David A. Schweidel, Peter S. Fader, & Eric T. Bradlow

Understanding Service Retention Within and Across Cohorts Using Limited Information

Service churn and retention rates remain central as constructs in marketing activities, such as valuation of service subscribers and resource allocation. Although extant approaches have been proposed to relate service churn to external factors, such as reported satisfaction, marketing-mix activities, and so on, 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 use a model of retention that accounts for (1) duration dependence, (2) promotional effects, (3) subscriber heterogeneity, (4) cross-cohort effects, and (5) calendar-time effects (e.g., seasonality). Then, they apply the framework 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, whereas 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. The authors use these results 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.

Keywords: service retention, proportional hazards model, valuation

Retention remains a key construct for contractual service providers because it is essential for determining the value of existing and future subscriptions and 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., Bolton, Kannan, and Bramlett 2000; Lewis 2005). Such research has furthered the 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 service retention patterns be modeled (and forecast)?

For example, consider the case of Shaun, a hypothetical analyst working for a large telecommunications firm, who has only standard billing information available for a particular service offered by a cable television 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 prior promotional programs 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, Shaun is aware of several factors that he can (and should) incorporate into his analysis. Prior research supports his initial observation that service churn decreases as a subscriber's tenure increases—that is, negative duration dependence (e.g., Hughes 2006; Reichheld 1996).

In his investigation, Shaun 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).

David A. Schweidel is Assistant Professor of Marketing, University of Wisconsin–Madison (e-mail: dschweidel@bus.wisc.edu). Peter S. Fader is Frances and Pei-Yuan Chia Professor and Professor of Marketing (e-mail: faderp@wharton.upenn.edu), and Eric T. Bradlow is K.P. Chao Professor, Professor of Marketing, Statistics, and Education, and Academic Director of the Wharton Small Business Development Center (e-mail: ebradlow@wharton.upenn.edu), The Wharton School, University of Pennsylvania. The authors express their appreciation to the anonymous firm that graciously provided the data used herein.

ISSN: 0022-2429 (print), 1547-7185 (electronic)

Shaun 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 and 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 because there may be systematic differences in its behavior compared with older cohorts (i.e., cross-cohort effects). Another factor of interest that he has identified from prior literature is seasonality in retention patterns (e.g., Danaher 2002; Radas and Shugan 1998).

To complicate his work further, Shaun has been asked to analyze separately the churn patterns for multiple different services. Although he has explained that his ability to do so with only subscriber counts is limited, he is interested in knowing whether the same set of factors affects retention of each service. If so, he can apply the same forecasting model to each of the company's services; if not, 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 data sets to address this empirical question and to understand the drivers of service retention. This is the main objective of the current research.

As a starting point, consider the retention of Service A by four cohorts of subscribers that first signed up between February 2002 and May 2002.[1] These service retention curves appear in Figure 1, Panel A, and the empirical hazard rates appear in Figure 1, Panel B. Thus, Figure 1, Panel A, shows the proportion of initial subscribers from each cohort who still have service after t months, and Figure 1, Panel B, shows the proportion of remaining subscribers at month t – 1 who discontinue service at month t.

From these curves, a pattern emerges in service retention. Specifically, subscribers seem less likely to discard service within the first three months. After this period, the likelihood of customers to discard service suddenly increases, before ultimately decreasing in later months. In this research, we explore five specific 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.[2] Because each of these factors may influence the hazard rate, we use a standard paradigm that allows for the incorporation of these factors—namely, the proportional hazard framework (e.g., Seetharaman and Chintagunta 2003). We now describe these factors briefly (a more formal technical description appears in the "Model Development" section).

Duration dependence allows the service churn rate for a subscriber to vary according to the length of time he or she 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 1, Panel A (and the increase in the hazard rate in Figure 1, Panel B). Thus, it is unlikely to be the sole contributing factor, but rather may be one of many.

Promotional effects, associated with a well-defined period of short-term marketing activity, can be seen clearly from the first portion of each cohort's retention curve. Throughout our data set, 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 1, Panel B.

