of past purchases, which often do not have an analog in contractual exchanges. Unlike a non-contractual exchange, where revenue is generated on purchase occasions until a customer defects, revenue in a contractual exchange is generated each period until a customer defects, which is observed. Thus, the models necessary to estimate retention (and hence value) in contractual and non-contractual exchanges fundamentally differ from each other.

As we have described, by no means are we the first to consider service retention, contractual service retention, or its antecedents or consequences. However, what is unique is our systematic exploration across multiple cohorts and services with limited information; albeit, that which is readily and commonly available.

Model Development

To forecast the number of subscribers retaining service in future periods, we propose a general specification using a parametric form for the "survival" probability, $S(t)$, i.e., the probability that a customer has maintained service until time t, and its complement, churn: $F(t) = 1 - S(t)$. That is, after calibrating the model on n periods of data, we can then forecast the likelihood of maintaining service until time t by calculating $S(t)$ for any $t>n$. This also allows us to estimate other related measures of interest to researchers and practitioners such as the predicted number of customers who still have the service at time t ($N_0*S(t)$, where N_0 is the number of customers who began service at time 0) and the number of customers expected to drop service between time t and t+1 ($N_0*(S(t) - S(t+1))$).

We focus on the hazard rate, the conditional rate of churn given that one has not yet already churned, which provides us with a well-established framework for duration models within which we can develop our general approach. We use a proportional hazards model as is commonly used to account for a variety of possible effects (e.g., Seetharaman and Chintagunta 2003; Jain and Vilcassim 1991) to incorporate the five factors listed above. We do so within a mixture model specification to incorporate heterogeneity, in which the probability that a randomly selected customer i has not dropped service by time t is given by:

(1) $$S(t) = \int S(t|\theta_i,\boldsymbol{\beta},\boldsymbol{X}(t))g(\theta_i)d\theta_i$$

where θ_i is an individual-specific set of latent parameters, $\boldsymbol{X}(t)$ is a vector of covariates at time t, and $\boldsymbol{\beta}$ is the effect of these covariates. In this manner, equation (1) may be considered a mixed-effects hazard model with both fixed and random components.

The mixture model in (1) consists of two main components: $S(t|\theta_i,\boldsymbol{\beta},\boldsymbol{X}(t))$ and $g(\theta_i)$. $S(t|\theta_i,\boldsymbol{\beta},\boldsymbol{X}(t))$ specifies the probability that a customer maintains service until time t, which can be written in terms of its hazard function, $h(t|\theta_i,\boldsymbol{\beta},\boldsymbol{X}(t))$, which we specify to incorporate the five components of our model:

(2) $$S(t \mid \theta_i, \mathbf{X}(t), \boldsymbol{\beta}) = e^{-\sum_{v=1}^{t}\left(\int_{v-1}^{v} h(u|\theta_i,\boldsymbol{\beta},\mathbf{X}(t))du\right)}$$

Rather than assuming that all subscribers are homogenous, the mixing distribution $g(\theta_i)$ allows for unobserved differences in subscribers' tendencies to discontinue service, as some may be inclined to do so after only a few months while others may be more reluctant. We next describe how each of the five factors that we consider (duration dependence, promotional activity, cross-cohort effects, calendar effects, and heterogeneity) is formulated within the mixture model presented in (1) and (2).

Duration Dependence

As noted earlier, the likelihood that a subscriber drops service may change based on the length of time for which they have had it. We therefore adopt the Weibull distribution for the baseline hazard, which is flexible and commonly used in proportional hazards models (e.g., Morrison and Schmittlein 1980; Seetharaman and Chintagunta 2003):

(3) $$h_0(t \mid \lambda_i, c) = c\lambda_i t^{c-1}$$

The Weibull distribution nests the "strawman" exponential distribution when c=1; if subscribers do not exhibit duration dependence in their likelihood to drop service, the model collapses to a constant hazard rate λ_i. Values of c>1 will yield an increasing hazard rate, implying that subscribers are more likely to discard service the longer they have had it. Conversely, c<1 leads to a decreasing hazard rate: subscribers become less likely to discard service as their tenure increases. Depending on the nature of duration dependence, managers may wish to allocate their marketing efforts towards "older" or "younger" subscribers.

Cross-Cohort Effects

The baseline hazard function given in (3) is independent of the time at which subscribers began service (or, the *cohort* to which they belong). To allow for systematic differences across cohorts, we incorporate a cohort-specific effect, $\ln(q(j))=\beta_1(j-1)+\beta_2(j-1)^2+\beta_3(j-1)^3$, as a covariate affecting the baseline hazard function:

$$h_0(t,j\,|\,\lambda_i,c,\beta_1,\beta_2,\beta_3)=c\lambda_i t^{c-1}e^{\beta_1(j-1)+\beta_2(j-1)^2+\beta_3(j-1)^3}\quad \text{for } j=1,2,3,\ldots \tag{4}$$

where t is the time that has elapsed since customers began service in month j of the observation period. We employ a third degree polynomial to allow for a range of possible patterns of cross-cohort effects, including a monotonically increasing or decreasing baseline hazard (as a function of the cohort j), a U-shaped (or inverse U-shaped) baseline hazard function, and a baseline hazard that increases, decreases, and then increases again. Note that for j=1, the baseline hazard function in (4) reduces to that given in (3), implying that the cross-cohort effect can be interpreted as a scaling of the baseline hazard function relative to that of the first cohort:[iii]

$$\frac{h(t,j\,|\,\lambda_i,c,\beta_1,\beta_2,\beta_3)}{h(t,1\,|\,\lambda_i,c,\beta_1,\beta_2,\beta_3)}=e^{\beta_1(j-1)+\beta_2(j-1)^2+\beta_3(j-1)^3}\quad \text{for } j=1,2,3,\ldots \tag{5}$$

This simple three-parameter model component, as given in (4), should be sufficient to capture (or at least approximate adequately) any cross-cohort dynamics present in our dataset.[iv] If cross-cohort effects are present, certain cohorts of subscribers may be of greater value to the provider and hence targeted in a different manner.

