

## ANALYZING DURATION TIMES IN MARKETING: EVIDENCE FOR THE EFFECTIVENESS OF HAZARD RATE MODELS

KRISTIAAN HELSEN AND DAVID C. SCHMITTLEIN  
*University of Chicago*  
*University of Pennsylvania*

Some statistical methods developed recently in the biometrics and econometrics literature show great promise for improving the analysis of duration times in marketing. They incorporate the right censoring that is prevalent in duration times data, and can be used to make a wide variety of useful predictions. Both of these features make these methods preferable to the regression, logit, and discriminant analyses that marketers have typically used in analyzing durations.

This paper is intended to fulfill three objectives. First, we demonstrate how decision situations that involve durations differ from other marketing phenomena. Second, we show how standard modeling approaches to handle duration times can break down because of the peculiarities inherent in durations. It has been suggested in recent marketing articles that an alternative to these conventional procedures, i.e., hazard rate models and proportional hazard regression, can more effectively handle duration type data. Third, to investigate whether these proposed benefits are in fact delivered for marketing durations data, we estimate and validate both conventional and hazard rate models for household interpurchase times of saline crackers. Our findings indicate the superiority of proportional hazard regression methods vis-à-vis common procedures in terms of stability and face validity of the estimates and in predictive accuracy.

**(Econometric Modeling; Estimation and Other Statistical Techniques; Pricing Research; Promotion; Regression and Other Statistical Techniques)**

### 1. Introduction

Many decision-making contexts in marketing suggest the analysis of duration time data, i.e., the time it takes for an event of interest to happen. A popular recent example is the attempt to assess a promotion's interpurchase acceleration effect (Neslin, Henderson and Quelch 1985). In this case the relevant duration time is the time between brand purchases. Table 1 lists a variety of such examples that illustrate both the pervasiveness of marketing duration times and the range of subdisciplines/decision-making issues where they are important.

When examining these durations, the decision-maker or analyst will typically have one (or more) of three objectives in mind:

#### (1) *Covariate Effects*

This interest is in measuring the effect of one or more covariate/explanatory variables on duration time, and testing the statistical significance of that effect. Such an objective

TABLE 1  
*Some Duration Times of Interest in Marketing*

Subdiscipline	Decision Contexts	Duration Times of Interest
Advertising	Copy testing Media selection Ad agency/client account management	Duration of advertising recall Timing of advertising exposures Duration of agency/client relationship
Pricing/Promotion	Timing of price changes or promotions; Measuring effect of promotions	Interpurchase duration; Timing of coupon redemption
Salesforce Management	Forecasting and managing salesforce turnover	Salesperson job duration
New Product Development	Forecasting trial, adoption, depth of repeat purchase	Duration time from new product introduction until initial trial; Interpurchase times
Marketing Strategy	Forecasting and managing success for new ventures	Time until failure for new ventures; Time until dissolution of joint venture
Marketing Research	Selecting purchase panel members; Forecasting future panel composition and new panelists needed Designing mail survey response incentives and response cutoff dates; Forecasting response rates Forecasting size and composition of firm's customer base	Time until panel dropout  Time until survey response  Time until customer becomes inactive or disaffected; Time until cancellation of subscription or service contract
Distribution Channels	Channel design	Duration of channel relationship
International Marketing	Developing cross-country entry strategies for new product launches	Time between sequential market entries

applies to the promotion example above, i.e., does the presence and size of a price promotion affect the expected interpurchase time, and, if so, by how much? Similarly, referring to Table 1, one might be interested in the effect of salesperson job satisfaction on job turnover, the impact of market conditions on the timing of price cuts, or the relation between advertising exposures and the adoption/trial time for a new product.

## (2) *Dynamics of Duration*

It is also common to focus on the way that the likelihood of the event occurring at one point in time depends on the amount of elapsed duration time. For example, salespersons' quitting rates have been seen to rise through their first couple years on the job and to decline thereafter (Schmittlein and Morrison 1983a). This decline in event rate as a function of elapsed duration is often called "inertia" and is also seen in interpurchase times (Schmittlein and Morrison 1983b; Dunn, Reader and Wrigley 1983), time between advertising exposures (Greene 1982), mail survey response time (Scott 1961; Cox 1966), and in the timing of failure for new business ventures (Sharma and Mahajan 1980). The inertia phenomenon in this last case is generally referred to as the "liability of newness"

(Freeman, Carroll and Hannan 1983; Singh, Tucker and House 1986). On the other hand, the new product adoption rate (i.e., proportion of current nonadopters who try a new product in this time period) is typically modeled as increasing monotonically with duration time (i.e., time since introduction) (Bass 1969; Schmittlein and Mahajan 1982). In each of these examples there has been an attempt to understand the dynamic pattern of duration times, typically abstracting the effect of any covariates on the process.

### (3) *Duration Time Forecasting*

In addition to examining cross-sectional covariate effects as in (1) or the longitudinal effects in (2), forecasts of duration times given some particular covariate/longitudinal profile may also be useful. We will see below that current approaches in marketing for accomplishing objectives (1) and (2) do not facilitate these forecasts. Predictions of interest may be either probabilistic (e.g., the probability that a salesperson currently earning \$50,000 with 2 years experience will still be with the firm in 5 years) or expectation-based (e.g., the expected additional time until next purchase of toothpaste for a household having 4 members, household income of \$31,000, whose last purchase of the product was 4 weeks ago).

Recent articles in the econometrics literature, and also two articles in marketing (Gupta 1991; Jain and Vilcassim 1991) suggest, sometimes implicitly, that these three modeling objectives can be accomplished within a single tractable class of duration time models, namely, proportional hazards regression. Our general objective is to investigate whether these models really do outperform conventional procedures (e.g., regression, logit, and probit analysis) that marketers have typically used in analyzing durations data. Neither Gupta nor Jain and Vilcassim have undertaken such a comparison. To our knowledge (and remarkably), this comparison has not been performed in the biometrics, econometrics, sociological methodology, or statistics literature either.

To this end, a first goal of the study is to pinpoint a number of features of duration times that make them distinctive from other kinds of marketing variables. In particular, we will touch on issues such as time varying explanatory variables and incompleteness of observations due to censoring or left truncation. Our second aim is to highlight the theoretical shortcomings of the traditional approaches taken to model duration times. We show, again theoretically, how a hazard function-based modeling framework can adequately resolve the issues pertinent to duration times.

Our third and primary purpose is the empirical evaluation of hazard function-based modeling in a representative marketing application via a comparison with standard procedures used to analyze duration type data. We examine interpurchase time decisions at the household level for the consumption of saltine crackers. We allow for right-censoring and time-varying covariates. Models estimated within the hazard function framework are contrasted with regression and probit models. Predictions are made for a holdout sample. Our findings indicate that, at least for the data set considered in the study, hazard rate models are more effective than commonly-used methods to calibrate duration time models. They outperform conventional models in terms of stability of the estimates, face validity of the parameter estimates, and predictive accuracy.

The next section describes the key characteristics that distinguish duration times from other phenomena studied in marketing. Section 3 provides a review of problems with standard econometric approaches applied to durations data in marketing, and a description of the proportional hazards regression approach. Estimation issues are covered in §4. Section 5 reports on the application of PHR to household interpurchase rates. Section 6 summarizes strengths and weaknesses of the hazard rate methodology and discusses managerial consequences and remaining research issues. Concluding remarks are given in §7.