Although 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 into play. Subscribers with high churn rates may drop service relatively quickly, leaving the firm with a smaller set of more stable subscribers in the long run and, thus, 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. Although both may lead to related retention patterns, they imply 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 in greater detail subsequently.

Next, Panels A and B in Figure 1 also show some variability across the four cohorts, which may be attributable to systematic differences across cohorts. Although no obvious trend emerges graphically from Figure 1, Panels A and B, we need to allow for the possibility of such a cross-cohort effect when we examine a broader set of cohorts.

Finally, 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 vary systematically for cohorts of all ages.

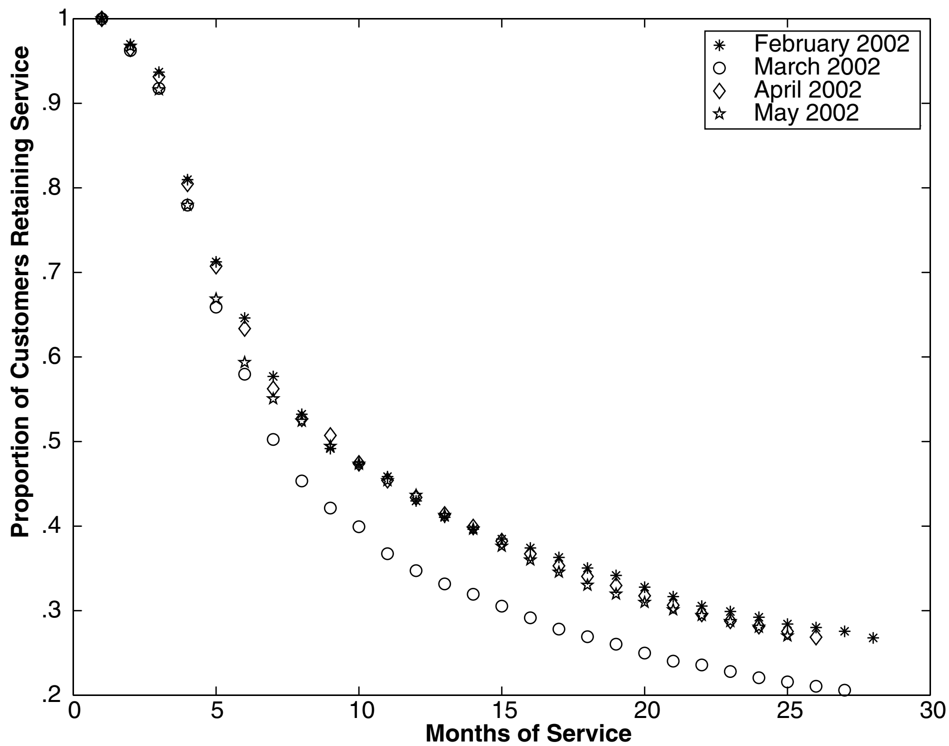
In summary, our objective is to build a general predictive model that can accommodate and sort out these competing explanations. By developing a predictive model, Shaun not only can forecast the future retention behavior of existing subscribers (in aggregate) but also can predict the retention patterns of future subscribers for which he currently has no information. The model itself is not revolu-

[1]The telecommunications company that provided the data preferred to remain anonymous; thus, we list services simply as A, B, C, and so on. We describe our data more fully throughout.