Promotional Activity

As noted above, we employ a proportional hazards model to incorporate time-varying covariates (e.g., promotional activity and calendar effects), where:

(6) $$h(t \mid \theta_i, \mathbf{X}(t), \boldsymbol{\beta}) = h_0(t \mid \theta_i) \exp[\boldsymbol{\beta}' \mathbf{X}(t)]$$

Our data provider (as previously described) indicated that a three-month promotional offer was standard practice. If promotional activity slows subscriber churn, as expected, then the hazard rate during the promotional period will be dampened in comparison to the non-promotional period. Thus, when the promotional period ends, subscribers may become much more likely to discard service, in a non-smooth way, reflecting the higher hazard rate. To capture promotional activity, we define the variable *Promo(t)* such that *Promo(t)*=1 for month t=1,2,3 and *Promo(t)*=0 otherwise. With an understanding of the impact of promotional activity, the service provider can determine if the activity is actually worth its cost based on the change in expected value with and without the promotion.

Calendar-time effects

Unlike duration dependence, which captures changes in service churn based on the length of time for which a customer has had service, calendar effects account for differences in the rate of service churn attributable to the time of year. For example, subscribers may be more inclined to discontinue certain services during the summer months if they will be on vacation. These effects will impact subscribers at different times in their service tenure, depending on the month in which they started each particular service. For example, subscribers beginning service in February will experience the "July effect" during their fifth month of service, while the June cohort will experience the same effect in its second month of service; thus, we can disentangle "calendar time" from "cohort time," to which duration dependence and promotional activity are linked. In doing so, marketers can predict the parts of the year in which they are most likely to experience larger decreases in the number of subscribers and decide if intervention is warranted.

Calendar-time effects are incorporated via month-specific shocks to the hazard function. Suppose that cohort j's t^{th} month of service occurs in calendar month k (January = 1, February = 2, etc.). Let $C(j,t) = \gamma_k$ for k=1,…,11 and $\gamma_{12} = 0$ (i.e., December is treated as a baseline). Combining promotional activity and calendar effects with the individual-level baseline hazard from (4) yields the conditional probability of a subscriber from cohort j of a particular service maintaining it until time t, where $\Theta=\{c,\varphi,\tau,\beta,\gamma\}$ is the set of parameters common across individuals, given by:

$$\text{(7)}\quad S(t,j\,|\,\lambda_i,\Theta,\text{Promo(t)}) = \exp\left(-\lambda_i\left(\sum_{v=1}^{t}\left(v^c-(v-1)^c\right)\exp\left(\ln(q(j\,|\,\beta_1,\beta_2,\beta_3))+\beta_4\,\Pr omo(v)+C(j,v)\right)\right)\right)$$
$$= \exp\left(-\lambda_i\cdot A(t,j\,|\,\Theta,\text{Promo(t)})\right)$$

where

$$\text{(8)}\quad A(t,j\,|\,\Theta,\text{Promo(t)}) = \sum_{v=1}^{t}\left(v^c-(v-1)^c\right)\exp\left(\ln(q(j\,|\,\beta_1,\beta_2,\beta_3))+\beta_4\,\Pr omo(v)+C(j,v)\right)$$

Subscriber Heterogeneity

To complete the proposed service retention model, we specify a mixture distribution, which accounts for unobserved subscriber heterogeneity in the likelihood of dropping a particular service. If all subscribers have the same propensities for discarding a service, the mixing distribution would simply be a spike ($g(\theta_i)=\lambda$), but this would be an extremely restrictive assumption. We instead allow for heterogeneity across service subscribers by assuming that each subscriber's λ_i is drawn from a gamma distribution:

(9) $$g(\lambda_i \mid r, \alpha) = \frac{\alpha^r \lambda_i^{r-1} e^{-\alpha\lambda_i}}{\Gamma(r)}$$

The gamma distribution is chosen not only for its flexibility, but also because it is the conjugate prior for the Weibull distribution and therefore is commonly used as a mixing distribution for this purpose (e.g., Morrison and Schmittlein 1980). As such, the marginal probability of a subscriber in cohort *j* surviving until time *t* is found by integrating (7) over the mixing distribution (9), as shown in detail in (1), and is given by:[v]

(10) $$S(t, j \mid r, \alpha, \Theta, \text{Promo}(t)) = \left(\frac{\alpha}{\alpha + A(t, j \mid \Theta, \text{Promo(t)})}\right)^r$$

We have presented our complete modeling framework for cohort-level service retention. Note that the framework can easily accommodate behavioral measures such as reported satisfaction and service usage as time-varying covariates if they were available. These factors would be incorporated in a similar fashion as *Promo(t)* and *C(j,t)* into the hazard function. However, even when such detailed information is not available, we can still forecast the number of subscribers retaining service at the cohort level, as we demonstrate next.

Empirical Analysis

The data used in our empirical analysis were provided by a major telecommunications provider that offers a broad range of services to its customers. To evaluate the importance of these five factors, we separately fit a series of 32 models (defined by the full factorial of all combinations of the inclusion and absence of each factor) to seven services offered by the company (which we denote as Services A-G). Some of them (Services A and C) are "base" services that do not require any of the others to operate and also require hardware, while the others are "add-on" services that create additional functionality. We begin by revisiting Service

A (initially explored in Figure 1). Then, to assess the robustness of our findings, we apply the same set of models to subscription data for the other services.

Subscription information was provided from a single regional billing center from January 2002 through May 2004, indicating (in aggregated monthly data) the number of households subscribing to each service at the end of each month. Thus, the first group that we observe from the time that it begins service is the February 2002 cohort. Figure 2 provides an illustrative example of the amount of data provided by each cohort – early cohorts are under observation for a longer period of time than later cohorts and consequently provide a larger number of observations for our analysis.[vi]

[Insert Figure 2]

The dashed line indicates the end of our chosen calibration period, at the end of February 2003 (T=12 months). As illustrated in Figure 2, each cohort for calibrating the model is utilized for a different length of time. While the cohorts that begin service in or after February 2003 are not observed during the calibration period and therefore do not provide information to calibrate the model, they are still used to assess the (out-of-sample) fit of the model, thus providing a rigorous test of the model's forecasting ability. Specifically, we forecast the service retention numbers from March 2003 until the end of our dataset in May 2004 for each cohort. It is the performance of the model(s) in this out-of-sample forecast period that we use to gauge the usefulness of the various model components.

