

**CONTENTS**

Linking to a Large-Scale Assessment: An Empirical Evaluation <b>Bruce Bloxom, Peter J. Pashley, W. Alan Nicewander,</b> and <b>Duanli Yan</b>	1
A Maximum Test for Scale: Type I Error Rates and Power <b>James Algina, R. Clifford Blair, and William T. Coombs</b>	27
It's Déjà Vu All Over Again: Using Multiple-Spell Discrete-Time Survival Analysis <b>John B. Willett and Judith D. Singer</b>	41
The Impact of BIB Spiraling-Induced Missing Data Patterns on Goodness-of-Fit Tests in Factor Analysis <b>David Kaplan</b>	69
Improved Repeated Measures Stepwise Multiple Comparison Procedures <b>H. J. Keselman and Lisa M. Lix</b>	83
Corrections to "Using Results From Replicated Studies to Estimate Linear Models" <b>Betsy Jane Becker</b>	100
<b>Book Reviews</b> Categorical Longitudinal Data: Log-Linear Panel, Trend, and Cohort Analysis, by Jacques A. Hagenaars <b>Scott R. Eliason</b>	103
A User's Guide to Principal Components, by J. Edward Jackson <b>Roger E. Millsap</b>	105

## It's Déjà Vu All Over Again: Using Multiple-Spell Discrete-Time Survival Analysis

John B. Willett and Judith D. Singer  
Harvard University

**Key words:** survival analysis, longitudinal data, event history analysis

*Multiple-spell discrete-time survival analysis can be used to investigate the repeated occurrence of a single event, or the sequential occurrence of disparate events, including: students' and teachers' entries into, and exits from, school; childrens' progress through stages of cognitive reasoning; disturbed adolescents' repeated suicide attempts; and so forth. In this article, we introduce and illustrate the method using longitudinal data on exit from, and reentry into, the teaching profession. The advantages of the approach include: (a) applicability to many educational problems; (b) easy inclusion of time-invariant and time-varying predictors; (c) minimal assumptions—no proportional-hazards assumption is invoked and so the effects of predictors can vary over time within, and across, spells; and (d) all statistical models can be fit with a standard logistic regression package.*

Social scientists often pose questions about sequences of repeatable events. For instance, clinical psychologists explore the occurrence of, and remission from, affective illness. Educational researchers investigate students' and teachers' entries into, and exits from, school. Sociologists examine recidivism among juvenile delinquents and convicted felons. Developmental psychologists follow children across stages of cognitive and moral reasoning.

Although survival analysis (Allison, 1982; Blossfeld, Hamerle, & Mayer, 1989; Cox, 1972; Kalbfleisch & Prentice, 1980; Miller, 1981) is ideal for examining patterns of event occurrence, researchers studying the occurrence of a sequence of one or more repeatable events face difficulty. Most technical work on survival methods has focused on *single nonrepeatable events* occurring in *continuous time* under the *proportional-hazards assumption* with *time-invariant predictors*. When lifetimes consist of many *spells* or *episodes*, methods for single episodes may not suffice. And often, time is measured discretely, the proportional-hazards assumption is untenable, and values of important predictors fluctuate with time.<sup>1</sup>

---

The order of the authors was determined by randomization. They would like to thank the American Statistical Association and the National Science Foundation for support during their research fellowships at the National Center for Educational Statistics, where this article was written. They are also indebted to Paul D. Allison, whose work set the stage for this article and whose private communications have provided much thoughtful insight while the article was being written. Comments and questions can be addressed to either author at the Harvard University Graduate School of Education, Appian Way, Cambridge, Massachusetts 02138.

Anyone in the population may experience the requisite target event—and hence terminate the current spell—in any time period. We define a dichotomous random variable,  $Y_{ijk}$ , whose values  $y_{ijk}$  indicate whether spell  $j$  is terminated in period  $k$  for individual  $i$ , where

$$y_{ijk} = \begin{cases} 0 & \text{if individual } i \text{ does NOT experience an event} \\ & \text{in period } k \text{ of spell } j, \\ 1 & \text{if individual } i \text{ DOES experience an event} \\ & \text{in period } k \text{ of spell } j. \end{cases} \quad (1)$$

Within a spell, some people may not experience the terminating event; their event histories are *censored* (Allison, 1984; Kalbfleisch & Prentice, 1980; Singer & Willett, 1993). Once censoring occurs, the person is not eligible to experience further spells. For instance, 1,734 of our special educators were censored in the first spell; they entered teaching and remained there throughout the data collection period, never experiencing the first event of interest (i.e., leaving teaching for the first time). Similarly, the careers of 1,471 special educators were censored in the second spell, 470 in the third, and 382 in the fourth.

#### Describing the Distribution of Event Occurrence Within Spell

If  $T_j$  is a discrete random variable denoting the time period, within spell  $j$ , when an event occurs, then the distribution of  $T_j$  can be described by the probability mass function

$$f_{ij}(k) = \text{Pr}_i\{T_j = k\} \quad (2)$$

denoting the probability that individual  $i$ 's  $j$ th spell will be terminated in the  $k$ th period. In response to the problem of censoring (Allison, 1984; Kalbfleisch & Prentice, 1980; Singer & Willett, 1993), attention is usually directed away from the probability mass function to the *survivor function* (the probability that individual  $i$ 's  $j$ th spell will be terminated after time period  $k$  of that spell),

$$S_{ij}(k) = \text{Pr}_i\{T_j > k\}, \quad (3)$$

and the *hazard function* (the probability that individual  $i$ 's  $j$ th spell will be terminated in the  $k$ th time period of that spell, given that they did not experience the event earlier),

$$h_{ij}(k) = \text{Pr}_i\{T_j = k | T_j \geq k\}. \quad (4)$$

Within the  $j$ th spell, individual  $i$ 's temporal profile of risk is represented by the population *hazard function*, a vector of hazard probabilities,  $h_{ij}(1), h_{ij}(2),$

In this article, we extend our earlier study of single-spell discrete-time survival analysis (Singer & Willett, 1993) to the multiple-spell case, thereby providing a method that permits the simultaneous analysis of data from all spells, the easy inclusion of time-invariant and time-varying predictors, no reliance on the proportional-hazards assumption, and parameter estimation that utilizes standard software for predicting binary outcomes. In the first section of the article, we define the multiple-spell discrete-time hazard function, derive the sample likelihood function, posit a model relating hazard to predictors, and establish estimation strategies. In the second section, we illustrate data-analytic approaches unique to the multiple-spell case using data on the career histories of 3,941 special educators.

#### The Multiple-Spell Discrete-Time Survival Process

In a population at risk of experiencing a sequence of events, each person's history can be divided into spells, each ending when the corresponding event in the sequence occurs. The methods described here apply to the repeated occurrence of a single event (e.g., repeated suicide attempts), or to the sequential occurrence of disparate events in an order that is fixed across people (e.g., progression through stages of moral development). In our data example, we explore the occurrence of a pair of alternating events in the lives of special educators—the events of leaving, and then returning to, teaching. We possess 12 years of longitudinal data describing up to four spells for each educator: (a) a *first spell* in teaching, (b) a *second spell* out of teaching, (c) a *third spell* in teaching, and (d) a *fourth spell* out of teaching.

Some population members may experience every event under study, others may not, but anyone who begins a later spell must have completed all earlier spells. Thus, progression through spells is conditional; experience of the first event qualifies a person to begin the second spell, and so on. In our example, 1,734 special educators experienced only a first spell; 1,471 experienced a first and a second spell; 470 a first, a second, and a third spell; and 266 experienced all four spells. We use index  $j$  to count spells,  $J_i$  to represent the number of spells experienced by the  $i$ th individual, and  $J$  to indicate the maximum number of spells experienced by anyone in the population ( $J = \text{Max}_i(J_i)$ ).

Within a spell, we partition time into discrete periods with the  $k$ th period of the  $j$ th spell,  $(t_{j,k-1}, t_{j,k}]$ , beginning immediately after time  $t_{j,k-1}$  and extending to, and including, time  $t_{j,k}$  (see Karlin & Taylor, 1975; Miller, 1981). This partition preserves the definition and continuity of the cumulative distribution function and the survivor function.<sup>2</sup> We denote the last time period when individual  $i$  was observed in spell  $j$  by  $K_{ij}$ ; the maximum number of periods experienced by any individual in spell  $j$  is  $K_j = \text{Max}_i(K_{ij})$ ; and, over all spells, the maximum number of periods experienced by any person in any spell is  $K = \text{Max}_j(K_j)$ .

...,  $h_{ij}(k)$ , ... These hazard probabilities are the fundamental parameters of the discrete-time survival process.<sup>3</sup>

Discrete-time hazard in (4) can be expressed using the notion of a risk set. Within a spell, the terminating event cannot occur twice, and so, in time period  $k$  of spell  $j$ , only a subset of the population remains eligible to experience the  $j$ th event—those who have not already experienced it. These *eligibles* make up the population risk set for period  $k$  of spell  $j$  and the corresponding population hazard probability is the proportion of the risk set that experiences the  $j$ th event in this period. Members of the period  $k$  risk set that experience the  $j$ th event in period  $k$  are removed from the risk set for period  $(k + 1)$ . Then, the hazard probability in period  $(k + 1)$  is the proportion of the newly reduced risk set experiencing the  $j$ th event in period  $(k + 1)$ . As time passes, and events occur, the risk set diminishes.