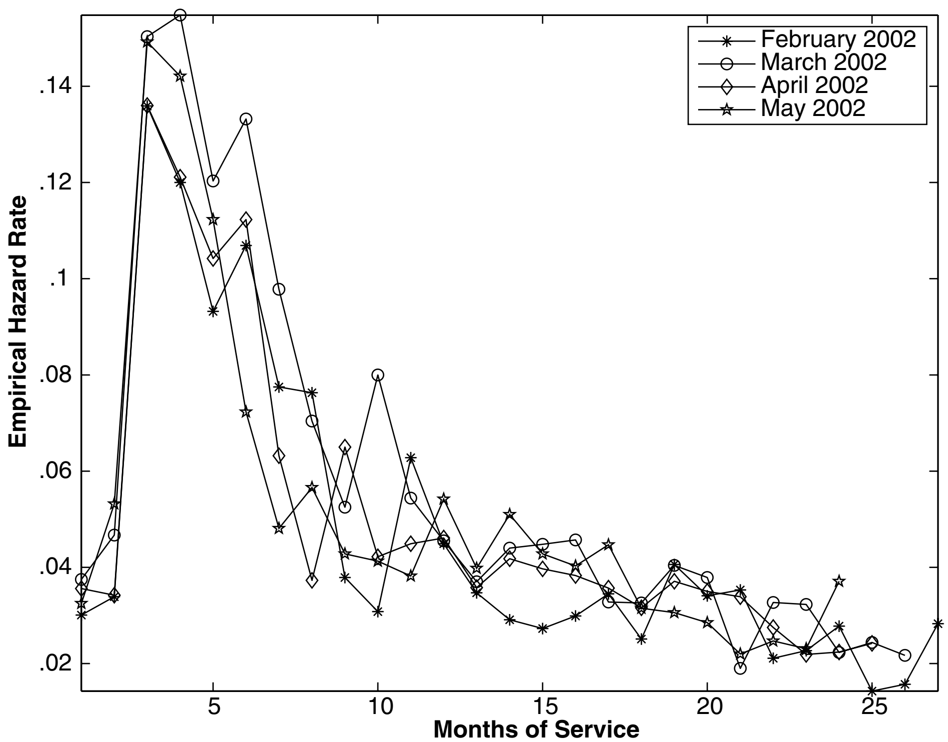
[2]Although these may not be the only possible effects, they represent the major factors described in the marketing literature, and our framework is flexible in its ability to accommodate other possible factors as well.

FIGURE 1
Cohort-Level Retention Behavior: A Selection of Cohorts

A: Cohort-Level Retention of Service A



B: Cohort-Level Empirical Hazard Rate of Service A



tionary; 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 relies heavily 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 also 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 Figure 1, Panels A and B) 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. Notably, 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 understand systematically 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 on the basis of subscriber counts.

A considerable body of research has examined the link between satisfaction with a service and the duration for maintaining 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 whether 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 greater 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. Apart from considering the effect of current satisfaction on subscribers' decisions to retain service, Lemon, White, and Winer (2002) incorporate expected future usage.

In addition to satisfaction, the link between service quality and service retention has received much attention in research. For example, Boulding and colleagues (1993) present a framework that links expectations of service quality to behavioral intentions. They 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 affects 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 business-to-business setting and find that subscribers are more likely to continue service after they experience 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., Bolton, Kannan, and Bramlett 2000; Verhoef 2003), 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 people 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 were service encounter failures (e.g., 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, as with the hypothetical Shaun, many service providers 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, because 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. Although 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 previously 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.

Note that extant research has also explored the link between defection in noncontractual (transactional) exchanges and customer value. In contrast to service retention, which we explore herein, the decision to defect in a noncontractual 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 number of prior purchases, which often do not have an analog in contractual exchanges. Unlike a noncontractual exchange, in which 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, thus, value) in contractual and noncontractual exchanges fundamentally differ from each other.

As we described, we are not the first to consider service retention, contractual service retention, or its antecedents and consequences. However, our systematic exploration across multiple cohorts and services with limited information is unique—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 forecast the likelihood of maintaining service until time t by calculating S(t) for any $t > n$. This also enables 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 $[N_0 \times S(t)]$, 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 $\{N_0 \times [S(t) - S(t + 1)]\}$.

We focus on the hazard rate, the conditional rate of churn given that the customer has not 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 various possible effects (e.g., Jain and Vilcassim 1991; Seetharaman and Chintagunta 2003), to incorporate the five factors listed previously. 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

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

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

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

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

Rather than assuming that all subscribers are homogeneous, the mixing distribution $g(\theta_i)$ allows for unobserved differences in subscribers' tendencies to discontinue service, because some may be inclined to do so after only a few months whereas others may be more reluctant. Next, we 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 Equations 1 and 2.

Duration Dependence

As we noted previously, the likelihood that a subscriber will drop service may change according to the length of time for which he or she has had it. Therefore, we 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):

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

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 $c > 1$ 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, such that subscribers become less likely to discard service as their tenure increases. Depending on the nature of duration dependence, managers may want to allocate their marketing efforts toward "older" or "younger" subscribers.