The provided subscription information, however, involves two forms of censoring that must be carefully incorporated. First, the data are left-censored, as we only observe customers who maintain service for at least one month. To account for this, we calculate the probability of continued service conditional on having maintained service through the first month:

(11) $$S^*(t,j) = P(T > t \mid t > 1) = \frac{S(t,j)}{S(1,j)}$$

where $S(t,j)$ is specified in (10). While the household data are left-censored, the cohorts are not left-censored; that is, we consider only those cohorts that begin service during the observation period. As such, households that began service prior to February 2002 are not included as we did not observe when they started service.

Next, each observation is interval-censored, as we only observe changes in the *number* of subscribers for each service at the end of each month. We therefore construct a dataset containing the number of households from cohort j with service through time t, denoted $N_{t,j}$. For each cohort j, households either maintain service through the calibration period T or discard service during the calibration period.

First, we consider households that maintained service throughout the entire calibration period. For a household starting service in month j (thus belonging to cohort j), keeping service through the calibration period implies that it maintained service for at least $T–j+1$ months. The probability that a household still has service at the end of the calibration period is therefore given by $S^*(T–j+1,j)$ and the number of households with service at the end of the calibration period from cohort j is $N_{T-j+1,j}$.

The other possibility is that a household discards the service during the observation period. A household from cohort j could discard service in any interval $(t,t+1]$ for t=1, 2, …,T-j. Given the interval-censored nature of the data, the probability of a household from cohort j churning during the interval $(t,t+1]$ is given by $S^*(t,j) – S^*(t+1,j)$ and the number of households discarding service during this interval is $N_{t,j}–N_{t+1,j}$.

From the households that maintain service and those that churn during the calibration period, we can construct the log-likelihood of the observed behavior of households from cohort j:

(12) $$LC(r,\alpha,\Theta \mid j) = N_{T-j+1,j} \log\left(S^*(T-j+1,j)\right) + \sum_{t=1}^{T-j}\left(\left(N_{t,j} - N_{t+1,j}\right)\log\left(S^*(t,j) - S^*(t+1,j)\right)\right)$$

where the first term accounts for households that maintain service through the entire calibration period and the second term accounts for households that churn during the calibration period, summing over all intervals in which the households could discard service. The log-likelihood for the full data set, consisting of all cohorts that begin service during the calibration period, is then:

(13) $$LL(r,\alpha,\Theta) = \sum_{j=1}^{T-1} LC(r,\alpha,\Theta \mid j)$$

Because the data are left-censored, we include only cohorts that begin service by T–1 so that we observe at least one month of behavior from each cohort. Parameter estimates were then obtained using maximum likelihood estimation.

As illustrated in Figure 2, each of the 32 models was calibrated using one year of data from the respective service, from the end of February 2002 through February 2003. To assess the overall forecasting ability of the models, we compare them based on mean absolute percentage error (MAPE) during the out-of-sample period (from February 2003 to May 2004) and BIC[vii] for the in-sample period.

Results

Results for Service A

Analysis of the results from the 32 models fit to Service A revealed that the inclusion of heterogeneity always led to improved performance on the basis of both in-sample (based on BIC) and out-of-sample (based on MAPE) criteria, regardless of the other model components. For the models with heterogeneity, incorporating promotional activity also improved both in-sample and out-of-sample performance. Therefore, for ease of exposition, we present the detailed results for the subset of eight models in which duration dependence, cross-cohort effects, and calendar-related effects are varied, but promotional activity and heterogeneity are always included in Table 1a.[viii] We first compare the relative performance of these eight models and then discuss differences in parameter estimates and resulting inferences across the models in Table 1b.[ix]

[Insert Tables 1a and 1b]

As noted above, the models incorporating heterogeneity and promotional activity perform significantly better than the corresponding models that omit these factors. While the in-sample BIC suggests that the cross-cohort effect is necessary, the performance of the models in the holdout period indicates that it is overfitting the calibration data. In the case of service A, we also find that duration dependence leads to systematic improvements in model performance (based on out-of-sample MAPE). Thus, it is the combination of unobserved heterogeneity, promotional activity, and duration dependence (Model 5) that leads to the model with the lowest error in holdout churn forecasts, albeit not dramatically so over the model with just heterogeneity and promotional activity.

Note that incorporating the cross-cohort effects and/or calendar-related effects both lead to reduced accuracy in out-of-sample forecasting. Thus, after accounting for promotional activity and subscriber heterogeneity, calendar-related and cross-cohort effects do not contribute to the forecasting ability of the model. While heterogeneity across subscribers must be taken into account, there do not appear to be systematic differences across households based on the time at which they begin Service A. Of course, the way in which one parameterizes the cross-cohort effect and calendar-effects, as well as the data set, could warrant their inclusion in other studies, and hence our findings may not be entirely general in this regard. But we have no reason to doubt the validity (or generalizability) of these observations, and test for this using our six other services.

We next examine the parameter estimates derived from the models which are presented in Table 1b (along with standard errors in parentheses), to understand the managerial impact of using different models. First, we see that there is positive duration dependence, as $\hat{c}$>1 in Models 5-8, indicating that households are increasing in their probability of discarding service as their duration of service increases. This seems to run counter to previous research that has observed increasing retention rates over time (e.g., Reichheld 1996), as well as the empirical hazard rates that we observe in Figure 1b. But, as noted by Follman and Goldberg (1988), omitting heterogeneity can lead to false conclusions about the effect of duration dependence, which is what we find here. Specifically, subscribers are not slowing down in the rate at which they drop service; rather, they are increasingly likely to drop service the longer they have subscribed to it, and it is the changes in the composition of remaining customers (those with higher retention rates maintain their subscriptions longer) that best explains the observed increase in the aggregate retention rate. This finding emphasizes the need for a well-specified modeling framework, as the type of duration dependence (positive or negative) can affect the way that marketers focus their retention efforts on different groups of subscribers..