Censoring also reduces the risk set, but the definition of population hazard—as a proportion of the population risk set—remains unchanged. At censoring, we lose track of a person's career; we do not know whether he or she experiences the event in the future. Censored people are removed from the population risk set at the point of censoring, contributing to hazard up until that time by being present in earlier risk sets. In any spell and period, we assume that the censored risk set is a random subpopulation of the unavailable "uncensored" risk set—i.e., that censoring is independent of the event history process.

### Estimation

#### Constructing the Sample Likelihood Function

Any sample of individuals can be viewed as a set of mutually exclusive subsamples where Subsample 1 contains  $n_1$  people who experienced only Spell 1, Subsample 2 contains  $n_2$  people who experienced Spells 1 and 2, subsample  $j$  contains  $n_j$  individuals who experienced spells 1 through  $j$ , and so on. The number of the spell in which individual  $i$  was last observed,  $J_i$ , corresponds to that individual's subsample membership number; for instance, if  $J_i = 1$ , then Spell 1 was individual  $i$ 's last and he or she is a member of Subsample 1, and so on. Since the maximum number of spells observed for any person is  $J$ , there are  $J$  subsamples in total, and overall sample size,  $n$ , is the sum of subsample sizes,  $n_1 + n_2 + \dots + n_J$ . In general, sample members can be arrayed in order of increasing subsample identification number.

Members of Subsample 1 experience only the first spell. Uncensored members experience the requisite event in the last time period  $K_{i1}$  in which they were observed. Their contribution to the sample likelihood equals the probability that they experienced the event in time period  $K_{i1}$ :

$$L_i^{(1)}(\text{uncensored}) = \Pr_i\{T_1 = K_{i1}\} = f_{i1}(K_{i1}) \quad (5)$$

where  $f_{i1}(K_{i1})$  is the value of the probability mass function of  $T_1$  in time period  $K_{i1}$ . Censored members of Subsample 1 do not experience the first event in any time period up to, and including, the last time period in which they were observed. Their contribution to the sample likelihood equals the probability that the first event will occur after time period  $K_{i1}$ .

$$L_i^{(1)}(\text{censored}) = \Pr_i\{T_1 > K_{i1}\} = S_{i1}(K_{i1}) \quad (6)$$

where  $S_{i1}(K_{i1})$  is the value of the survivor function of  $T_1$  in time period  $K_{i1}$ . We can combine (5) and (6) by defining a dichotomous censoring indicator,  $C_i$ , with values

$$c_{ij} = \begin{cases} 0 & j\text{th spell of individual } i \text{ is NOT censored,} \\ 1 & j\text{th spell of individual } i \text{ IS censored.} \end{cases} \quad (7)$$

Now the contribution to the likelihood by any member of Subsample 1 is

$$L_i^{(1)} = [f_{i1}(K_{i1})]^{1-c_{i1}} [S_{i1}(K_{i1})]^{c_{i1}}. \quad (8)$$

Members of Subsample 2 experience both first and second spells. To enter the second spell, they must experience the first event and not be censored. Once *qualified*, their second spell can terminate by event occurrence or by censoring. For an uncensored person (for whom  $c_{i2} = 0$ ), the second event occurs at the end of the second spell in time period  $K_{i2}$ . But, to account for the qualifying event at the end of Spell 1, the contribution to the sample likelihood is a product of (a) the probability of terminating Spell 1 with an event in time period  $K_{i1}$ , and (b) the probability of terminating Spell 2 with an event in time period  $K_{i2}$ . For a censored person (for whom  $c_{i2} = 1$ ), the second spell terminates without event occurrence. But, to account for the occurrence of the qualifying event at the end of Spell 1, the contribution to the sample likelihood is the product of (a) the probability of terminating Spell 1 with an event in time period  $K_{i1}$ , and (b) the probability of terminating Spell 2 without experiencing the second event in, or before, time period  $K_{i2}$ . Thus,

$$L_i^{(2)} = [\Pr_i\{T_1 = K_{i1}\}] \{[\Pr_i\{T_2 = K_{i2}\}]^{1-c_{i2}} [\Pr_i\{T_2 > K_{i2}\}]^{c_{i2}}\} \\ = [f_{i1}(K_{i1})][f_{i2}(K_{i2})]^{1-c_{i2}} [S_{i2}(K_{i2})]^{c_{i2}}$$

where  $f_{i2}(K_{i2})$  and  $S_{i2}(K_{i2})$  are the values of the probability mass function and the survivor function of  $T_2$  in time period  $K_{i2}$  for individual  $i$ , respectively. Because experience of the first event at the end of Spell 1 is a precursor to

The values of the probability mass function and the survivor function in the last time period of any spell can be expressed in terms of hazard probabilities. The value of the probability mass function in the last time period of spell  $j$  for individual  $i$ ,  $f_{ij}(K_{ij})$ , is a product of terms, one per time period within spell  $j$ , describing the conditional probabilities that the event did not occur in periods 1 through  $K_{ij} - 1$ , but did occur in period  $K_{ij}$ :

$$f_{ij}(K_{ij}) = \Pr_i\{T_j = K_{ij} | T_j \geq K_{ij}\} \Pr_i\{T_j \neq K_{ij} - 1 | T_j \geq K_{ij} - 1\} \dots \Pr_i\{T_j \neq 1 | T_j \geq 1\} \\ = h_{ij}(K_{ij}) \prod_{k=1}^{K_{ij}-1} (1 - h_{ij}(k)) \tag{15}$$

Similarly, the value of the survivor function for individual  $i$  in the last time period of spell  $j$ ,  $S_{ij}(K_{ij})$ , is a product of terms, one per period, describing the conditional probabilities that the event did not occur in any observed period, 1 through  $K_{ij}$ :

$$S_{ij}(K_{ij}) = \Pr_i\{T_j \neq K_{ij} | T_j \geq K_{ij}\} \Pr_i\{T_j \neq K_{ij} - 1 | T_j \geq K_{ij} - 1\} \dots \Pr_i\{T_j \neq 1 | T_j \geq 1\} \\ = \prod_{k=1}^{K_{ij}} (1 - h_{ij}(k)) \tag{16}$$

Substituting from (15) and (16) into (14), we have

$$L = \prod_{i=1}^n \prod_{j=1}^{J_i} \left[ h_{ij}(K_{ij}) \prod_{k=1}^{K_{ij}-1} (1 - h_{ij}(k)) \right]^{1-c_{ij}} \left[ \prod_{k=1}^{K_{ij}} (1 - h_{ij}(k)) \right]^{c_{ij}} \\ = \prod_{i=1}^n \prod_{j=1}^{J_i} \left[ \frac{h_{ij}(K_{ij})}{1 - h_{ij}(K_{ij})} \right]^{1-c_{ij}} \left[ \prod_{k=1}^{K_{ij}} (1 - h_{ij}(k)) \right] \tag{17}$$

In (17), the observed data are represented by values of the censoring indicator,  $C$ , and values of  $C$  are connected to values of the event indicator,  $Y$ . Within any spell for a given individual, the sequence of  $y$ -values can take on one of two patterns. If the requisite event did not occur in spell  $j$  and individual  $i$  was censored, then the  $y_{ik}$  are zero in all periods observed during the spell,  $k = 1, \dots, K_{ij}$ . Alternatively, if there was no censoring and the requisite event occurred in the last time period of the spell,  $K_{ij}$ , observed for

participation in Spell 2, the value of the Spell 1 censoring indicator,  $c_{i1}$ , must be zero for all members of Subsample 2. Thus,

$$f_{i1}(K_{i1}) = [f_{i1}(K_{i1})]^{1-c_{i1}} [S_{i1}(K_{i1})]^{c_{i1}} \quad (i \in \text{Subsample 2}) \tag{10}$$

and substitution into (9) leads to

$$L_i^{(2)} = [f_{i1}(K_{i1})]^{1-c_{i1}} [S_{i1}(K_{i1})]^{c_{i1}} [f_{i2}(K_{i2})]^{1-c_{i2}} [S_{i2}(K_{i2})]^{c_{i2}} \tag{11}$$

This process of constructing the contribution of each member to the total sample likelihood can continue until all  $J$  subsamples are exhausted. For instance, a member of subsample  $j$  must have terminated all earlier spells with requisite events and so, generalizing from (11), his or her contribution must be

$$L_i^{(j)} = [f_{i1}(K_{i1})]^{1-c_{i1}} [S_{i1}(K_{i1})]^{c_{i1}} [f_{i2}(K_{i2})]^{1-c_{i2}} [S_{i2}(K_{i2})]^{c_{i2}} \dots [f_{ij-1}(K_{ij-1})]^{1-c_{ij-1}} [S_{ij-1}(K_{ij-1})]^{c_{ij-1}} [f_{ij}(K_{ij})]^{1-c_{ij}} [S_{ij}(K_{ij})]^{c_{ij}} \tag{12}$$