## 2. Duration Times Dynamics

### 2.1. Key Characteristics of Duration Times

We first briefly describe the features that distinguish durations from other phenomena studied in marketing. One important difference stems from the way duration times are measured. Usually, events are recorded within the boundaries of an observation window, that is, the investigator starts recording events at some fixed calendar time  $C_L$  and finishes the observation process at time  $C_R$ . As a result, for those individuals for whom the event occurs after  $C_R$ , the only piece of information is that their duration process lasted at least until  $C_R$ . Such individuals are right-censored. Likewise, for individuals whose duration processes started before  $C_L$ , the exact duration time will be unknown. Such observations are said to be left-censored. The response in marketing has tended to either ignore censored observations, or to model the censoring event separately.<sup>1</sup>

Even in the absence of the sample selection biases discussed above, duration times may pose a problem when the effect of covariates is to be measured. Often times, the value of at least one of the predictor variables in which the investigator is interested may change during the event spell. It is not clear how to handle this issue properly using the “standard” methods (e.g., regression). In accurately assessing longitudinal or cross-sectional effects on durations, it is important to control for the impact of sample selection biases and the nonstationary nature of covariates. We will argue that the hazard function framework offers a better apparatus to tackle these issues.

### 2.2. Duration Process for an Individual

We view the duration time for individual  $i$ ,  $T_i$ , as a random variable having some p.d.f.  $f(t)$  and c.d.f.  $F(t)$ . In characterizing the dynamics of duration times, it is convenient to consider the *hazard rate*  $h(t) \equiv f(t)/(1 - F(t))$ , which is the conditional likelihood that the event of interest occurs at duration time  $t$ , given that it has not occurred in the duration interval  $(0, t)$ .

The rate  $h(t)$  offers several benefits in modeling duration time dynamics. First, unlike  $f(t)$ , which is shape-constrained by the restriction that it integrate to 1, and unlike  $F(t)$ , which must be nondecreasing in  $t$ ,  $h(t)$  can take any shape as a function of  $t$ , needing only to be nonnegative.

Second, the slope of the hazard function  $h(t)$  represents the important qualitative features of “duration time dynamics” that one typically has in mind. For example, “no dynamics” is generally taken to mean that the event of interest happens randomly in time, which in turn means that the likelihood of the event happening at time  $t$ , given that it has not happened yet, does not depend on how long we have been waiting. Thus, “no dynamics” is equivalent to a hazard rate  $h(t)$  that is constant over time ( $dh(t)/dt = 0$ ). This lack of memory is a property of the exponential distribution

$$f(t) = \lambda e^{-\lambda t}, \quad \lambda > 0; \quad t > 0, \quad (1)$$

for which it is easy to verify indeed that hazard rate is constant, i.e.,

$$h(t) = \lambda. \quad (2)$$

A hazard rate that increases with time (i.e., positive duration dependence  $dh(t)/dt > 0$ ) is often associated with a “snowballing” phenomenon, as has been proposed for the new product adoption process (Schmittlein and Mahajan 1982). Conversely, a de-

<sup>1</sup> Duration modeling may be hampered by another sample selection bias, called left-truncation. This arises, for example, when a longitudinal data set is purged periodically of completed durations (e.g., panelists who dropped out prior to a certain date are removed from a purchase panel file). Schmittlein and Helsen (1989) study the issues surrounding left-truncated data in marketing.

clining hazard rate (i.e., negative duration dependence  $dh(t)/dt < 0$ ) is often referred to as “inertia.” That is, the longer we have waited for the event to occur, the less likely it is to happen in the near future. As suggested in the introduction, this phenomenon of inertia has been observed in mail survey response times, interpurchase times, media viewing behavior, and new or joint venture survival.

When representing duration dynamics in a parametric model, it is thus useful to choose a specification that can admit either increasing, decreasing, or constant hazard rates. The most popular option meeting this criterion is the 2-parameter Weibull distribution, with p.d.f.

$$f(t|\lambda, \gamma) = \lambda\gamma t^{\gamma-1} e^{-\lambda t^\gamma} \quad \lambda, \gamma > 0; \quad t > 0, \quad (3)$$

and hazard rate

$$h(t|\lambda, \gamma) = \lambda\gamma t^{\gamma-1}. \quad (4)$$

To model situations for which the hazard rate pattern is likely to be nonmonotonic (e.g., salesforce turnover; see Schmittlein and Morrison 1983a), the researcher can use a (quadratic) Box-Cox formulation for the hazard function (Flinn and Heckman 1982):

$$h(t|\gamma_0, \gamma_1, \gamma_2) = \exp\{\gamma_0 + \gamma_1(t-1) + \gamma_2[(t^2-1)/2]\}. \quad (5)$$

Other hazard function specifications can be found in Cox and Oakes (1984), Kalbfleisch and Prentice (1980), or Lawless (1982).

A final benefit in choosing to model hazard rates is their direct link to other useful characteristics of the duration time distribution. We saw above that the hazard rate is determined by  $F(t)$ . The p.d.f. is in turn determined by the hazard function

$$f(t) = h(t) \exp\left(-\int_0^t h(x) dx\right) \quad (6)$$

(see, e.g., Cox and Oakes 1984). A shift upward (downward) in  $h(t)$  is also easily seen to reduce (increase) the mean and median duration time, assuming these latter quantities are finite. Indeed, the models in the next section view the covariates of interest as producing just such a shift in the hazard rate.

At this point we hope the reader is convinced that modeling hazard rates is an appealing general approach to capturing longitudinal and cross-sectional effects in duration times. Little is lost relative to modeling, e.g., mean duration times. As per the previous paragraph, a directional effect on the hazard rate will simply have the opposite effect on the mean duration. And the exact numeric effect on the mean duration of such a hazard rate shift can be calculated via (6). Further,  $h(t)$  can identify dynamics of duration such as inertia that are not captured by the mean. Lastly, we will see that hazard rate models are particularly tractable in facilitating the desired estimates of the effect of covariates on duration times.

### 3. Models for Prediction and Explanation

In this section we describe proportional hazards regression models as a way of incorporating cross-sectional factors in duration time models. Before doing so we will discuss briefly the two typical current approaches for analyzing marketing or social science durations. To motivate the usage of hazard rate models, we will focus on the fallacies of these more conventional methods.

#### 3.1. Duration Time Regression

One common approach involves simply regressing the observed duration time for each individual  $T_i$  on a set of covariates  $x_{1i}, \dots, x_{ki} \equiv x_i$  (see, e.g., Neslin et al. 1985; Lucas

et al. 1987). Relying on standard regression methods to study covariate effects on duration times may create two potentially severe problems. The first problem involves the sample selection bias to which we alluded in §2. The second issue arises when some of the predictor variables are time varying in the course of a spell. We will now discuss each of these problems in turn.

Use of duration time regression in the face of censoring may lead to biased estimates of the covariate effects. To illustrate the nature of this bias, we consider a simple example.<sup>2</sup> Suppose that the duration time for the event of interest can be adequately described by an exponential distribution with mean  $E(t|x) = \beta x$ . Furthermore, assume that the  $x$ 's are fixed (i.e., they do not vary over time) and that each member in the sample starts at time 0 (e.g., mail survey response, coupon redemption). Suppose now that the observation period ends at  $t = C$ . A standard practice is to focus on completed spells (see e.g., Neslin et al. 1985; Lucas et al. 1987). One can show that the expected duration time then becomes

$$E(t|x, C, t < C) = \beta x \left\{ 1 - \frac{(C/\beta x) \exp(-C/\beta x)}{1 - \exp(-C/\beta x)} \right\}. \quad (7)$$

Therefore, regressing  $t$  on  $x$  will clearly generally produce biased estimates of  $\beta$  towards 0. Even if unbiased, such estimates would be inefficient.