Second, with regard to the parameters governing the mixing distribution, the models that ignore duration dependence (Models 1-4) reflect a greater degree of homogeneity than the models with duration dependence (Models 5-8) (The coefficients of variation, the ratio of mean to standard deviation, are around 1.2 for Model 2-4 and around 1.7 for Model 6-8).[x] Thus, ignoring duration dependence may lead one to infer erroneously that subscribers behave more similarly in their tendency to discard Service A than is actually demonstrated by the data. This effect isn't very large, but it is systematic as it recurs for the other services, as well. As Gupta, Lehmann and Stuart (2004) demonstrate changes in retention rates can lead to large changes in customer value. Thus, underestimating the impact of heterogeneity among subscribers' baseline service retention rates can lead to errors in calculating the expected value of a subscription. This is explored more fully and described later in Table 2.

The final finding of interest concerns the effect of promotional activity. In comparison to Models 5-8, the effect of promotional activity is stronger in Models 1-4. Because promotional activity occurred during the first three months of service, to get a "clean" estimate of the effect of the promotion, we must also take into account changes in the propensity to discard service that are attributable to duration dependence. Since $\hat{c}>1$ for Models 5-8, subscribers are more likely to discard service as the length of time they have had service increases. Thus, by ignoring the effect of duration dependence, one may overestimate the effect of early promotional activity. If the provider were to not offer the promotion, under Model 1 (no duration dependence), 78.2% of subscribers would be expected to still have service after three months, a reduction of 15.6% compared to when the promotion is offered. Under Model 5, 82.9% of subscribers are expected to remain, a reduction of 10.4%. By ignoring duration dependence in the baseline hazard function, managers may incorrectly attribute the reduction in churn solely to promotional activity, thus overestimating its effectiveness. Detecting such a difference can help managers to avoid needless additional spending on promotional activities.

Generalized Findings for Other Services

Having found that the combination of promotional activity, subscriber heterogeneity, and duration dependence led to the best performing model, based on out-of-sample analysis, for Service A, we applied the same analytic procedure to the subscriber data for the remaining six services to assess the robustness of our findings. Table 2 summarizes the components of the “winning” models based, once again, on out-of-sample MAPE performance for each service (although we looked at in-sample statistics such as BIC as well).

[Insert Table 2]

We find the need for at least two of the highlighted factors for each of the seven services. For every service, we observe that promotional activity is always an important component of the best model, while the cross-cohort effect is not present in any of the “winning” models, despite its intuitive appeal. For most services, we see that heterogeneity, duration dependence, and calendar-time effects also contribute significantly to model performance.[xi] When we look across the “winning” models of different services, no more than two services share the same model specification, highlighting the need for a generalized modeling framework.

Interestingly, calendar-time effects are present in the “winning” models for all of the “add-on” services offered by the provider. While services A and C are base services that require hardware, services B and D-G provide additional functionality that can be added or dropped by simply calling the service provider. In contrast, disconnecting a base service may require the customer to return hardware to the provider. One reasonable explanation posited by our data provider was that customers would disconnect these “add-on” services based on the time of year because of vacation patterns or the content of the service, which changed cyclically.

To highlight the managerial and economic importance of capturing the modeling components that affect service retention, we computed the expected value of a subscription, a measure akin to customer lifetime value, for each service (e.g., Berger and Nasr 1998). This measure can be calculated by multiplying the price less cost associated with the service by the expected duration for which the subscriber retains the service (e.g., Bolton 1998).[xii] We perform this calculation for the eight models described in Table 2, where promotional activity is always present, the cross-cohort effect is always omitted, and the inclusion of the remaining three factors is varied.[xiii] We present the expected value of a subscription under the "winning" model for the service, as well as the range of the expected value of a subscription under the eight model specifications. We note these calculations could be performed based on the full-factorial design with 32 models, but our primary interest is in assessing the combination of components present in the "winning" models[xiv]

For the services in which heterogeneity is present in the "winning model," the expected value of a subscription is closer to the high end of the range, while the converse is true for the services that do not require heterogeneity in the "winning" models. In general, omitting subscriber heterogeneity (when it is necessary) leads to a downward effect in estimates of subscription value, further emphasizing the need for a general, robust modeling framework for subscription duration.

In summary, we find that promotional activity, heterogeneity, duration dependence, and calendar effects are common elements in many of the models that yield the highest accuracy during the holdout period across the seven services considered. Furthermore, there are interactions among them: for instance, omitting duration dependence appears to inflate the estimated effect of promotional activity, as well as cause one to underestimate the degree of heterogeneity across subscribers. While detailed subscriber-level information such as usage and satisfaction would provide us with more insights into what affects subscribers' retention decisions, we are able to forecast the retention of current and future subscribers based on limited cohort-level data.

Discussion and Conclusions

The motivation for this research was to develop a modeling framework that could forecast the number of subscribers retaining service with limited information, as well as understand the factors contributing to retention behavior. While much research has established the link between reported satisfaction, quality, and other antecedents, limited work has afforded such an understanding without this information. The flexible framework that we present is based on the extant modeling literature and can easily incorporate this external information when available. However, even without it, forecasts of the number of remaining subscribers from a given cohort are obtainable.

We examine five specific factors for their impact on retention: duration dependence, promotional activity, subscriber heterogeneity, cross-cohort effects, and calendar-time effects. Looking across seven different services, there is not a single set of factors that always leads to the best performing model. In fact, there is no overall model specification that "wins" for more than two of the seven datasets. On one hand, the lack of a single winning specification may seem like a limitation of the analysis presented here; but, on the other hand, the general framework is very easily implemented (it takes less than a minute to run each model and can be done in widely available software such as Microsoft Excel) and allows for differences across services to be revealed. This systematic exploration is just what managers need to do in order to understand their different sets of subscribers, as a "one size fits all" model might tend to take managers' attention away from the critical issues that we have covered. One area that warrants future consideration is understanding the characteristics of services that are impacted by particular factors, such as calendar effects and duration dependence.