Since individuals are drawn randomly from the population, the net sample likelihood is a product of the contributions of all individuals in all  $J$  subgroups:

$$L = \left( \prod_{i=1}^{n_1} [f_{i1}(K_{i1})]^{1-c_{i1}} [S_{i1}(K_{i1})]^{c_{i1}} \right) \\ \times \left( \prod_{i=n_1+1}^{n_1+n_2} [f_{i1}(K_{i1})]^{1-c_{i1}} [S_{i1}(K_{i1})]^{c_{i1}} [f_{i2}(K_{i2})]^{1-c_{i2}} [S_{i2}(K_{i2})]^{c_{i2}} \right) \\ \times \dots \times \left( \prod_{i=n_1+\dots+n_{j-1}+1}^{n_1+\dots+n_j} [f_{i1}(K_{i1})]^{1-c_{i1}} [S_{i1}(K_{i1})]^{c_{i1}} [f_{i2}(K_{i2})]^{1-c_{i2}} [S_{i2}(K_{i2})]^{c_{i2}} \right) \\ \dots [f_{ij}(K_{ij})]^{1-c_{ij}} [S_{ij}(K_{ij})]^{c_{ij}} \tag{13}$$

which, because subsample membership is determined by the spell number  $J_i$  of the last spell for which an individual was observed, can be written as:

$$L = \prod_{i=1}^n \prod_{j=1}^{J_i} [f_{ij}(K_{ij})]^{1-c_{ij}} [S_{ij}(K_{ij})]^{c_{ij}} \tag{14}$$

The Person-Spell-Period Data Set

The equivalence of (19) and the likelihood for an independent Bernoulli trials model underwrites a simple algorithm for analyzing multiple-spell discrete-time survival data. Each of the  $n$  people in the original sample contributes multiple independent observations to a Person-Spell-Period data set in which there are  $N$  lines of data—one observation for every time period in every spell for which each person in the original sample was observed. Table 1 presents a Person-Spell-Period data set containing the histories of two special educators. There are multiple data lines for each teacher, corresponding to the years of the career.<sup>4</sup> Teacher 1 was observed for 11 years and out of teaching again. Teacher 2 has 12 records: spells of duration 8, 1, and 3 years, respectively. The entire Person-Spell-Period data set for the 3,941 special educators in our example has 38,672 records.

TABLE 1  
Structure and contents of the special educator Person-Spell-Period data set

Teacher	SP NUM	PD NUM	FEMALE	ASSIGN	Y
1	1	1	1	1	0
1	1	2	1	0	0
1	1	3	1	0	1
1	2	1	1	•	1
1	3	1	1	0	0
1	3	2	1	0	0
1	3	3	1	0	0
1	3	4	1	0	0
1	3	5	1	0	1
1	4	1	1	•	0
1	4	2	1	•	1
2	1	1	0	0	0
2	1	2	0	0	0
2	1	3	0	0	0
2	1	4	0	0	0
2	1	5	0	0	0
2	1	6	0	1	0
2	1	7	0	1	0
2	1	8	0	1	1
2	2	1	0	•	1
2	3	1	0	0	0
2	3	2	0	0	0
2	3	3	0	0	0
2	3	3	0	0	0

individual  $i$ , then the  $y_{jk}$  will all be zero except for the last, which will be one. Then,

$$\prod_{k=1}^{k_{ij}} \left[ \frac{h_{ij}(k)}{1 - h_{ij}(k)} \right]^{y_{jk}} = \left[ \frac{h_{ij}(K_{ij})}{1 - h_{ij}(K_{ij})} \right]^{1 - y_{jk}} \tag{18}$$

Substituting from (18) into (17) replaces the values of the censoring indicator by dichotomous realizations of the event-history process,  $y_{jk}$ , and the sample likelihood becomes

$$L = \prod_{i=1}^n \prod_{j=1}^{I_i} \prod_{k=1}^{K_{ij}} (h_{ij}(k))^{y_{jk}} (1 - h_{ij}(k))^{1 - y_{jk}} \tag{19}$$

which is only a function of the data,  $y_{jk}$ , and the population hazard probabilities,  $h_{ij}(k)$ .

Recognizing the Sample Likelihood

The expectation of the dichotomous event indicator  $Y_{jk}$  equals the probability that terminating event  $j$  occurs to individual  $i$  in time period  $k$  of spell  $j$ :

$$E(Y_{jk}) = 1 \cdot \Pr\{Y_{jk} = 1\} + 0 \cdot \Pr\{Y_{jk} = 0\} = \Pr\{Y_{jk} = 1\} \tag{20}$$

But  $Y_{jk}$  can equal 1 in period  $k$  only if individual  $i$  did not experience the requisite event in all earlier periods of spell  $j$ . So, if  $Y_{jk}$  equals 1, then  $Y_{ji}$  through  $Y_{j,k-1}$  must equal 0, and

$$\begin{aligned} E\{Y_{jk}\} &= \Pr\{Y_{jk} = 1\} = \Pr\{Y_{jk} = 1 | Y_{j,k-1} = 0, \dots, Y_{ji} = 0\} \\ &= \Pr\{T_j = k | T_j \geq k\} \\ &= h_{ij}(k) \end{aligned} \tag{21}$$

Thus, the values of the event indicator are observed realizations of hazard probability—for individual  $i$  in spell  $j$  and time period  $k$ ,  $Y_{jk}$  equals 1 with probability  $h_{ij}(k)$ .

Recognize that the multiple-spell likelihood function in (19) is identical to the likelihood function for  $N$  independent Bernoulli trials—each a realization of a dichotomous random variable  $Y_{jk}$  with “success” probability parameter  $h_{ij}(k)$ —where  $N$  equals the total number of discrete time periods observed across all spells for all individuals. This multiple-spell extension of the single-spell likelihood (see Allison, 1982; Brown, 1975; Laird & Oliver, 1981; Singer & Willett, 1993) is analogous to the multiple-spell generalization also found in continuous time (Hamerle; 1989; Petersen, 1991).

A Person-Spell-Period data set contains: (a) *spell and time period identifiers*, (b) *predictor variables*, and (c) *the event indicator variable*. The spell and time period identifiers distinguish among the multiple lines of data for each person. In Table 1, the columns SP NUM and PD NUM specify the spell number  $j$  and the time period  $k$  of each record. For person  $i$ , SP NUM takes on values ranging from 1 to  $J_i$  and PD NUM from 1 to  $K_{ij}$ . Over time within a spell for each person, PD NUM increments until the last observed period of that spell is reached and then, when the next spell begins (SP NUM = 2), the time period counter restarts (PD NUM is reset to 1). Teacher 1, for example, was observed for four spells. Her first three records refer to Time Periods 1, 2, and 3 of her first spell and hence SP NUM remains at 1 for all three of these records, while PD NUM takes on the values 1, 2, and 3. Her next record refers to the first period of her second spell and hence SP NUM takes on the value 2 and PD NUM is reset to 1. Remaining values of SP NUM and PD NUM are entered similarly.

The predictor variables distinguish among individuals in substantively interesting ways. In the record for the  $k$ th period of the  $j$ th spell, we enter attributes of the  $i$ th individual appropriate for that spell and period. Columns 5 and 6 of Table 1 display two predictors: FEMALE and ASSIGN.<sup>5</sup> FEMALE records the sex of the special educator (0 = male, 1 = female). It is *time-invariant*, with identical values in all time periods for all spells. The predictor ASSIGN, in contrast, is *time-varying*—it records the special educator's job assignment in each discrete period. When ASSIGN = 0, the special educator is a classroom teacher; when ASSIGN = 1, he or she provides support-related services outside the classroom. For instance, Teacher 1 spent the first year of her first spell as a member of the support staff, and the second and third years as a classroom teacher. When the special educator is out of teaching, the undefinable values of ASSIGN are missing. This absence of value for time-varying predictors in some spells is a common, but surmountable, nuisance in multiple-spell survival analyses (see later).

The event-indicator values in the last column of Table 1 indicate whether individual  $i$  experienced an event of interest in time period  $k$  of spell  $j$ . Teacher 1 experienced all four events. She left teaching for the first time in the third year of her first spell, returned in the first year of her second spell, left teaching again in the fifth year of the third spell, and returned in the second year of her fourth spell. Teacher 2, in contrast, experienced only two events. He entered the third spell during the tenth year of data collection by returning to teaching after a one year break. When data collection ended three years later, he had not yet experienced the third event—he had not yet left teaching again. Hence  $Y$  is zero for every year of his third spell, including the last, in which his history was censored.

