*Time Is of the Essence: Event History Models in Political Science**

Janet M. Box-Steffensmeier, *Ohio State University*
Bradford S. Jones, *University of Arizona*

Many questions of interest to political scientists may be answered with event history analysis, which studies the duration and timing of events. We discuss the statistical analysis of event history data—data giving the number, timing, and sequence of changes in a variable of interest. These methods are illustrated by examining three substantive political science problems: overt military interventions, challenger deterrence, and congressional career paths; many other applications are possible. Our article is intended to provide a better understanding of the growing number of applications that currently exist in political science and to encourage greater use of these models by showing why event history models are useful in political science research and explaining how one specifies and interprets these models.

Introduction[1]

Why do some states adopt certain kinds of public policies while other states do not? Why are some governments more stable than others? How is tenure in a political office related to the odds of losing that office? The answer to each of these questions bears, at least in part, on some implicit assumption that *when* some event occurs is as important as *if* some event occurs. Time plays a key role in politics and a class of econometric models, known collectively as event history analysis, can provide researchers leverage on the issue of the timing of political change.[2] Event history analysis allows researchers to answer a more extensive set of questions than conventional analyses by using information on the number, timing, and sequence of changes in the dependent variable.

In event history analysis, we are concerned with patterns and causes of change (Yamaguchi 1991). We are interested in knowing how the duration spent in one social state affects the probability some entity will make a transition to another social state. Political science theories have become increasingly focused on change processes; and temporal data are becoming widely available, yet the vast majority of empirical research focuses on static rela-

*Authors are listed alphabetically. We would like to thank Laura Arnold, Neal Beck, Barry Burden, Bill Dixon, David Kimball, Miller McPherson, Krishnan Namboodiri, David Patterson, Lynn Smith-Lovin, and Richard Tucker for helpful comments and suggestions.

[1]The data, program code, and documentation may be obtained by anonymous ftp from ICPSR's publication-related archive under the authors' names or title of the article.

[2]We recognize that these models belong to the broader class of probabilistic models. Readers interested in a more mathematical treatment are referred to Tuma and Hannan (1984) or Greene (1993).

American Journal of Political Science, Vol. 41, No. 4, October 1997, Pp. 1414–1461

tionships (i.e., at one point in time, typically cross-sectional studies). But as Tuma and Hannan (1984, 3) point out, even when time-series or panel data are analyzed, the temporal structure is often ignored and the data are treated as though they are cross-sections with some additional methodological complications involving autocorrelations. In addition, we strongly believe that as appropriate methods for studying change—event history methods—are better understood, more dynamic analyses will be conducted.

Bartels and Brady (1993) argue that political methodology should be directed toward more wide-spread use of models for event history. In an attempt to further this goal, we discuss why event history models are useful in political science research, explain how one specifies and estimates models, how one interprets the results, and how one chooses between competing models. We illustrate our discussion with empirical examples and the Appendix discusses software available for estimation. To understand the potential utility of event history models, one must first understand the problems associated with traditional regression-based methods.

Traditional Regression-Based Methods and Event History Data

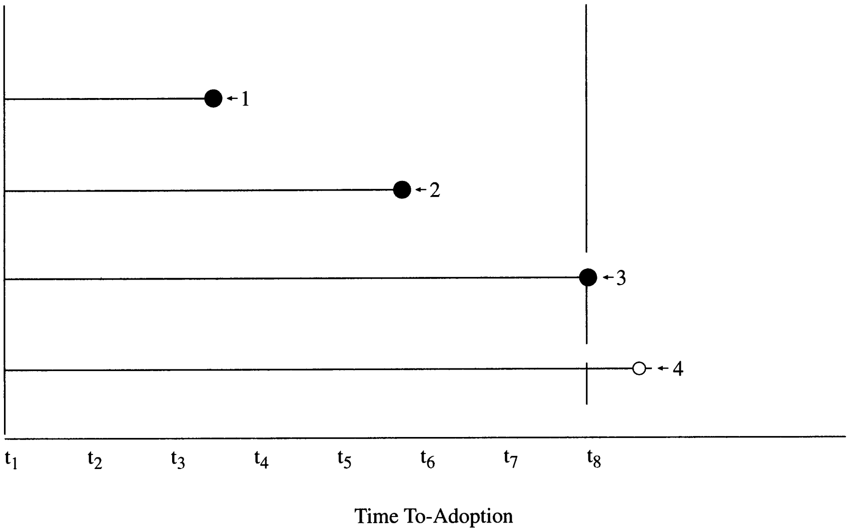
Suppose we are interested in understanding why some U.S. states are more apt to adopt a certain kind of public policy while other states are not. One possible strategy would be to record the number of legislative sessions the policy was considered until its adoption. Presumably those states quicker to adopt had stronger preferences for the policy while other states which adopted later had weaker preferences for the policy. One indicator of issue innovation and diffusion might be to record the duration of time before adoption occurs. In this simple example, our dependent variable would be an indicator of time elapsed before adoption. We can denote this as t_i. In addition to collecting data on duration, data could be collected on some exogenous variables theorized to influence t_i. To test hypotheses concerning these exogenous variables, a regression could be estimated:

$$t = \beta'X + \varepsilon \qquad [1]$$

where $\beta'X$ is a matrix of exogenous variables and associated parameters and ε is a random disturbance. While this design seems plausible, at least two problems plague this approach: right-censoring and time-varying covariates.

To illustrate the problem of right-censoring, consider Figure 1. In this figure, four hypothetical cases of state policy adoption are considered. In case one, the duration until adoption is three "time units." In case two, the duration prior to adopting is five units. Case three represents a state that adopted the policy during the last period of data collection (eight units). Finally, case four depicts a state that by the last time period observed, had yet

Figure 1. Hypothetical Durations for Time-to-Adoption



Each time unit represents an observation period; ● denotes the occurrence of an event and ○ denotes nonoccurrence of an event)

to adopt a policy. In this figure, cases 1–3 present no special analytic problems. We see that the duration prior to adoption is two-thirds longer in state two than in state one. And the duration in state three is 60% longer than in state two. The variation in these durations could possibly be explained by the exogenous variables. Case four, however, is problematic: by the end of the observation period this state still had not adopted the policy, it is considered right-censored. We do not know when state four will end.

The problem with the regression model is its failure to distinguish state three from state four although the two states are quite distinct; one has adopted and the other has not. Inclusion of right-censored observations in the model implicitly treats them as having experienced the event (policy adoption) when in fact they have not. And since we cannot foretell the future, we do not know how "much longer" (if ever) censored observations would go before experiencing an event. One "solution" to the censoring problem would be to eliminate all censored observations from the data set. Unfortunately, this cure is worse than the ill. If the factors producing censoring are *completely unrelated* to factors promoting an event's occurrence, then truncating the sample may be a solution; however, censored observations are often influenced by precisely the same factors uncensored observations are.

Under such conditions, truncating the sample to include only uncensored observations would produce a biased sample because only observations initially prone to experience an event would be included.[3] States with "staying power" would be eliminated from the sample. Truncating the sample would induce selection bias into the data and the implications of this problem are well-known (Achen 1986; Geddes 1990; King, Keohane, and Verba 1994).

Another alternative to avoiding the censoring problem would be to create a dichotomous indicator denoting whether or not a state experienced the event within the time frame of the analysis. As Petersen (1991) notes, however, this "solution" belies the logic of studying time-dependency in the first place. An indicator variable cannot capture the variability in duration time a state spends prior to adoption—precisely the effect we are trying to understand. In short, a dummy indicator could not discriminate between varying times-to-adoption. However, running logit or probit models and losing information on when an event occurs is quite common in the literature. This does not cause bias or inconsistency, but it gives estimates that are inefficient, i.e., that have larger variances, relative to the estimates from event history analysis (Chung, Schmidt, and Witte 1991).

An additional problem involves incorporation of time-varying predictors. The regression model must treat all exogenous variables as fixed. Unfortunately, many political factors related to policy adoption vary across time. For example, the partisan distribution in a state assembly might affect the odds of adoption. Because the partisan distribution in an assembly varies, it would be preferable to use a methodology that could account for this. Thus, the traditional regression approach breaks down in an important way; we have a dynamic process, but we do not have a dynamic model. A methodology is needed that can adequately address these problems.[4]

The Event History Approach

The underlying premise of a duration or event history model is that the "duration" of some social process is being modeled. In political research, scholars may be interested in the duration of time before countries engage in military conflict, the "survival" of parliamentary governments, or the length of congressional careers. The event history approach may also prove useful to political psychologists. Duration models could be one way to model reaction times in cognitive experiments. For example, Robinson and Smith-Lovin (1990) have employed event history methods in their studies

[3]The bias occurs when there is a lot of right-censoring; it is not relevant when N is large or right-censored cases are few and start mostly near the end of the observation period.

[4]Alt and King (1995) provide an enlightening discussion of coding and analysis schemes (binary, count, and duration). Tuma, Hannan, and Groeneveld (1979) contrast event history analysis with cross-sectional, count, and panel data analysis.

of conversational dynamics. In general, the event history approach can lead to insights on the full span of a social process. This approach is obviously preferable to typical cross-sectional designs and even panel designs. With respect to cross-sectional designs, the dynamics of a process cannot possibly be modeled. Additionally, we implicitly make the unlikely assumption that the process being modeled with the cross-sectional data is in equilibrium (Tuma and Hannan 1984). Panel study designs clearly elicit greater leverage on the problem of understanding change; however, as Singer and Spilerman (1976a, 1976b) noted long ago, the change process observed in typical panel studies is often consistent with many dynamic specifications. Furthermore, depending on the spacing of panels, a panel design may lead to inaccurate conclusions about the rates and timing of change. An event history design can avoid these problems.

To begin the discussion, we start with three elementary concepts: the survivor function, the occurrence of an event, and the hazard rate. The survivor function, $S(t)$, expresses the probability that the duration, T, has survived beyond, or has not ended by time t:

$$S(t) = P\,(T \geq t) \tag{2}$$

So if we were modeling the duration of time a government stays in power, each government still in power at the time of our observation would be considered a survivor. The second concept involves the realization of an event, which refers to the occurrence in time, of some event or outcome. Often we collect data or model processes where events can occur continuously in time. Therefore, we think of the probability density function of an event occurring within some differentiable area:

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t + \Delta t > T \geq t)}{\Delta t} \tag{3}$$

The term $f(t)$ represents the probability density function of the duration ("duration density") and may be interpreted as the instantaneous probability of the occurrence of an event T at time t. The cumulative distribution function of the duration may be expressed as:

$$F(t) = \int_0^t f(u)du \tag{4}$$

The third key concept underlying event history analysis may be defined as the hazard function. The hazard function or *hazard rate* may be expressed as the following:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t + \Delta t > T \geq t \mid T \geq t)}{\Delta t} \qquad [5]$$

The hazard rate reflects the rate at which a duration or episode ends in the interval $[t, t + \Delta t]$, given that the duration has not terminated prior to the beginning of this interval. Most analyses using duration data tend to model the hazard rate, $h(t)$, rather than $S(t)$ or $f(t)$. The reason for this is straightforward. One may interpret the hazard rate as reflecting the risk an object incurs at any given moment in time, given an event has not yet occurred. From a substantive point of view, political scientists can ask interesting questions such as, "what is the risk a cabinet government will fall?" (c.f. Warwick 1992), or "how long do members of Congress stay on a congressional committee?" (c.f. Katz and Sala 1996). Such questions beckon a time dimension; and by modeling the hazard rate, we gain substantial leverage on answering such questions. An additional reason why we focus on the hazard rate is that mathematically, $h(t)$ possesses several desirable properties.