Cross-Cohort Effects

The baseline hazard function given in Equation 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:

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

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 Equation 4 reduces to that given in Equation 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:[3]

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

$$\text{for } j = 1, 2, 3, \ldots.$$

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

Promotional Activity

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

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

Our data provider (as previously described) indicated that a three-month promotional offer was standard practice. If promotional activity slows subscriber churn, as is expected, the hazard rate during the promotional period will be dampened compared with the nonpromotional period. Thus, when the promotional period ends, subscribers may become much more likely to discard service in a nonsmooth 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 whether the activity is actually worth its cost on the basis of the change in expected value with and without the promotion.

Calendar-Time Effects

Unlike duration dependence, which captures changes in service churn according to 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 affect 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, whereas 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 whether intervention is warranted.

Calendar-time effects are incorporated through month-specific shocks to the hazard function. Suppose that cohort j's tth month of service occurs in calendar month k (January = 1, February = 2, and so on). 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 Equation 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

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

$$= \exp\left\{-\lambda_i \times A[t, j|\Theta, \text{Promo}(t)]\right\},$$

where

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

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. Instead, we allow for heterogeneity across service subscribers by assuming that each subscriber's λ_i is drawn from a gamma distribution:

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

We chose the gamma distribution 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 Equation 7 over the mixing distribution (Equation 9), as shown in detail in Equation 1, and is given by[5]

[3]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. Note also that an intercept is not needed in the polynomial, because c serves in that role.

[4]Although 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 way.

[5]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 βs). 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.

$$S[t, j|r, \alpha, \Theta, \text{Promo}(t)] = \left\{\frac{\alpha}{\alpha + A[t, j|\Theta, \text{Promo}(t)]}\right\}^r. \tag{10}$$

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 are available. These factors would be incorporated in a similar way 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 the 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 the company offers (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, whereas 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.

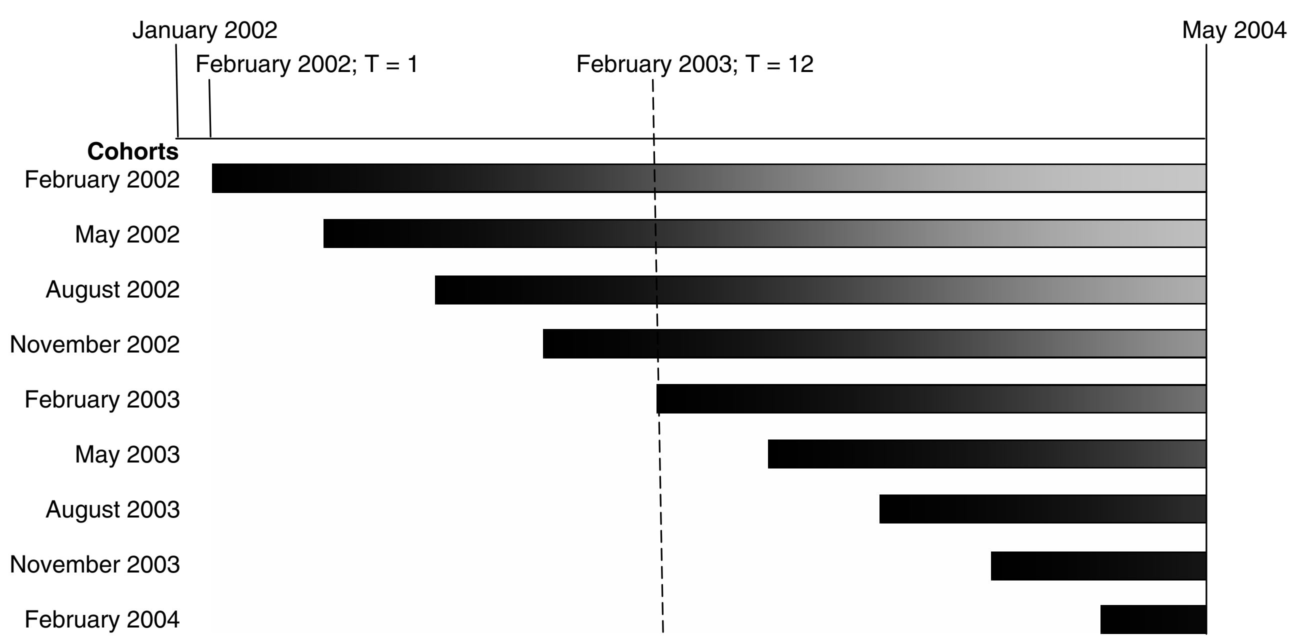
A single regional billing center provided subscription information from January 2002 to May 2004, indicating (in aggregated monthly data) the number of households subscribing to each service at the end of each month. Thus, the first group that we observe from the time that it begins service is the February 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 than later cohorts and consequently provide a larger number of observations for our analysis.[6]