#### Specifying a Statistical Model for Hazard

We use a generalized linear model (Nelder & Wedderburn, 1972) to represent the event indicator as a function of characteristics of individuals, spells, and time periods:

$$E(Y_{i\#k} = 1) = h_{ij}(k) = f(g(\text{COVARIATES}, \text{SPELL}, \text{PERIOD})) \quad (22)$$

where  $f(\ )$  indicates that the function linking the two sides of (22) has not been specified, and  $g(\ )$  indicates that the model is expressed in terms of *classes of variables* (denoted by COVARIATES, SPELL, and PERIOD), not as a linear combination of predictors and slope parameters.

In fact, the likelihood function in (19) is agnostic with respect to the functional relationship in (22) and so any link function,  $f(\ )$ , suitable for modeling a binary outcome as a function of predictors is acceptable. The complementary log-log and the logistic are two popular link functions (Collett, 1991; Cox & Snell, 1989). The former function yields a discrete-time hazard model that is implied by the proportional-hazards model for continuous-time data but which is asymmetric around hazard probabilities of .5. In our work, we follow the advice of Cox and Snell and use the logistic link function:

$$E(Y_{i\#k} = 1) = h_{ij}(k) = \frac{1}{1 + e^{-g(\text{COVARIATES}, \text{SPELL}, \text{PERIOD})}} \quad (23)$$

Given a specific parameterization of COVARIATES, SPELL, and PERIOD, we can estimate all slope parameters implicit in (23) with standard logistic regression software. Working in the Person-Spell-Period data set, event indicator  $Y$  is treated as a dichotomous dependent variable and is regressed upon predictors. The equivalence of the multiple-spell discrete-time and independent-Bernoulli-trials likelihoods ensures that appropriate estimates result.

### Doing the Data Analysis

#### Inspecting Sample Hazard Functions

Before logistic regression analyses begin, we suggest that estimates of the spell-by-spell hazard functions be inspected to identify trends, suggest parameterizations of spell and time period, and identify predictors for subsequent inclusion in statistical models. In Figure 1, we display sample hazard functions estimated separately by spell for our special educators. Notice the shape and level of the risk profiles. Overall, they share a common shape. In each spell, risk is highest initially; each spell is most likely to end immediately after it begins. Teachers tend to leave shortly after taking a job; ex-teachers tend to return shortly after leaving. But within each spell, risk declines over time. Experienced teachers are less likely to leave; ex-teachers who have been out of the classroom for some time are less likely to return.<sup>6</sup>

There are also dissimilarities in profile across spell. In the in-teaching spells, the risk of leaving teaching is higher earlier, but risk differentials across time within each spell are relatively small. In out-of-teaching spells, temporal risk differentials are more pronounced—initial risks are higher, later risks lower. Comparing the two sets of profiles suggests that, in the early

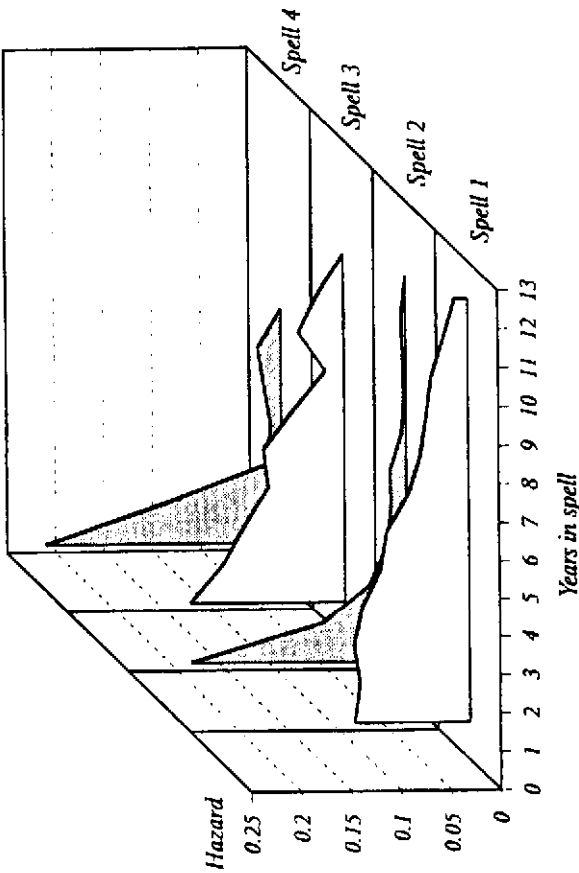


FIGURE 1. Sample hazard functions for the special educators, by spell. Spell 1: first spell in teaching; Spell 2: first spell out of teaching; Spell 3: second spell in teaching; Spell 4: second spell out of teaching.

years, an out-of-teaching spell is more likely to end than an in-teaching spell, after a few years, the reverse is true. Finally, within spell type, there is also a small difference, depending on whether the spell is *initial* or *repeat*. Repeat spells are more extreme than their initial partners: early risks for repeat spells are higher and later risks for repeat spells are lower, regardless of whether the spell is in- or out-of-teaching.

There is less data on event occurrence in later periods and spells, for structural and sampling reasons. The structural shortfall arises because many research designs limit observation in later time periods of later spells. This may occur when data collection ends at a fixed calendar date. Our study of special educators, for example, ended in 1985 regardless of a teacher's year of entry, and so we have up to 12 years of data to describe first spells, a maximum of 11 years of data to describe second spells, and so on. The analytic impact of the structural shortfall is similar to that of structural zeros in contingency table analysis (Agresti, 1990; Bishop, Fienberg, & Holland, 1975). The sampling shortfall arises because the size of the risk set shrinks as events and censoring occur—fewer people experience later periods of a given spell; fewer people survive to see later spells. Our study of special educators yields data on 3,941 first spells, 2,207 second spells, 736 third spells, and 266 fourth spells. When no one experiences an event in a particular period, there is a sampling shortfall because nonoccurrence has an impact similar to that of sampling zeros in contingency table analysis. These shortfalls

suggest that when we fit statistical models of hazard, we should be wary of empty cells, and that when we examine sample hazard plots, we should not emphasize small differences that may be due to sampling variation.

*An Initial Model for Hazard: The Effects of SPELL and PERIOD*

We recommend that the first step in a multiple-spell discrete-time survival analysis be the fitting of models to investigate the relationship between hazard and the effects of SPELL and PERIOD. These initial models for the effects of time include no substantive predictors, but capture the shape of the hazard profile across time periods within a single spell and across spells. They provide a standard against which later models can be compared. Currently, in the Person-Spell-Period data set, SPELL and PERIOD are represented by the counting variables SP NUM and PD NUM. However, to represent the potentially complex shapes of the hazard profiles, initial models for the effects of time need a general specification of SPELL and PERIOD. We suggest reparameterizing SP NUM and PD NUM as sets of dummy variables (as do others, in the single-spell case: Allison, 1982; Brown, 1975; D'Agostino et al., 1990). Across everyone in the sample, the maximum values of SP NUM and PD NUM are *J* and *K*, respectively, and so SPELL can be denoted quite generally by *J* spell indicators (*S*<sub>1</sub>, *S*<sub>2</sub>, . . . , *S*<sub>*J*</sub>) and PERIOD by *K* period indicators (*P*<sub>1</sub>, *P*<sub>2</sub>, . . . , *P*<sub>*K*</sub>). In the *j*th spell, all spell indicators are coded zero except for the *j*th, which is set to 1; in the *k*th period, all time period indicators are set to zero except for the *k*th, which is set to 1. In our data example, there are 4 spell indicators and 12 period indicators. For Teacher 1, *S*<sub>1</sub> equals 1 in the first three lines of data, *S*<sub>2</sub> equals 1 in the fourth line, and so on. The period indicators are defined similarly; in the three records for the first spell of Teacher 1, for instance, *P*<sub>1</sub> equals 1 in the first year of the spell, *P*<sub>2</sub> equals 1 in the second year, and *P*<sub>3</sub> equals 1 in the third and last year.

Given this new parameterization, we recommend that a taxonomy of three initial discrete-time hazard models be fitted. Model A includes only the main effect of PERIOD,

$$\text{logit}(h_{ij}(k)) = [\alpha_1 P_1 + \alpha_2 P_2 + \dots + \alpha_{12} P_{12}], \quad (24)$$

constraining the population logit-hazard profiles to be identical across spells. Estimates of the  $\alpha$  parameters describe the magnitudes of the constrained period-by-period hazard probabilities (notice that all time period dummies are retained in the model, but the intercept term is omitted to avoid complete linear dependency). Model B adds the main effect of SPELL,

$$\text{logit}(h_{ij}(k)) = [\alpha_1 P_1 + \alpha_2 P_2 + \dots + \alpha_{12} P_{12}] + \{[\beta_2 S_2 + \beta_3 S_3 + \beta_4 S_4], \quad (25)$$

including all the spell dummies but one and thereby allowing the logit-hazard profiles to differ in *level* across spells, but not in *shape*. Estimates of the  $\beta$  parameters describe the vertical displacement in hazard profile between the first spell and each subsequent spell, respectively. Model C adds 27 cross product terms to represent the two-way interaction of SPELL and PERIOD:

$$\begin{aligned} \text{logit}(h_{ij}(k)) = & [\alpha_1 P_1 + \alpha_2 P_2 + \dots + \alpha_{12} P_{12}] \\ & + [\beta_2 S_2 + \beta_3 S_3 + \beta_4 S_4] \\ & + \sum_{m=1}^{10} \gamma_m (S_2 \times P_m) + \sum_{m=1}^9 \gamma_{(m+10)} (S_3 \times P_m) \\ & + \sum_{m=1}^8 \gamma_{(m+19)} (S_4 \times P_m) \end{aligned} \quad (26)$$

thereby permitting the logit-hazard profiles to differ not only in level, but also in shape, from spell to spell. Later spell-by-period cross product terms have been omitted in order to avoid linear redundancy and to account for the presence of structural zeros in the design.

