DOCUMENT RESUME

ED 373 633 HE 027 619

AUTHOR Ronco, Sharron L.
TITLE Meandering Ways: Studying Student Stopout with Survival Analysis. AIR 1994 Annual Forum Paper.
PUB DATE May 94
NOTE 23p.; Paper presented at the Annual Forum of the Association for Institutional Research (34th, New Orleans, LA, May 29-June 1 1994).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Persistence; College Students; Dropout Research; Dropouts; Higher Education; Hispanic Americans; Longitudinal Studies; Predictive Measurement; *Reentry Students; *Stopouts; Student Attrition
IDENTIFIERS *AIR Forum; Survival Ratio Models; University of Texas El Paso

ABSTRACT

This study applied newly-developed methods of survival analysis with repeated events to a longitudinal data set in order to illustrate stopout hazards for college students. Survival analysis incorporates cases of students returning to college and not yet returning to college in a single analysis. This study applied a multiple-spell survival analysis to stopout data on a cohort of first-time entering freshmen in fall 1983 at University of Texas at El Paso for 20 terms through Spring, 1993. Spells were defined as "enrolled" or "not enrolled." Data collection ended either when the student received a first baccalaureate degree or in spring, 1993. Results found that students' greatest risk of leaving was after the second term of enrollment. Hazard again peaked after the fourth term, then leveled off after the sixth term. Students who left were more likely to return after only one or two terms out. Once they had been gone for six terms, the odds of returning were virtually nil. Return enrollment and stopout spells share basically the same shape as initial spells. Hispanic students were 1.7 times more likely than White non-Hispanic students to return to the institution after a stopout period. (Contains 20 references.) (JB)

MEANDERING WAYS:

STUDYING STUDENT STOPOUT WITH SURVIVAL ANALYSIS

by

Sharron L. Ronco
Director, Institutional Studies

University of Texas at El Paso
500 University Ave.
El Paso, Texas 79968-0512
(915) 747-5207
ae05%utep@utepvm.ep.utexas.edu

Paper presented at the Annual Forum of the Association for Institutional Research, New Orleans, LA, May 1994

AIR

for Management Research, Policy Analysis, and Planning

This paper was presented at the Thirty-Fourth Annual Forum of the Association for Institutional Research held at The New Orleans Marriott, New Orleans, Louisiana, May 29, 1994 - June 1, 1994. This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of Forum Papers.

Jean Endo
Editor
Forum Publications

MEANDERING WAYS:

STUDYING STUDENT STOPOUT WITH SURVIVAL ANALYSIS

ABSTRACT

Up to one-third of college students voluntarily interrupt their postsecondary enrollment for one or more terms. Stopout affects not only students, but also educators, planners and researchers. Time to degree completion, required for many accountability reports, cannot be a useful piece of consumer information unless it is interpreted in the context of actual enrollment behavior. In this paper, newly-developed methods of survival analysis with repeated events are applied to a longitudinal data set in order to illustrate stopout hazard for several subgroups of students. These promising techniques have the ability to address questions such as: When are students most likely to stop out? How long before they return? How do subsequent episodes of enrollment or stopout differ from initial ones? How is the risk of stopping out related to student characteristics such as ethnicity, or number of credit hours attempted?

"Stopout", a term coined by the Carnegie Council in their 1980 report on Policy Studies in Higher Education, describes the voluntary interruption by students in their enrollment in postsecondary education for one or more terms. The Commission noted that the legitimization of stopout was one of the most radical changes made in higher education in the 1960's and 1970's, as colleges encouraged students to stop out before receiving their degrees to spend time working, travelling or engaging in some other constructive activity. In 1969, 17% of undergraduates in U.S. colleges and universities had stopped out; by 1976, the proportion had reached 26% (Carnegie Council, 1980). More recent estimates of student stopout range from 10% to 33% (Porter, 1989; Tichenor and Cosgrove, 1992).

Students today stop out for a variety of reasons. Nontraditional students, who have been enrolling in colleges in larger numbers, generally rely on their own resources to fund their postsecondary education. Those resources are increasingly unable to cover the cost of college tuition and fees, requiring them to fill in with episodes of full-time employment. For these students, family responsibilities can also mean a semester or two out of school. Cutbacks in institutional budgets restrict the number of courses that can be made available on a semester basis, delaying fulfillment of degree plans. Students whose grade point averages fail to meet academic standards may have to sit out for a semester, or repeat a course at the community college. But students recognize the value of a college degree, and when they do return to college, their major incentive is to acquire the training and preparation that will provide careers with higher levels of reward and satisfaction. (Smart and Pascarella, 1987.)

How does stopping out affect colleges? In a controversial paper on productivity in higher education, SUNY Chancellor D. Bruce Johnstone argued that the interruption and subsequent resumption of learning is enormously costly, both for the faculty and facilities required for a typical undergraduate degree, and for the student, who may be kept from a better and higher-paying job for longer than is necessary. Market and political forces, which are demanding more productivity from all forms of education, are fueling proposals to reconfigure education

so that learning and therefore graduation can take place faster and with greater efficiency. But there are strong arguments that the industrial efficiency model does not apply to the higher education of the future, and that lifelong learning will replace the model of a 'front-loaded' education where students in the pipeline flow under constant pressure without interruption toward a degree. Stopout may even turn out to be the more 'productive' enrollment behavior, since re-entry students, with time and experience on their side, tend to be more purposeful and motivated (see "Comments" in Johnstone, 1993).

Despite the fact that students are increasingly interacting with postsecondary education in a nontraditional fashion, most accountability reporting still focuses on the completion rates of full-time first-time freshmen who are assumed to be continuously enrolled until graduation or dropout. Although time to completion is certainly a useful piece of consumer information, it must be evaluated in a realistic context. This requires additional information on the rate of interrupted enrollment present at the institution (Ewell and Jones, 1991). For many institutions, information on enrollment patterns may be difficult to obtain. Those who are able to track cohort stopout are still faced with the overwhelming task of making sense of stopout patterns.