The dashed line indicates the end of our chosen calibration period, at the end of February 2003 (T = 12 months). As Figure 2 illustrates, each cohort for calibrating the model is used for a different length of time. Although the cohorts that began 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, for each cohort, we forecast the service retention numbers from March 2003 until the end of our data set in May 2004. We use the performance of the models in this out-of-sample forecast period to gauge the usefulness of the various model components.

However, the provided subscription information involves two forms of censoring that we must carefully

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

FIGURE 2
Diagram of the Cohort Analysis



incorporate. First, the data are left censored because 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:

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

where S(t, j) is specified in Equation 10. Whereas the household data are left censored, the cohorts are not left censored; that is, we consider only cohorts that began service during the observation period. As such, we did not include households that began service before February 2002 because we did not observe when they started service.

Next, each observation is interval censored because we observe changes only in the number of subscribers for each service at the end of each month. Therefore, we construct a data set that contains the number of households from cohort j with service through time t, denoted as $N_{t,j}$. For each cohort j, households either maintain service through the calibration period T or discard service during the calibration period.

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. Therefore, the probability that a household still has service at the end of the calibration period is 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:

$$(12) \qquad LC(r, \alpha, \Theta|j) = N_{T-j+1,j}\log[S^*(T-j+1, j)] + \sum_{t=1}^{T-j}\Big\{(N_{t,j} - N_{t+1,j})\log[S^*(t, j) - S^*(t+1, j)]\Big\},$$

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 began service during the calibration period, is as follows:

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

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

As Figure 2 illustrates, we calibrated each of the 32 models 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 on the basis of mean absolute percentage error (MAPE) for the out-of-sample period (from February 2003 to May 2004) and Bayesian information criterion (BIC) for the in-sample period.[7]

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, in Table 1, Panel A, we present the detailed results for the subset of 8 models in which duration dependence, cross-cohort effects, and calendar-related effects are varied but promotional activity and heterogeneity are always included. A comparison of the full series of 32 models appears in the Appendix. We first compare the relative performance of these 8 models and then discuss differences in parameter estimates and resultant inferences across the models in Table 1, Panel B.[8]

As we noted previously, the models incorporating heterogeneity and promotional activity perform significantly better than the corresponding models that omit these factors. Although 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, the combination of unobserved heterogeneity, promotional activity, and duration dependence (Model 5) 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. We provide an illustration of the predictive ability of this model in Figure 3.

Note that incorporating the cross-cohort effects and/or the calendar-related effects leads to reduced accuracy in out-of-sample forecasting. Thus, after we account for promotional activity and subscriber heterogeneity, calendar-related and cross-cohort effects do not contribute to the

[7]The 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).

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

TABLE 1
Model Results for Service A

A: Model Performance for Service A

Model	Duration Dependence	Cross-Cohort	Calendar-Related	Log-Likelihood	BIC	MAPE
1				–45,946	45,961	10.8%
2			✓	–45,679	45,752	11.4%
3		✓		–45,927	45,959	19.5%
4		✓	✓	–45,624	45,714	46.8%
5	✓			–45,895	45,916	10.7%
6	✓		✓	–45,641	45,720	11.0%
7	✓	✓		–45,884	45,921	13.7%
8	✓	✓	✓	–45,612	45,706	46.1%