Figure 2 summarizes the fitting of these (and other) discrete-time hazard models in our example. Each box in the chart summarizes a fitted model listing the included predictors, the deviance ( $-2\log$ -likelihood) statistic, and the number of parameters (predictors) in the model. Arrows between nested models are labeled with the difference in the deviance statistic ( $\Delta\chi^2$ ), a statistic for testing the null hypothesis that constraints imposed in the reduced model are correct.<sup>7</sup> Comparing Models A and B, we reject the null hypothesis of no difference in risk across spells ( $\Delta\chi^2 = 76.82$  on 3 *df*,  $p < .0001$ ). Comparing B and C, we reject the null hypothesis that the effect of PERIOD on logit-hazard is the same across spells ( $\Delta\chi^2 = 568.29$  on 27 *df*,  $p < .0001$ ) and the ensuing interaction between SPELL and PERIOD indicates that the temporal shape of the logit-hazard profile differs from spell to spell.

Fitting the fully interactive SPELL by PERIOD model in (26) illustrates an interesting and important property of the sample likelihood function in (19). Taking logarithms, we can partition the multiple-spell log-likelihood in (19) into independent additive components, one per spell, each identical to the corresponding single-spell log-likelihood that would be obtained if the multiple-spell process were broken up into its constituent single-spell processes. This additive property of the multiple-spell log-likelihood provides an interesting choice: The within-spell hazard functions can be estimated either *simultaneously in one multiple-spell analysis* that applies Model C, or *sequentially in several separate spell-by-spell analyses* using single-spell methods applied repeatedly across spells. The two approaches lead to identical parameter estimates and standard errors, and the deviance statistic for the multiple-

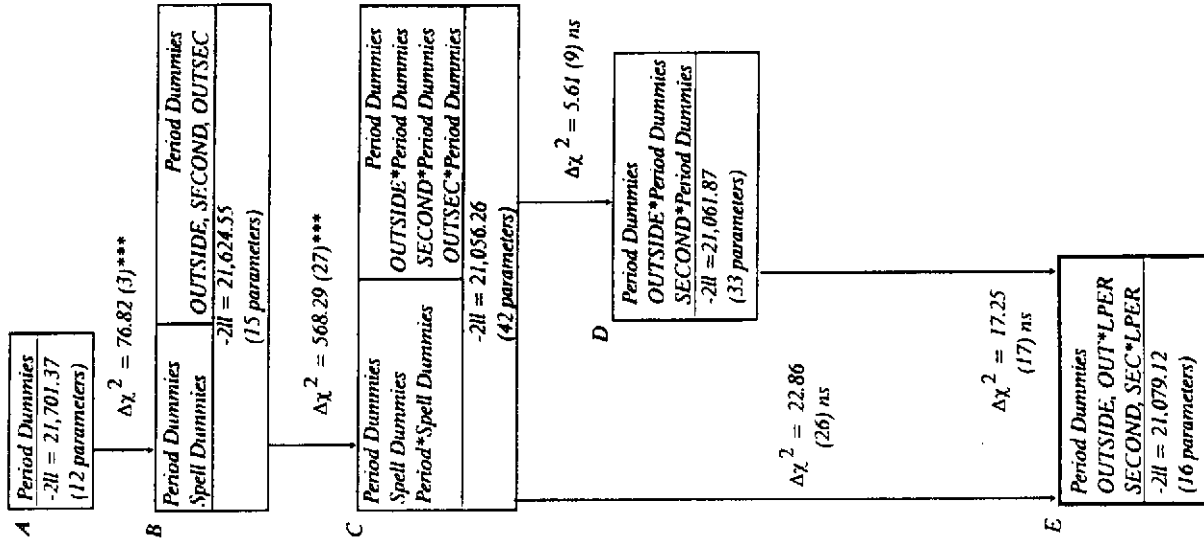


FIGURE 2. A taxonomy of multiple-spell discrete-time hazard models fitted to the special educator data.

spell analysis is simply the sum of the deviance statistics obtained in the separate single-spell analyses. For the special educators, for instance, independent single-spell analyses yield within-spell deviance statistics of 14,584, 4,373, 1,643, and 456, respectively, totaling to 21,056, which is the value of the overall multiple-spell deviance statistic reported for Model C in Figure 2. All other parameter estimates and standard errors are identical.<sup>8</sup> Advantages of the multiple-spell approach over the sequential single-spell approach include (a) the practical efficiency of the simultaneous analysis, (b) the ability to test for differences in the impact of substantive predictors across spells (by including predictor-spell interactions), and (c) the ability to test for the lagged impact of a predictor from one spell to the next.

Nevertheless, a more parsimonious and substantively appealing model can usually be found for the initial effect of time. The simpler model is preferred because, when a fully interactive model is parameterized completely in terms of spell and period dummies (i.e., Model C), sampling zeros may be encountered in later spells and periods. In our example, despite a large sample of 3,941 teachers, the fitting of Model C led to seven sampling zeros, four in the fourth spell alone. This can often be avoided by choosing smooth functions to represent PERIOD and SPELL, either as main effects or interactions (Hosmer & Lemeshow, 1989). First, consider SPELL. Our earlier inspection of sample hazard plots suggests that two dichotomous predictors—OUTSIDE (distinguishing in- and out-of-teaching spells), and SECOND (distinguishing initial and repeat spells)—are, with their cross product OUTSEC, an appealing replacement for the original spell dummies. Models B and C have identical deviance statistics under the new parameterization (see Figure 2). Model D, in which the two-way interaction OUTSEC and the three-way interaction between OUTSEC and the period dummies have been eliminated, is preferred over Model C ( $\Delta\chi^2 = 5.61, 9 df, p > .50$ ). Thus, the shape of the logit-hazard profile differs by whether the spell is initial or repeat and, independently, by whether it is in- or out-of-teaching. And, fortuitously, the number of sampling zeros is reduced from seven to three between Models C and D.

Second, we examine reparameterizations of PERIOD. Given the jagged profiles evident in most sample hazard plots, we usually choose to retain the general dummy variable parameterization for the main effect of PERIOD, but to test whether interactions between PERIOD and other predictors can utilize smooth functions of PD NUM. Often, inspection of estimates of the interaction parameters under the dummy parameterization suggest a particular linear, quadratic, or logarithmic transform of PERIOD to be used in interaction with other predictors. In our example, Model E includes the natural logarithm of PD NUM (called LPER in Figure 2) in interaction with OUTSIDE and SECOND. Model E fits almost as well as Models C and D, but with greater parsimony (E vs. D:  $\Delta\chi^2 = 17.25, df = 17, p > .40$ ; E vs. C:  $\Delta\chi^2 = 22.86, df = 26, p > .50$ ). Thus, Model E is a parsimonious initial model for the effects of time:

$$\begin{aligned} \text{logit}(h_{ij}(k)) = & [\alpha_1 P_1 + \alpha_2 P_2 + \dots + \alpha_{12} P_{12}] \\ & + [\beta_1 \text{OUTSIDE} + \beta_2 (\text{OUT} \cdot \text{LPER})] \\ & + \beta_3 \text{SECOND} + \beta_4 (\text{SEC} \cdot \text{LPER})] \end{aligned} \quad (27)$$

To confirm that the model accurately portrays the dependence of risk across periods and spells, and to communicate with a broad audience, we plot fitted hazard functions (Figure 3) for comparison with sample hazard profiles (Figure 1). Notice that, while the jagged character has been smoothed, the essential features are retained.

Parameter estimates for Model E, in Table 2, quantify these effects. The coefficients on  $P_1, \dots, P_{12}$  describe the logit-hazard profile for the first spell in teaching, when OUTSIDE and SECOND are zero. Their magnitudes suggest that risks of leaving are relatively stable during the first few years but then decline steadily over time. Notice that the corresponding standard errors also increase over time, reflecting the decreasing precision that derives from the diminishing risk set. By substituting estimated coefficients into (27) and solving for  $h_{ij}(k)$ , we recaptured the fitted first-spell hazard profile in Figure 3.