USING SURVIVAL ANALYSIS TO STUDY STOPOUT

Institutional researchers are often interested in studying events that happen to students: Their enrollment in a particular institution, their retention, transfer, stopout, dropout, graduation. Whether or not these events occur is a simple matter of tabulating data, which then surfaces in institutional fact books and reports to consumers, state legislators, federal bureaucrats, and others. More interesting and informative are questions associated with the duration of events: How long is the average student retained? When are students at greatest risk for dropout, stopout, transfer, graduation? What are the predictors of these events? Are greater risks associated with certain factors or particular subgroups? Do certain policies have an effect on whether these events will occur?

75

Research questions about time pose unique design and analytic difficulties (Singer and Willett, 1991). Data collection must end at some arbitrarily-defined period without some of the subjects having experienced the target event. Should the researcher assume that the dropout will not return at some later time, or that the persister will never graduate? If the individual does not experience the event during the data collection period, that individual's data is said to be censored; it is not known whether or when that individual will eventually experience the event. These censored durations have the potential to greatly inform our understanding of enrollment behavior, but also greatly complicate our statistical analysis of the data.

One analytic strategy for dealing with the problem posed by duration data has been survival analysis. Survival analysis is a technique that incorporates both censored and uncensored cases in a single analysis, and thus is an ideal method for studying the occurrence of events. Event history methods, which include survival analysis or hazard models, were originally developed by biostatisticians studying clinical lifetime data, and have been extended by engineers, economists and sociologists. Increasingly, educational statisticians are beginning to adapt these methods to the study of educational phenomena (DesJardins, 1993; Mensch and Kandel, 1988; Moore, 1992; Murnane, Singer and Willett, 1988; Rosenfeld and Jones, 1987; Willett and Singer, 1991a). Filling the void for straightforward, understandable and adaptable methods of survival analysis for educational researchers, Judith Singer and John Willett have pioneered a number of applications of survival methods, particularly discrete-time methods, to social science data.

Some events are irreversible: Once they occur, they cannot occur again. First-time enrollment in college and graduation are two such events. Other events are repeatable: They occur more than once for at least some subjects in the sample (Yamaguchi,1991). Repeated events include multiple spells, the duration intervals each of which correspond to a distinct occurrence of the event. Repeated spell data is complex and approaches to dealing with it have had shortcomings. The first entry into a particular state, such as marriage, can have quite different characteristics from re-entries. In their recent analysis of the careers of

16

special education teachers, Willett and Singer (1993) found that the risk of exiting and re-entering teaching differs by whether the action is an initial or repeat spell. Spells can be analyzed separately, but often the sample size for later spells becomes too small for meaningful analysis. Pooling all spells into a single data set and using the number of spells as a predictor of risk is another strategy but is flawed because the failure to link each individual's repeated spells inflates the degrees of freedom and underestimates the standard errors of parameter estimates (Singer and Willett, 1991).

These flaws led Willett and Singer to propose an extension of discrete-time survival analysis of event occurrence in a single spell to the case of multiple spells. The method allows data from all spells to be analyzed simultaneously. Predictors can be included which are both constant and time-varying. With multiple-spell survival analysis, the main effects of each predictor can be explained, along with all possible interactions between predictors, including interactions with time and spell. These methods can be used to study the repeated occurrence of a single event, or the sequential occurrence of disparate events (Willett and Singer, 1993).

APPLICATION OF MULTIPLE-SPELL SURVIVAL ANALYSIS TO STOPOUT DATA

Method

To illustrate the application of multiple-spell survival analysis to stopout data, the cohort of first-time entering freshmen in fall 1983 at this University was followed for 20 long terms, through spring, 1993. Summer terms were excluded from the analysis. Spells are defined as "enrolled" or "not enrolled". For example, spell 1 consists of the number of continuous long terms that the student was enrolled for the first time. For students who did not return in spring 1984, for example, spell 1 consisted of only one term. Spell 2 consists of the number of continuous long terms that the student was **not** in attendance at this institution. For those who dropped out after one semester and never returned, spell 2 = 19 terms. Spell 3 occurs when the student returns to school, spell 4 when the student drops out again, and so on. Table 1 shows the stopout profile of the fall 1983 entering cohort.

TABLE 1
Stopout profile for first-time freshmen in fall 1983

Spell Number	Description	Risk Set	Graduated	Censored Persist	Censored Drop?
1	First enrolled	1790	281	1	---
2	First out	1508	---	---	884
3	Second enrolled	624	123	30	---
4	Second out	471	---	---	265
5	Third in	206	36	23	---
6	Third out	147	---	---	94
7	Fourth in	53	5	15	---
8	Fourth out	33	---	---	26
9	Fifth in	7	1	4	---
10	Fifth out	2	---	---	2

For this cohort, data collection ended either when the student received a first baccalaureate degree, or if no graduation occurred, in spring 1993. Those students who did not transition from one spell to the next; i.e., remained enrolled or remained not enrolled, are censored in that spell. Censoring can occur in any spell, but once it occurs, the individual is no longer eligible to experience further spells (Willett and Singer, 1993). Individuals who are censored in "in-school" spells (1, 3, 5, 7 or 9) were enrolled in spring 1993. Individuals censored in "out-of-school" spells (2, 4, 6, 8 and 10) have not returned and may have dropped out permanently.