B: Model Parameters for Service A

Model	Duration Dependence	Cross-Cohort	Calendar-Related	(r, α)	$(\beta_1, \beta_2, \beta_3)$	β_4	c
1				(.68, 3.53)	—	–1.49	1
				(.01, .05)		(.02)	
2			✓	(.77, 6.43)	—	–1.44	1
				(.01, .08)		(.02)	
3		✓		(.66, 2.84)	(–.10, .02, –.00)	–1.48	1
				(.01, .04)	(.00, .00, .00)	(.02)	
4		✓	✓	(.34, .22)	(–1.01, .16, –.01)	–1.36	1
				(.00, .00)	(.00, .00, .00)	(.02)	
5	✓			(.32, 7.91)	—	–1.14	1.90
				(.00, .13)		(.03)	(.01)
6	✓		✓	(.37, 12.93)	—	–1.13	1.80
				(.00, .20)		(.02)	(.01)
7	✓	✓		(.33, 6.24)	(–.09, .01, –.00)	–1.16	1.83
				(.00, .10)	(.00, .00, .00)	(.02)	(.01)
8	✓	✓	✓	(.25, .28)	(–1.05, .16, –.01)	–1.19	1.43
				(.00, .01)	(.11, .02, .00)	(.02)	(.09)

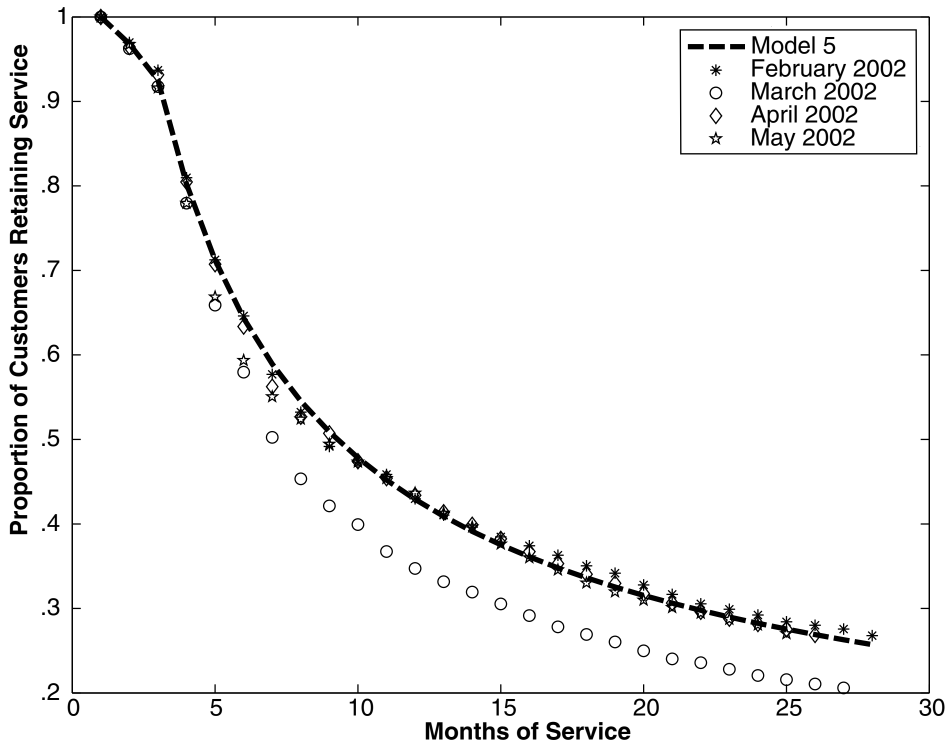
forecasting ability of the model. Although heterogeneity across subscribers must be account for, there do not appear to be systematic differences across households based on the time at which they begin Service A. The way the cross-cohort and calendar-time effects are parameterized, as well as the data set, could warrant their inclusion in other studies, and thus our findings may not be entirely general in this regard. However, we have no reason to doubt the validity (or generalizability) of these observations, and we test for this using our six other services.