But how does this profile of risk differ across spells? Consider the coefficients for OUTSIDE and OUT\**LPER*, which compare in-teaching and out-of-teaching spells. Because the logarithm of unity is zero, the coefficient on OUTSIDE (0.74) estimates the difference between the fitted in- and out-of-teaching logit-hazard probabilities in Period 1. Thus, antilogging, the estimated odds of returning to teaching after one year out are 2.1 times the odds

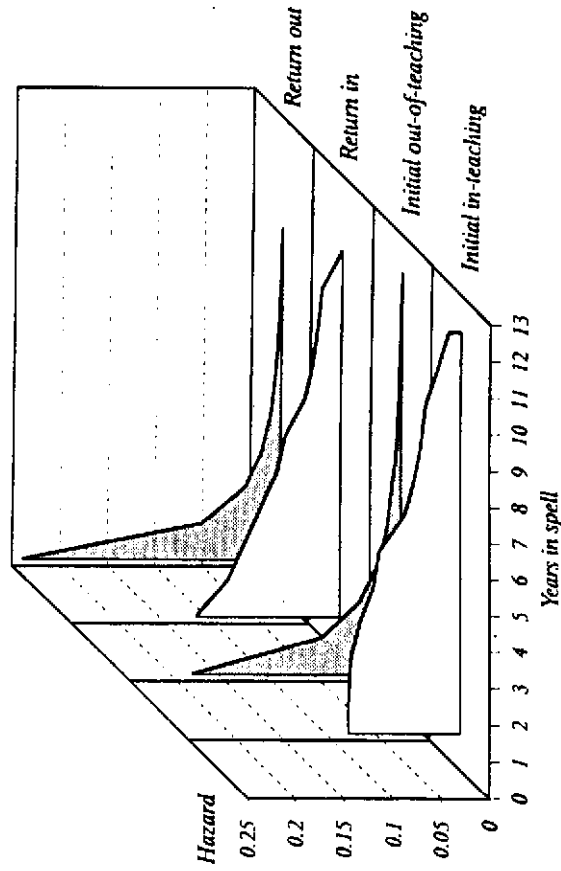


FIGURE 3. Fitted hazard functions for the special educators, by spell, based on Model E in Table 2.

TABLE 2  
Parameter estimates, standard errors, and goodness-of-fit statistics for three discrete-time multiple-spell hazard models fitted to the special educator data. Model E: an initial model for spell and time period; Model F: a model for the time-invariant predictor FEMALE; Model G: a model for the time-varying predictor SUPPORT.

Predictor	Parameter estimates (SE)		
	Model E	Model F	Model G
P <sub>1</sub>	-2.04 (.05)	-2.25 (.09)	-2.05 (.05)
P <sub>2</sub>	-2.04 (.05)	-2.36 (.07)	-2.06 (.05)
P <sub>3</sub>	-2.06 (.05)	-2.44 (.07)	-2.08 (.05)
P <sub>4</sub>	-2.15 (.06)	-2.57 (.08)	-2.16 (.06)
P <sub>5</sub>	-2.31 (.07)	-2.76 (.10)	-2.32 (.07)
P <sub>6</sub>	-2.38 (.07)	-2.85 (.11)	-2.40 (.07)
P <sub>7</sub>	-2.77 (.09)	-3.27 (.13)	-2.79 (.09)
P <sub>8</sub>	-2.98 (.11)	-3.50 (.15)	-3.00 (.11)
P <sub>9</sub>	-3.14 (.14)	-3.67 (.17)	-3.16 (.14)
P <sub>10</sub>	-3.25 (.17)	-3.79 (.20)	-3.27 (.17)
P <sub>11</sub>	-3.68 (.25)	-4.23 (.28)	-3.70 (.25)
P <sub>12</sub>	-4.35 (.45)	-4.89 (.46)	-4.36 (.45)
OUTSIDE	0.74 (.06)	0.74 (.06)	0.76 (.06)
OUT*LPER	-3.79 (.19)	-3.81 (.19)	-3.79 (.19)
SECOND	0.30 (.08)	0.28 (.08)	0.30 (.08)
SEC*LPER	-0.92 (.22)	-0.92 (.23)	-0.92 (.22)
FEMALE		0.25 (.09)	
FEM*LPER		0.40 (.18)	
SUPPORT			0.16 (.07)
-2LL	21079.12	21018.75	21073.47
Change in -2LL (df)		60.37 (2)	5.65 (1)
p		<.0001	<.05

of leaving teaching after one year in. Because the estimated coefficient on the interaction is large and of opposite sign to the main effect, this differential reverses in subsequent time periods—current teachers are more likely to leave than former teachers are to return. For instance, after 7 years in a spell, current teachers are nearly 12 times more likely to leave than are former teachers likely to return.

Now focus on initial and repeat spells—whether in- or out-of-teaching—described by the coefficients on the predictors SECOND and SEC\*LPER. The antilog of 0.30, 1.35, indicates that the odds that a repeat spell of either type will end in the first time period are 35% higher than the odds that an initial spell will end in the same period. The negative sign on the interaction coefficient suggests that, over time, this differential also reverses. By Period

3, repeat and initial spells are approximately equally likely to end but, beyond this, the odds ratio drops below 1. Thus, in the early years of a teaching spell, returnees are at greater risk of leaving than their novice counterparts. Over time, however, the differentials reverse. A similar pattern holds among those out of teaching. Former teachers who left before are more likely to return in the first year or two after leaving than are teachers who left for the first time.

Time-Invariant Predictors

The relationship between hazard probability and time-invariant predictors can be investigated by adding them to the initial model and comparing model goodness-of-fit, before and after. In our example, the addition of the main effect of the time-invariant predictor, FEMALE, improved goodness-of-fit substantially, confirming that the hazard profile for women differed substantially from that for men ( $\Delta\chi^2 = 55.00, df = 1, p < .0001$ ).

But predictor effects may differ across periods and spells. For instance, in our example, we detected a statistically significant interaction between FEMALE and LPER. While comparing Models E and F in Table 2 reveals the total improvement in fit resulting from the simultaneous inclusion of both the main effect of FEMALE and its interaction with log-period, when the interaction term was subsequently omitted from Model F on its own, we showed that the sex differential did indeed differ across time periods within spell ( $\Delta\chi^2 = 4.84, df = 1, p = .0281$ ). The estimated coefficient for FEMALE, .25, captures the difference in risk between women and men in their first year of a spell. After antilogging, the odds that a woman who is teaching will leave, or that a woman who is not teaching will return, are 1.28 times the odds for a man of similar employment status. And, the positive sign on the interaction coefficient indicates that the sex differential escalates over time within spell. By Year 7, the fitted odds ratio has increased to 1.79, illustrating that the odds that a woman will leave or return are 80% higher than they are for men, and illustrating that women, who make up three quarters of all special educators, are a very mobile group.

Time-Varying Predictors

Identical methods can evaluate the impact of time-varying predictors on hazard. However, some time-varying predictors are defined only in certain spells. Here, for instance, ASSIGN is missing in out-of-teaching spells (see Table 1). Reconceptualizing the offending predictor usually resolves this problem. In our example, ASSIGN can be recoded as a trichotomy with values of *support*, *in classroom*, and *not-in-teaching* and, as with any trichotomy, can be represented by two dummy variables.

We have already been using a predictor that represents one of these categories—the variable OUTSIDE, with value 1 for out-of-teaching spells and 0 otherwise. Recoding ASSIGN so that its former missing values are zero, we

can test the impact of job assignment on hazard by adding the recoded variable, now called SUPPORT, to a model that already includes OUTSIDE (see Model G, Table 2).<sup>9</sup> Then, antilogging the coefficient on SUPPORT (0.16) shows that, in any period, the odds that a support-service provider will leave are 1.17 times those for a classroom teacher—a differential that applies only during in-teaching spells.

Nevertheless, because SUPPORT is a time-varying predictor, its effects must be interpreted with care. The additional risk associated with providing support services can only "kick in" when the teacher is working in these settings. For this reason, the presence of time-varying predictors complicates interpretation of fitted hazard functions used to display findings because there are potentially as many hazard profiles as there are temporal combinations of the values of the time-varying predictor. We recommend that a small number of fitted profiles be displayed at substantively interesting temporal combinations of time-varying predictor values. One approach is to present extreme functions—one for those who were classroom teachers for their entire school careers and one for those who provided support services without break. Such extremes provide an "envelope" encompassing the fitted profiles of all special educators, and those whose job assignments fluctuated over time will possess a hazard profile that falls intermittently at one or other of the extremes.

### Conclusion

In this article, we review multiple-spell discrete-time survival analysis, an extension of the single-spell analyses that have been proposed by ourselves (Singer & Willett, 1993) and others (Allison, 1982; Efron, 1988; Laird & Oliver, 1981). We define the discrete hazard probabilities that are the fundamental parameters of the multiple-spell discrete-time survival process, and we show how maximum-likelihood estimates of these parameters can be obtained by application of standard methods for the analysis of dichotomous outcomes in a Person-Spell-Period data set. We illustrate our approach with a data-based example and comment on data-analytic strategies that are specific to the multiple-spell case.