A further advantage of the formal modeling framework that we have laid out is its ability to aid managers in their evaluation of the effectiveness of marketing activity. As shown in Table

2, the model can provide estimates of the expected value of a service subscription. Managers can subsequently use this measure as a guide for determining how much they should be willing to spend on activities to induce a customer to subscribe to a particular service. In addition, by "turning off" the promotional activity (i.e., setting X(*t*)=0 for all *t*), the same approach could be utilized to determine the expected increase in revenue from the specific promotion. Managers can use the calculation of expected value of a service subscription to tailor more effective marketing activities by changing features such as the time at which offers are available, the amount of the discount offered, or its duration (e.g., Lewis 2005). Through the proposed framework, by linking marketing activities to a financial metric (in this case, the expected value of a subscription), marketers can more effectively allocate their resources, as well as make their marketing expenditures financially accountable (e.g., Rust, Zahorik and Keiningham 1995).

While this research focused on one service at a time, a key area of future research involves delving into the possible interplay among different products/services. Since many firms such as our data provider offer multiple contractual services, the development of integrated models for the adoption and retention of multiple services may assist managers in understanding multi-dimensional retention and churn issues, including cross-selling and subscriber valuation in a multi-service context. Unlike the model presented here, such a framework may require individual-level subscription information rather than cohort-level information. While such information could allow for individual-level marketing activities and more precise predictions of individuals' behavior, those companies that have not implemented sophisticated systems can still benefit from using cohort-level data, as we demonstrate here.

Table 1a
Model Results for Service A

Model	Duration Dependence	Cross-Cohort	Calendar-Related	LL	BIC	MAPE
1				-45946	45961	10.8%
2			✓	-45679	45752	11.4%
3		✓		-45927	45959	19.5%
4		✓	✓	-45624	45714	46.8%
5	✓			-45895	45916	10.7%
6	✓		✓	-45641	45720	11.0%
7	✓	✓		-45884	45921	13.7%
8	✓	✓	✓	-45612	45706	46.1%

Table 1b
Model Parameters for Service A

Model	Duration Dependence	Cross-Cohort	Calendar-Related	(r,α)	$(\beta_1,\beta_2,\beta_3)$	β_4	c
1				(0.68,3.53) (0.01,0.05)	--	-1.49 (0.02)	1
2			✓	(0.77,6.43) (0.01,0.08)	--	-1.44 (0.02)	1
3		✓		(0.66,2.84) (0.01,0.04)	(-0.10,0.02,-0.00) (0.00,0.00,0.00)	-1.48 (0.02)	1
4		✓	✓	(0.34,0.22) (0.00,0.00)	(-1.01,0.16,-0.01) (0.00,0.00,0.00)	-1.36 (0.02)	1
5	✓			(0.32,7.91) (0.00,0.13)	--	-1.14 (0.03)	1.90 (0.01)
6	✓		✓	(0.37,12.93) (0.00,0.20)	--	-1.13 (0.02)	1.80 (0.01)
7	✓	✓		(0.33,6.24) (0.00,0.10)	(-0.09,0.01,-0.00) (0.00,0.00,0.00)	-1.16 (0.02)	1.83 (0.01)
8	✓	✓	✓	(0.25,0.28) (0.00,0.01)	(-1.05,0.16,-0.01) (0.11,0.02,0.00)	-1.19 (0.02)	1.43 (0.09)

Table 2
Summary of "Winning Models"

Service	Heterogeneity	Duration Dependence	Calendar-Related	Winning Model Value	Minimum Value	Maximum Value
A	✓	✓		\$1380	\$216	\$1380
B	✓		✓	\$141	\$56	\$192
C		✓		\$2376	\$856	\$9472
D	✓	✓	✓	\$151	\$54	\$198
E	✓	✓	✓	\$143	\$53	\$203
F	✓		✓	\$150	\$51	\$208
G		✓	✓	\$74	\$60	\$290