Next, we examine the parameter estimates derived from the models, which appear in Table 1, Panel B (along with standard errors in parentheses), to understand the managerial impact of using different models. First, we observe 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 1, Panel B. As Follman and Goldberg (1988) note, however, 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 the changes in the composition of remaining customers (those with higher retention rates maintain their subscriptions longer) best explain the observed increase in the aggregate retention rate. This finding emphasizes the need for a well-specified modeling framework because the type of duration dependence (positive or negative) can affect the way 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 approximately 1.2 for Models 2–4 and 1.7 for Models 6–8.[9]) Thus, ignoring duration dependence may lead to the erroneous inference that subscribers behave more similarly in their tendency to discard Service A than is actually demonstrated by the data. This effect is not very

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

FIGURE 3
Model 5 Fit for Service A



large, but it is systematic because 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.

The final finding of interest concerns the effect of promotional activity. The effect of promotional activity is stronger in Models 1–4 than in Models 5–8. Because promotional activity occurred during the first three months of service, to obtain a "clean" estimate of the effect of the promotion, we must also account for changes in the propensity to discard service that are attributable to duration dependence. Because $\hat{c} > 1$ 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, the effect of early promotional activity may be overestimated. If the provider were not to 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 with when the promotion is offered. Under Model 5, 82.9% of subscribers would be 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 avoid needless additional spending on promotional activities.

Generalized Findings for Other Services

Having found that, on the basis of out-of-sample analysis, the combination of promotional activity, subscriber heterogeneity, and duration dependence led to the best-performing model 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, again on the basis of out-of-sample MAPE performance for each service (though we also examined in-sample statistics, such as BIC).

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, whereas the cross-cohort effect is not present in any of the winning models, despite its intuitive appeal. For most services, heterogeneity, duration dependence, and calendar-time effects also contribute significantly to model performance.[10] Across the winning models of different services, no more than two services

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

TABLE 2
Summary of Winning Models

Service	Heterogeneity	Duration Dependence	Calendar-Related	Winning Model Value ($)	Minimum Value ($)	Maximum Value ($)
A	✓	✓		1,380	216	1,380
B	✓		✓	141	56	192
C		✓		2,376	856	9,472
D	✓	✓	✓	151	54	198
E	✓	✓	✓	143	53	203
F	✓		✓	150	51	208
G		✓	✓	74	60	290

share the same model specification, highlighting the need for a generalized modeling framework.

Notably, calendar-time effects are present in the winning models for all the add-on services the provider offers. Whereas 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. A reasonable explanation posited by our data provider was that customers might disconnect these add-on services according to 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). We can calculate this measure 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).[11] We perform this calculation for the eight models (see Table 2); promotional activity is always present, the cross-cohort effect is always omitted, and the inclusion of the remaining three factors is varied.[12] 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. Note that these calculations could be performed using a full-factorial design with 32 models, but our primary interest is in assessing the combination of components present in the winning models.[13]

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, whereas 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 example, omitting duration dependence appears to inflate the estimated effect of promotional activity and causes underestimation of the degree of heterogeneity across subscribers. Although detailed subscriber-level information, such as usage and satisfaction, would provide more insights into what affects subscribers' retention decisions, we can forecast the retention of current and future subscribers using 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 who retain service with limited information and to understand the factors that contribute to retention behavior. Although much research has established the link among 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. Across seven different services, there is not a single set of factors that always leads to the best-performing model. Indeed, there is no overall model specification that wins for more than two of the seven data sets. On the one hand, the lack of a single winning specification may seem like a limitation of our analysis; on the other hand, the general framework is easily implemented (it takes less than a minute to run each model and can be done in widely avail-

[11]Because 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.

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

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

able software, such as Microsoft Excel) and allows differences across services to be revealed. This systematic exploration is what managers need to do to understand their different sets of subscribers, because a "one-size-fits-all" model might take managers' attention away from the critical issues that we have covered. An area that warrants further consideration is understanding the characteristics of services that are affected by particular factors, such as calendar effects and duration dependence.

A further advantage of our formal modeling framework is its ability to aid managers in their evaluations of the effectiveness of marketing activity. As Table 2 shows, 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 used 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 and make their marketing expenditures financially accountable (e.g., Rust, Zahorik, and Keiningham 1995).