To illustrate these properties, we need to consider in more detail the linkage between the survivor function, the duration density, and the hazard rate. First, define the cumulative hazard function as follows:

$$H(t) = \int_0^t h(u)du \qquad [6]$$

$H(t)$ can be thought of as the summation of the hazard experiences from the beginning of time until t. Knowing $H(t)$, we can now reexpress the survivor function shown in [1] as a function of the cumulative hazard function:

$$S(t) = \exp\left[-\int_0^t h(u)d(u)\right] = \exp^{-H(t)} \qquad [7]$$

while the probability density function of the duration T, $f(t)$, shown in [2] may be equivalently reexpressed as:

$$f(t) = h(t) \cdot \exp\left[-\int_0^t h(u)d(u)\right] = h(t) \cdot \exp^{-H(t)} \qquad [8]$$

From [7] and]8], it is clear that the following relationship holds:

$$f(t) = h(t) \cdot S(t) \qquad [9]$$

Now we can restate the hazard rate, *h(t)* from [4] in the following manner:

$$h(t) = \frac{h(t) \cdot \exp^{-H(t)}}{\exp^{-H(t)}} \qquad [10]$$

Consequently, the hazard rate can be expressed as the following (from Equations 7 and 9):

$$h(t) = \frac{f(t)}{S(t)} \qquad [11]$$

Finally, rearranging terms in [11] leaves us with the following result:

$$S(t) = \frac{f(t)}{h(t)} \qquad [12]$$

This shows that the three core concepts of event history modeling—*S(t), f(t), h(t)*—are mathematical functions of each other. Knowing h(t) allows one to derive the others, which is why h(t) is so useful. Because of these important relationships, we are in a better position to understand the usefulness of modeling hazard rates.

Often, practitioners of event history models make assumptions about the hazard rate's dependency on time. Does the hazard rate systematically vary over time or is the probability of duration termination time-invariant? The assumptions determine which distributions may be used. Suppose, as it often is, that the hazard rate is assumed to be time-invariant. Time-invariance means that the rate of termination is neither a decreasing nor increasing function of time. The most commonly used distribution under such assumptions is the exponential. The hazard rate for the exponential distribution serves as a baseline for comparison because it is a constant:

$$h(t) = h \qquad [13]$$

The hazard rate at any given point in time is equivalent to the hazard rate at any other point in time. Hence, a graphical depiction of *h* with respect to time would yield a flat line. From [6], it can be shown that the cumulative hazard function for the exponential distribution is:

$$H(t) = h \cdot t \qquad [14]$$

Consequently, the survivor function for the exponential distribution is (from equations 7 and 14):

$$S(t) = \exp^{-h \cdot t} \qquad [15]$$

and the duration density function, $f(t)$, is (from equations 8 and 15):

$$f(t) = h \cdot \exp^{-h \cdot t} \qquad [16]$$

The exponential distribution is used here to show that knowledge of *only* the hazard rate function permits calculation of the survivor and duration density functions. In general, knowing the hazard rate function or making some assumptions about the distribution of the hazard rate function (and knowing the relationship of the hazard rate function to the survivor and duration density functions shown in Equations [7]–[12]) allows computation of the survivor and distribution density functions. Hence, for mathematical purposes, event history analyses routinely focus on the hazard rate function. This, coupled with the nice interpretative qualities of the hazard rate, makes it an ideal quantity to model in political science. Later in the paper, we consider distributions other than the exponential; however, before discussing estimation topics, we examine issues of data.

Event History Data

A practical concern for estimation of an event history model is the structure of the data set. Hypotheses generated from processes at least partially dependent on time duration presuppose the idea that the duration spent in one social state is related to the probability of experiencing some event. Hence, event history data, at a minimum, consist of the length of time a unit spends in a state before experiencing an event and an indicator denoting whether or not the observation is right-censored. For most political applications, it is unlikely the researcher will be satisfied in only examining the effect of time on the probability of experiencing an event.[5]

Consequently, if one believes covariates are related to the likelihood of an event's occurrence, then these data for each unit in the sample will also be included. If some covariates are time-varying, then it is necessary to have data for each unit at each observation period. Finally, event history data traditionally "track" units at some prespecified starting point until each unit experiences an event. If a unit does not experience an event by the time the last observation period ends, the unit is right-censored.

It is critical for the researcher to have a theoretically sound reason for hypothesizing when a social process for an observation can begin, i.e., when

[5]In fact, early applications of these types of models tended to only examine the duration of time until failure or death of an entity or object. Life-table analysis, a precursor to event history models, generally only examined the distributional nature of death or failure-times. Absent the specification of covariates, any model we discuss in this paper can be thought of as models solely as a function of time. Classic estimators of survivor and hazard functions like the Kaplan-Meier Statistic are not dealt with in this paper (Cioffi-Revilla 1985; Cioffi-Revilla and Lai 1995).

does time start? This involves the notion of "being at risk." Observations at risk are simply the sample of units that begin a social process; they become "at risk" of experiencing some event. The "risk set" therefore, is the set of units in some time interval that are at risk of experiencing an event. In so-called discrete time models, the unit is at risk of experiencing the event at predefined times (for example, election day). A continuous-time model presupposes event occurrence can happen at any point in time. In either formulation, once the event is experienced, the unit exits the risk set. Hence, at each observation period, the risk set progressively dwindles until, by the end of the observation plan, no units are at risk (as each has experienced the event) or they are right-censored.[6]

In some applications, the natural starting point for a unit to enter the risk set is straightforward: governments are only at risk of being overthrown once they take control. In other applications, however, defining the entry point or beginning point of a social process is not obvious. For example, when does a conflict, such as the Vietnam War, begin?[7] Or when does an individual first acquire a partisan identification? The left-censoring problem occurs when observations of individual cases begin *after* the case has already entered the risk set. Unobserved histories raise special problems. If one's prior history influences the probability of experiencing an event, coefficients generated from data that are heavily laden with left-censored observations will produce incorrect estimates of duration effects and could result in biased estimation of covariates. Time-varying covariates are especially problematic (Tuma and Hannan 1979; Yamaguchi 1991). Ham and LaLonde (1996) discuss the problems of initial conditions on studying duration data

[6]Standard event history models require the assumption that all individuals will eventually experience the event; split population models relax this assumption and estimate the probability of the event occurring (Schmidt and Witte 1988, 1989). Schmidt and Witte point out that an advantage of the split population model is they do not imply unreasonably high failure rates. Swaim and Podgursky (1994) generalize and apply the split-population model by allowing for right-censoring. Harris, Kaylan, and Maltz (1981) develop a related model where subgroups have hazard rates of θ_1 and θ_2, but neither is zero. Split population models also allow one to identify any differences that may exist between the determinants of failure and determinants of the timing of failure. In some applications, such as the timing of PAC (political action committee) contributions (Box-Steffensmeier and Radcliffe 1996), the possibility arises that an event, i.e., a contribution is given, and the timing of the event depend on different forces. If the population is split into two groups, one composed of members of Congress that received a contribution and the other composed of members that did not, then whether the factors that explain receiving a contribution are different from the factors that explain the timing of receiving a contribution can be explored.

[7]The problem is not that the starting points are unmeasurable but rather that they may be controversial or are subject to the researchers discretion. One way that the onset of conflict has been measured is by using the Correlates of War project data and defining a war to begin when military engagements produce fatalities in the first year of greater than one thousand (Small and Singer 1982).

and argue that ignoring these conditions leads to misleading inferences. In fact, there exist few readily accessible methods to account for the problem of left-censoring.[8] From a practical point of view, researchers should try their best to avoid left-censored data.

Let us turn attention now to the formulation of the dependent variable. In what Petersen (1991, 295–7) calls the "event history" approach, the dependent variable is an indicator variable denoting whether or not an event occurred. At each observation point, an individual is coded either a "0" denoting no event occurred, or a "1," denoting the occurrence of an event. Therefore, there are as many 0 and 1 codings in the data set as there are person-period observations. The second form of the dependent variable in this class of models is what Petersen (1991) calls the "duration" approach. In this formulation, the dependent variable records the amount of time that has passed before an event is experienced. Both approaches result in the same specification, estimation, and interpretation of the models.

Estimating an Event History Model

In this section, we consider various alternative approaches to estimation of a typical event history model. As noted earlier, the discrete-time formulation of the event history model presumes change only occurs at discrete, often predetermined times (for example, election day).[9] Many conceivable social processes do not behave so predictably: transitions from one state to another may occur *anywhere* in time. Such processes may be referred to as continuous-time processes. Unfortunately, while many processes may be absolutely continuous, our techniques for observation and/or measurement may fail to approximate the continuous nature of change. Data for continuous-time processes often are collected at discrete intervals. Examples might include data collected by fiscal month, quarter, or year. Change may occur anywhere in the interval, but the data are only "observed" at predefined periods. So while the continuous-time process presumes knowledge of when change occurred in time, we only have an approximate guess as to when the change or transition actually occurred.

In contrast, some longitudinal processes may conceivably be continuous-time processes, but the explicit knowledge of when in time change occurred is largely unimportant. To illustrate, consider an example of state adoption of public policy. Presumably, a legislature could adopt a policy anytime within a legislative session. Because state legislatures routinely record votes, we could easily discern precisely when change occurred. In

[8]We should note that left-censoring is not a problem endemic only to event history analysis. Indeed, all cross-sectional models are plagued with left-censored observations.

[9]See Beck (1995) for an interesting view on the discrete versus continuous-time methods for analyzing duration data.

most analyses of state policy adoption, however, the crucial issue is *not* knowing exactly when adoption (the "event") occurred within a legislative session, but rather when adoption occurred relative to other states. In such analyses, the year in which a policy was adopted may be sufficient to demarcate the occurrence of an event. Therefore, while policy adoption may be a continuous-time process, a discrete-time view is probably an accurate view of the process *given the nature of the research question.* Thus, work by Berry and Berry (1990, 1992, 1994), Mintrom (1994) and others who have used discrete-time models for an otherwise continuous-time process are correctly and accurately modeling the time path of policy adoption. As another example, the events of retirement and ambition for members of Congress certainly can occur anywhere in time, however, these decisions are roughly bounded, occurring near the onset of an election cycle (Jones 1994).

The adequacy of the discrete-time formulation largely hinges on the scope of one's research question. If data are only observed at discrete intervals *and* transitions from one state or event to another can be approximated by the discrete nature of the data with little loss of relevant information, then the discrete-time formulation is likely to be a reasonable approach toward modeling the process. More formally, the discrete-time approximation of a continuous-time process improves if the intervals between observations are small, i.e., they are only good approximations if the conditional probabilities of an event at each period are small (Yamaguchi 1991).

An advantage of the discrete-time formulation is its ability to handle "ties" (Yamaguchi 1991). Because data are often only gathered at researcher-defined periods, the likelihood many units in the risk set experience an event at the same time is high. Standard continuous-time event history models such as the Cox model, cannot easily handle the problem of co-occurrence of events (Arjas and Kangas 1992; Yamaguchi 1991). Estimation of the Cox proportional hazards model on data containing many ties yields biased parameter estimates (a more detailed discussion of the Cox model is provided in the section on continuous-time models). The discrete-time formulation does not produce bias in the parameter estimates. Thus, if a researcher is dealing with a continuous-time process but has data that are observed at crude intervals, then use of the discrete-time formulation may actually improve inferences made about the process.