Analysis of stopout data begins with the creation of a person-spell-term data set. Table 2A shows how data entry begins. To conserve space, this illustration presumes that data collection lasted only eight terms instead of the 20 under study. Person A was enrolled for the first time for four terms; hence, spell=1 and term takes on the value of 1, 2, 3 and 4. Person A then stopped out for one semester (spell 2, term 1); returned for two terms (spell 3, terms 1, 2); then left and has not returned to date. Person B was enrolled for four terms, then left and has not returned to date. Since the individuals' outcomes are known for each spell except the last, only the last spell is censored. A new variable is created, Y, which becomes the dependent variable in the regression analysis. It is coded '1' for the term in which the individual experienced a transitioning event. Person A left school for the first

98

TABLE 2A
Contents of the Person-Spell-Term Data Set

Student ID	Spell	Term	ETHNIC	STATUS	Censored	Y
A	1	1	0	1	0	0
A	1	2	0	1	0	0
A	1	3	0	0	0	0
A	1	4	0	1	0	1
A	2	1	0	.	0	1
A	3	1	0	1	0	0
A	3	2	0	0	0	1
A	4	1	0	.	1	0
B	1	1	1	1	0	0
B	1	2	1	1	0	0
B	1	3	1	1	0	0
B	1	4	1	1	0	1
B	2	1	1	.	1	0
B	2	2	1	.	1	0
B	2	3	1	.	1	0
B	2	4	1	.	1	0

NOTE: Ethnic: 0 = White, non-Hispanic; 1 = Hispanic
Status: 0 = Part-time; 1 = Full-time

time after term 4, so Y is coded '1'. Y for the second term is coded '1' in spell 2 since Person A again experienced a transitioning event - returning to school. At the end of data collection (spell 4, term 1) no event of returning to school had yet occurred, so Y = '0'.

In order to investigate the effects of variables on the risk of stopout and return from stopout, one predictor was included whose value is a constant (ethnicity), along with a predictor whose values vary by semester - full or part-time attendance. Ethnicity included the two major groups at this university, Hispanic and white non-Hispanic. Since few students at this institution are permanently dismissed for academic reasons - they may be placed on academic probation or be suspended for a semester - it was believed that reinstatement after academic failure was still largely under the control of the student, and therefore stopouts for academic reasons were considered voluntary. These variables are included in the person-spell-term data set. Students cannot have status indicators for terms in which they were not enrolled,

99

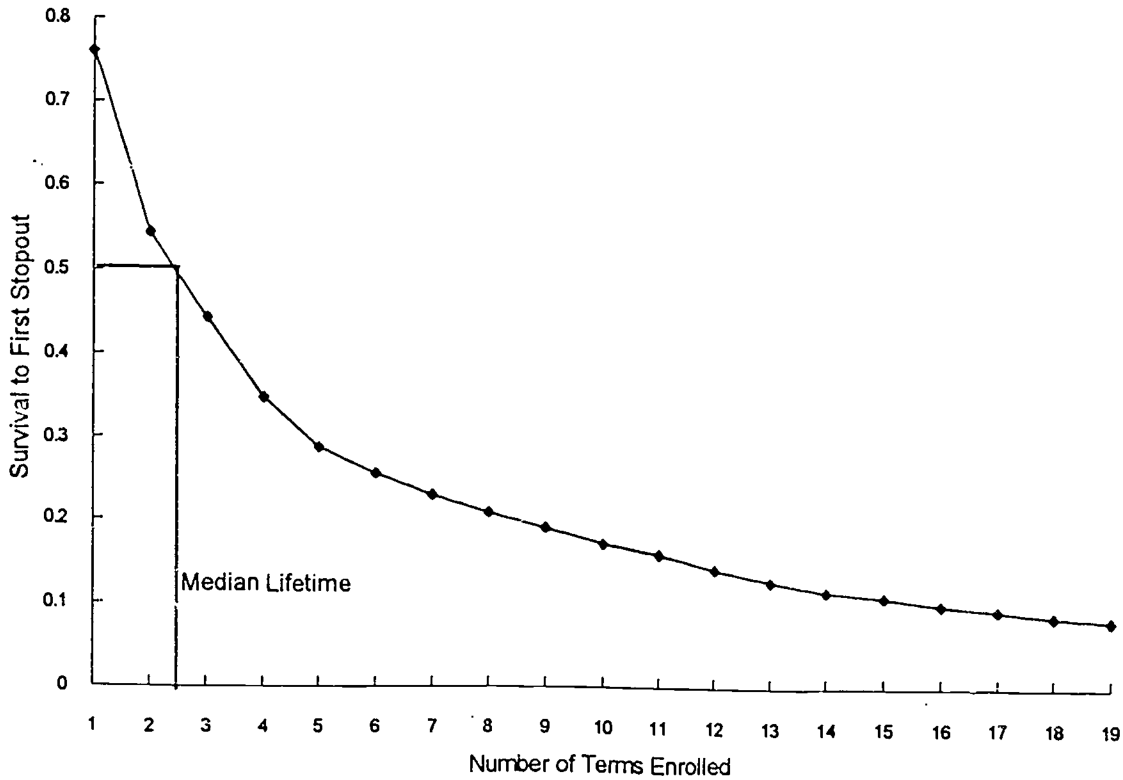

Figure 1. Survival Function For First Spell Of Enrollment

and so information for these terms is missing. This creates a minor nuisance for data analysis, which can be overcome with a procedure described below.

Survival analysis begins with the survivor function, which is observed as the proportion of an initial population which will survive through each of several successive time patterns. At a survivor function of .50 half of the sample has experienced the target event, half has not. A useful concept which is derived from the survivor function is the median lifetime, which indicates how much time passes before half of the sample experiences the target event. Figure 1 illustrates the survivor function and median lifetime for the first spell of enrollment. The average length of first spell enrollment for the fall 1983 cohort was between two and three terms. Since the survivor function maintains a consistent shape regardless of the distribution of risk, it is generally more informative to examine the hazard function. The hazard is the number of new events occurring during a time period, expressed as a proportion of the number of individuals at risk. When time is measured discretely (i.e., terms or years) rather than continuously, hazard is a probability. Although it is an

/02

unobserved variable, hazard controls both the occurrence and timing of events. As such, it is the fundamental dependent variable in an event history model and forms the cornerstone of survival analysis (Allison, 1984; Singer and Willett, 1993).