One might speculate on the presence of censoring biases in the marketing examples we referred to earlier. In their salesforce turnover study, Lucas et al. (1987) present duration time regression estimates for 3 different data sets. The amount of right-censoring (i.e., percentage of salespeople still with the company at the end of the observation period) goes from moderately large (data sets 1 and 2; see their Table 6) to heavy (>90%, data set 3). The regression results reported in their Table 7 reveal several anomalies: (1) only a few of the variables are significant (none is significant across the 3 data sets); (2) several sign reversals occur; (3) none of the hypotheses is supported; and (4) the overall explanatory power of the models is low (e.g., F-test nonsignificant for data set 3). Conceivably, some of these outcomes may manifest censoring biases. The lack of significant results in the purchase acceleration study of Neslin et al. (1985) might also reflect a sample selection bias.<sup>3</sup>

Regression methods are also inappropriate when the values of the explanatory variables change over time. When time-varying variables are included in the model specification (e.g., most marketing mix variables), the appropriate functional form will depend on the time path of the explanatory variables. To illustrate this point we return to the previous example. Assume, to simplify the exposition, that the data are not censored. We now consider an individual for whom the value of  $x$  during the spell changes at  $\tau$  from  $x_1$  to  $x_2$ . Consequently, the conditional expectation of the "spell duration"<sup>4</sup> will be

$$E(t|x_1, x_2, \tau) = \beta x_1 + \exp(-\tau/\beta x_1)(\beta x_2 - \beta x_1), \quad (8)$$

which substantially differs from  $E(t|x) = \beta x$ . Ad hoc procedures such as taking the within-spell average value for  $x$  fail to recognize that different time paths in the predictor variables require different functional forms. As a result, reliance on ad hoc procedures to cope with time-trended variables can produce very pathological estimates (see, e.g., Flinn and Heckman 1982, p. 69, Table 1, for an empirical illustration). Given the prev-

<sup>2</sup> The example is taken from Flinn and Heckman (1982).

<sup>3</sup> With regard to the Neslin, Henderson and Quelch (1985) study, it is difficult to judge the potential of censoring biases since no information is given on the amount of censoring. It would be useful in future studies that analyze duration times to report such information.

<sup>4</sup> In the terminology of duration time modeling, for our application each individual's successive "spells" refer to the time periods between purchases. The "spell duration" is therefore the interpurchase time.

absence of censoring and time-varying predictor variables, duration time regression does not seem an appealing modeling approach. Of course, we will have more to say about the actual performance of such regression models in our empirical results below.

### 3.2. Modeling the Event/No Event Outcome

A second common alternative for duration time analysis models whether or not the event of interest has occurred by some prespecified cutoff duration time  $T_c$  (see, e.g., Sharma and Mahajan 1980; Lucas et al. 1987; Peterson, Albaum and Ridgeway 1989). The dependent variable in such analyses can be represented for individual  $i$  as

$$E_i \equiv \begin{cases} 1 & \text{if } T_i > T_c, \\ 0 & \text{if } T_i \leq T_c, \end{cases}$$

and  $E_i$  can be modeled using any of the approaches commonly applied to binary dependent variables, e.g., logit, probit, or discriminant analysis.

Besides its failure to handle time-varying predictor variables, this type of analysis has several shortcomings. First, like duration time regression above, it does not incorporate all of the information available to the analyst. Second, in many marketing situations there is no natural time interval for which the event studied may take place. For example, Peterson et al. (1989) used a 12-month cutoff period in their study of the demographic profiles of direct sales customers. There is no a priori reason that justifies a 12-month cutoff period within this setting of consumer decision making.

Third, estimates for such a model do not allow the researcher to make predictions about either the expected duration times or the probability of the event happening to individual  $i$  for time intervals that are not integer multiples of  $T_c$ . Finally, one cannot even consistently apply these models to a succession of times  $T_{c,1}, T_{c,2}, \dots$ . That is, if a probit model is calibrated for a particular cutoff point  $T_c$ , its parameter estimates and interpretation do not carry over to time intervals of different length.

These issues restrict the appropriateness of these binary duration event models to a very limited set of scenarios: namely to the case where the *only* event of interest is whether an individual exceeds the prespecified duration time  $T_c$ . Again, we will investigate the empirical performance of such a binary model in the results below. As for the regression model findings, they will indicate whether some of these points raised concern merely theoretical niceties, or rather represent serious practical drawbacks.

### 3.3. Proportional Hazards Regression

A model proposed by Cox (1972, 1975) in the biometrics literature provides an appealing alternative to the foregoing approaches. Let  $h(t|x_t)$  denote the hazard rate at time  $t$  for an individual having covariate values  $x_{1t}, x_{2t}, \dots, x_{kt} \equiv x_t$  at time  $t$ .<sup>5</sup> Thus, the covariate values may vary over time for any individual. This hazard rate is assumed to take the form

$$h(t|x_t) = h_0(t)\phi(x_t, \beta), \quad (9)$$

where  $\beta_j$  indicates the effect of covariate  $x_{jt}$  on the hazard rate, and  $h_0(t)$  is the baseline hazard function. Thus the model has two multiplicative components. The first,  $h_0(t)$ , captures the longitudinal regularities in duration time dynamics discussed in §2. The second,  $\phi(x_t, \beta)$ , adjusts  $h_0(t)$  up or down proportionately to reflect the effect of the measured covariates. In light of this proportional adjustment of the baseline hazard rate, estimation of the  $\beta$ -vector in (9) is termed *proportional hazards regression* (PHR).

<sup>5</sup> Notice that our notation implicitly assumes that the starting point for time is 0. If this is not the case,  $x_t$  should be replaced by  $x_{t+\tau}$ , where  $t + \tau$  is the calendar time. To avoid an overabundance of notation, we ignore this technicality.

In most applications  $\phi$  is formulated as an exponential function:

$$h(t | x_i) = h_0(t) e^{\beta' x_i}, \tag{10}$$

which renders the estimation of  $\beta$  easier given that no constraints need to be imposed to ensure nonnegativity of  $\phi$ . For this special case:

$$\partial \ln h(t | x_i, \beta) / \partial x_i = \beta, \tag{11}$$

Thus  $\beta_j$  can be viewed as the constant proportional effect of  $x_j$  on the hazard rate. To interpret the coefficients associated with the predictor variables, one can look at the following transformation of  $\beta_j$ :  $100 \times (\exp(\beta_j) - 1)$ . This form of the coefficients can be given a quasi-elasticity interpretation as it indicates the percentage change in the hazard rate for a unit change in the corresponding explanatory variables  $x_j$ .

To parameterize (10) one need not formulate an explicit functional form for the baseline hazard,  $h_0(t)$ . In the semiparametric version of (10),  $h_0(t)$  is imagined to take any shape whatsoever over time. This point will be addressed further in §4 when we discuss estimation of (10) via a “partial likelihood.”

As an alternative, “parametric models” have also been proposed and estimated in biometric applications (Pike 1966; Glasser 1967; Peto and Lee 1973; Breslow 1974), in which a specific functional form is selected for  $h_0(t)$ . The two most common representations are those associated with the exponential or Weibull distribution. In the exponential model the constant baseline hazard of equation (1) is substituted for  $h_0(t)$  in (10), yielding

$$h(t | \lambda, \beta, x_i) = \lambda e^{\beta' x_i}. \tag{12}$$

Similarly, the Weibull regression model is formed by substituting (4) in (10).

#### 4. Estimation Issues

This section discusses two alternative approaches for estimating a hazard model. In particular, we will look at semiparametric and parametric estimation procedures.

##### 4.1. Semiparametric-Partial Likelihood

For duration time processes, the usual (“total”) likelihood has as the event of interest the fact that individuals  $i$ ’s duration time (i.e., the random variable  $T_i$ ) took on the observed value  $T_i = t$  (or, in the case of right censoring, the event that  $T_i > t_R$ , the right censor point) for individuals  $i = 1, \dots, N$ .