The Discrete-Time Formulation

In discrete-time event history analysis, the statistical model is used to derive estimates of the underlying hazard probability of a unit experiencing an event. Whether or not a unit experiences some event is indicated by the dependent variable. To illustrate the discrete-time event history estimator, let us first consider the formulation of the hazard probability. Since an event can occur only at discrete times, it is assumed the probability of event T exactly occurring at t is observable. Therefore, the discrete time hazard probability is a straightforward expression:

$$\lambda(t) = P(T = t \mid T \geq t) \qquad [17]$$

We use $\lambda(t)$ as the notation for the discrete-time hazard function to distinguish it from $h(t)$, the continuous-time hazard function. It is useful to contrast [17] with [4], which is the continuous time version of the hazard rate. In the discrete-time case, we make the assumption the event is observable (and measured) at some exact point in time. Hence, the interpretation of $\lambda(t)$ is quite appealing. The discrete-time hazard may be interpreted as the probability that a unit experiences an event at t, given the event has yet to be experienced. The results we derived for $h(t)$ in the previous section also hold for $\lambda(t)$.[10]

To this point, we have only thought about the hazard rate (or hazard probability as it is properly called in the discrete-time case) as a function of time; however, most researchers are interested in how the hazard probability varies as a function of independent variables, or covariates. To account for covariates, [17] can be reexpressed as the following:

$$\lambda(t) = P(T = t \mid T \geq t;\ \alpha,\ \beta'X) \qquad [18]$$

The α term represents the baseline probability, i.e., the probability of an event occurring when the covariates equal zero. The $\beta'X$ term represents a matrix of covariates and their associated parameters. A convenient feature of the discrete time formulation lies in the fact that λ's are probabilities. In a classic article, Cox (1972) demonstrated this probability can be parameterized to have a logistic dependence on the covariates (including the baseline parameter). To obtain logistic dependency, we rewrite [18]:

$$\lambda(t) = \frac{1}{1 + \exp^{-[\alpha+\beta'X]}} \qquad [19]$$

From this result, we can place the hazard probability into the logistic form by taking the logistic transformation on both sides of [19]:

$$\ln \frac{\lambda(t)}{1 - \lambda(t)} = \alpha + \beta'X \qquad [20]$$

This formulation illustrates the conditional log-odds of an event occurring at t is dependent both on the baseline term, α, and the covariates, X. By

[10]Of course in the discrete-time case, we no longer would use the calculus notation.

expressing $\lambda(t)$ in this fashion, [20] is estimable with logit—a technique widely used by political scientists. The discrete-time logit model, however, differs in interpretation from traditional logit models. This differing interpretation stems from the fact that the data used to estimate this model (or any event history model) are duration data. Later in the paper, we focus on interpretation of event history parameters.[11]

The adequacy of the discrete-time logit event history model is solely dependent on the process under study. As previously discussed, discrete-time approaches are only feasible under two conditions. First, if change only occurs at discrete times (on election day, for example), then the model illustrated in [20] may be the appropriate estimator. Second, if one is modeling a continuous-time process, but the intervals of measurement are close, then a discrete-time approximation is often an adequate approximation.[12] Furthermore, as we will see in the next section, the distinction between discrete-time and continuous-time models becomes more blurred when we consider the parameterization of the baseline hazard.

Our discussion of the discrete-time model has been expressed in terms of the logit estimator; however, estimation of the discrete-time event history model may be accomplished in a number of ways. Recent work by Sueyoshi (1995) presents alternative estimation techniques for the discrete-time model. Beck (N.d.) provides a nice discussion and overview of some of these techniques. We also point the reader to recent works by Alt, King, and Signorino (1996) and Beck and Tucker (1996) for examples and discussions of estimation of the discrete-time model.

The Continuous-Time Formulation

When the times of events are not constrained to predefined periods, we have a continuous-time process. Generally speaking, the dependent variable in a continuous-time formulation of the event history model reflects the duration or time spent in a social state. Consequently, the dependent variable is most often thought of as a continuous random variable and is measured as some metric of time.[13] The hazard rate for the continuous-time event history

[11]See Allison (1982, 1984); Yamaguchi (1990); and Singer and Willet (1993) for a more detailed discussion of using logit to estimate event history parameters.

[12]To see this, note that logit can be interpreted as a ratio of two odds: the odds of an event occurring (i.e., the realization of a "1") relative to the baseline odds of an event not occurring (i.e., the realization of a "0"). As Yamaguchi (1991) notes, the ratio of two odds approaches the ratio of two *rates* as the interval between observations gets smaller. Obviously, the antithesis to this is that if one is trying to understand a continuous-time process but has intervals of observed data distantly or irregularly separated in time, then the discrete-time logit model is probably inappropriate.

[13]Examples of a "time metric" may include days, months, or even years spent in some social state until an event occurs (or until transition is made from one state to another). In some applications of continuous-time models, the time metric is measured in seconds or even finer-grained measures (see for example Robinson and Smith-Lovin 1990).

model accounts for the fact that transitions can occur anywhere in time. We have already presented the continuous-time hazard rate in [4], but it bears repeating:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t + \Delta t > T \geq t \mid T \geq t)}{\Delta t} \qquad [21]$$

Again, the quantity $h(t)$ reflects the instantaneous probability that a duration or episode ends in the interval $[t + \Delta t]$, given the duration has not terminated prior to the beginning of the interval. Or analogously, the hazard rate can be interpreted as the instantaneous probability that an event occurs given that the event has not yet occurred.

The hazard rate can be modeled as a function of both a baseline rate, α, and covariates. To account for covariates, we can reexpress the hazard rate as the following:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t + \Delta t > T \geq t \mid T \geq t\,; \alpha, \beta' X)}{\Delta t} \qquad [22]$$

There exist a wide variety of estimators for [22]. The major issue involved with the estimation of [22] (and [21] for that matter) involves the parameterization (or nonparameterization) of the baseline hazard rate. To understand this, bear in mind that the baseline hazard rate can be thought of as the time path that durations follow if the effects of all covariates are zero. The base–line hazard, then, reflects time dependence (or independence). Occasionally, researchers have some idea of what the theoretical hazard function should look like. For example, the hazard of human death is initially steep (at birth) due to birth defects and infant diseases, the death rate then drops and flattens out during childhood and early adulthood, and begins to rise as humans get older due to cumulative wear and tear. Heckman and Willis (1977) find a "bathtub shaped" distribution for the participation probabilities of women in the labor force as well. This bathtub function indicates that the hazard rate is not only time dependent, but follows some known distribution. By parameterizing the baseline hazard rate we can explicitly account for time dependence. Parameterization means specifying a distributional form for the baseline hazard. And to that end, there are numerous distributions to choose from. We consider some of these distributions in turn.

The Exponential Model

We have already considered some of the properties of the exponential distribution in our discussion of the hazard rate; however, some elaboration on this distribution will be useful. If one parameterizes the baseline hazard function as following the exponential distribution, the assumption being made is that the hazard rate is invariant to time. Graphically, we would observe a hazard function as a flat line with an intercept equal to the hazard rate. Time-invariant baseline hazards are commonly assumed, particularly when one includes many covariates. The exponential model takes the following form:

$$h(t) = \exp(\beta_0 + \beta_k' X_i) \quad [23]$$

The baseline hazard in this model is tantamount to a constant term and is represented in [23] as the β_0 term. The term $\beta_k' X_i$ represents a matrix of coefficients for the k covariates. Interpreting [23] is straightforward. When all covariates are set to zero, the hazard rate is a constant, thus implying time-invariance. All the "movement" of the hazard rate comes from the covariates. This model may be estimated by maximum likelihood. Many standard statistics packages can estimate the exponential model (see Appendix A for a discussion of software).

The Weibull Model

A probability distribution that plays a central role in the analysis of event history data is the Weibull distribution. The Weibull model is frequently used by event history analysts because of the flexibility in specifying different functional forms. One example of why, theoretically, a scholar may use the Weibull model rather than the exponential model is if he or she posits that the dependent variable shows time dependence. Specifically, when studying the mortality (cessation or merger) of political parties, someone may posit that new organizations fail at higher rates than old ones. A Weibull model would allow the mortality rate to vary by the age of the party. Comparing the fit of exponential and Weibull models would allow one to draw conclusions about the null hypothesis of age independence.

The hazard function for the Weibull distribution takes the following form:

$$h(t) = h\alpha(ht)^{\alpha - 1} \quad [24]$$

The terms are the shape parameters for the distribution. When $\alpha = 1$, the Weibull hazard is equivalent to the exponential hazard shown in [13]. Hence, the exponential distribution is a special case of the Weibull distribution. When $\alpha < 1$, the baseline hazard rate is a strictly decreasing function. When $\alpha > 1$, the baseline hazard rate is a strictly increasing function.

The Weibull model is typically estimated in the following fashion:

$$h(t) = \exp(\beta' X + \alpha \ln t) \quad [25]$$

The last term on the right-hand side of [25] represents the shape parameter of the Weibull model and is estimated directly from the data. The coefficients, $\beta'X$, in conjunction with α, can be interpreted as reflecting how covariates increase or decrease the risk, or hazard, of experiencing an event at some point in time. Because the shape parameter is estimated, one can use the Weibull specification to test hypotheses about different "shapes" of the hazard function. We present an application of the Weibull parameterization later.

Piecewise Constant Hazard Rates

A general method toward modeling the hazard rate is to model $h(t)$ as a so-called piecewise function. In stark contrast to the exponential model where the baseline rate is time-invariant, treating $h(t)$ as a piecewise function is tantamount to saying the hazard rate varies from time period to time period. To estimate a piecewise hazard, one simply includes separate constant terms for each t. Hence, the model would take the following form:

$$h(t) = \exp(\alpha_t + \beta_k' X_i) \qquad [26]$$

The term α_t represents separate constant terms for each period t. These parameters provide information on how the baseline hazard rate increases or decreases per some defined time period.[14] Clearly, the piecewise formulation differs in an important way from the continuous-time models previously discussed. In this formulation, the researcher is making the assumption that the hazard rate is not constant across time intervals (as in the formulation shown in Equation 23). Inclusion of time dummies captures this. The parameter estimates of these time dummy variables simply informs the researcher on how the hazard function changes in the time interval captured by the indicator variable.

Other Distributional Models

The distributions described will not necessarily provide a satisfactory model for survival times in all circumstances. Any continuous distribution for nonnegative random variables could be used, such as the normal (which allows for negativity), log-normal, log-logistic, gamma, extreme-value, and Gompertz distributions. We refer interested readers to sources such as Kalbfleish and Prentice (1980), Cox and Oakes (1983), and Petersen (1985) for further discussion of alternative parametric models.[15]

[14]One advantage of this approach is that periodic heterogeneity is explicitly modeled. To that end, this approach is roughly similar to Stimson's (1985) notion of accounting for time effects.

[15]In addition to the technical references cited, Cioffi-Revilla (1984) and Cioffi-Revilla and Lai (1995) provide more graphs (theoretical and observed) than can be presented here.

The Issue of Censored Observations

Each of the model specifications we have considered thus far all can be estimated by maximum likelihood. Understanding the likelihood function permits an understanding of how (right) censoring is accounted for in event history analysis. Recall that observations remaining in the risk set at the time of the last observation are considered right-censored. A regression model implicitly makes no distinction between censored and uncensored observations (see Figure 1, cases three and four) and treating them as equivalent can lead to biased estimates of covariates and incorrect inferences.

With maximum likelihood estimation, we account for censoring through the likelihood function. Each observation, conceptually, can be thought of as a "success" or "failure." A failure would be any observation that *did* experience the event (or had a duration that terminated) within the observation period (that is, the event occurred and we observed it). A success is any surviving observation, i.e., an observation not experiencing an event within the observation period. Fortunately, we have already discussed two mathematical expressions that succinctly capture the notion of failure and success: $f(t)$ and $S(t)$. The likelihood for all uncensored observations (the "failures") is $f(t)$ for the distribution specified, while the likelihood for all censored observations (the "successes") can be expressed as $S(t)$ for the distribution. So, for example, the contribution to the likelihood of uncensored observations for the exponential distribution takes the following form:

$$f(t) = h \cdot \exp^{-h \cdot t} \qquad [27]$$

and the contribution to the likelihood of censored observations for the exponential distribution is the following:

$$S(t) = \exp^{-h \cdot t} \qquad [28]$$

Hence, using maximum likelihood estimation provides some leverage on the issue of censoring. Censored observations are treated differently from uncensored observations, as they should be. Furthermore, all information on the length of the duration is used with maximum likelihood estimation. Thus, we avoid the problems associated with traditional regression-based techniques.

The Issue of Heterogeneity

The parametric models discussed previously are based on an assumption of homogeneity of the distributions of the dependent variable across individuals.[16] Explanatory variables are included in the model to control for heterogeneity. Control may be incomplete, however, if some explanatory variables are inappropriately left out, the functional form is misspecified, or when unobservable variables are important; thus the problem of heterogeneity arises. Gourieroux, Monfort, and Trognon (1984) show that if the assumption of homogeneity is incorrect, the parameter estimates will be inconsistent and/or inferences will be based on inappropriate standard errors.