These functions can easily be extended to multiple spells. If i denotes an individual, j a particular spell, and k a particular time period, then the survivor function is defined as the probability that individual i's jth spell will be terminated after time period k of that spell. The hazard function is defined as the probability that individual i's jth spell will be terminated in the kth period, given that individual i did not experience the event in a prior time period of that spell.

In order to understand the effects of time on hazard, the event indicator Y is regressed on the predictor representing spell and term in the person-spell term data set. Since hazard may very well differ by spell and/or time period, it is recommended that a very general representation of spell and period be used in order not to force conformance of the data to an inappropriate shape. This can be achieved by defining dummy variable specifications for spell and term. Table 2B shows the result of transforming the person-spell-term data set. For Person A, S1 is coded 1 only when the record pertains to the first spell, and is 0 otherwise. Likewise, term 1 is set to 1 whenever the record pertains to the first term of a given spell, and is set to 0 otherwise. Although there were a total of ten spells produced by this data set, it was decided to limit investigation to the first four, since the risk set diminished greatly after the fourth spell.

Following the authors' recommendation, two additional dichotomous predictors are created: OUTSIDE and SECOND. OUTSIDE indicates whether the spell is an in-school or out-of-school spell. Spells 1 and 3 are in-school and so OUTSIDE is coded '0' for these spells, and '1' for spells 2 and 4. SECOND indicates whether the spell is an initial or return spell for that type. Spells 1 and 2 are initial spells for in and out of school. Spells 3 and 4 are return spells. Capturing the effect of spell with these two new variables is especially helpful when the spells being analyzed terminate in different kinds of events, in this case, leaving school vs. returning to school.

TABLE 2B
Transformed File using Dummy Predictors

ID	S1	S2	S3	S4	T1	T2	T3	T4	...	T19	ETHNIC	STATUS	OUTSIDE	SECOND	Y
A	1	0	0	0	1	0	0	0		0	0	1	0	0	0
A	1	0	0	0	0	1	0	0		0	0	1	0	0	0
A	1	0	0	0	0	0	1	0		0	0	0	0	0	0
A	1	0	0	0	0	0	0	1		0	0	1	0	0	1
A	0	1	0	0	1	0	0	0		0	0	.	1	0	1
A	0	0	1	0	1	0	0	0		0	0	1	0	1	0
A	0	0	1	0	0	1	0	0		0	0	0	0	1	1
A	0	0	0	1	1	0	0	0		0	0	.	1	1	0
B	1	0	0	0	1	0	0	0		0	1	1	0	0	0
B	1	0	0	0	0	1	0	0		0	1	1	0	0	0
B	1	0	0	0	0	0	1	0		0	1	1	0	0	0
B	1	0	0	0	0	0	0	1		0	1	1	0	0	1
B	0	1	0	0	1	0	0	0		0	1	.	1	0	0
B	0	1	0	0	0	1	0	0		0	1	.	1	0	0
B	0	1	0	0	0	0	1	0		0	1	.	1	0	0
B	0	1	0	0	0	0	0	1		0	1	.	1	0	0

The person-spell-term data set now includes dummy variables that identify spell and term, predictor variables, both constant and time-varying, and an outcome variable which describes how the spell ended. Because the outcome is dichtomous, the relationship between outcome and predictors can be investigated via regular statistical software; in this case, SAS PROC LOGISTIC.

Analysis and Results

Initial Model for the Effects of Time

The initial analytic step in survival analysis involves describing a baseline hazard, or the distribution of risk across time with no other predictors included to distinguish individuals. With multiple-spell survival analysis, we are interested in knowing not only whether hazard varies across time periods, but also across spells. Initially, a model is fitted which includes only the time periods T1 - T19 (Term 20 was dropped from the analysis since only one individual occupied that term, presenting convergence complications for the logistic regression procedure). The -2 log likelihood (-2ll) statistic produced by PROC LOGISTIC = 16,237 with 19 predictors. Adding the main effects of spell to the model resulted in -2ll

= 15,332.9 with 22 predictors. The differences between these two models are distributed as chi-square with df equal to the difference in the number of predictors between models. In this case, $\Delta \chi^2 = 904.9$ (3), $p < .001$, indicating that the effect of hazard is not the same across spells. Adding to the model the interaction between term and spell (-2ll = 15,118, 55 predictors) again resulted in a significantly better fit, indicating that the effect of term on hazard is not the same across spells ($\Delta \chi^2 = 214.5$ (33), $p < .001$). This difference is graphically displayed in Figure 2, which models hazard functions by spell. The magnitude in each term indicates the risk of terminating the spell in that term. If the four spells under study were not significantly different from each other, they would all share the same profile. If there were not a significant interaction with term effect, spells would differ in elevation but not in shape. The out-of-school spells (2 and 4) share a similar profile in that for both, the 'risk' of returning to school drops precipitously by term 3. It is evident from inspection of the four spells, however, that profiles differ by both spell and term.

Even though the model using term, spell and the ss-products presents a best fit so far for this data, Willett and Singer suggest a simpler and more parsimonious representation which will both capitalize on the type of spell (enrolled vs. not enrolled, initial vs. repeat) as well as overcome the problem of a shrinking data set for later terms. By expressing spell as two main effects (OUTSIDE and SECOND) and interaction with term as the base-10 log of the term, the variation is captured. Code for making this conversion is given by Willett and Singer (1993).