The partial likelihood also focuses on the observed durations  $t_1, t_2, \dots$ , but considers them in a different way. Imagine that individual  $i$  has an uncensored duration  $T_i = t$ . At this duration time  $t$ , a number of other individuals were “at risk,” i.e., had not yet experienced the duration event (the “risk set”). Of all those at risk, individual  $i$  is the one who actually experienced the duration at  $t$ , and it is this selection event that the partial likelihood considers. Thus, the partial likelihood is the likelihood that individual  $i$  is the one, of those at risk, who has the duration of  $t$ , given that *someone* is known to have a duration of  $t$ .

Since the hazard rate  $h(t)$  measures the likelihood of the duration event happening at  $t$  for those who have made it up to time  $t$  without experiencing an event, this rate determines the odds of selection in the partial likelihood for each individual at risk. Thus, for an observed time  $t$  at which individual  $i$  experiences a duration ( $T_i = t$ ), the partial likelihood that this duration indeed happened to individual  $i$  (and not to one of the other individuals at risk) is

$$L(i | t, j_1, \dots, j_{n(t)}) = \frac{h_i(t)}{\sum_{k=1}^{n(t)} h_{j_k}(t)}, \tag{13}$$



where  $n(t)$  is the number of individuals at risk at  $t$ , and these individuals are denoted  $j_1, \dots, j_{n(t)}$ .

Substituting the proportional hazards model (10) in (13) yields

$$L(i|t, j_1, \dots, j_{n(t)}) = \frac{h_0(t)e^{\beta'x_{it}}}{\sum_{k=1}^{n(t)} h_0(t)e^{\beta'x_{kt}}}, \quad (14)$$

for which the longitudinal effect  $h_0(t)$  cancels, leaving

$$L(i|t, j_1, \dots, j_{n(t)}) = \frac{e^{\beta'x_{it}}}{\sum_{k=1}^{n(t)} e^{\beta'x_{kt}}}. \quad (15)$$

The partial likelihood estimate of  $\beta$  is obtained by maximizing the product of expression (15) over all observed duration times. Note that, unlike the usual duration time regression models described in §4, the right-censored observations *do* enter the partial likelihood (15), i.e., these individuals, each having some covariate vector  $x_{jt}$  were at risk at  $t$  but did not experience the duration. The information in this event relevant for the response coefficient is appropriately taken into account in (15).

To summarize, the only thing “partial” about the partial likelihood is in the event it chooses to model. The total likelihood is concerned with the total duration event, i.e., “When will the duration occur for each individual?” The partial likelihood considers only part of the total duration event, namely, “Given that a duration occurred to someone at a specific time, *which* individual, of those still at risk, experienced it?” Since the answer to this latter question hinges on the relative riskiness of various individuals all measured at the same duration time, it comes as no surprise that the longitudinal effects  $h_0(t)$  drop out in (14), leaving (15) dependent only on the desired response coefficients  $\beta$ . Efron (1977) and Oakes (1977) provide evidence indicating that maximizing the partial likelihood results in very efficient estimates of  $\beta$ . Tsiatis (1981) shows that under general conditions the partial MLE is consistent and asymptotically normal.

#### 4.2. Parametric

Joint estimation of the baseline hazard and covariate effects can be accomplished through maximum likelihood. In constructing the likelihood function, we should distinguish between censored and uncensored observations. Further, we presume that the censoring mechanism is “noninformative,” that is, knowledge of the fact that an observation is censored does not convey any information except that the duration exceeds the censoring date. Sample data with a completed spell contribute a term  $f(t|\Theta)$  to the likelihood function. For individuals that are (right) censored at  $t$ , the only piece of information that we have is that their duration time is at least  $t$ . Such individuals contribute a term  $S(t|\Theta)$ , i.e., the probability of survival beyond  $t$ , to the likelihood function. The log-likelihood function is then:

$$\log l = \sum_i \delta_i \log f(t_i|\Theta) + \sum_i (1 - \delta_i) \log S(t_i|\Theta), \quad (16)$$

where the indicator variable  $\delta_i = 1$  if the observation is uncensored and  $\delta_i = 0$  otherwise. We can reformulate the log-likelihood in terms of the hazard rate. Remember that the hazard rate  $h(t|\Theta) = f(t|\Theta)/S(t|\Theta)$  and the integrated hazard rate

$$H(t|\Theta) = \int h(u|\Theta) du = -\log S(t|\Theta) \quad (\text{from (6)}).$$

Substituting these quantities into (16), one gets

$$\log l = \sum_i \delta_i \log h(t_i|\Theta) - \sum_i H(t_i|\Theta). \quad (17)$$

Under a set of regularity conditions, consistent estimates of the parameter vector  $\theta$  can be found via maximum likelihood estimation of (17). Further aspects of estimation of hazard models (e.g., estimation in the presence of unobserved heterogeneity) are covered in, e.g., Lancaster (1990, Chapters 8 and 9). A review of survival analysis software packages can be found in Goldstein et al. (1989).

## 5. Application to Interpurchase Times

We now illustrate and evaluate the proposed methodology with an analysis of households' interpurchase time decisions for premium brand saltine crackers. Jain and Vilcassim (1991) also apply a hazard rate modeling approach to interpurchase time decisions. Their focus is on unobserved heterogeneity. The application reported here differs in two major respects: (i) we allow for time-varying covariates, and (ii) we validate the hazard rate model on a holdout sample, using "standard" models (i.e., regression, probit) as a benchmark. Another departure is that our models ignore unobserved heterogeneity.

### 5.1. Data

UPC scanner panel data for saltine crackers, collected by IRI (Information Resources, Inc.), were used for estimating and validating our models. The data set covers approximately two years. A purchase file stores the purchase history of each panel member. A store file maintains weekly records of store shelf prices and promotions for each brand size. We restricted ourselves to premium brand purchases. Among the premium brands selected were the four major brands that account for about 65% of all saltine cracker purchases: Nabisco Premium NST (8.2%), Nabisco Premium (41.9%), Keebler Zesta (8.3%), and American Brands' Sunshine Krispy (6.0%). The calibration sample consisted of all households in the Williamsport, Pennsylvania market that satisfied the following two criteria: (i) at least two purchases were made, and (ii) at least 90% of the purchases contained one of the four premium brands selected. In this manner we retained 364 households. The same screening test was applied to the Rome, Georgia market to construct the validation sample, yielding 264 households.<sup>6</sup> In both samples we concentrated on the first interpurchase time spell observed for the panelists.

### 5.2. Variables

Three variables were selected to parameterize the models:

- $RP_t$  = the unit regular price,<sup>7</sup>
- $PPC_t$  = the promotional price cut, and
- AIT = average interpurchase time of a household (in weeks).

These variables were also considered by Gupta (1988) in his study of the IRI coffee panel data. The IRI store data records the shelf price. We constructed the regular price and promotional price cut from the history of shelf prices, using a procedure resembling the rule advocated by Abraham and Lodish (1989): a price decline is considered to be a promotional price cut when the decline is at least 5%, and the price is raised by at least 3% within four weeks. For each household, the regular price and discount were computed as a weighted average of the prices or discounts, respectively, in the store where most of the household's transactions occurred. The weights for each brand were based on the share of purchases accounted for by that brand within the household's total purchases.<sup>8</sup>

<sup>6</sup> The holdout data set was constructed from a market different from the market for the calibration sample. This was done to extend the external validity of our comparisons. Also note that possible biases due to differences across markets would impact all models.

<sup>7</sup> Given that all brands have the same size, no adjustment is needed of the unit price.

<sup>8</sup> Household specific weights were also used by Gupta (1988) in his study of purchase timing decisions.

The value of a household's average interpurchase time is computed across the household's entire purchase history.