The most common problem resulting from unobserved variables is that the estimated hazard rate becomes biased toward negative duration dependence (Heckman and Singer 1984a). The intuition can be illustrated by considering the topic of challenger entry in House races with an incumbent. Suppose that hazard rates differ across incumbents, i.e., some incumbents are more prone to failure (a challenger entering) than others. Of course, incumbents with higher hazard rates tend to fail earlier, on average, than incumbents with lower hazard rates. So the average hazard rate of the surviving incumbents will decrease over time (as the primary date approaches) because those incumbents most failure-prone have been selectively removed. This is true even if the hazard rate is constant for all incumbents but the hazard rate varies across them.[17]

There are two main solutions to addressing the problem of heterogeneity. Note first that the dependent variable in event history analyses can always be expressed as a sum of its mean and an error term, i.e., the hazard rate equals $e^{(\beta X + \delta\mu)}$. The first of the two main solutions involves imposing a specific distribution on μ and is referred to as the random-effects procedure. In the example above, it would mean imposing a specific distribution such as normal, lognormal, or gamma, on failure-prone incumbents and a different distribution for those not "vulnerable."[18] In terms of estimation, one first derives the likelihood for the observed history on an individual, conditional on observed and unobserved variables. Then one uses the imposed distribution of the unobservable to compute the mean of the likelihood when the

[16]See Heckman and Singer (1984b) and Kiefer (1988) for a detailed discussion of heterogeneity, Hougaard (1991) for a discussion of the choices of distributions for the unobserved covariates, and Vaupel and Yashin (1985), for an illustration of how two subgroups or strata within a population can have dramatically different hazard rates, which can lead to a single misleading hazard rate for the entire data set if not properly accounted for.

[17]See Proschan (1963) for a formal proof.

[18]If heterogeneity is suspected to be a problem, one should first try to incorporate individual characteristics into the model, which should reduce the unobservable component of heterogeneity and thus the arguments for a decreasing hazard rate due to heterogeneity (Tuma, Hannan, and Groeneveld 1979). If the decreasing hazard rate is due to state dependence, including individual characteristics will not remove a decreasing hazard rate. State dependence refers to actual change in behavior over time, at an individual level.

unobserved are not taken into account. This procedure is repeated for all individuals in the sample and then maximized.[19]

The problem is that neither theory nor data provides much guidance for imposing a specific distribution. Heckman and Singer (1984b, 1985) criticize the imposition of a specific distribution and develop an estimator of the hazard rate that is nonparametric with respect to the distribution of the different types. It appears, however, that their estimator is sensitive to the parametric form of the hazard chosen for the general model and to the number and choice of explanatory variables.[20]

The second approach, the fixed-effects approach, treats μ as a fixed variable. The advantage of this approach is that few assumptions are imposed on μ. The disadvantages, however, are that the fixed-effects approach applies only to processes where the event is repeated over time, at least two transitions have been observed on some of the observations, and only the effects of covariates that change over time can be estimated (so effects of race or gender cannot be estimated).

Proportional Hazard Rate Models

The specification of the event history model has, to this point, dealt with the issue of parameterizing time-dependency. The discussion of the Weibull model shows that the hazard rate may take numerous monotonically increasing or decreasing shapes. Distributions like the log-logistic or Gompertz (among others) can elicit hazard functions that are both increasing and decreasing with respect to time. And of course, interpretation of covariates in any of these types of models hinges on the shape of the hazard function. The parameterization a researcher chooses for a particular process may have a substantial impact on the inferences one makes about the process. Consequently, estimating an event history model *without* having to specify or parameterize time-dependency would be useful.

The most commonly used model of this type is the Cox proportional hazards model. Because of its widespread usage in social science, we present an extended discussion of the Cox model.[21] This model allows one to estimate the effects of individual characteristics on survival time without having to assume a specific parametric form for the distribution of time until an event occurs. For an individual with a vector of characteristics, X, the proportional hazards model assumes a hazard rate of the form:

[19]See Lancaster (1979) for elaboration.

[20]See Trussel and Richards (1985), Kiefer (1988), and Hoem (1990).

[21]Political science examples include Bienen and van de Walle (1989, 1992) who investigate leadership duration and Box-Steffensmeier, Arnold, and Zorn (1997) who look at the timing of strategic position taking by members of the United States House of Representatives.

$$h(t \mid x) = h_0(t)\exp(\beta_k' X_i) \qquad [29]$$

where $h_0(t)$ is the baseline hazard function, which is estimated nonparametrically due to Cox's partial likelihood estimation; and where X_i represents the covariate values that may depend on time as well.

The term "proportional hazards" refers to the effect of any covariate having a proportional and constant effect that is invariant to when in the process the values of the covariate changes. That is, each individual's hazard function follows exactly the same pattern over time, but there is no restriction on what this pattern can be. To test the assumption that the effects of the covariates are proportional, one can specify a particular functional form of interaction effects between a covariate and time. The most straightforward example would be including the explanatory variable x_2, where $x_2 = x_1 t$ and where x_1 is already in the model. A test of the hypothesis that the coefficient on $x_2 = 0$ is a test of the assumption of proportional hazards. Another approach to testing the proportionality assumption involves using a set of time-varying dummy variables to contrast distinct time segments against the baseline segment (Yamaguchi 1991, 107). By interacting covariates with the dummy variables, we can see if proportionality holds over time. The significance testing can be assessed with chi-square tests.[22]

The model with time-varying covariates, often referred to as a Cox regression model, is an extension of the proportional hazards model developed by Cox (1972, 1975). Since the values of the variables, $x(t)$, vary over time t, so does the *relative* hazard, $h(t)/h_0(t)$. This means that the assumption of proportional hazards no longer holds (Collett 1994, 224). A simplified model with two explanatory variables, one constant and one varying over time, may be written:

$$h(t) = h_0(t)e^{\beta_1 x_1 + \beta_2 x_2(t)} \qquad [30]$$

In this model, the hazard at time t, $h(t)$, depends on the value of x_2 at the same time t.

Cox regression models do not have constant terms. Instead the constant is absorbed into the baseline hazard. Signs of the coefficients from a hazard rate model indicate whether some particular variable increases or decreases the hazard rate. The standard errors can be used to determine statistical significance, i.e., test statistics calculated by dividing a coefficient by its standard error, as in the more familiar ordinary least squares context. To

[22]See Yamaguchi (1991) and Box-Steffensmeier and Zorn (1996) for further discussion of these approaches and Gill and Schumacher (1987) for an alternative test based on a comparison of generalized rank estimators of the relative risk.

understand the magnitude of the effect, the percentage change in the risk of experiencing the event is useful. For a dichotomous independent variable, the percentage change in the risk of experiencing the event is:

$$100\left[e^{(\beta_k*1)} - e^{(\beta_k*0)}\right] / e^{(\beta_k*0)} \quad [31]$$

Negative coefficients produce values of $e^{(\beta_k*1)}$ that are less than one, and therefore produce negative percentage changes. The interpretation for a continuous independent variable is similar:

$$100\left[e^{(\beta_k*(x+\delta))} - e^{(\beta_k*x)}\right] / e^{(\beta_k*x)} \quad [32]$$

This gives the percentage change in the hazard rate for a δ unit change in the independent variable, x.[23]

Yamaguchi (1991, 102–3) outlines four disadvantages of the Cox method that users should be aware of and thus we repeat them here. First, the precision of the parameter estimates compared to those of maximum likelihood can be much less when the sample size is small because Cox's method only uses information about the relative order of durations. Thus, this method should not be used with small samples. Second, ties, which occur when more than a single observation exits at the same time, are problematic because only the order matters, not the exact numerical values of the failure times or of the censoring times. Computer programs employ a standard approximation to the exact partial likelihood for generating results since the exact partial likelihood is inestimable in the presence of ties. Prentice and Farewell (1986, 14) state that as a rule of thumb, if no more than 5% of the observations fail at one time the resulting bias is not a concern. If more than 5% of the observations are tied, Yamaguchi (1991, 103) suggests using maximum likelihood methods with discrete-time models. Third, if there is interest in the *form* of time-dependence, this is not the appropriate method. Fourth, Yamaguchi points out that users should realize the theoretical foundations for the maximum likelihood method are stronger than those of the Cox model. Specifically, "although the major asymptotic properties of parameter estimates are known [for the Cox model], caveats are necessary for the procedure of model selection" (Yamaguchi 1991, 103). Recently, Sueyoshi (1992) has developed an alternative semiparametric estimator for duration data that avoids the problems with ties and uses maximum likelihood estimation. Diermeier and Stevenson (1994) discuss and apply this method to political data.

[23]See Namboodiri and Suchindran (1987) and Teachman and Hayward (1993) for further elaboration.

The Issue of Model Construction

To help choose among the parametric models, it is useful to fit the Cox model and examine the shape of the baseline hazard function in addition to examining one's theory for guidance. If a specific shape is suggested, a parametric model may be suitable because the estimated parameters would be more precisely estimated (Collett 1994). Hazard plots will be useful here. Political science theories may not provide specific guidance about what particular distribution is needed to relate observable covariates to rates of events. Instead, scholars typically choose basic, tractable distributions that agree qualitatively with the substantive arguments (Hannan 1989, 357).

Recall that the exponential model is a special case of the Weibull with the scaling parameter, which indicates the shape of the failure time distribution, set to one. Therefore, a test of the exponential model versus the Weibull model is to fit a Weibull model and see if the value of the parameter is statistically distinguishable from one. Choosing between other parametric models is not as clear. The fit and coefficients for each of the models should be compared.[24]

In nonlinear models with explanatory variables, the chi-square test for the joint hypothesis that all coefficients besides the constant are zero gives an indication of how much the explanatory variables jointly contribute to the fit of the model. The chi-square statistic is analogous to the F-statistic in a linear regression model. A high chi-square value does not mean, however, that the model is not satisfactory. Residual plots and diagnostics are more helpful here.

Residual analysis, which is an informal method of checking specification, helps to assess a chosen specification, just as in the linear regression model. Residual plots show departures from a hypothesized model and may suggest ways to improve the specification. Another assessment tool for specification is to split the sample based on values of the explanatory variables and estimate the model separately between groups. If the specification is correct, the estimated parameters should agree up to the estimated error.

Checking for unobserved heterogeneity is also advisable. For example, since the vulnerability of incumbents (in terms of having a challenger enter) could differ due to variables not included in the model, such as scandals or the incumbent's position on his or her career path (see Jones 1994), accounting for unobserved heterogeneity could be important.

One advantage of parametric models is predicting what will happen beyond the "follow-up" period of the data. For example, in studies of when states enact a particular law, it is useful to be able to predict which states

[24]Hannan (1989) provides an excellent discussion of choosing among distributions and their fit. Chung, Schmidt, and Witte (1991) is also useful.

may do so beyond the time frame of the observation period. Another advantage of the parametric models is that it smooths the data; the disadvantage is that it can be the wrong model.

Among the models that use time-varying covariates, the Cox regression is usually the preferred model. Collett (1994) points out that discrimination between a Cox and Weibull model is difficult. If the standard errors for the Weibull are substantially smaller than those for the Cox model, the Weibull model would be preferred because of efficiency. If the standard errors are similar, the Cox model is preferred because of its less restrictive assumptions.

Complicated Events in Event History Analysis

Sometimes, political or social processes involve complicated event structures. In this section, we consider the special case where an event may be repeatable as well as the issue of multiple events, or competing risks.