The results of applying this last model to the stopout data are displayed in Table 3. The parameter estimate associated with a predictor indicates the difference, or vertical 'shift' from the entire baseline hazard function which is produced by the predictor. The odds ratio derived from this coefficient indicates the likelihood of a terminating event. For example, 0.31 in T1 conveys the odds of ending the initial period of enrollment after the first term. The odds ratio is produced by SAS PROC LOGISTIC and is computed by antilogging the parameter estimate associated with the predictor; $0.31 = \exp^{(-1.16)}$. The odds ratio indicates how much more likely a stopout is when the predictor variable is 1 (for students in their

/33

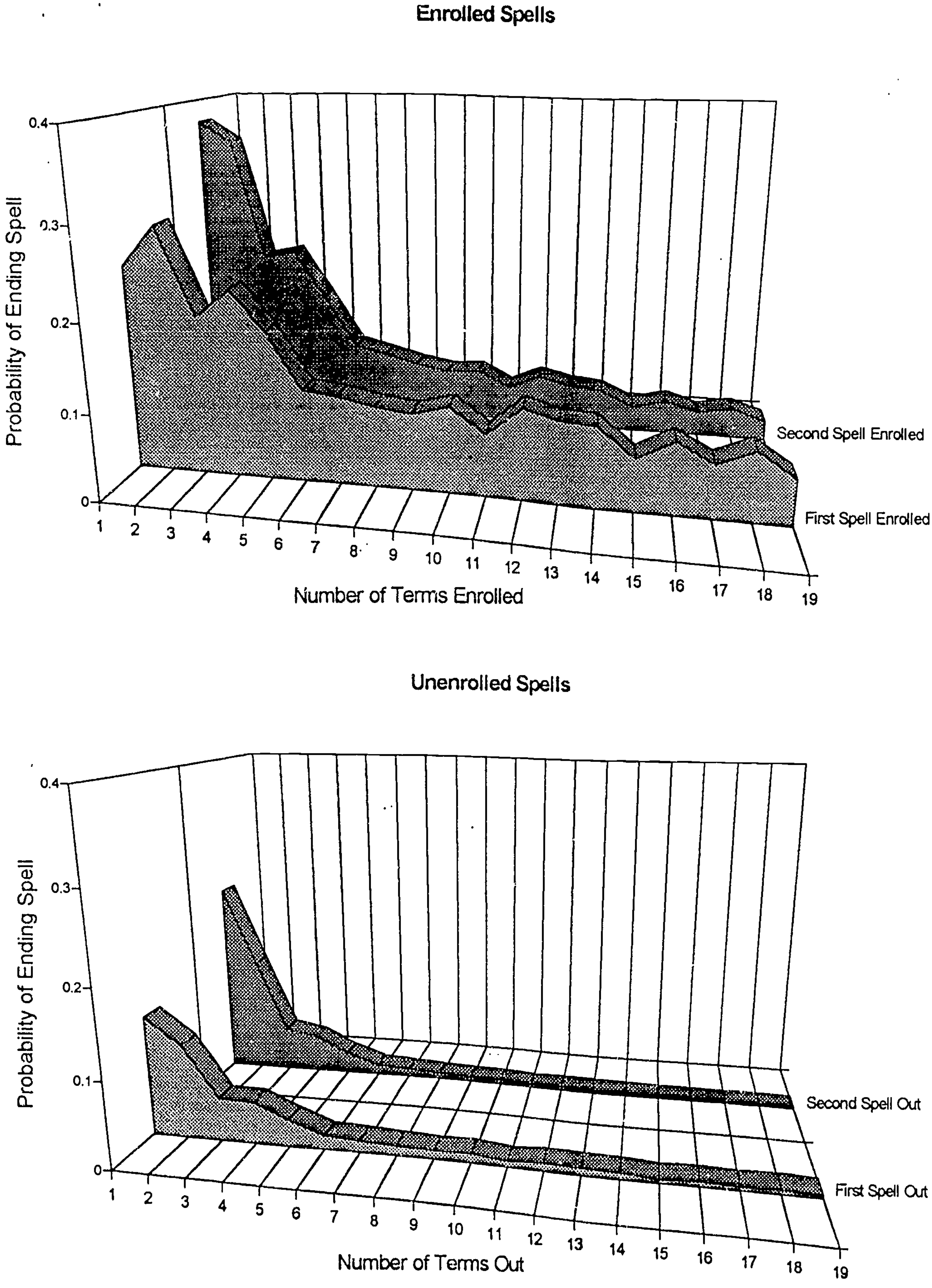

Figure 2. Hazard Functions for Enrolled and Unenrolled Spells

TABLE 3
Estimates for the Initial Model for the Effects of Time

Predictor	Parameter Estimate[1]	Odds of leaving Spell 1	Odds of returning after N terms out	Odds of ending a repeat spell
T1	-1.16 (.05)	0.31	0.52	1.90
T2	-0.92 (.05)	0.40	0.31	1.34
T3	-1.48 (.07)	0.23	0.23	1.10
T4	-1.28 (.07)	0.28	0.19	0.95
T5	-1.57 (.07)	0.21	0.16	0.85
T6	-2.11 (.11)	0.12	0.14	0.78
T7	-2.16 (.13)	0.12	0.12	0.72
T8	-2.26 (.14)	0.10	0.11	0.67
T9	-2.32 (.16)	0.10	0.10	0.63
T10	-2.19 (.17)	0.11	0.09	0.60
T11	-2.54 (.23)	0.08	0.09	0.57
T12	-2.13 (.22)	0.12	0.08	0.55
T13	-2.24 (.27)	0.11	0.08	0.53
T14	-2.26 (.30)	0.11	0.07	0.51
T15	-2.72 (.42)	0.07	0.07	0.49
T16	-2.39 (.40)	0.09	0.07	0.47
T17	-2.72 (.51)	0.07	0.06	0.46
T18	-2.44 (.51)	0.09	0.06	0.45
T19	-2.87 (.72)	0.06	0.06	0.44
OUTSIDE	-0.66 (.07)	0.52		
OUTLTRM	-1.71 (.14)	0.18		
SECOND	0.64 (.07)	1.90		
SECLTRM	-1.15 (.17)	0.32		

[1] All parameter estimates are significant at $p<.001$

first term of spell 1) rather than 0. Thus, the risk of leaving the spell of first enrollment peaks after the second term, and declines thereafter. The increasing standard errors are the result of a declining risk set; although 1790 students were enrolled for the first term, only 596 survived through term 5 without interruption in their enrollment. By term 19, only 3 remained.