Other covariates that are sometimes used in the literature were not included for various reasons. The presence of features and displays was very highly correlated with price cuts. Couponing was not very prominent in this market. Household inventory was not included in the models since such effects can be captured via the duration dependence terms and previous purchase occasion volume (Jain and Vilcassim 1991). However, since volume purchases seldom amounted to more than one unit, we did not include a volume covariate.<sup>9</sup>

### 5.3. Analysis and Results

In analyzing the data we were guided by the following three questions: (i) how do the estimates and accuracy of the forecasts evolve as we shift the right censoring point (i.e., the amount of time for which households are observed); (ii) how do the estimates of hazard rate models compare with estimates obtained using conventional procedures in terms of, e.g., stability and face validity; and (iii) to what extent do hazard rate models generate better forecasts? To answer the first question, we manipulated the right censoring point by stretching it from 8 to 16 weeks.<sup>10</sup> So a household that bought 9 weeks after the first purchase would have been right censored in the first estimation round but not in the second round. To address the second and third questions, we will compare the hazard modeling approach with two more common procedures used to analyze durations, namely duration time regression and probit.

Hazard rate models were estimated using the fully parametric approach. The CTM package developed by Yi et al. (1987) was used to estimate two special cases of the following hazard function:

$$h(t) = \exp\left\{\beta_0 + \beta_1 \text{RP}(t + \tau) + \beta_2 \text{PPC}(t + \tau) + \beta_3 \text{AIT} + \sum_{i=1}^2 \gamma_i (t^{\lambda_i} - 1) / \lambda_i\right\}; \quad (18)$$

(i) A Weibull baseline hazard ( $\lambda_1 = 0$ ,  $\gamma_2 = 0$ ), and (ii) a quadratic baseline hazard ( $\lambda_1 = 1$ ,  $\lambda_2 = 2$ ), where  $\tau$  is the starting point of the interpurchase time spell for a given household, that is, the time period at which the last purchase was made, and  $t + \tau$  is the calendar time. As benchmarks we calibrated a probit model and a regression model that consider completed spells only.<sup>11</sup> For the probit models, the dependent variable is a 0/1-variable that indicates whether (=1) or not (=0) a household repurchased saltine crackers by 8 (16 for the second cutoff point) weeks after its last purchase occasion. The regressors are identical to the ones used in the hazard model. The following specification was used for the regression models:

$$\log(t) = \alpha_0 + \alpha_1 \text{RP} + \alpha_2 \text{PPC} + \alpha_3 \text{AIT} + u, \quad (19)$$

where  $\log t$  is the natural logarithm of the household's interpurchase time interval and  $u$  is a disturbance term. Note that we have two time-varying covariates, the regular price and the promotional price cut. Only the hazard model (18) is able to handle time dependent variables. The other models were estimated using the within-spell average value for these variables.

<sup>9</sup> Note also that, to measure household inventories, several assumptions need to be made which themselves may introduce biases.

<sup>10</sup> Considering the median interpurchase time, the latter case (16 weeks) gives a large number of completed spells, compared to the former (8 weeks). If relative performance is related to the duration of the observation period, this comparison should reveal it.

<sup>11</sup> Deletion of censored observations was motivated by the fact that the other studies in marketing that have used duration time regressions also dropped all censored data.

#### 5.4. Discussion of Parameter Estimates

For the three variables that are included in all three methodologies, the anticipated sign of the coefficients is:

Variable	Model		
	Hazard Rate	Probit	Regression
Regular Price	Negative	Negative	Positive
Discount	Positive	Positive	Negative
Avg. Interpurchase Time	Negative	Negative	Positive

In the hazard model the “dependent variable” can be viewed as the hazard rate, i.e., the likelihood of a repurchase in the next time period, given that the repurchase in question has not yet occurred. This repurchase likelihood should be depressed in the presence of high regular prices, increased by the magnitude of discounts, and lower for households with larger average interpurchase times. Similarly, for the probit model the probability that a repurchase occurs during the observation period (8 weeks or 16 weeks) should be related negatively to the regular price, positively to the magnitude of discounts, and negatively to the household’s average interpurchase time. Finally, for the regression model the dependent variable (i.e., duration of the interpurchase time) should be increased by higher regular prices, decreased by discount magnitude, and positively related to the average household interpurchase time.

The estimates for the parametric hazard models are reported in Table 2. We first concentrate on the longitudinal effects. The duration dependence effects for the Weibull and quadratic baseline hazard specifications are portrayed in Figure 1. The Weibull result indicates a monotonically increasing hazard rate: as time elapses, a household is more likely to make its next transaction. The quadratic specification, however, reflects an initially increasing hazard rate reaching a peak at about four weeks (10 weeks for the 16-week cutoff period) and then declining.

TABLE 2  
Parameter Estimates: *Hazard Rate Model—Calibration Data Set*

	Censoring Point = 8 Weeks		Censoring Point = 16 Weeks	
	Weibull	Quadratic	Weibull	Quadratic
Intercept	1.989 (1.22)*	1.570 (1.28)	0.440 (0.86)	0.232 (0.92)
Ln ( $t$ )	0.513 (0.15)	.	0.423 (0.10)	.
$t$	.	0.891 (0.22)	.	0.242 (0.07)
$t^2$	.	-0.198 (0.05)	.	-0.024 (0.01)
Regular Price	-2.850 (1.07)	-2.669 (1.11)	-1.831 (0.76)	-1.637 (0.79)
Discount	0.757 (1.06)	0.815 (1.08)	1.299 (0.85)	1.184 (0.85)
Average Interp. Time	-0.195 (0.02)	-0.192 (0.02)	-0.140 (0.01)	-0.137 (0.01)
Negative Log Likelihood	417.3	412.6	717.6	720.0

Number of observations = 364.

\* Standard errors between parentheses.

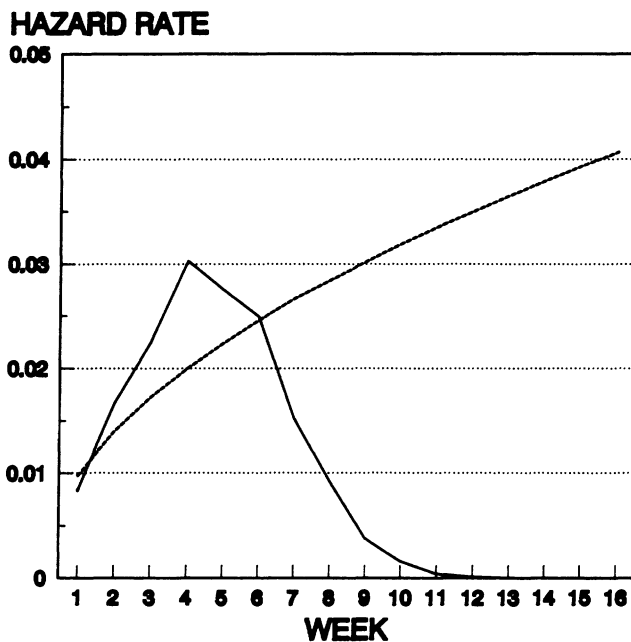


FIGURE 1. Hazard Rate—Quadratic versus Weibull Model.  
Censoring Point at 8 Weeks.  
Quadratic Specification: ———.  
Weibull Specification: - - - -.

Note that the two curves are close to one another up to about two weeks beyond the peak of the quadratic form. Given that the nonmonotonic pattern is probably more realistic, and is supported by Jain and Vilcassim's (1991) work, the quadratic specification will form the basis for making our forecasts.

The signs of the parameter estimates make intuitive sense: higher prices lead to longer interpurchase time intervals; households with longer average interpurchase times are expected to wait longer until their next purchase. For example, looking to the price coefficient in the last column of Table 2, we note that, by computing the quantity  $100(\exp(\beta) - 1)$ , a one dollar regular price increase will lower the hazard by about 84%. The coefficient for the average household interpurchase time tells us that each one-week increment in AIT will lead to a 14% decline in the hazard rate. The discount variable has the anticipated sign but is not significant. We also note that, as we move the censor point from 8 to 16 weeks, the directional effects do not change but the standard errors become smaller.<sup>12</sup> Comparing the price variable effects across the different censoring points, we observe that the magnitudes of the effects go down as we increase the censoring cutoff point. One possible explanation for this finding goes as follows. As the observation period lengthens, we pick up a larger group of households (with completed spells) that buy crackers very infrequently. If price matters less to these households than for heavy users (e.g., given that they have little information to form a strong reference point), this may explain the pattern observed.