Repeated Events

Up to this point we have discussed only single-state nonrepeatable events. That is, there is a single-state that can be occupied only once, i.e., entry into the first elected political office held (provided one does not distinguish between different offices such as state representative, mayor, etc.). We now turn to repeatable event processes, which means an individual can occupy a state several times.[25] Political office holding histories fall within this class of processes. Researchers would focus on the amount of time an individual spent in each political office (not just the first political office held and still not distinguishing offices). That is, consider an individual who had held *m* political positions with durations $t_1, t_2, \ldots, t_j, \ldots, t_m$, where the last duration may be censored.

There are two procedures to specify the hazard rate of leaving political position *j*. First, one assumes that the shape of the rate and parameters of the rate are the same for all positions. Dependence on previous history may be captured with explanatory variables. All the jobs of each individual can be pooled, and the parameters from the data on all of the jobs can be estimated. Second, one assumes that the rate and its parameters differ from position to position, or at least between subsets of positions, such as early political positions and late political positions. So if we think that the hazard rate and its parameters vary between jobs, depending on, for example, position number (first, second, and so on), then we need to estimate the parameters separately for each position number. The parameters for position *j* are

[25]See Hannan (1989) for a more detailed discussion and Olzak (1987, 1989) for applications of repeatable events.

estimated from the durations in position j.[26] To do either estimation, the data are arranged such that there is one record of data for each position an individual held.[27]

Competing Risks Models

In competing risks problems, an individual may fail because of two or more reasons, e.g., leaving the Senate by retirement, defeat, or death in office (Jones 1994); government transition due to a call for new elections, vote of no confidence, coup, etc. (see Alt and King 1994; King et al. 1990; and citations therein for applications concerning government transitions). Most of the time in a competing risks framework, the occurrence of an event means that there is no longer a risk of experiencing another event(s), assuming that the causes are mutually exclusive.[28] Chung, Schmidt, and Witte (1991) provide a concise, intuitive statement about what competing risks models are: "The competing risks model is designed to provide estimates of the effects of explanatory variables on the cause-specific hazard functions, so that the effect of a variable on the timing of one type of failure can be distinguished from its effect on the timing of other types of failures" (1991, 90).

Chung, Schmidt, and Witte (1991) go on to state that the important feature of competing risks models is that the likelihood function factors into a separate component for each type of failure (1991, 91). The particular parameterization of the hazard can be any of those discussed above, e.g., Cox model, log-normal, or Weibull. When assuming that events are mutually exclusive, the estimation is straightforward since every other type of event is treated as censoring.[29] Namboodiri (1996) states this more clearly, ". . . a

[26]See Petersen (1991) for further elaboration and Blossfeld and Hamerle (1989) for an example.

[27]No assumptions about independence between positions on the same person are made.

[28]Namboodiri (1996) points out the following contrasts. First, an individual may continue to be at risk of experiencing another event, e.g., the risk of a challenger from Party B entering continues even after a challenger from Party A enters in a multi-party system. Second, an event may occur and the individual may no longer be observed even though the individual may be at risk for another competing risk. Namboodiri (1996) provides an example of migration to illustrate this situation. If one is studying internal migration within a Latin American country and the individual leaves the country (international migration), the individual is still at risk of internal migration within the new country he or she migrated to, but the researcher is unlikely to be able to study this individual further. Third, an individual may experience one event but this may not affect the risk of any other event or affect whether the individual will continue to be observed. Finally, if an individual experiences one event, it may affect the hazard rate of experiencing other events, e.g., a small incumbent victory margin in the past election may enhance the hazard rate of defeat.

[29]See Han and Hausman (1990) for their development of a flexible parametric proportional competing risks model that permits unrestricted correlation among the risks, Sueyoshi (1992) for extensions to include time-varying covariates, and Diermeier and Stevenson (1994) for an application in the area of cabinet duration. Work by Hill, Axinn, and Thornton (1993) examines the assumption

crucial question is whether a different set of regression parameters is appropriate for each cause of failure. If it is, each cause of failure can be analyzed separately, by treating all failures other than the focal one as being censored times. If the occurrence of one of the competing events affects the hazard rate of experiencing another event, then the former should be entered as a time-varying covariate in the model for the latter."

Interpreting Event History Models

Parameters in event history models provide considerable information about the risk that units incur as they progress through time. As we have already discussed, estimation of baseline hazard rates yields some notion of how a social process is or is not time-dependent. Furthermore, the baseline hazard gives us some idea of when in time observations are most at risk. For many applications, however, time-dependency may not be of primary concern. Often we are most interested in how external factors, or covariates, amplify or dampen the risk of experiencing an event. To use a substantive example, we may be less interested in the time-dependency of the duration of overt military interventions than we are in learning how factors like dependency or intervenor power accelerates or decelerates the hazard rate. To that end, researchers need to take care in substantively interpreting event history parameters for covariates.

Typically, we think of two kinds of covariates: time-varying and time-invariant. Examples of the latter generally involve demographic factors such as race, gender, or geographic region. However, *any* covariate can be thought of as time-invariant if its value does not change over the course of the duration. Consequently, the partisanship of a Congress member, dependency status of a state, or other political variables that can change, but generally do not change, are often treated as time-invariant factors. Parameter estimates for such factors represent how the relative risk of experiencing an event increases or decreases for a unit that possesses the attribute the covariate is measuring.

This enhancement or diminution of the hazard rate is generally treated as relative and proportional to the baseline hazard rate. The proportionality interpretation means that the hazard rate's increase (or decrease) due to the value of some covariate is proportional to the baseline hazard by the value of

that alternative risks are stochastically independent, which rules out individual-specific unmeasured risk factors that are shared by two or more alternatives (referred to as SURF). They then develop and apply a generalized standard discrete-time competing risks model that allows for the types of stochastic dependencies resulting from SURF. Their Monte Carlo work suggests that the biases introduced by violations of the temporal independence assumption primarily affect time-varying covariates (Hill, Axinn, and Thornton 1993, 245–6). See also Crowder (1991), Bagai and Rao (1992), and Kuk (1992) on the well-known problem of identifiability in competing risks.

the covariate. This proportionality is treated as constant across time unless one hypothesizes (and of course models and tests) interactions of the covariate with the baseline hazard function, which was described previously. To illustrate this point with a pedagogical example, suppose we estimated an event history model using logistic regression (see Equation 20) and obtained a parameter of −1.5. Because the metric of the logit estimator is in terms of log-odds, the interpretation of this number would be that the log-odds of experiencing an event at t would be proportionately lowered by $-1.5*X$. When $X = 0$, the baseline hazard rate would be obtained (assuming only one covariate was included in the model). When $X = 1$, the log-odds of the baseline hazard would decrease by 1.5. Furthermore, this proportional decrease would be constant across time. That is to say, at any t, the log-odds of the hazard would be lowered by −1.5.

The use of time-invariant covariates can provide considerable insights on social processes. The parameters of such covariates generate novel information by giving us an estimate of how risk varies by demographic, political, or social characteristics. The researcher must take care in interpreting such covariates. Since the data reflect an underlying longitudinal process, parameter estimates for covariates reflect how risk increases or decreases across time for some units relative to other units. The inclusion and interpretation of time-invariant covariates provide no special analytical problems for event history analysis. Inclusion of time-varying parameters, however, leads to more complex modeling and interpretation.

As the name suggests, a covariate is time-varying if its value can change across time. For example, campaign spending in congressional elections is not time-invariant. Other examples might include public opinion measures, fatalities in military conflict, or economic indicators. Interpretation of time-varying covariates involves considerable care. As for time-invariant covariates, time-varying covariates reflect how the hazard rate changes as the values of the covariates are realized; however, unlike time-invariant parameters, the effects of time-varying parameters are substantially influenced by when in time the value of a covariate changes (Teachman and Hayward 1993). To understand this, note that the hazard rate is discontinuous at the point at which the value of the covariate changes (Teachman and Hayward 1993). That is, after experiencing a change in the value of some covariate, the hazard rate increases or decreases proportional to the parameter value of the covariate. This process of change is also called a jump process (Petersen 1991). The hazard "jumps" by some amount at the point the covariate changes value. So, for example, if by becoming a chair of a House committee, the likelihood diminishes for a House member to leave the House to seek higher office, then the hazard of political ambition drops by some amount. Once the hazard jumps to a new level, however, that rate *remains* *proportional* across time until the unit experiences another change in the value of the covariate. This is a major distinction between varying and invariant predictor variables. Time-invariant parameters reflect change that is constant and proportional across all time periods. The change exhibited in the hazard rate for time-varying parameters is only "in effect" for the time periods in which the given level of the covariate is realized.

Obviously, inclusion of time-varying predictors can generate novel conclusions about longitudinal processes by providing estimates of how risk moves as conditions in the process change. Nevertheless, researchers using time-varying parameters must understand that interpretation is conditional on the value of the covariate at any time point, and further, that the hazard rate may be highly variable, particularly if conditions rapidly change over time. Another obvious implication of the inclusion of time-varying parameters involves the structure of the event history data set. If time-varying predictors are used, each unit in the sample will necessarily have multiple records in the data file. Traditionally, this involves collecting observations on each unit in time and measuring and recording the value of the covariate at each time point. Consequently, the dependent variable becomes a series of zeroes, denoting the nonoccurrence of an event for a unit, culminating in a one when the duration is terminated. As we noted earlier, this form of the dependent variable is what Petersen (1991) calls the event history approach. We now turn to some applications of event history models, which more fully explore the basic ideas of interpretation discussed in this section.[30]

Applications of Event History Models

In this section, we present three distinct applications of event history methodology. The first application involves the duration of overt military interventions using a Weibull model with one time-invariant covariate. In the second application, a Cox regression model of challenger deterrence in House elections is presented. The final application involves estimation of a discrete-time, competing risks event history model of congressional careers.

Duration of Overt Military Interventions: A Weibull Model

At issue in this application is modeling the hazard rate of overt military interventions (OMIs). An OMI is military intervention by one state into another independent state, dependent country, or region (see Tillema 1991, for a fuller discussion of OMIs). An OMI may last a single day or last several thousand days (such as Vietnam). The question we address is, do OMIs exhibit time-dependency, and if so, what is the form of that dependency? The

[30]See Teachman and Hayward (1993) for further discussion of interpreting event history models.

data set used in this application involves all OMIs from 1945 to 1991.[31] The dependent variable is the duration of the OMI measured in days. A single, time-invariant predictor variable is included in the model. The variable, denoted as "dependency status," is a binary variable coded 1 if the occupied target is dependent upon another state and 0 if not. A Weibull model was selected for this application. Recall from our earlier discussion that a Weibull model is a flexible method for modeling the shape of the hazard function. Additionally, we can include covariates and estimate a Weibull regression model. The parameter estimates and standard errors are presented in Table 1.

Two columns of parameter estimates are presented for the Weibull model. Typically, Weibull estimates are presented in terms of one of two parameterizations. The first type of parameterization is known as the "accelerated failure time" parameterization. This parameterization is shown in column 2. The second type of parameterization is known as the "relative hazard" parameterization. This parameterization is presented in column 3. The relative hazard parameterization simply reflects a mathematical transformation of the accelerated failure time parameters. To see this, let β correspond to the accelerated failure time parameters, β^* correspond to the relative hazard parameterization, and σ correspond to the inverse of α, the scaling parameter of the Weibull distribution (see Equation 24).[32] The β^* parameters are obtained by dividing the β parameters by $-\alpha$, that is $-(\beta/\sigma)$.

Because the two parameterizations differ, the interpretation of a Weibull model is contingent on which parameterization is employed. Roughly speaking, the accelerated failure time parameterization presents coefficients in terms of their relationship to expected failure times. A negative sign on a coefficient using the accelerated failure time parameterization implies that the duration is "shortened" by some value per unit change in the covariate. That is, the expected time-to-failure is sooner rather than later. Consequently, it is important to note that a negative coefficient implies an *increase* in the hazard rate, while a positively signed coefficient implies a decrease in the hazard. In terms of the OMI data, we see from column 2 in Table 1 that the coefficient for dependency status is .39. This positive and significant coefficient simply suggests that interventions into targets that have dependency status tend to last

[31] Data on OMIs comes from Tillema (1991, 1994).