Figure 1a
Cohort–Level Retention

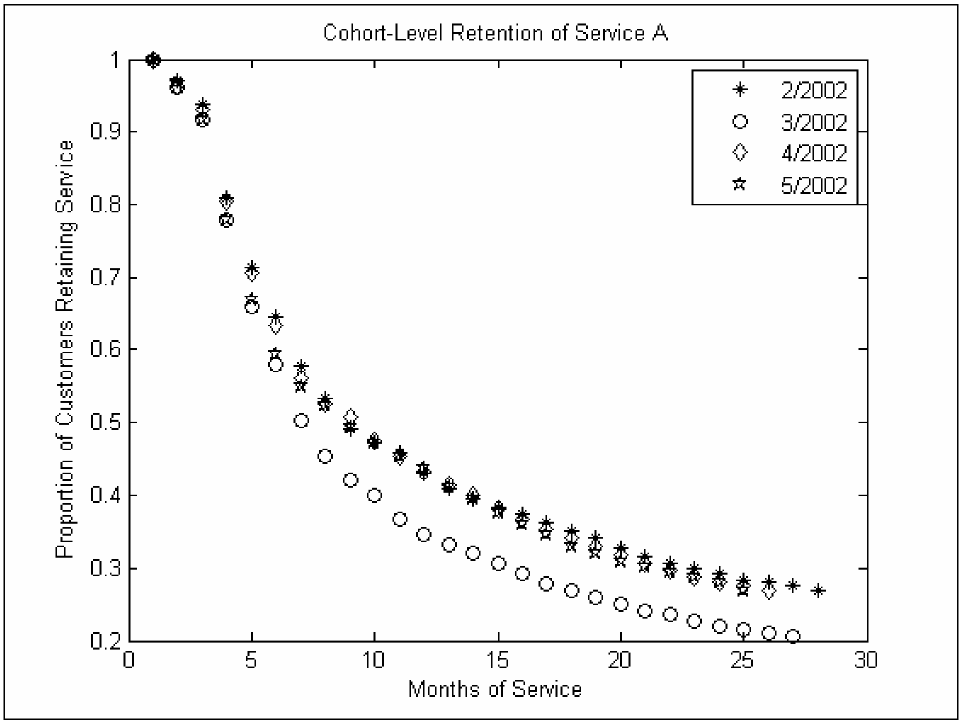


Figure 1b
Cohort-Level Empirical Hazard Rate

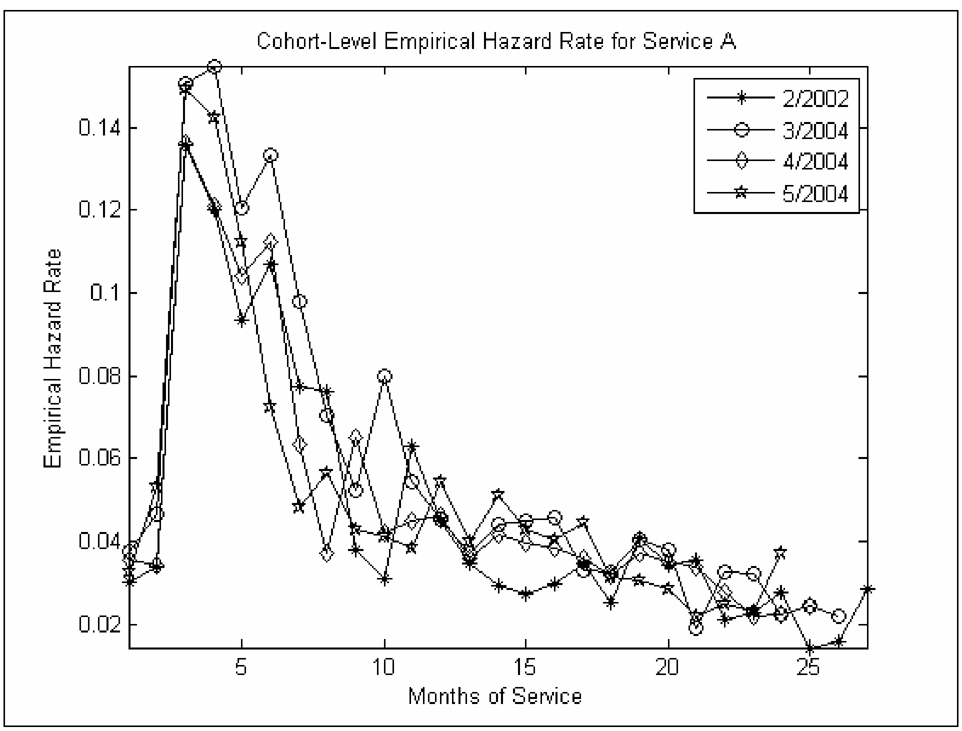


Figure 2
Diagram of the Cohort Analysis

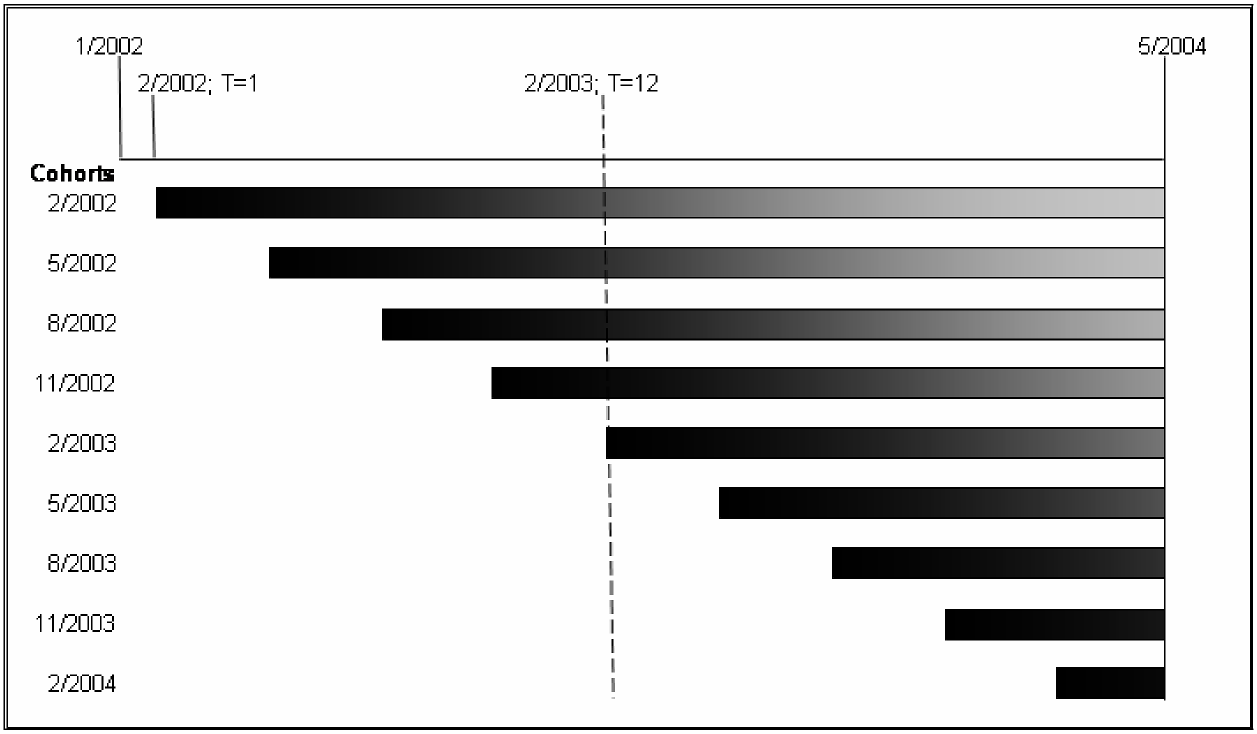
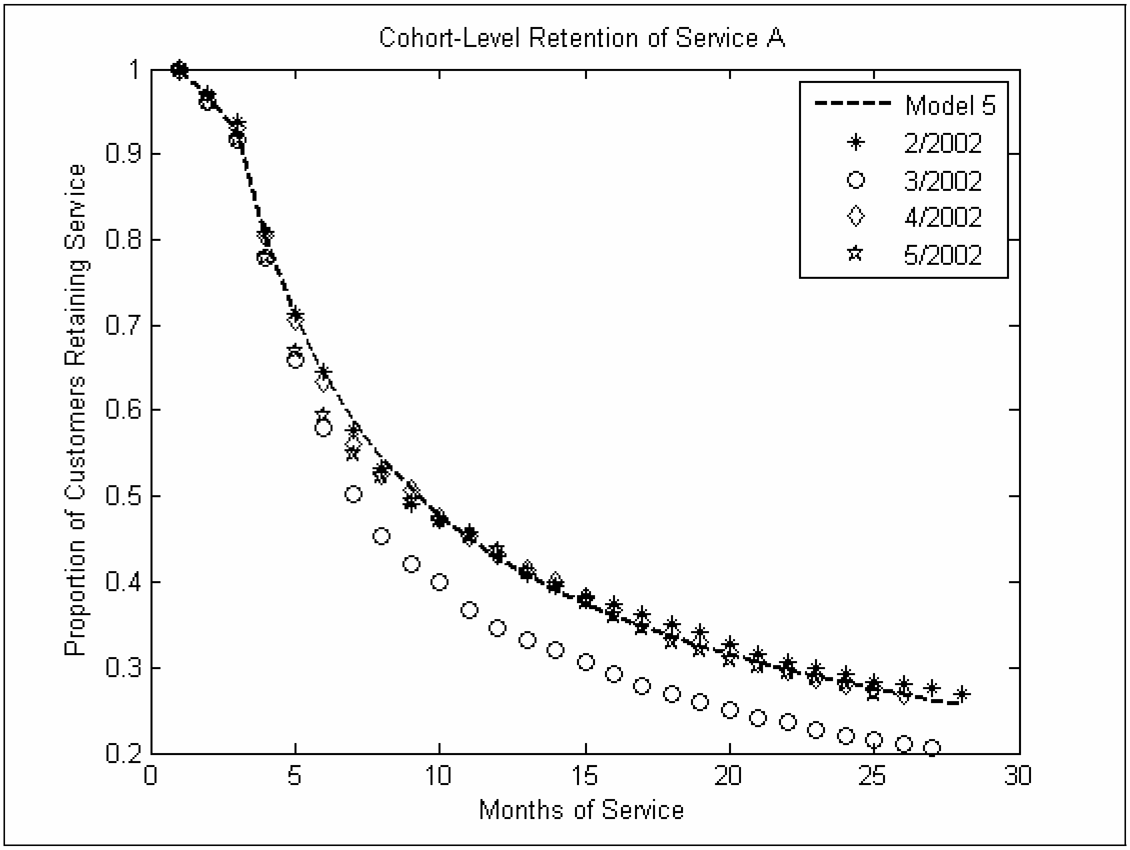