The odds ratio associated with the variable OUTSIDE (.52) indicates that the risk of ending an out-of-school spell is only 52% the risk of ending an in-school spell ($\exp^{(-0.66)}$). In other words, students are only about half as likely to return to school after one term out as those who are enrolled are likely to leave after one term in. The interaction term, OUTLTRM, shows that this pattern accelerates over time. The longer the student is not enrolled, the less likely he is to return. Column 4 of Table 3 shows the odds of ending an outside spell, that is, returning to school after N terms out ($\exp^{(-0.66-1.71(\log10(\text{term number}))}$). The declining odds

indicate that the longer a student is not enrolled, the less likely he is to return. After 9 terms, for example, nonenrolled students are only 9% as likely to return as enrolled students are to leave; or to invert for easier interpretation, enrolled students are over ten times more likely to leave than nonenrolled students to return.

The coefficient on the variable SECOND indicates whether the spell is an initial or repeat spell. The odds-ratio of 1.90 for T1 indicates that the risk of ending a return spell, whether in or out of school, is almost twice the risk of ending an initial spell. However, the coefficient associated with the interaction with time term (SECLTRM = -1.15) shows that this differential reverses itself as time goes on ($\exp^{(0.64-1.15(\log10(\text{term number})}$). In other words, in the early years of a repeat spell, returning students (or returning stopouts) are more likely to end their return spells than their counterparts in an initial spell of enrollment or stopout. Over time, however, the returnees are more likely to remain in their spell. Again, this pattern can be better visualized by referring to Figure 1, where hazard profiles are displayed. Plotting of hazard functions is preferred to odds-based interpretations in discrete time survival analysis (Singer and Willett, 1993). Hazard is defined as $1/(1 + \exp^{(-(x)})$ where x is a parameter estimate.

Adding Constant and Time-Varying Predictors

One of the clear advantages of survival analysis is the ability to incorporate meaningful predictors and examine their effect on hazard. In this study, two additional predictor variables, ethnicity (ETHNIC) and full- or part-time attendance (STATUS) were added to the initial model for the effects of time. When using multiple-spell data, it is of interest to know not only whether the effect of predictors are constant over time, but also whether they are constant across spells (Willett and Singer, 1993). To test the main effect of ETHNIC, this variable was added to the initial model and its contribution to the goodness-of-fit was assessed in the same manner as described above for the initial model. The results ($\Delta \chi^2 = 0.04$ (1), $p > .10$) indicated that the hazard profile for Hispanic students in their first spell of enrollment did not differ from that of white, non-Hispanic students. When an interaction