[32] Note that the σ parameter and α parameter are simply the inverse of one another. That is $\sigma = \alpha^{-1}$ and $\alpha = \sigma^{-1}$. Consequently, when $\alpha > 1$ then $\sigma < 1$. Both results indicate a monotonically *increasing* hazard. When $\alpha < 1$ then $\sigma > 1$, thus implying a monotonically *decreasing* hazard. Obviously, when $\alpha = 1$ then $\sigma = 1$ and the exponential model is obtained. Statistical programs like SAS' Proc LIFEREG report the scaling parameter in terms of σ and the default coefficients are expressed in terms of the accelerated failure time parameterization. STATA allows computation of either the accelerated failure time parameterization or the relative hazard parameterization; however, STATA presents the scaling parameter in terms of σ.

Table 1. A Weibull Regression Model of OMI Duration

Dependent Variable: Duration of OMI (measured in days)

Parameter	Accelerated Failure-Time Estimates	Relative Hazard Estimates	Exponential Model Estimates
Duration Dependence	$\sigma = 2.37$ (.07)‡	$\alpha = .42$ (.07)	
Constant (β_0)	5.37 (.10)	–2.27 (.04)	6.33 (.04)
Dependency Status (β_1)	.39 (.23)	–.17 (.10)	.40 (.10)
Log-likelihood:	–1675	–1675	–2277
N = 690			

‡Note that $\alpha = .42$ so $\sigma = .42^{-1}$.
Estimates are maximum likelihood estimates of a Weibull and exponential duration model. Standard errors are in parentheses. Data are from Tillema (1991, 1994). See text for interpretation.

longer (i.e., the hazard is decreased) than interventions into nondependent targets. Or similarly, OMIs into nondependent targets "fail" sooner than OMIs into dependent targets. If we exponentiate the parameter estimate ($e^{.39}$), a natural interpretation of the coefficient is found: OMIs into dependent targets "survive" about 48% longer than OMIs into nondependent targets.

If one presents the coefficient estimates in terms of the relative hazard parameterization, the interpretation of Weibull coefficients differ. Here the values of the coefficients are expressed in terms of the *baseline* parametric hazard and do not reflect expected failure times (as in the case of the accelerated failure time parameterization). Consequently, a negatively signed coefficient reflects how "much longer" the duration will last relative to the baseline hazard. So in the case of the relative hazard parameterization, a negatively signed coefficient implies a decrease in the hazard while a positively signed coefficient implies an increase in the hazard. Turning again to Table 1 and looking at the third column of estimates, we find that the coefficient for dependency status is –.17. This negatively signed coefficient suggests that the hazard rate for OMIs into dependent targets is decreased relative to the baseline hazard of OMIs into independent targets. Or to put it another way, the failure rate of OMIs into dependent targets is 85% of the failure rate of OMIs into nondependent targets.

We now focus on interpreting the scaling parameter. Recall that the scaling parameter estimated in the Weibull models provides us with information about the shape of the hazard function. If $\alpha > 1$, the hazard is monotonically rising, if $\alpha < 1$, the hazard is monotonically decreasing, and if $\alpha = 1$, the hazard is flat, thus implying an exponential distribution. The coefficient for the scaling parameter shown in Table 1 is .42, indicating the hazard rate of OMIs is a decreasing function of time. The null hypothesis of interest in this instance is that $\alpha = 1$. A test of this hypothesis can be conducted by using the following formula (Blossfeld, Hamerle, and Mayer 1989, 240):

$$z = \frac{\alpha - 1}{s(\alpha)} \qquad [33]$$

We see that applying [33] yields a $z = -8.29$. This value is significant beyond the .05 level and we can safely reject the null hypothesis. However, of greater interest is the substantive interpretation of α. One issue we can address is how the "risk" of terminating an OMI changes with respect to time. A convenient way to do this is to assess how the magnitude of the hazard rate changes for different points in time. The logic of this analysis is to evaluate the hazard rate at one point in time and compare it to the hazard rate at some later point. So, for example, suppose we wanted to compare the baseline hazard for an OMI that has lasted 100 days with an OMI that has lasted one thousand days. The following equation establishes how we evaluate the hazard rate at a specific duration (denoted as d):

$$h(t = d) = \exp^{\frac{\beta_0}{\alpha}} \cdot \alpha \cdot (\exp^{\frac{\beta_0}{\alpha}} \cdot d)^{\alpha - 1} \qquad [34]$$

In this equation, d can be set to any duration time. To determine the percentage increase or decrease in the hazard rate across two points in time, the following equation can be used:

$$\frac{h(t = d_2) - h(t = d_1)}{h(t = d_1)} \cdot 100\% \qquad [35]$$

In our application, we see the value of $\alpha = .42$ and the coefficient for $\beta_0 = -2.27$ (using the relative hazard parameterization estimates from Table 1, column 3). To compare the hazard rate for OMIs with a duration of 100 days to OMIs with a duration of one thousand days, we first substitute the relevant numbers into [34] to derive the individual hazard rates. For durations of 100 days, we find that [$h(t = 100) = .0045 \cdot .42 \cdot (.0045 \cdot 100)^{-.58} = .003$]. For durations of one thousand days, we find that [$h(t = 1000) = .0045 \cdot .42 \cdot (.0045 \cdot 1000)^{-.58} = .0008$]. Substituting these figures into [35], we obtain [$(.0008 - .003)/.003 \cdot 100\% = -73\%$]. Substantively this means that for a nation or state that has been involved in an OMI for one thousand days, it is 73% less likely to terminate the intervention than a nation or state involved in an OMI for 100 days.

In Table 2 under the column "Percent Change in Baseline Risk," we examine how the risk of ending an OMI diminishes as time passes. Once an

OMI achieves a duration of one year, a nation is about 30% less likely to terminate the intervention than it would with a duration of six months. Comparing OMIs of one versus three years, we see the risk of termination diminishes by about 47%. Substantively, this is a potentially interesting conclusion. This analysis indicates that once a country becomes enmeshed in a military intervention, terminating the intervention progressively becomes more and more difficult. From a foreign policy perspective, the debate on whether to intervene in another country or not is obviously a politically explosive matter. Taking the United States as an example, in recent years, we have observed ambivalence toward military interventions in Somalia, Haiti, and Bosnia on the grounds that once troops are sent, extraction of them becomes difficult. Congressional leaders have expressed fear that a sort of "Vietnam syndrome" will inevitably emerge. These results, admittedly presented here for pedagogical purposes, provide some initial estimates of the "Vietnam syndrome."[33] Once nations become extricated in another nation or state, the difficulty of terminating the intervention increases over time. We should not be surprised that when the president or congressional leaders assure the public that an intervention will only last six months or a year, it is common for the intervention to extend far beyond the initial ending date.

However, what time-dependency actually means is not so clear-cut from a statistical point-of-view. As Beck (N.d.) notes, time-dependency or duration-dependence frequently indicates an inadequately specified model. That is, if we could account for the covariates that are related to increasing or decreasing duration, then in principle, duration-dependence would "disappear." In this sense, duration-dependence is really a nuisance and the research question requires further investigation (Beck N.d.). So should we forgo the Weibull model in favor of, say, an exponential model where the parameterization of duration-dependence yields no dependency? The answer is simply "no." Since the exponential model is a special case of the Weibull, the test of the null hypothesis that $\alpha = 1$ is analogous to testing whether or not the exponential model is the "appropriate" model. Clearly, in this simple, pedagogical example, we prefer the specification of the Weibull distribution over the exponential. In Table 1 in the fourth column, we presented the results of the OMI model using the exponential model. Clearly, the coefficient estimates of the constant term and of the dependency status variable are roughly similar to the accelerated failure time parameterization. We find, for example, that OMIs into dependent targets tend to last about 49% longer than OMIs into nondependent targets ($e^{.40}$). This inference is

[33]Clearly, there are concerns aside from troop extraction that lead to decisions to intervene militarily. We are only noting that the concerns about extracting troops is a legitimate and empirically demonstrable concern.

Table 2. The Effect of Time on the Termination of OMIs

Duration	%Δ Baseline Risk
6 Months vs. 1 Year	–33
1 Year vs. 2 Years	–33
1 Year vs. 3 Years	–47
1 Year vs. 4 Years	–55
1 year vs. 5 Years	–61

The second column reflects the percentage change in the hazard rate of OMI termination for one duration time vs. another duration time (i.e., OMIs lasting one year are 33% less likely to terminate than OMIs lasting six months). These percentages were calculated using the relative hazard parameters shown in Table 1, column 2. Data are from Tillema (1991, 1994).

quite similar to that made from the Weibull; nevertheless, if we were to choose which specification is appropriate, we would select the Weibull because the scaling parameter significantly departs from 1.

Challenger Deterrence: A Cox Regression Model[34]

The key to using parametric models is to have a sound idea of the distributional nature of the hazard function. In the previous pedagogical example, we made the assumption that shape of the hazard function is a monotonic function of time, and hence, the Weibull model was a good estimator. Had we specified another distribution that could elicit a nonmonotonic hazard function, then the parameter estimates may have been widely different. Researchers often cannot specify the distributional parameterization of time-dependency and thus use non or semiparametric methods of estimation, such as the Cox model.

The Cox model with time varying covariates is applied to the problem of challenger entry. The issue is whether war chests (campaign finance reserves) deter challenger entry. Time-varying covariates are critical because they allow the temporal dynamics of war chests to be incorporated into the model.

Table 3 shows the estimation results of a Cox's regression model with time-varying covariates in which an event is defined as the entry of a high quality challenger. The overall fit of the model is good; we may reject the null hypothesis that the coefficients are jointly zero at the 0.001 level. The signs of the coefficients tell us that incumbents from the South are more

[34]This is taken directly from Box-Steffensmeier (1996).

Table 3. Factors Influencing the Timing of Entry by High Quality Challengers

Variable	Coefficient	P-value	Percent Change in the Hazard Rate
South	–0.44 (0.42)	0.30	–35.5
Party	0.23 (0.32)	0.47	26.4
Prior Vote	–6.97 (1.66)	0.00	-6.7
War Chest	–3.01 (1.39)	0.03	–26.0

Note: Units for the war chest variable in the third column is $100,000 and for prior vote, 1%.
Log-likelihood = –197.39
Chi-Square(4) = 34.11
($p < 0.001$)
N = 397

likely to enjoy a race without a high quality challenger for a longer time period. Republican incumbents are also more likely to not face a high quality challenger early on. Finally, having a large war chest and prior vote margin helps deter high quality challengers. Only the war chest and prior vote variables are statistically significant, thus the subsequent discussion focuses on the interpretation of these two variables.

The fourth column of Table 3 shows the percent change in the hazard rate. The results tell us that for a 1% increase in the prior vote, the hazard rate decreases by almost 7%. That is, the percentage change in the hazard of challenger entry at any time t for two incumbents whose war chests differ by 1% in prior vote and who have the same values for the other independent variables is approximately 7%. A one standard deviation increase in the prior vote, which is 14.2%, decreases the hazard rate by 62.9%. A 5 and 10% increase results in a 29.4 and 50.2% decrease, respectively.

Each $100,000 increase in an incumbent's war chest decreases the hazard of a high quality challenger entering by 26%. That is, the percentage change in the hazard of challenger entry at any time t for two incumbents whose war chests differ by $100,000 and who have the same values for the other independent variables is 26%.[35] A one standard deviation increase in

[35]The effect of a time-varying covariate, $x_k(t)$, on the overall survivor function depends on when the change with respect to the status of x_k occurs. The intuition is straightforward; the earlier the increase in the incumbent's war chest occurs, the larger the effect is on the overall probability of not having a high quality challenger enter.

an incumbent's war chest, which is \$239,000, results in a 51.3% decrease in the hazard rate. The effect of the increase in the war chest is nonlinear. For example, if the increase is \$100,000, the hazard rate decreases 26%; if the increase is \$200,000, the hazard rate decreases by 45%. So this \$100,000 differential increase (\$200,000 – \$100,000) results in a decrease of 19% (–45.0% – (–26.0%)). In contrast, a \$100,000 differential increase between \$900,000 and \$1,000,000 results in a decrease of only 1.8% (–95.1% – (–93.3%)). Thus, there are diminishing returns.