Figure 3
Model Fit for Service A



References

Berger, Paul D. and Nada I. Nasr (1998), "Customer Lifetime Value: Marketing Models and Applications," *Journal of Interactive Marketing*, 12 (1), 17-30.

Bolton, Ruth N. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: the Role of Satisfaction," *Marketing Science*, 17 (1), 45-65.

Bolton, Ruth N. and Katherine N. Lemon (1999), "A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction," *Journal of Marketing Research*, 36 (May), 171-186.

Bolton, Ruth N., Katherine N. Lemon and Matthew D. Bramlett (2004), "The Effect of Service Experiences Over Time on a Supplier's Retention of Business Customers," Marketing Science Institute Working Paper Series, Issue 4, Report No. 04-119.

Bolton, Ruth N., P.K. Kannan and Matthew D. Bramlett (2000), "Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value," *Journal of the Academy of Marketing Science*, 28 (1), 95-108.

Boulding, William, Ajay Kalra, Richard Staelin, and Valarie Zeithaml (1993), "A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions," *Journal of Marketing Research*, 30 (February), 7-27.

Danaher, Peter (2002), "Optimal Pricing of New Subscription Services: Analysis of a Market Experiement," *Marketing Science*, 21 (2), 119-138.

Fader, Peter S. and Bruce G.S. Hardie (2007), "How to Project Customer Retention," forthcoming, *Journal of Interactive Marketing*.

Fader, Peter S., Bruce G.S. Hardie and Ka Lok Lee (2005), "RFM and CLV: Using Iso-value Curves for Customer Base Analysis," *Journal of Marketing Research*, 42 (November), 415-430.

Fader, Peter S., Bruce G.S. Hardie and Robert Zeithammer (2003), "Forecasting New Product Trial in a Controlled Test Market Environment," *Journal of Forecasting*, 22, 391-410.

Follman, Dean A. and Matthew S. Goldberg (1988), "Distinguishing Heterogeneity from Decreasing Hazard Rates," *Technometrics*, 30 (4), 389-396.

Gupta, Sunil, Donald R. Lehmann and Jennifer Ames Suart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (1), 7-18.

Hughes, Arthur M. (2006), *Strategic Database Marketing*. New York, NY: McGraw-Hill.

Jain, Dipak C. and Naufel J. Vilcassim (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, 10 (1), 1-23.

Keaveney, Susan M. (1995), "Customer Switching Behavior in Service Industries: An Exploratory Study," *Journal of Marketing*, 59 (April), 71-82.

Lemon, Katherine N., Tiffany Bartnett 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.

Lewis, Michael (2004), "The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention," *Journal of Marketing Research*, 51 (August), 281-292.

Lewis, Michael (2005), "Incorporating Strategic Consumer Behavior into Customer Valuation," *Journal of Marketing*, 69 (October), 230-238.

Morrison, Donald G. and David C. Schmittlein (1980), "Jobs, Strikes, and Wars: Probability Models for Duration," *Organizational Behavior and Human Performance*, 25 (2), 224-251.

Neslin, Scott A., Sunil Gupta, Wagner Kamakura, Junxiang Lu and Charlotte Mason (2006), "Defection Detection: Improving Predictive Accuracy of Customer Churn Models," *Journal of Marketing Research*, 43 (2), 204-211.