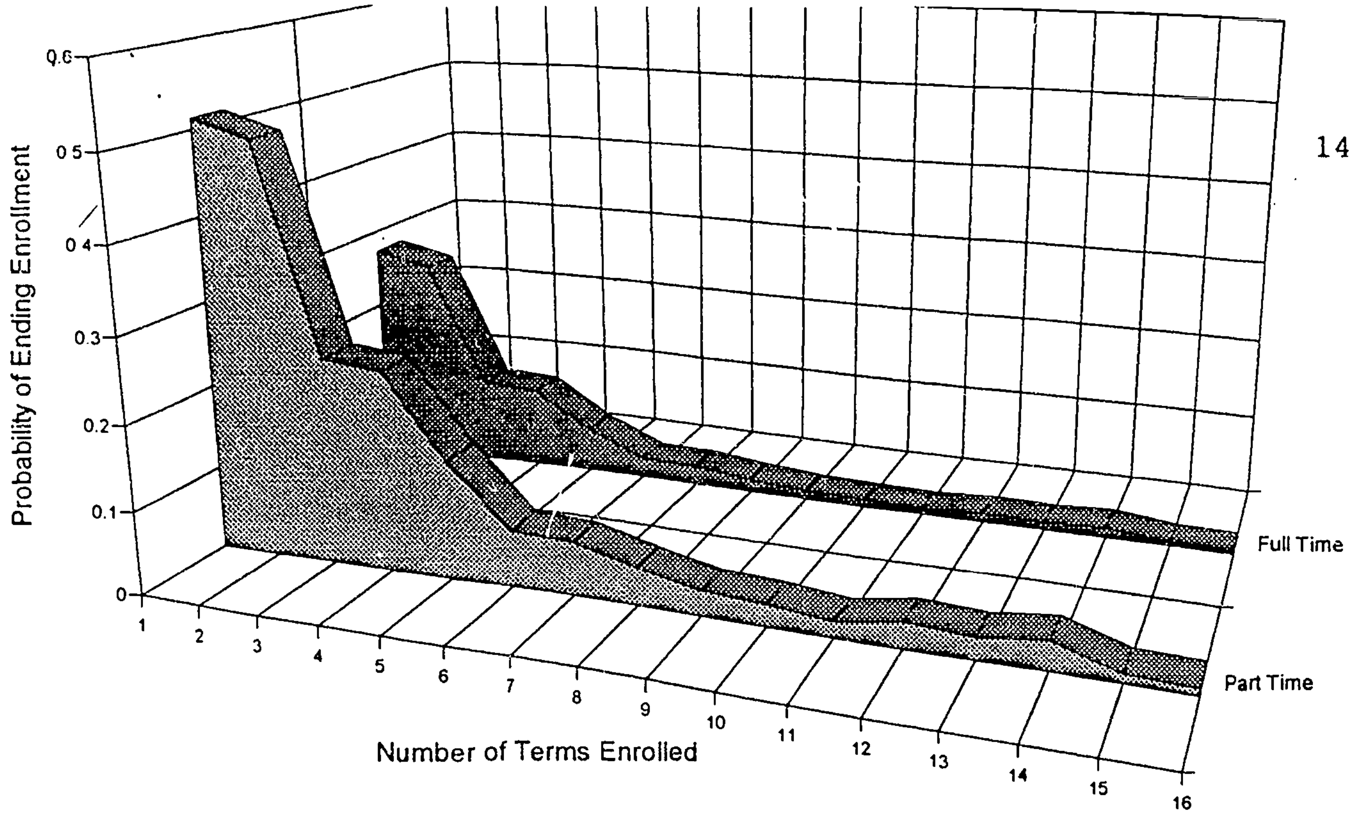

Figure 3. Hazard Functions for Full & Part Time Enrollments

term between ethnic and outside was added to the model, however, the resulting odds ratio indicated that Hispanic students are 1.7 times more likely than white non-Hispanic students to return to this institution after a period of stopout. The interaction with the log-period was also significant, indicating that this differential increased with time. The 2-way interaction between ETHNIC and SECOND did not add to the model's goodness of fit; therefore, Hispanics and non-Hispanic white students did not differ in the hazard rates of their return spells. The parameter estimates and their associated odds ratios are given in Table 4.

Assessing the effect of time-varying predictors in multiple-spell survival analysis is more complicated since values for out-of-school spells are missing from the data set. Willett and Singer (1993) propose a trichotomization of these variables so that the dummy variables will now represent membership in each of three status categories: Full-time enrolled, part-time enrolled, and not enrolled. By adding STATUS to the baseline model for the effects of time, we conclude that the odds of ending a first spell enrolled are almost three times as likely for part-time as for full-time students (Figure 3). The interaction with time, however, was not significant, nor was the interaction with SECOND; thus, ending a second episode of enrollment was equally likely for full and part-time students.

TABLE 4
Estimates for predictors ETHNIC and STATUS and their interactions with time

Predictor	Parameter Estimate[1]	Odds ratio	Parameter Estimate[1]	Odds ratio	Parameter Estimate[1]	Odds ratio
T1	-1.05	0.35	-1.05	0.35	-0.65	0.52
T2	-0.81	0.44	-0.90	0.41	-0.02	0.98
T3	-1.36	0.26	-1.51	0.22	-0.69	1.50
T4	-1.16	0.31	-1.35	0.26	-0.44	0.65
T5	-1.45	0.23	-1.68	0.19	-0.84	0.43
T6	-1.99	0.14	-2.24	0.11	-1.50	0.22
T7	-2.03	0.13	-2.31	0.10	-1.39	0.25
T8	-2.14	0.12	-2.43	0.09	-1.62	0.20
T9	-2.20	0.11	-2.51	0.08	-1.92	0.15
T10	-2.07	0.13	-2.40	0.09	-1.96	0.14
T11	-2.41	0.09	-2.76	0.06	-2.17	0.11
T12	-2.00	0.14	-2.37	0.09	-1.60	0.20
T13	-2.11	0.12	-2.49	0.08	-1.70	0.18
T14	-2.13	0.12	-2.52	0.08	-1.34	0.26
T15	-2.59	0.08	-2.99	0.05	-2.46	0.09
T16	-2.26	0.11	-2.66	0.07	-2.65	0.07
T17	-2.58	0.08	-2.99	0.05	**	**
T18	-2.31	0.10	-2.73	0.07	**	**
T19	-2.74	0.07	-3.16	0.04	**	**
OUTSIDE	-0.99	0.37	-0.67	0.51	**	**
OUTLTRM	-1.71	0.18	-1.67	0.19	**	**
SECOND	0.64	0.90	0.66	1.90	0.70	2.02
SECLTRM	-1.16	0.31	-1.17	0.31	-2.36	0.10
ETHCD	-0.19	0.82	-0.20	0.82		
ETHOUT	0.57	1.70				
ETHLTRM			0.51	1.66		
STATUS					-1.04	0.36

[1] All parameter estimates are significant at $p<.01$ except underlined.
** Empty cells.

Results for time-varying variables must be interpreted with care, however, since the risk associated with being a part-time student is only present during the terms when the student is attending part-time. Since a continuously-enrolled student may have terms of both full- and part-time enrollment in a single spell, interpretation of the risk profile is not at all straight-forward. Willett and Singer (1993) suggest that to ease interpretation, only individuals who were of one status for each spell be compared against those who were of the other status for the entire spell. These extremes will form the boundaries within which all students with "mixed spells" will fall.

DISCUSSION

Survival analysis provides a tool for analyzing enrollment behavior that is more meaningful and informative than simply tallying outcomes. From this limited analysis of stopout data using multiple-spell methods, we concluded that students' greatest risk of leaving is after the second term of enrollment. Hazard again peaks after the fourth term, then levels off after the sixth term. Students who leave are more likely to return after only one or two terms out. Once they have been gone for six terms, their odds of returning are virtually nil. Return enrollment and stopout spells share basically the same shape as initial spells, except that risks of ending return spells are more pronounced in the early terms, and risks tend to decline more quickly in later terms. Students who return after stopout are at particular risk for dropout during their first two terms back.

Hispanic students at this institution are a particularly persistent group. Although their risk of first stopout is about the same as that for non-Hispanic white students, they tend to return to school more often after stopping out. As time goes on, they become even more likely to re-enroll, relative to non-Hispanic white students.