Incorporating the dynamics of war chests through the use of event history methods allowed us to measure empirically a previously elusive effect. It is important to look at the effect of incumbents' war chests on the entry of high quality challengers because the size of an incumbent's war chest is controllable to the extent that additional effort results in a larger war chest and because incumbents can control the size of their war chests to a much greater degree than the other covariates in the model.

Congressional Careers: A Competing Risks Model

In this section, a discrete-time, competing risks model is applied to the problem of Congressional careers: how, when, and why do incumbent members of the House of Representatives exit office? Full or near-full career path data were collected on every member of the House from each freshman class elected from 1950 to 1976.[36] Each incumbent in the data set was tracked from his or her first reelection bid until his or her last term served in office. A member initially elected in 1950, then, does not enter the risk set until the election cycle of 1952. At each subsequent election, the incumbents are observed and either experience a terminating event or are reelected. Once a terminating event is experienced, the incumbent exits the risk set and is no longer observed. All election cycles from 1952 up to and including 1992 are covered in this data set. The last freshman class on which data were collected was 1976. This decision was made in order to avoid severe right-censoring problems.

Two competing risks models are estimated in this application. The first models the risk of voluntary career termination in the United States House. The events in this competing risks model are the decision to retire and the decision to seek alternative office (denoted as "ambition" in the tables). Both of these outcomes are compared to the baseline category of running for reelection. Hence, the dependent variable in this model is trichotomous: 0 = seek reelection, 1 = retire, and 2 = higher ambition. The second competing risks model investigates electoral success. Given an incumbent has decided

[36] The core data set used to construct the events history of incumbent members is the *Roster of U.S. Congressional Officeholders*. Additional data were added to this core data set (see Jones 1994 for a fuller explanation).

to run for reelection, was he or she successful? The outcome variable in this model is also trichotomous, where 0 = winning the election, 1 = losing in the general election, and 2 = losing in the primary election. Traditionally, scholars have ignored House primary elections because not much occurs in these types of races. That is, incumbents virtually always win. As we will see, however, with the event history parameters, we obtain estimates of the risk certain profiles of incumbents incur in House primaries. Obviously, the electoral outcome model is "nested" within the first model: in order to win or lose an election, you first have to enter the race. Thus, the two processes were modeled separately.[37]

Several covariates are used as predictors of career termination. Because the purpose of this application is to illustrate a variant of the competing risks models, we leave it to Jones (1994) for a fuller explanation of the covariates. The following covariates are included in the competing risks models:

Duration: the duration of time (measured as terms served) the incumbent has spent in Congress prior to the election cycle.

Incumbent Age: the incumbent's age (in years) at each election cycle.

Southern Democrat Status: two binary covariates are used to indicate southern Democrats. The first, denoted as *S. Dem. Before 1970* is coded 1 if an incumbent was a southern Democrat prior to 1970, and 0 if the incumbent was not. The second, denoted as *S. Dem. After 1970* is coded 1 if the incumbent was a southern Democrat after 1970, and 0 if not. The inclusion of these covariates is used to model the changing nature of southern politics during and after the Solid South Period.[38]

Post 1966 Cohort: a binary indicator denoted as 1 if the member was elected in the freshman class of 1966 or after, and 0 if the incumbent was elected prior to 1966. The aim of this variable is to capture the effects of the increased perquisites of office incumbents allocated to themselves beginning in the mid-Sixties (Jacobson 1992).[39]

Prior Vote Margin: the incumbent's margin of victory in his or her previous election.

Redistricting: this is a binary variable coded 1 if the incumbent's district was substantially redistricted, and 0 if not.

[37]We do not formally estimate a nested model in this application (in this sense of a nested logit). The estimators used here are two separate multinomial logits.

[38]Obviously the Solid South did not discretely cease to exist in 1970. Rather, the change was gradual through the early 1970s and even into the 1980s. This measure is clearly a crude, but as we shall see, effective way to capture these changes.

[39]Picking the 1966 class as the "beginning" of the mid-Sixties boost in office perquisites is a bit arbitrary. After all, incumbents serving prior to 1966 surely reaped some of the same benefits as the post-1966 incumbents. What this variable will allow us to assess, then, is how, beginning with the 1966 freshman, how the relative risks of career termination increased or decreased compared to earlier freshman classes.

Scandal: a dummy variable coded 1 if an incumbent was involved in an ethical or sexual misconduct scandal or when the incumbent was under criminal investigation, and 0 otherwise.[40]

Open Gubernatorial Seat/Open Senatorial Seat: coded 1 if there is an open gubernatorial and/or open Senatorial seat available in the incumbent's state, and 0 if not.

Constituency Size: this is the reciprocal of the number of Congressional districts in the state. Hence, it measures the proportion of the state the incumbent's district encompasses.[41]

Reform Era: coded 1 for the election cycles of 1968, 1970, and 1972, and 0 otherwise. Hibbing (1982a, 1982b) has found that the years of House reform were substantially related to members' explanations for voluntarily exiting Congress.

Prestige Position: coded 1 if a member is in the House leadership and/or is a chair of a standing House committee, and 0 otherwise.

Republican Status: coded as 1 if the incumbent is a Republican and 0 otherwise.

Same Party as President: coded 1 if the incumbent's party affiliation is the same as the President's.

Presidential Approval: this variable ranges (theoretically) from –1 to 1 and reflects the Gallup Presidential Approval rating in the month prior to the election.[42]

*Approval*Inc. Party*: this is an interaction term between presidential approval and whether or not the incumbent is of the party of the president.

Change in RDI: this variable measures the change in real disposable income from the last quarter of the year preceding the November election to the first quarter of the election year.

*Change in RDI*Inc. Party*: this is an interaction term between RDI change and whether or not the incumbent is of the president's party.

Watergate and *the Golden Parachute*: the *Watergate* variable is coded 1 for Republicans in the 1974 election cycle, and 0 for all other cases. *Golden Parachute* is coded 1 for the 1992 election cycle (the last year incumbents could keep war chests as personal cash), and 0 for all other election cycles.

In addition to these covariates, four interactions with the duration covariate were included. The basis for these interactions is to assess timing

[40]Data for this variable primarily comes from *Congressional Quarterly* (1992) and various editions of the *New York Times*.

[41]The rationale for this variable as well as the open gubernatorial and open Senate seat variables comes from Rhode (1979) and Brace (1984, 1985) who find them positively related to political ambition.

[42]A 0 denotes a 50% approval rating, a 1 (would) denote 100% approval, and –1 (would) denote 0% approval. Within these bounds, approval obviously varies.

of change. In particular, we are asking whether or not it makes any difference *when* in an incumbent's career he or she is redistricted, when (or if) he or she is involved in a scandal, or when the opportunity for higher office occurs. We can address these issues systematically through interactions of these covariates with the duration measure.

To estimate the parameters of the competing risks model, a multinomial logit estimator (MNL) is used.[43] The results of the two models are shown in Table 4. The second and third columns of coefficients in Table 4 correspond to the voluntary termination model, and the last two columns refer to the electoral termination model. Because the coefficients are MNL estimates, they may be interpreted as reflecting by how much the log-hazard of career termination increases or decreases. With respect to duration (measured as the number of terms in office), the log-hazard of retiring increases by .06 with each term served. What ".06" means substantively is not obvious. A convenient way of interpreting logistic regression coefficients is through the use of odds ratios.[44] The odds ratio of the coefficient is the antilog of the coefficient, i.e., $e^{\beta k}$. Thus, a coefficient of .06 elicits an odds ratio of 1.06. Odds ratios have a convenient interpretation. Ratios equal to 1 indicate that as a covariate's value changes, the marginal increase or decrease in the hazard is 0.[45] Ratios greater than 1 imply that the hazard is increasing as the value of the covariate increases by a unit. In contrast, ratios less than 1 imply the hazard is decreasing as the value of the covariate increases by a unit.

If one wishes to assess the percentage change in the hazard of career termination, this calculation is straightforward. Suppose a covariate changes by some unit Δ. The odds of this change is then $e^{\beta\Delta}$ (Farole, Levine, and Morgan 1995). To convert this factor change into a percentage change, the following formula can be used:

$$\%\Delta = (e^{\beta_k \Delta} - 1) * 100 \qquad [36]$$

[43]The MNL estimator works under the assumption of independence of irrelevant alternatives (IIA). Researchers employing the multinomial logit estimator for event-history data (or any type of data for that matter) should not naturally assume the IIA property holds. Hausman and McFadden (1984), Small and Hsiao (1985), and others have developed specification tests for the IIA property. Zhang and Hoffman (1993) provide a concise overview of these tests. If the IIA property does not hold, the solutions are less accessible than the tests. Alvarez and Nagler (1995) discuss the multinomial probit (MNP) estimator. The MNP does not assume IIA (see also Hill, Axinn, and Thornton 1993). For these data, the Small-Hsiao specification test was run and the IIA assumption is reasonable. Consult Small and Hsiao (1985) for details of this estimation.

[44]See Farole, Levine, and Morgan (1995) for a nice discussion of odds ratio analysis.

[45]An odds ratio of 1 implies a 1:1 bet. You pay \$1 to make \$1 dollar. The overall impact on your income is 0.

Table 4. A Competing Risks Model of Congressional Careers

	Voluntary Termination		Electoral Termination	
	Retire vs. Running	Ambition vs. Running	Losing (Primary) vs. Winning	Losing (Gen.) vs. Winning
Duration	.06**	.10**	–.09***	.03
Incumbent Age	.08**	–.05***	.04***	.04***
Southern Dem. Status				
S. Dem. Before 1970	1.26***	–1.54**	–.62*	1.24***
S. Dem. After 1970	.51**	–.53*	–.19	.53
Personal Vulnerability				
Post 1966 Cohort	–.03	.41**	–.45***	–.05
Prior Vote Margin	–.01***	.007**	–.07***	–.01***
Redistricting	–1.78	1.66**	2.28***	.16
Timing of Redist.	.34**	–.09	–.004	.21
Scandal	1.33*	—	3.12***	4.42***
Timing of Scandal	–.10	—	.01	–.17*
Opportunity Structure				
Open Gov. Seat	–.19	.94***		
Timing of Gov.Seat.	.03	–.11**		
Open Sen. Seat	.27	.70**		
Timing of Sen. Seat	–.02	.08		
Constituency Size	.44	2.41***	.85**	1.48**
Institutional Desirability				
Reform Era	.73**	.23		
Prestige Position	–.33	–1.50*	–.54	—
Republican Status	.27*	.10	–.81**	.39
National Conditions				
Same Party as Pres.	.36**	.08	1.51***	.12
Pres. Approval	–1.40	.54	4.56***	2.58
Approval*Inc. Party	–.36	.26	–9.3***	–3.78
%ΔRDI 1 Quarter	.04	.06	.60**	–.44**
%ΔRDI 1*Inc. Party	.01	–.003	–1.63***	.10
Watergate Election	.27	–.76	–.37	–.89
Golden Parachute	.80**	–.23		
Constant	–8.12**	–2.11**	–3.67**	–6.55**
	Log Likelihood = –1689 Chi-Squared(50) = 494.63 $p < .0001$ N = 5508		Log Likelihood = –1207.90 Chi-Squared(36) = 638.94 $p < .0001$ N = 5036	

***p < .01, **p < .05, *p < .10.
Numbers are multinomial logit coefficients.

We interpret the results of the competing risks model in terms of percentage increases or decreases in the hazard. Table 5 presents an analysis of the change in hazards for selected covariate profiles. Given the number of parameters estimated in the two models and because of space considerations, we only consider a few covariates.