Table 3 shows the estimates for the probit and regression models. For the probit estimates, the significant coefficients (i.e., the effect of average interpurchase time, AIT) show essentially the same directional impact of the covariates on interpurchase times as the hazard rate models. The effect of the regular price variable, however, becomes insignificant in the probit model, unlike the hazard model specification. More disturbingly,

<sup>12</sup> As the number of uncensored observations increases with the length of the cutoff period, the parameter estimates become more precise due to the gain in information on interpurchase time lengths.

TABLE 3  
*Parameter Estimates: Probit and Regression Models—Calibration Data Set*

Variable	Censoring Point			
	8 Weeks		16 Weeks	
	Probit	Regression Excl. Cens.	Probit	Regression Excl. Cens.
Intercept	3.981 (1.52)*	-0.718 (0.92)	3.100 (1.40)	-1.990 (0.79)
Regular Price	-2.273 (1.34)	1.409 (0.82)	-1.341 (1.23)	2.808 (0.70)
Discount	-2.694 (2.53)	-1.672 (0.92)	4.677 (2.98)	-1.057 (1.01)
Average Interpurchase Time	-0.132 (0.02)	0.042 (0.01)	-0.092 (0.01)	0.045 (0.01)
Adj. $R^2$	.	0.122	.	0.234

\* Standard errors between parentheses.

for the 8-week censor point the probit model indicates that discounts lengthen interpurchase times—a very unlikely result and one at odds with all of the hazard model results.<sup>13</sup> In contrast, the directional effects of the variables in the regression models are consistent with the estimates obtained for the hazard rate models.

### 5.5. Validation

As discussed earlier, we validated the models on a holdout sample composed of panel members that are located in another market. To check the stability of the estimates, we first re-estimated all models on the validation sample. The estimates for the hazard rate models are presented in Table 4. They are remarkably consistent with those for the calibration sample (see Table 2). (In the only notable discrepancy, for the 16-weeks cutoff case, the estimate for the regular price became nonsignificant.) Table 5 reports the parameter estimates for the benchmark models. Comparing the estimates in Table 5 with those for the calibration sample in Table 3, we can trace several departures. For the probit model, the sign of the regular price variable reverses, in the validation sample, relative to the calibration sample. In the regression model, the regular price variable is nonsignificant in all cases in contrast to significant effects for the calibration sample. Conversely, the price cut variable's effect becomes positive and significant in one case, after having been negative throughout all calibration cases (significant in one of them). As for probit above, finding a positive effect of price cut on interpurchase time would be highly counter-intuitive. We conclude that the stability of the estimates produced by the hazard rate models is not matched by the more conventional probit and regression models.

The second part of our validation compares the hazard rate model's predictions with those of the conventional models. We first briefly describe how forecasts for the hazard rate models were generated. The hazard rate is assumed to remain constant within each week. An estimate for the survival probability after one week is then  $\hat{S}(1) = \exp[-\hat{h}(1)]$ ; after two weeks  $\hat{S}(2) = \hat{S}(1) \exp[-\hat{h}(2)]$ , and so on. To contrast the hazard rate model with a probit model, we checked whether the reliability function  $\hat{S}(c)$  at the probit model's cutoff point was above or below 0.5. If above 0.5, the interpurchase time is predicted to have exceeded the probit's cutoff duration. (Recall that this is the only forecast feasible with probit-type models, unlike the case for hazard models.)

<sup>13</sup> This coefficient for discount, while not significant at the usual cut-off level, is never less 1.06 times its standard error and has an implausible sign.

TABLE 4  
*Parameter Estimates: Hazard Rate Model—Validation Data Set*

	Censoring Point = 8 Weeks		Censoring Point = 16 Weeks	
	Weibull	Quadratic	Weibull	Quadratic
Intercept	2.112 (1.07)*	1.774 (1.12)	0.043 (0.85)	-0.021 (0.90)
Ln ( <i>t</i> )	0.703 (0.131)	.	0.511 (0.09)	.
<i>t</i>	.	0.946 (0.20)	.	0.327 (0.07)
<i>t</i> <sup>2</sup>	.	-0.199 (0.05)	.	-0.036 (0.01)
Regular Price	-2.829 (0.97)	-2.668 (1.01)	-0.970 (0.77)	-0.956 (0.82)
Discount	1.510 (1.28)	1.312 (1.29)	0.285 (0.94)	0.287 (0.97)
Average Interp. Time	-0.256 (0.02)	-0.247 (0.02)	-0.202 (0.01)	-0.192 (0.01)
Negative Log Likelihood	403.1	401.7	576.1	580.0

Number of observations = 264.  
 \* Standard errors between parentheses.

To generate hazard rate model forecasts of the median duration (for comparison with the standard regression models), we compute the time point at which the reliability function drops below 0.5. Linear interpolation was then used to produce a forecast median duration.

Table 6 shows the hit rates (i.e., percent of households correctly classified) for the probit and quadratic hazard rate models for both samples. The hazard rate model predicts better than probit, although the gain is marginal. The probit model is *calibrated* to perform well in making this single specific prediction, i.e., whether the interpurchase time takes longer or shorter than *C* time units. The hazard rate model is calibrated to capture as well as possible *all* possible cutoff times, not the particular cutoff time *C*. One would thus expect it to perform less well than probit, but it does not.

TABLE 5  
*Parameter Estimates: Probit and Regression Models—Validation Data Set*

Variable	Censoring Point			
	8 Weeks		16 Weeks	
	Probit	Regression Excl. Cens.	Probit	Regression Excl. Cens.
Intercept	2.639 (1.47)*	1.436 (0.69)	-0.998 (1.86)	0.405 (0.70)
Regular Price	-0.713 (1.49)	-0.731 (0.68)	3.381 (1.97)	0.317 (0.69)
Discount	-0.940 (3.30)	1.128 (0.98)	-2.877 (4.83)	2.171 (1.10)
Average Interpurchase Time	-0.169 (0.02)	0.058 (0.01)	-0.138 (0.02)	0.073 (0.01)
Adj. <i>R</i> <sup>2</sup>	.	0.132	.	0.279

\* Standard errors between parentheses.

TABLE 6  
*Hit Rate—Probit Versus Hazard Rate Model*

Censoring Point	Model	
	Probit	Hazard Rate
8 weeks		
Calibration	78.7%	79.3%
Validation	76.7%	77.5%
16 weeks		
Calibration	79.3%	80.7%
Validation	84.7%	84.7%

In Table 7 we contrast the predictive performance of the quadratic hazard rate model vis-à-vis common regression models. Table 7 reflects two sets of predictions. When predicting with the hazard rate model, the reliability function for some cases (e.g., households with relatively large average interpurchase times) never drops below 0.5.<sup>14</sup> As we shift the censoring point, the number of cases in this set reduces drastically. Thus the first two rows in Table 7 are based on a common data set that covers all households for which the reliability function drops below 0.5 for the two different censoring points. The bottom row is based on the complete set of such households for the 16-weeks cutoff point. We readily see that the hazard model outperforms regression on all three validation measures. Another noteworthy observation is that the predictive performance does not substantially increase by lengthening the estimation period beyond eight weeks.

To summarize, we can state the following conclusions. Hazard models:

- (i) predict much better than regression models and perform at least as well as probit. Regarding the latter, they are also more flexible as one is not constrained to making predictions at the cutoff point for which the model was calibrated;
- (ii) provide more stable estimates (in terms of split sample reliability) than either probit or regression;
- (iii) apparently yield estimates with greater face validity (e.g., the positive signs for the discount variable coefficient for the results in Table 5).