In our presentation, we justify the creation of the Person-Spell-Period data set and the ensuing estimation of hazard probabilities by the exact equivalence of sample likelihood functions for the multiple-spell discrete-time survival process and for the independent Bernoulli trials model. This required equivalence of likelihoods holds only if the multiple-spell discrete-time hazard model is valid. In practice, as Allison (1982) points out, the discrete-time hazard model may be flawed because it assumes that heterogeneity among individuals is completely explained by their values on the predictors included in the model. If such unobserved heterogeneity among individuals exists and is temporally stable, then the occurrence of an individual's subsequent events will not be independent of prior events (even when conditioned on the measured predictors).

This potential lack of independence should not lead the empirical investigator to set these methods aside, however, because currently there are no readily accessible replacements for them that resolve the problem. The best approach, therefore, may be to continue in the explicit knowledge that the problem exists. Furthermore, the lack of independence is a problem that is not unique to multiple-spell discrete-time survival analysis. In fact, the same dependence also occurs across consecutive time periods within a single spell. Hence, both discrete-time and continuous-time survival methods designed for the analysis of single spells—including the popular Cox continuous-time proportional-hazards model—suffer from exactly the same drawback. Specifically, any analytic method—whether single- or multiple-spell—that ignores the presence of unobserved heterogeneity is likely to provide inefficient parameter estimates with negatively biased standard errors (Allison, 1982).

Allison (1982) suggests that the conditional independence assumption built into the multiple-spell discrete-time hazard model can be relaxed most easily by simple modifications to model specification. He states that it may be sufficient to introduce two predictors "representing the dependency of the hazard rate on the individual's previous history" (p. 93) into the discrete-time hazard model. He suggests as candidate predictors: (a) the number of events that have occurred prior to the current spell, and (b) the duration of the prior spell. In the models that we have described, the inclusion of spell dummies as predictors is tantamount to controlling for the number of previous events, and the duration of a prior spell can easily be included as an additional time-varying predictor.

Nevertheless, the search for explicit methods that resolve the problem of unobserved heterogeneity in both discrete-time and continuous-time hazard models is ongoing. Some authors have proposed including random individual-level "frailty" parameters in hazard models and have offered methods of estimation under a variety of distributional assumptions on the random effects, although the results of the estimation appear to be sensitive to the distributional assumptions made (see Aalen, 1988; Oakes, 1992; Vaupel, Manton, & Stallard, 1979). Therneau (1993), on the other hand, surveys methods of accounting for the within-person dependence of events by applying corrected variance estimates in the estimation of model parameters (see also Blackstone, 1989; Naftel, 1993). Finally, in a personal communication, Allison (1993) has outlined random-effects and fixed-effects approaches to representing unobserved heterogeneity in discrete-time hazard models, and has offered programming suggestions. But easy access for the empirical researcher still appears remote.

Despite their shortcomings, we continue to believe that the cautious application of multiple-spell discrete-time survival methods in education and the social sciences has much utility because these methods can be used to address a broad spectrum of important research questions that continue to plague researchers. We have illustrated our article with an example of employee

turnover—the repeated passage of educators into and out of work. However, the methods easily apply in many other domains. They can be used, for instance, to examine students' educational careers from entry into high school through ultimate graduation from a doctoral program. They can tell us when students will graduate from high school, enter college, complete their degrees, when they will dropout, return, or complete a GED, and how these risks depend on student characteristics, perhaps lagged or varying over time and spell.

Multiple-spell discrete-time survival methods are also suitable for analyzing the progress of children through a predefined sequence of developmental milestones. We might ask, for example: At what ages are children most likely to transit between adjacent stages of moral, interpersonal, or cognitive development? Do these risks depend on time-varying aspects of a child's life—those critical attributes of the home, school, and peer group that impact the temporal spacing of developmental transitions? And, although not discussed here, extensions of the basic multiple-spell discrete-time model lead to powerful alternative analytic avenues. If the temporal distribution of risk within each developmental spell was hypothesized to have a particular shape—perhaps low initially and rising exponentially to a peak after a certain number of years—this could be tested by a suitable reparameterization of PERIOD. And these hypotheses need not be identical from spell to spell, nor across children.

Whatever the domain, if the research question is concerned with the repeated occurrence of an event, or the occurrence of a known sequence of events, then—within well-defined limits—multiple-spell discrete-time survival analysis is a simple and straightforward method that can be applied. It permits complex longitudinal processes to be modeled, parsed, and interpreted. We recommend it for its utility, flexibility, and interpretability.

### Notes

<sup>1</sup>We recommend discrete-time methods only when discretization unavoidably stems from the data collection procedures themselves and the width of the time period is fixed by the fineness of the temporal information. We do not recommend that continuous-time data be arbitrarily discretized simply to facilitate the use of discrete-time methods. When continuous-time data are available, the methods of Allison (1984), Hamerle (1991), Kalbfleisch and Prentice (1980), Petersen (1991), and Yamaguchi (1991) can be used. Nevertheless, for continuous-time Poisson processes at least, Efron (1988) showed that “the information loss due to discretization goes to zero very quickly” (p. 423), suggesting that discrete- and continuous-time analyses converge rapidly as the width of the discrete intervals decreases.

<sup>2</sup>Other authors (Cox & Oakes, 1984; Kalbfleisch & Prentice, 1980) partition and index time periods differently, with the  $k$ th period of the  $j$ th spell,  $[t_{j,k}, t_{j,k+1})$ , beginning at  $t_{j,k}$  and extending to, but excluding,  $t_{j,k+1}$ . Cox and Oakes note that this simplifies some formulae but is an “unusual convention” (p. 13). The results are unaffected by choice of definition.

<sup>3</sup>Following Miller (1981, p. 2), we define the survivor function  $S(k)$  as  $\Pr\{T > k\}$  so that the cumulative distribution function (CDF),  $F(k) = \Pr\{T \leq k\}$ , equals  $1 - S(k)$ . This is consistent with our partition of time and ensures that the CDF is right-continuous (Mood, Graybill, & Boes, 1974, pp. 54–56), and the survivor function left-continuous. Other authors (e.g., Cox & Oakes, 1984) define the survivor function as  $\Pr\{T \geq k\}$ , which, in conjunction with their alternative partition of time, leads to left-continuity for the CDF. Whichever definition and time partition are used, the survivor function covers identical regions of the time axis (assuming  $\Pr\{t \leq T < t + \Delta t\} \rightarrow 0$  as  $\Delta t \rightarrow 0$ ) and our results hold.

<sup>4</sup>For researchers with data in a *one person, one record* format, we provide illustrative code for data conversion (SAS, Version 6.07) in the Appendix.

<sup>5</sup>The predictors can be continuous or dichotomous, as in regression analysis. We have included two dichotomous predictors here for ease of description and because earlier research suggested that these predictors were related to the risk of terminating a first spell in teaching (Murnane, Singer, Willett, Kemple, & Olsen, 1991; Singer, 1992, 1993).

<sup>6</sup>An anonymous reviewer suggested that these declining hazards may be a result of unobserved heterogeneity (see Vaupel & Yashin, 1985).

<sup>7</sup>All analyses were conducted with PROC LOGISTIC in SAS, Version 6.

<sup>8</sup>An anonymous reviewer suggested that separate single-spell analyses could be conducted to check the findings of the multiple-spell analysis. The additive property of the multiple-spell log-likelihood demonstrates that findings are automatically identical under either approach, providing the underlying assumptions of the multiple-spell model are valid (i.e., that an individual's subsequent events are independent of prior events once the hazard probability is conditioned on the measured covariates).

<sup>9</sup>We reached this main-effects specification for SUPPORT after finding no interactions between it and SECOND, the 12 period dummies, or LPER. We need not, indeed cannot, test the interaction between OUTSIDE and SUPPORT.

### APPENDIX

#### Constructing the Person-Spell-Period data set (using SAS Version 6)

The following code converts a dataset with one record per person into a Person-Spell-Period data set for fitting multiple-spell discrete-time hazard models. The data set has variables indicating:

NWAVE_1	the number of waves of data collected for the individual
WORK1-WORK13	a series of work indicator variables, one per wave
ASSIGN1-ASSIGN13	a series of job assignment variables, one per wave
FEMALE	teacher sex

If all cases have the same number of waves of data, NWAVE\_1 is constant. The array of work indicator variables is used to divide each person's event history into a series of spells. The array of assignment indicator variables is used to construct a time-varying predictor, SUPPORT. FEMALE is used to illustrate the inclusion of a time-invariant predictor.

## APPENDIX (continued)

The user of this code must obviously substitute variable names relevant for his or her data application. We have used macro variables `NSPELLS` and `NPERIODS` to indicate that the user can limit analysis to whatever number of spells or periods is desired.