As expected, part-time students have a greater risk than full-time students of stopping out, although the odds equalized for students who re-enrolled after stopout. Perhaps the part-timers returned better equipped to handle the demands of college after resolving whatever competing priorities led to their first stopout. It is also possible that students whose second enrollment spells had terms of part-time enrollment actually began their academic careers as full-time students; thus, they were more invested in their education by the time they switched to part-time attendance.

What implications do these results have for the University? A previous survival analysis using the fall 1986 entering cohort showed that the main reason for dropping out is academic failure; students in good academic standing are at relatively low risk of dropout (Ronco, 1993). The reasons why students fail are many, but primarily concern lack of preparation for college, competing family responsibilities and employment. Clearly, whatever the

109

University can do to ease the transition to college, help build good learning skills, alleviate the financial burden of going to college, and provide a conducive framework for students whose lives involve more than going to school, will favorably impact retention. Students have a much greater chance of success if they can remain continuously enrolled.

The first two terms after dropout are the best time to try to recapture lost students. Perhaps some effort could be made to reach these students before they are gone permanently. Again, recognizing that the first two terms after re-enrollment are critical for retention, special attention to the progress of these students is warranted.

Although the methods of multiple-spell survival analysis seem a bit esoteric at first exposure, once the data set is properly configured, analyses are straightforward and relatively simple to run with the usual statistical software. The interpretation of results is not quite as simple or straightforward, particularly as the number of spells increases or when time-varying predictors are added. In addition, key assumptions for linearity, homogeneity and proportionality should be checked. There are several excellent resources, particularly the studies of John Willett and Judith Singer, which can guide researchers through the mechanics of the process.

Multiple-spell survival analysis offers many advantages over more traditional methods, such as the ability to graphically examine the shape of hazard in several spells over time. Although this example used only two substantive predictors, extensions using other variables such as term GPA or receipt of financial aid, easily come to mind. Many universities like this one have instituted special programs in math and the sciences to retain and graduate minority students. These interventions can also become predictors to examine the participants' longitudinal enrollment behaviors. Outcomes such as graduation and transfer to other institutions were not considered in this study, but could be used as 'competing risks' of ending enrollment, resulting in an even more informative profile.

Since students no longer march lockstep through four years of college toward a degree, our methods of analyzing their progress must also change pace. Multiple-spell survival analysis is an important step in that direction.

REFERENCES

Allison, Paul D. (1984). Event history analysis: Regression for longitudinal event data. Newbury Park, CA: Sage Publications.

Carnegie Council on Policy Studies in Higher Education (1980). San Francisco: Jossey-Bass Publishers.

DesJardins, S.L. (1993, May). Using hazard models to study student careers. Paper presented at the 33rd annual forum of the Association for Institutional Research, Chicago, IL.

Ewell, Peter T. & Jones, Dennis P. (1991). Assessing and reporting student progress: A response to the new accountability. National Center for Higher Education Management Systems. ERIC Document Reproduction Services No. 337 112.

Hosmer, David W. & Lemeshow, Stanley (1989). Applied logistic regression. New York: John Wiley & Sons.

Johnstone, D. Bruce (1993). Enhancing the productivity of learning. AAHE Bulletin, 4, 3-7.

Mensch, Barbara S. & Kandel, Denise B. (1988). Dropping out of high school and drug involvement. Sociology of Education, 61, 95-113.

Murdnane, Richard J., Singer, Judith D. & Willett, John B. (1988). The career paths of teachers: Implications for teacher supply and methodological lessons for research. Educational Researcher, 17, 22-30.

Porter, Oscar F. (1989). Undergraduate completion and persistence at four-year colleges and universities: Completers, persisters, stopouts and dropouts. National Institute of Independent Colleges and Universities. ERIC Document Reproduction Service No. ED 319 343.

Ronco, Sharron L. (1993, Feb.). Getting started with survival analysis: An application to retention data. Paper presented at the Texas Association for Institutional Research Annual Meeting, College Station, TX.

Rosenfeld, Rachel A. & Jones, Jo Ann (1987). Patterns and effects of geographic mobility for academic men and women. Journal of Higher Education, 58, 493-515.

Singer, Judith D. & Willett, John B. (1991). Modeling the days of our lives: Using survival analysis when designing longitudinal studies of duration and timing of events. Psychological Bulletin, 110, 268-290.

Singer, Judith D. & Willett, John B. (1993). It's about time: Using discrete-time survival analysis to study duration and timing of events. Journal of Educational Statistics, 18, 155-195.

Smart, John C. & Pascarella, Ernest T. (1987). Influences on the intention to re-enter higher education. Journal of Higher Education, 58, 306-322.

Tichenor, Richard & Cosgrove, John (1992). Enrollment and academic progress of fall 1986 new students: Fall 1986 - spring 1991. ERIC Document Reproduction Service No. 346 892.

Willett, John B. & Singer, Judith D. (1988). Doing data analysis with proportional hazards models: Model building, interpretation and diagnosis. ERIC Document Reproduction Service No. ED 293 899.

Willett, John B. & Singer, Judith D. (1991a). How long did it take? Using survival analysis in educational and psychological research. In L. Collins and J. Horn (Eds.). Best Methods for the Analysis of Change: Recent Advances, Unanswered Questions, Future Directions (pp. 310-327). Washington DC: American Psychological Association.

Willett, John B. & Singer, Judith D. (1991b). From whether to when: New methods for studying student dropout and teacher attrition. Review of Educational Research, 61, 407-450.

Willett, John B. & Singer, Judith D. (1993). It's deja-vu all over again: Using multiple-spell discrete time survival analysis. Submitted to the Journal of Educational Statistics. Copy available from authors.

Yamaguchi, Kazuo (1991). Event history analysis. Newbury Park, CA: Sage Publications.

1:2