TABLE 7  
*Validation Measures—Regression Model Versus Hazard Rate Model (Quadratic)*

Cutoff Point	Model					
	Regression Excl. Cens.			Hazard Rate Model		
	MAE*	MSE**	Prod. Mom.***	MAE	MSE	Prod. Mom.
8 weeks	3.664	35.977	0.513	2.880	23.780	0.676
16 weeks	3.355	30.986	0.430	2.886	21.880	0.654
16 weeks (complete set)	4.823	76.413	0.483	4.154	55.183	0.508

\* Mean absolute error of predicted duration time.

\*\* Mean squared error of predicted duration time.

\*\*\* Product moment correlation of predicted with actual duration.

<sup>14</sup> This can be explained by the fact that when  $\gamma_2 < 0$  (i.e., the coefficient associated with the squared duration dependence term), quadratic duration distributions are defective, that is,  $\lim_{t \rightarrow \infty} S(t) > 0$ , so that, in principle, the event may never take place.



## 6. Summary and Areas for Future Research

This paper examined the effectiveness of hazard rate modeling as a framework for analyzing duration times in marketing. We evaluated the approach for household inter-purchase times of a packaged good by comparing it with more conventional procedures. The empirical evidence favors the proposed methodology. In all fairness, we should add that hazard modeling also entails several limitations. The major weaknesses are:

- parameter estimates are not as readily interpretable as, say, regression estimates (In hazard models, the parameters have a simple proportionate effect on the hazard rate. Their effect, however, on various purchase probabilities and expected duration times is nonlinear and, while directionally straightforward, not simple to summarize numerically.);
- hazard models are generally calibrated via maximum likelihood which means that the estimation may fail to converge or produce local optima;
- predictions of duration times are somewhat more complicated to compute than for standard regression models; and,
- software availability and ease of use has been limited (but is improving).

Nevertheless, we feel that these weaknesses are counterbalanced by the benefits derived from hazard modeling. The key strengths are:

- ability to handle sample selection biases such as censoring;
- ability to incorporate time-varying covariates;
- no need to restrict oneself to a given time horizon when making inferences or predictions regarding the duration time process being studied; and
- ability to identify cross-sectional and longitudinal effects.

Making accurate inferences of covariate effects on duration times is important to marketers in many circumstances. For instance, in a new product pricing context, managers may face the issue of whether the price policies should differ for triers versus repeat buyers. A comparison of the price responsiveness of repeat buyers and triers may shed light on this issue (Helsen and Schmittlein 1990). Good insights in duration time processes also matter when managers desire to make forecasts of certain durations. As an example, accurate information on the timing of coupon redemptions matters to both manufacturers and retailers. For the manufacturer it is important to know when he becomes liable for coupon redemptions. The retailer, on the other hand, may be interested when possible advertising support for a coupon drop should take place (Blattberg and Neslin 1990, 299–301).

The methodology that we presented here may also prove useful for analyzing frequency processes. For example, it could be used as a tool to analyze single source data to track how the number of advertising exposures over a given period relates to the number of purchases of the advertised product during that time span (Gullen and Johnson 1986/1987).

In all these decision contexts we described (as well as many others—see Table 1), it is important to get accurate estimates of the covariate effects on the durations (or frequencies) of interest. To avoid the inaccuracies associated with the more conventional methods, the analyst may consider hazard models.

The research described here can be extended naturally in two directions. First, proportional hazards regression models as described and adapted above can be applied to other marketing durations data, and can address key substantive issues regarding those durations. Second, there is a great deal of room for further development of proportional hazards models themselves.

### 6.1. *Application to Substantive Issues in Marketing Decision Situations*

A wide variety of marketing durations have been described in the Introduction and summarized in Table 1. We will not repeat that discussion here; obviously the techniques

of §§4 and 5 can be used with any of these duration phenomena. But we do want to stress the link between modeling those durations and resolving substantive issues in marketing.

Interpurchase times provide a good example of this link. To what practical use can a longitudinal/cross-sectional model of interpurchase times be put?

(i) *Promotion Evaluation*. Do promotions accelerate purchases? Translated into the concepts developed in this paper, do promotions shift up the hazard rate for interpurchase times? In fact, the hazard rate paradigm suggests that other questions of interest here ought to be:

- do promotions affect the hazard rate  $h(t)$  in the same way at all interpurchase times  $t$ ?

- do promotions interact with other covariates to determine the hazard function  $h(t)$ ?

(ii) *Customer Valuation*. How valuable is a customer known to have a certain demographic/psychographic profile and a certain purchase history? This is a key question in:

- strategic planning, i.e., how much sales volume is expected from the current customer base (Schmittlein, Morrison and Colombo 1987)?

- merger and acquisition decisions, i.e., how much is a business worth?

- customer prospecting, i.e., how much should the firm be willing to pay to acquire a new customer?

The answer to each of these questions is provided by a customer-level model predicted when purchases will occur (e.g., proportional hazards regression), possibly combined with a model of purchase quantities.

(iii) *Timing Customer Contacts*. What is the ideal timing of direct marketing, sales calls, or advertising? “Reminder” type advertising/direct marketing should be timed to coincide with a drop in the customer’s hazard rate. On the other hand, “sale closing” type contacts should be targeted during times when the hazard rate is relatively high.

(iv) *Customer File Management*. When do customers become “inactive,” and thus merit being dropped from a firm’s file of active customers? Again, the hazard rate for a future purchase is the key quantity.

(v) *Forecasting New Product Success*. After trial (also amenable to proportional hazards models), the timing of repeat purchase is an important determinant of long-run market share.

The reader can no doubt imagine additional examples. The point is that an understanding of the cross-sectional and longitudinal patterns of these durations is useful in making marketing decisions.

## 6.2. Methodological Developments

Methodological/modeling issues deserving attention include:

- estimation in the presence of left-censored durations;
- development of appropriate parametric models that capture important qualitative characteristics of specific marketing duration processes (e.g., “loyalty,” “inertia”);

- linking these duration time models with models for other related marketing phenomena (e.g., linking models for interpurchase time and quantity purchased; see Lancaster 1985); and

- design trade-offs in collecting marketing durations data. Under what conditions would it be better to observe a smaller sample of individuals for a longer time (i.e., extending the right censor date  $C_R$ ) versus watching a larger sample for a shorter time?

## 7. Conclusions

Many types of durations are clearly of interest to marketers. We hope that those encountering these data will take away three general points from this article.

1. “*Durations are different.*”

Most importantly, the left and right censoring and left truncation uniquely associated with durations data generally make the “usual” econometric analyses inappropriate. More is at stake than niceties of estimation efficiency: those econometric techniques do not allow one to make the major predictions of interest with durations data, and do not even provide consistent estimates of the effects that they attempt to model.

2. “*Proportional hazards regression is a natural approach for modeling durations.*”

As seen in §§3, 4, and 5, the hazard rate paradigm and resulting model does allow the analyst to answer the key questions associated with longitudinal and cross-sectional characteristics of duration times.

3. “*Proportional hazards regression tends to give more reliable estimates, with greater face validity, than conventional methods. Moreover, there is evidence the forecasts generated by PHR are more accurate.*”

This principle was illustrated empirically in the previous section with an application to interpurchase times. With marketers’ increased interest in investigating both the longitudinal and cross-sectional characteristics of marketing phenomena, the methods in this paper provide both a natural paradigm and tool for understanding these characteristics.<sup>15</sup>

**Acknowledgment.** The first author acknowledges financial support from the Graduate School of Business, The University of Chicago.

<sup>15</sup> This paper was received in June 1989 and has been with the authors 10 months for 2 revisions. Processed by Imran Currim, former Area Editor.