In Section A, we go spell by spell through the waves to determine: (a) the wave number for the first period of the spell; (b) the wave number for the last period of the spell; (c) whether the spell is censored. The length of the spell is then computed on the basis of the wave numbers corresponding to each of these demarcations. Notice that we do not subscript these variables. This allows the variables representing these constructs to take on the same name regardless of spell in the Person-Spell-Period data set.

In Section B, we construct variables representing the effects of spell. We illustrate the ideas here using the two different systems presented in the article: (a) a series of spell dummies `S1-S4`; and (b) a pair of spell characteristics, `OUTSIDE` and `SECOND`. Other methods of expressing the effects of spell may be appropriate in other applications. Simply substitute the relevant classification scheme here. When choosing another scheme, however, be sure to retain the spell dummies, `S1-S4`, if you would like to test whether effects of predictors differ across spells.

In Section C, we create the Person-Spell-Period data set. We first create the event indicator `Y` on the basis of the length variable (which tells us whether the event occurred in that period) and the censoring indicator. We then loop through the periods to create the time-varying predictor support, appropriately keying the value of the variable to the correct wave number corresponding to the period of the spell. Finally, we create the period dummies, `P1-P12`, through a second loop within the loop. It is here that any interactions between period and predictors would be created. Because we parameterized the interaction between period and each of `OUTSIDE`, `SECOND`, and `FEMALE` in terms of a smooth function based on taking the log 10 of `PD NUM`, we use that form of interaction here. By similar logic, however, you could substitute interactions with the period dummies here as well.

In Section D, we use PROC LOGISTIC with the Person-Spell-Period data set to fit multiple-spell discrete-time hazard models.

```
%LET NSPELLS = 4;
%LET NPERIODS = 12;
DATA PSP;
  SET SPED.MSPED;
  ARRAY S1(&NSPELLS);
  ARRAY P1(&NPERIODS);
```

```
*SECTION A*;
DO SPELL = 1 TO &NSPELLS;
  IF SPELL = 1 THEN FIRST = 1; ELSE FIRST = LAST + 1;
  DO W=(FIRST+1) TO NWAVE_1;
    IF (WORK[W] NE WORK[W-1]) THEN DO;
      LAST = W-1;
      CENSOR = 0;
      GO TO OUT1;
    END;
  END;
```

## APPENDIX (continued)

```
LAST = NWAVE_1;
CENSOR = 1;
OUT1:
LENGTH = LAST - (FIRST-1);

*SECTION B*;
DO J = 1 TO &NSPELLS;
  IF SPELL = J THEN S[J] = 1; ELSE S[J] = 0;
END;
IF SPELL = 2 OR SPELL = 4 THEN OUTSIDE = 1; ELSE OUTSIDE = 0;
IF SPELL > 2 THEN SECOND = 1; ELSE SECOND = 0;
```

```
*SECTION C*;
DO PERIOD = 1 TO MIN(LENGTH, &NPERIODS-SPELL+1);
  IF PERIOD = LENGTH AND CENSOR = 0 THEN Y = 1; ELSE Y = 0;
  SUPPORT = ASSIGN((FIRST-1)+PERIOD);
  SECLPER = SECOND*LOG10(PERIOD);
  OUTLPER = OUTSIDE*LOG10(PERIOD);
  FEMLPER = FEMALE*LOG10(PERIOD);
  DO K = 1 TO &NPERIODS;
    IF PERIOD = K THEN P[K] = 1; ELSE P[K] = 0;
  END;
  OUTPUT;
END;
IF CENSOR = 1 THEN GO TO OUT2;
END;
OUT2:;
```

```
*SECTION D*;
PROC LOGISTIC DATA=PSP NOSIMPLE;
  TITLE2 "AN INITIAL MODEL FOR THE EFFECTS OF TIME";
  MODEL Y = P1-P12 OUTSIDE OUTLPER SECOND SECLPER/NOINT;
PROC LOGISTIC DATA=PSP NOSIMPLE;
  TITLE2 "INTERACTION BETWEEN FEMALE AND LOG-PERIOD";
  MODEL Y = P1-P12 OUTSIDE OUTLPER SECOND SECLPER FEMALE
    FEMLPER/NOINT;
PROC LOGISTIC DATA=PSP NOSIMPLE;
  TITLE2 "MAIN EFFECT OF SUPPORT";
  MODEL Y = P1-P12 OUTSIDE OUTLPER SECOND SECLPER SUPPORT/NOINT;
```

## References

- Aalen, O. O. (1988). Heterogeneity in survival analysis. *Statistics in Medicine*, 7, 1121-1137.
- Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.
- Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 61-98). San Francisco: Jossey-Bass.
- Allison, P. D. (1984). *Event history analysis: Regression for longitudinal event data* (Sage University paper series on quantitative applications in the social sciences, Series No. 07-046). Beverly Hills, CA: Sage.

- Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis*. Cambridge, MA: MIT Press.
- Blackstone, E. H. (1989). Research methods for decision-making in cardiac surgery: Inferences from repeating morbid events. *Heart Surgery*, 301-309.
- Blossfeld, H. P., Hamerle, A., & Mayer, K. U. (1989). *Event history analysis: Statistical theory and application in the social sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Brown, C. C. (1975). On the use of indicator variables for studying the time-dependence of parameters in a response-time model. *Biometrics*, 31, 863-872.
- Collett, D. (1991). *Modeling binary data*. London: Chapman and Hall.
- Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society*, 34, 187-202.
- Cox, D. R., & Oakes, D. (1984). *Analysis of survival data*. London: Chapman and Hall.
- Cox, D. R., & Snell, E. J. (1989). *Analysis of binary data* (2nd ed.). London: Chapman and Hall.
- D'Agostino, R. B., Lee, M. L., Belanger, A. J., Cupples, L. A., Anderson, K., & Cannel, W. B. (1990). Relation of pooled logistic regression to time-dependent Cox regression analysis: The Framingham heart study. *Statistics in Medicine*, 9, 1501-1515.
- Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American Statistical Association*, 83, 414-425.
- Hamerle, A. (1989). Multiple-spell regression models for duration data. *Applied Statistics*, 38, 127-138.
- Hamerle, A. (1991). On the treatment of interrupted spells and initial conditions in event history analysis. *Sociological Methods and Research*, 19, 388-414.
- Hosmer, D., & Lemeshow, S. (1989). *Applied logistic regression*. New York: Wiley.
- Kalbfleisch, J. D., & Prentice, R. L. (1980). *The statistical analysis of failure time data*. New York: Wiley.
- Karlin, S., & Taylor, H. M. (1975). *A first course in stochastic processes*. New York: Academic Press.
- Laird, N., & Oliver, D. (1981). Covariance analysis of censored survival data using log-linear analysis techniques. *Journal of the American Statistical Association*, 76, 231-240.
- Miller, R. G. (1981). *Survival analysis*. New York: Wiley.
- Mood, A. M., Graybill, F. A., & Boes, D. C. (1974). *An introduction to the theory of statistics* (3rd ed.). New York: McGraw Hill.
- Murnane, R. J., Singer, J. D., Willett, J. B., Kemple, J. J., & Olsen, R. J. (1991). *Who will teach?: Policies that matter*. Cambridge, MA: Harvard University Press.
- Naftel, D. C. (1993, August). *The parametric risk factor analysis of repeating events (rejection and infection) in cardiac transplantation patients*. Paper presented at the joint statistical meetings of the American Statistical Association, San Francisco, CA.
- Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society*, A135, 370-384.
- Oakes, D. (1992). Frailty models for multiple event times. In J. P. Klein & P. K. Goel (Eds.), *Survival analysis: State of the art* (pp. 371-379). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Petersen, T. (1991). The statistical analysis of event histories. *Sociological Methods and Research*, 19, 270-323.
- Singer, J. D. (1992). Are special educators' career paths special?: Results from a 13-year longitudinal study. *Exceptional Children*, 59, 262-279.
- Singer, J. D. (1993). Once is not enough: Former special educators who return to teaching. *Exceptional Children*, 60, 58-72.
- Singer, J. D., & Willett, J. B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics*, 18, 155-195.
- Therneau, T. M. (1993, August). *Using multiple-events Cox model*. Paper presented at the joint statistical meetings of the American Statistical Association, San Francisco, CA.
- Vaupel, J. W., Manion, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16, 439-454.
- Vaupel, J. W., & Yashin, A. I. (1985). Heterogeneity's ruses: Some surprising effects of selection on population dynamics. *The American Statistician*, 39, 176-185.
- Willett, J. B., & Singer, J. D. (1991). From whether to when: New methods for studying student dropout and teacher attrition. *Review of Educational Research*, 61, 407-450.
- Yamaguchi, K. (1991). *Event history analysis*. Newbury Park, CA: Sage.

#### Authors

JOHN B. WILLETT and JUDITH D. SINGER are Professors, Harvard University Graduate School of Education, Appian Way, Cambridge, MA 02138. They eat lunch together every Tuesday at the Ta Sheng Restaurant in Harvard Square and specialize in quantitative methods, with an emphasis on design and data analysis for longitudinal research.

Received December 2, 1992  
Accepted March 28, 1994