Although this research focused on one service at a time, a key area for further research involves delving into the possible interplay among different products and services. Because 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 multidimensional retention and churn issues, including cross-selling and subscriber valuation in a multiservice context. Unlike the model we present, such a framework may require individual-level subscription information rather than cohort-level information. Although such information could allow for individual-level marketing activities and more precise predictions of individuals' behavior, companies that have not implemented sophisticated systems can still benefit from using cohort-level data, as we demonstrate herein.

APPENDIX
Complete Model Results for Service A

Model	Heterogeneity	Duration Dependence	Promotional Activity	Cross-Cohort	Calendar-Related	Log-Likelihood	BIC	MAPE
1						–47,643	47,648	17.2%
2					✓	–47,220	47,283	17.4%
3				✓		–47,515	47,536	41.0%
4				✓	✓	–47,047	47,126	42.3%
5			✓			–46,419	46,430	24.2%
6			✓		✓	–46,015	46,084	26.6%
7			✓	✓		–46,400	46,427	30.6%
8			✓	✓	✓	–45,680	45,764	55.3%
9		✓				–47,327	47,338	27.3%
10		✓			✓	–46,877	46,945	29.8%
11		✓		✓		–47,277	47,304	43.4%
12		✓		✓	✓	–45,974	46,058	96.4%
13		✓	✓			–46,071	46,086	12.0%
14		✓	✓		✓	–45,784	45,857	13.1%
15		✓	✓	✓		–46,057	46,089	21.1%
16		✓	✓	✓	✓	–45,639	45,728	87.6%
17	✓					–47,643	47,653	17.2%
18	✓				✓	–47,220	47,288	17.4%
19	✓			✓		–47,515	47,541	41.5%
20	✓			✓	✓	–46,825	46,909	51.7%
21	✓		✓			–45,946	45,961	10.8%
22	✓		✓		✓	–45,679	45,752	11.4%
23	✓		✓	✓		–45,927	45,959	19.5%
24	✓		✓	✓	✓	–45,624	45,714	46.8%
25	✓	✓				–46,253	46,269	11.1%
26	✓	✓			✓	–45,971	46,044	11.9%
27	✓	✓		✓		–46,247	46,278	15.0%
28	✓	✓		✓	✓	–45,930	46,019	48.6%
29	✓	✓	✓			–45,895	45,916	10.7%
30	✓	✓	✓		✓	–45,641	45,720	11.0%
31	✓	✓	✓	✓		–45,884	45,921	13.7%
32	✓	✓	✓	✓	✓	–45,612	45,706	46.1%

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.

———, 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.

——— 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), 171–86.

———, ———, 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 Report No. 04-119.

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 Experiment," *Marketing Science*, 21 (2), 119–38.

Fader, Peter S. and Bruce G.S. Hardie (2007), "How to Project Customer Retention," *Journal of Interactive Marketing*, 21 (1), 76–90.

———, ———, and Ka Lok Lee (2005), "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis," *Journal of Marketing Research*, 42 (November), 415–30.

———, ———, and Robert Zeithammer (2003), "Forecasting New Product Trial in a Controlled Test Market Environment," *Journal of Forecasting*, 22 (5), 391–410.

Follman, Dean A. and Matthew S. Goldberg (1988), "Distinguishing Heterogeneity from Decreasing Hazard Rates," *Technometrics*, 30 (4), 389–96.

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

Hughes, Arthur M. (2006), *Strategic Database Marketing*. New York: 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), 281–92.

——— (2005), "Incorporating Strategic Consumer Behavior into Customer Valuation," *Journal of Marketing*, 69 (October), 230–38.

Morrison, Donald G. and David C. Schmittlein (1980), "Jobs, Strikes, and Wars: Probability Models for Duration," *Organizational Behavior and Human Performance*, 25 (2), 224–51.

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 (May), 204–211.

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

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

Reinartz, Werner, Jacquelyn S. Thomas, and V. Kumar (2005), "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," *Journal of Marketing*, 69 (January), 63–79.

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

———, ———, 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.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *Annals of Statistics*, 7 (2), 461–64.

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), 368–82.

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 (October), 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.