Radas, Sonja and Steven M. Shugan (1998), "Seasonal Marketing and Timing of New Product Introductions," *Journal of Marketing Research*, 35 (3), 296-315.

Reichheld, Frederick F. (1996), *The Loyalty Effect*. Boston, Massachusetts: Harvard Business School Press.

Rust, Roland T. and Anthony J. Zahorik (1993), "Customer Satisfaction, Customer Retention, and Market Share," *Journal of Retailing*, 69 (2), 193-215.

Rust, Roland T., Anthony J. Zahorik and Timothy L. Keiningham (1995), "Return on Quality (ROQ): Making Service Quality Financially Accountable," *Journal of Marketing*, 59 (April), 58-70.

Schmittlein, David C., Donald G. Morrison and Richard Colombo (1987), "Counting Your Customers: Who are they and what will they do next?" *Management Science*, 33 (January), 1-24.

Schmittlein, David C. and Robert A. Peterson (1994), "Customer Base Analysis: An Industrial Purchase Application," *Marketing Science*, 13 (1), 41-67.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *Annals of Statistics*, 7 (2), 461-464.

Seetharaman, P.B. and Pradeep K. Chintagunta (2003), "The Proportional Hazard Model for Purchase Timing: A Comparison of Alternative Specifications," *Journal of Business and Economics Statistics*, 21 (3), 368-382.

Vanhuele, Marc, Marnik G. Dekimpe, Sunil Sharma and Donald G. Morrison (1995), "Probability Models for Duration: The Data Don't Tell the Whole Story," *Organizational Behavior and Human Decision Processes*, 62 (1), 1-13.

Verhoef, Peter C. (2003), "Understanding the Effects of Customer Relationship Management Efforts on Customer Retention and Customer Share Development," *Journal of Marketing*, 67 (4), 30-45.

Zeithaml, Valarie A., Leonard L. Berry, and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, 60 (April), 31-46.

Appendix
Complete Model Results for Service A

Model	Heterogeneity	Duration Dependence	Promotional Activity	Cross-Cohort	Calendar-Related	LL	BIC	MAPE
1						-47643	47648	17.2%
2					✓	-47220	47283	17.4%
3				✓		-47515	47536	41.0%
4				✓	✓	-47047	47126	42.3%
5			✓			-46419	46430	24.2%
6			✓		✓	-46015	46084	26.6%
7			✓	✓		-46400	46427	30.6%
8			✓	✓	✓	-45680	45764	55.3%
9		✓				-47327	47338	27.3%
10		✓			✓	-46877	46945	29.8%
11		✓		✓		-47277	47304	43.4%
12		✓		✓	✓	-45974	46058	96.4%
13		✓	✓			-46071	46086	12.0%
14		✓	✓		✓	-45784	45857	13.1%
15		✓	✓	✓		-46057	46089	21.1%
16		✓	✓	✓	✓	-45639	45728	87.6%
17	✓					-47643	47653	17.2%
18	✓				✓	-47220	47288	17.4%
19	✓			✓		-47515	47541	41.5%
20	✓			✓	✓	-46825	46909	51.7%
21	✓		✓			-45946	45961	10.8%
22	✓		✓		✓	-45679	45752	11.4%
23	✓		✓	✓		-45927	45959	19.5%
24	✓		✓	✓	✓	-45624	45714	46.8%
25	✓	✓				-46253	46269	11.1%
26	✓	✓			✓	-45971	46044	11.9%
27	✓	✓		✓		-46247	46278	15.0%
28	✓	✓		✓	✓	-45930	46019	48.6%
29	✓	✓	✓			-45895	45916	10.7%
30	✓	✓	✓		✓	-45641	45720	11.0%
31	✓	✓	✓	✓		-45884	45921	13.7%
32	✓	✓	✓	✓	✓	-45612	45706	46.1%

[i] The telecommunications company that provided the data preferred to remain anonymous; hence, services are simply listed as A, B, C, etc. Our data is described more fully throughout.

[ii] While these may not be the only possible effects, they do represent the major factors described in the marketing literature, and our framework is quite flexible in its ability to accommodate other possible factors as well.

[iii] The cross-cohort effect is incorporated in a manner consistent with the proportional hazard framework, where ln(q(j)) is treated as a stationary covariate. Also note that an intercept is not needed in the polynomial, since c serves in that role.

[iv] While we employ an "agnostic" third degree polynomial in our empirical analysis, any functional form could be incorporated into the proportional hazard framework in a similar fashion.

[v] An acknowledged limitation of the proposed mixture modeling framework is that customers are assumed to have the same responsiveness to covariate effects, i.e., constant β's. In addition, other baseline hazard specifications could be chosen, though we use the Weibull to exploit its conjugacy with the Gamma mixing distribution, allowing for a parsimonious model specification.

[vi] For ease of presentation, we show every third cohort in Figure 2. In our data, though, a new cohort comes under observation each month.

[vii] BIC is a commonly used penalization method that tries to prevent overfitting by penalizing the likelihood for each parameter added to the model. Lower values of BIC indicate better fitting models. For more details on BIC, see Schwarz (1978).

[viii] A comparison of the full series of 32 models can be found in the appendix.

[ix] The small standard errors in Table 1b are, in part, due to our large sample size. Small perturbations in the parameter estimates result in changes sufficiently large in the log-likelihood.

[x] Models 4 and 8 yield coefficients of variation that are greater in magnitude than that of Models 1-3 and Models 5-7, respectively, but the directional relationship holds as the coefficient of variation of Model 4 is less than that of Model 8.

[xi] In the models that incorporate heterogeneity, we find positive duration dependence ($\hat{c} > 1$); those models that do not warrant the inclusion of heterogeneity demonstrate negative duration dependence ($\hat{c} < 1$).

[xii] As price and cost information were not provided to us, for the purposes of our demonstration, we assume a cost of zero and prices that are consistent with the provider's current pricing scheme.

[xiii] Calculations for the full set of 32 models for each service are available upon request.

[xiv] We compute the expected value for each of the 27 cohorts in our analysis and present the median result.