## References

- Abraham, M. N. and L. M. Lodish (1989), “PROMOTIONSCAN: A System for Improving Promotion Productivity for Retailers and Manufacturers Using Scanner Store and Household Panel Data,” Working Paper, Marketing Department, The Wharton School, University of Pennsylvania.
- Amemiya, T. (1985), *Advanced Econometrics*, Cambridge, MA: Harvard University Press.
- Bass, F. M. (1969), “A New Product Growth Model for Consumer Durables,” *Management Science*, 15, 215–227.
- Blattberg, R. C. and S. A. Neslin (1990), *Sales Promotions: Concepts, Methods and Strategies*, Englewood Cliffs, NJ: Prentice-Hall.
- Breslow, N. E. (1974), “Covariance Analysis of Censored Survival Data,” *Biometrics*, 30, 89–99.
- Chamberlain, G. (1984), “Heterogeneity, Omitted Variable Bias and Duration Dependence,” in J. Heckman and B. Singer (Eds.), *Longitudinal Analysis of Labor Market Data*, New York: Academic Press, 3–38.
- Cox, D. R. (1972), “Regression Models and Life Tables (with discussion),” *Journal of the Royal Statistical Society Series B*, 34, 187–220.
- , (1975), “Partial Likelihood,” *Biometrika*, 62, 269–276.
- and D. Oakes (1984), *Analysis of Survival Data*, London: Chapman and Hall.
- Cox, W. E., Jr. (1966), “Response Patterns to Mail Surveys,” *Journal of Marketing Research*, 3, 392–397.
- Dunn, R., S. Reader and N. Wrigley (1983), “An Investigation of the Assumptions of the NBD Model as Applied to Purchasing at Individual Stores,” *Applied Statistics*, 32, 249–259.
- Duwors, R. E. Jr. and G. H. Haines Jr. (1990), “Event History Analysis of Brand Loyalty,” *Journal of Marketing Research*, 27, 485–493.
- Efron, B. (1977), “Efficiency of Cox’s Likelihood Function for Censored Data,” *Journal of the American Statistical Association*, 72, 557–565.
- Flinn, C. J. and J. Heckman (1982), “Models for the Analysis of Labor Force Dynamics,” in R. L. Bassman and G. F. Rhodes Jr. (Eds.), *Advances in Econometrics*, Greenwich, CT: JAI Press Inc., 3–34.
- Freeman, J., G. R. Carroll and M. T. Hannan (1983), “The Liability of Newness: Age Dependence in Organizational Death Rates,” *American Sociological Review*, 48, 692–710.
- Glasser, M. (1967), “Exponential Survival with Covariance,” *Journal of American Statistical Association*, 62, 561–568.
- Goldstein, R., J. Anderson, A. Ash, B. Craig, D. Harrington and M. Pagano (1989), “Survival Analysis Software on MS/PC-DOS Computers,” *Journal of Applied Econometrics*, 4, 393–414.

- Greene, J. D. (1982), *Consumer Behavior Models for Non-Statisticians*, New York: Praeger.
- Gullen, P. and H. Johnson (1986/87), "Relating Product Purchasing and TV Viewing," *Journal of Advertising Research*, 27, 9–19.
- Gupta, S. (1988), "Impact of Sales Promotions on When, What and How Much to Buy," *Journal of Marketing Research*, 25, 342–355.
- (1991), "Stochastic Models of Interpurchase Time with Time-Dependent Covariates," *Journal of Marketing Research*, 28, 1–15.
- Heckman, J. and B. Singer (1982), "The Identification Problem in Econometric Models for Duration Data," in W. Hildenbrand (Ed.), *Advances in Econometrics*, Cambridge: Cambridge University Press, 39–77.
- and ——— (1984), "Econometric Duration Analysis," *Journal of Econometrics*, 24, 63–132.
- Helsen, K. and D. C. Schmittlein (1990), "Understanding Price Effects for New Nondurables: How Price Responsiveness Varies Across Depth-of-Repeat Classes and Types of Consumers," *European Journal of Operational Research*, forthcoming.
- Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge: Cambridge University Press.
- Jain, D. C. and N. Vilcassim (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, 10, 1–23.
- Kalbfleisch, J. D. and R. L. Prentice (1980), *The Statistical Analysis of Failure Time Data*, New York: Wiley.
- Kiefer, N. M. (1988), "Economic Duration Data and Hazard Functions," *Journal of Econometric Literature*, 26, 649–679.
- Lancaster, T. (1985), "Simultaneous Equations Models in Applied Search Theory," *Journal of Econometrics*, 28, 113–126.
- (1990), *The Econometric Analysis of Transition Data*, Cambridge: Cambridge University Press.
- Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data*, New York: John Wiley and Sons.
- Lucas, G. H., Jr., A. Parasuraman, R. A. Davis and B. M. Enis (1987), "An Empirical Study of Salesforce Turnover," *Journal of Marketing*, 51 (July), 34–59.
- Morrison, D. G. and D. C. Schmittlein (1980), "Jobs, Strikes and Wars: Probability Models for Duration," *Organizational Behavior and Human Performance*, 25, 224–251.
- Neslin, S. A., C. Henderson and J. Quelch (1985), "Consumer Promotions and the Acceleration of Product Purchases," *Marketing Science*, 4, 147–165.
- Oakes, D. (1977), "The Asymptotic Information in Censored Survival Data," *Biometrika*, 64, 441–448.
- Peterson, R. A., G. Albaum and N. M. Ridgway (1989), "Consumers Who Buy from Direct Sales Companies," *Journal of Retailing*, 65, 273–286.
- Peto, R. (1972), Contribution to the discussion of paper by D. R. Cox, *Journal of the Royal Statistical Society B*, 34, 205–207.
- and P. Lee (1973), "Weibull Distributions for Continuous Carcinogenesis," *Biometrics*, 29, 457–470.
- Pike, M. C. (1966), "A Method of Analysis of Certain Class of Experiments in Carcinogenesis," *Biometrics*, 22, 142–161.
- Schmittlein, D. C. and K. Helsen (1989), "A Method for Analyzing Left-Truncated Marketing Durations with an Application to Panel Dropout," Working Paper, Marketing Department, The Wharton School, University of Pennsylvania.
- and V. Mahajan (1982), "Maximum Likelihood Estimation for an Innovation Diffusion Model of New Product Acceptance," *Marketing Science*, 1, 57–78.
- and D. G. Morrison (1983a), "Modeling and Estimation Using Job Duration Data," *Organizational Behavior and Human Performance*, 32, 1–22.
- and ——— (1983b), "Prediction of Future Random Events with the Condensed Negative Binomial Distribution," *Journal of the American Statistical Association*, 78, 449–456.
- Schmittlein, D. C., D. G. Morrison and R. Colombo (1987), "Counting Your Customers: Who Are They and What Will They Do Next?" *Management Science*, 33, 1–24.
- Scott, C. (1961), "Research on Mail Surveys," *Journal of the Royal Statistical Society, Series A*, 124 (Part 2), 143–205.
- Sharma, S. and V. Mahajan (1980), "Early Warning Indicators of Business Failure," *Journal of Marketing*, 44, 80–89.
- Singh, J. V., D. J. Tucker and R. J. House (1986), "Organizational Legitimacy and the Liability of Newness," *Administrative Science Quarterly*, 31, 171–193.
- Tsiatis, A. A. (1981), "A Large Sample Study of Cox's Regression Model," *The Annals of Statistics*, 9, 93–108.
- Vaupel, J. W. and A. I. Yashin (1985), "Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics," *The American Statistician*, 39, 3, 176–185.
- Yi, K. M., B. Honoré and J. Walker (1987), *A Program for Estimation and Testing of Continuous Time Multi-State Multi-Spell Models*, Chicago: ERC/NORC.