With respect to duration, we find that as each term passes, the risk associated with retirement increases by about 6%. In contrast, the risk associated with losing a general election decreases by about 8%. Thus, seniority in Congress has a mild impact on termination. The longer an individual stays in Congress, the less likely he or she is to exit office electorally and the more likely he or she is to leave office on his or her own terms. Institutional attributes also affect the retirement decision. We find that the risk associated with retirement was about 107% greater for incumbents serving through the reform era in the House of Representatives. As Hibbing (1982a, 1982b) notes, the House reforms altered the desirability for serving in the House. Seniority rule was substantially undermined and junior members accrued more power (relative to earlier periods). For incumbents serving prior to this era, the House was evidently a more desirable institution in which to serve. After the reforms, however, the hazard of voluntary career termination through retirement greatly increased relative to incumbents who did not serve in the House during these reforms. In addition to the reform period, we see that Republican incumbents were much more likely to exit office by retirement than Democrats. The hazard of retirement is about 31% greater for Republicans than Democrats. This difference is probably attributable to the Republicans' status as the minority party during much of the time period considered in this analysis.

Turning attention to ambition decisions, we find that the opportunity structure for higher office is strongly related to the odds of terminating the House seat to seek alternative office. Ambition is voraciously related to the availability of open gubernatorial or open Senatorial seats. These huge increases in the hazard of ambition substantively suggest that when the likelihood of winning alternative office is relatively high (open seats imply no running incumbent), incumbents harboring ambition are quite likely to forgo the House seat in search of another office. These results largely confirm the findings of Rhode (1979) and Brace (1984). In addition to this confirming finding, use of event history data permits a finer understanding of how the attractiveness of alternative office changes across the span of a career. What we find is that the attractiveness of a gubernatorial seat *diminishes* as one progresses through a career. As each term passes, the attractiveness of an open gubernatorial seat decreases by about 11%. Interestingly, this pattern does not hold for Senate seats. Regardless of where in a career an open Senate seat emerges, ambitious incumbents are just as likely to seek this type of

Table 5. Change in Hazards for Selected Career Covariate Profiles

	The %Δ in the Odds of			
	Retire vs. Run	Ambition vs. Run	Losing in Gen. vs. Winning	Losing in Prim. vs. Winning
As Duration (Terms served) increases by 1 Term:	+6%	+10%	–8%	—
As an Incumbent's Age increases by 1 Unit:	+8%	–5%	+4%	+4%
As the Incumbent's Prior Electoral Margin increases by 1%:	–1%	+1%	–7%	–1%
The Attractiveness of an Open Gov. Seat compared to Non-open seat:	—	+157%	—	—
For each term served, the attractiveness of a Gov. Seat changes by:	—	–11%	—	—
The Attractiveness of an Open Sen. Seat compared to Non-open seat:	—	+100%	—	—
As the Proportion of the State Represented Increases [a]:	—	—	+9%	+16%
For Incumbents of the President's Party, as Pres. Approval Increases[b]:	—	—	–38%	—
For Incumbents of the President's Party, as Pres. Approval Decreases[b]:	—	—	+61%	—
For Incumbents of the President's Party, as the Economy Improves[c]:	—	—	–41%	—
For Incumbents of the President's Party, as the Economy Worsens[c]:	—	—	+67%	—
Post 1966 Incumbents compared to Pre 1966 Incumbents:	—	+50%	–36%	—
For Incumbents serving during the Reform Era compared to those not:	+107%	—	—	—
Having a prestigious institutional position compared to not having one:	—	–78%	—	—
For Republican Incumbents compared to Democratic Incumbents:	+31%	—	—	—
For Incumbents serving during the "Golden Parachute" Election:	+122%	—	—	—

[a] The numbers in this row can be thought of as the relative difference in the hazard for incumbents from small states vs. incumbents from large states. The 8% increase in the odds of losing vs. winning simply means that for an incumbent who represents a portion of the state one-tenth larger than another incumbent, the risk of his or her losing the general election increases by 8%. The unit increment to calculate the percentages in this row is a one-tenth change.

[b] The coding of presidential approval is such that it ranges from –1 to 1. Hence, a 1 unit increase (decrease) is equal to one-tenth .

[c] These percentages are based on a 1/2% change in RDI in the first quarter. Percentages in the cells reflect how much the odds of experiencing a given kind of event change as the value of the covariates change. See text for further details.

office. Institutional prestige is also related to the hazard of ambition. As one accrues a position in the leadership or one acquires a full committee chairmanship, the hazard of ambition is 78% lower than an incumbent not holding a prestigious House position.

The risk of losing an election (and thus ending one's career) is substantially related to the incumbent's visibility, vulnerability, and national conditions. Comparing across incumbents, we find that as the proportion of the state an incumbent represents increases by one-tenth, the hazard of losing in the general election increases by about 9%. Substantively, this suggests that electoral vulnerability of the incumbent (as measured through risk of defeat) is greater in low populous states than high populous states. In states like Montana, Vermont, and Alaska, where the incumbent House member may be as well-known as the Senator, competition for the House seat is likely to be greater than in states like New York, California, or Ohio, where the visibility and possibly the desirability of the House seat is much less. In addition, we find a strong relationship between visibility and loss in the primary election. The risk of primary defeat increases by about 16% as we move from incumbents in high-populous states to incumbents in low-populous states.

National conditions seem to also exert considerable leverage on the hazard of electoral loss. Presidential popularity is largely related to winning or losing. When the incumbent is of the same party as the president *and* the president's approval rating increases positively by one-tenth[46], the hazard or risk of electoral defeat diminishes by about 38%. In stark contrast, when the incumbent is of the president's party and the president's approval rating drops by one-tenth, the risk associated with losing increases dramatically by 61% (compared to incumbents not of the president's party). A similar effect emerges for national economic conditions (measured as change in real disposable income). When the first quarter RDI changes positively by 1/10th % (thus improving) and the incumbent is of the president's party, the risk of defeat drops by 41%. Conversely, when the economy is faring poorly and the incumbent is of the president's party, the odds of defeat increase by 67%. These findings, we believe, are indicative, in part, of the strategic politician theory (Jacobson 1992; Jacobson and Kernell 1983). When the economy, early in the election cycle, is doing poorly, strong challengers tend to emerge. These results provide some basis for this contention. The risk of defeat is strongly related to first quarter economic conditions. It is precisely during this time of an election cycle that most candidacy decisions must be made. If strong challengers are emerging in response to an unfavorable

[46]Recall that the coding for this variable ranges from –1 to 1. An increase of one-tenth corresponds to a 1% increase in approval.

economy, then the hazard of electoral defeat should increase. In fact, it increases substantially.

Some final results to consider include the relationship between electoral security and changes during the mid-Sixties in House office perquisites. Incumbents elected in or after 1966 have a risk of electoral defeat about 36% lower than incumbents elected prior to 1966. Furthermore, these incumbents also display a greater risk toward ambition. They are about 50% more likely to seek alternative office than incumbents in earlier periods.

The event history approach to understanding congressional careers has elicited both novel findings and confirmation of previous research. Event history parameters generate politically relevant information concerning change processes. In the application here, as well as the two applications discussed earlier, we have seen how longitudinal information adds a new perspective on processes that have either been dealt with cross-sectionally or failed to exploit the longitudinal nature of the data.

Conclusions

Inferences from event history models can be powerful and represent a substantial improvement over traditional cross-sectional models or panel models containing few and/or distantly spaced waves of data. Many processes of interest to political scientists are longitudinal; we are often interested in the dynamics underlying change, or when change occurs. The event history framework is ideally suited for such research questions. We have presented an overview of some central themes in analyzing event histories and provided three applications to demonstrate the potential usefulness of the methodology. Our aim has been to keep the discussion at a fairly "applied" level. We have conveniently omitted comprehensive discussion of some of the more technical aspects of the event history model and instead refer interested readers to the references. However, readers should have some sense of both how to estimate and interpret event history models and when to use such models. Like any method, the event history model is not a cure-all. The data requirements for proper estimation of such models are important to recognize. Nevertheless, given the powerful inferences that can be drawn from such models, data collection efforts should be undertaken with these sorts of models in mind. There is a wide scope of application of event history models in political science. As more political scientists understand event history methodology, new questions may be asked and old questions may more appropriately be answered by having more "realistic" models of processes under study and full use of the data that are available.

Manuscript submitted 13 March 1996.
Final manuscript received 21 November 1996.

APPENDIX A

A wide variety of statistical packages, of varying degrees of familiarity and sophistication, will perform common types of event history analysis without the need to write more than a basic set of instructions. We briefly review the capabilities of five of the most common programs.

SPSS for Windows, version 6.1 : two of the distinguishing features of SPSS are the ease of use even for those with little familiarity with statistic packages (windows, point and click format) and the extensive diagnostics for the Cox Proportional Hazards model, which can include time-varying covariates. SPSS is a good starting package, however, it is unable to compare the Cox and parametric models, which is important because if a parametric model is appropriate, the parametric model is more efficient. SPSS has an informative web site that can be reached at: http://www.spss.com. The discussion list address is: comp.soft-sys.spss. SPSS can be reached at: 444 N. Michigan Avenue, Chicago, IL 60611 or by 1-800-543-2185.

SAS, version 6: the SAS System has several modules that allow estimation of many of the models discussed in this paper. Three proc statements comprise the bulwark of SAS's event history capabilities: PROC PHREG, PROC LIFETEST, and PROC LIFEREG. The PHREG statement permits estimation of the Cox proportional hazards model. Time-varying covariates can be used through PHREG. The LIFEREG procedure is a general estimator that allows estimation of the exponential, Weibull, log-normal, gamma, and Gompertz duration models. The LIFEREG procedure computes such statistics as Kaplan-Meier functions. The SAS system is good at providing diagnostic options, including residual plots. Additionally, Allison (1995) has published a compendium that presents detailed examples of the SAS System for event history models. The web page for SAS is: http://www.sas.com.

STATA, version 5.0: nonparametric, semiparametric, and parametric models with time-varying covariates (but only exponential and Weibull distributions) can all be estimated. Additionally, repeating events models are estimable in STATA. Cox proportional hazard models with time-varying covariates can be estimated (LIMDEP does not have a canned procedure for this). STATA will estimate Kaplan-Meier survival curves with confidence intervals and compare two survival curves for similarities. STATA has excellent graphics. For example, after estimating a Cox model, STATA will generate variables for baseline hazards and survivor functions as well as observation-specific survival estimates (but only if the model does not have time-varying covariates, unfortunately). STATA will also estimate a Cox regression with left-censored data or with stratification. The STATA manual requires particular mention since it is not particularly user-friendly (good on-line help is some consolation). STATA has a discussion list that sees a lot of action and is a great resource of other users and STATA developers: listproc@dsg.harvard.edu. STATA also has one of the better web pages, which can be reached at: http://www.stata.com. One of the unique things about the STATA web page is that you can take introductory through advanced courses (for a fee).

LIMDEP, version 7.0: Limdep provides a comprehensive set of programs for analyzing event history data, including nonparametric approaches, such as life tables, semiparametric models, and finally parametric models. Distributional assumptions that are built-in include exponential, Weibull, normal, logistic, gamma, and Gompertz.

Time-varying covariates are available for most of these models. A Weibull (and therefore exponential) model with gamma heterogeneity, split population model, and models in which the truncation point is not zero can all easily be estimated in LIMDEP. In contrast to the STATA manual, the LIMDEP manual is a wonderful resource containing technical information and valuable citations of where to start to learn more about econometric issues concerning a particular model. Another valuable asset of LIMDEP is the discussion list and the fact that William Greene, who developed LIMDEP and wrote *Econometric Analysis*, 1993, replies to inquiries so quickly. To subscribe to the discussion list send a message to: limdep-1@gsb.usyd.edu.au. A copy of the most recent LIMDEP manual can be found at http://wuecon.wustl.edu/limdep/limdep.html.

Finally, Altman and Stavola (1994) and Collett (1994) provide comparisons and code for SAS, SPSS, and STATA, among others, in the context of estimating a Cox model with time-varying covariates. The data, code, and output for the examples used in this paper are available, see note 1. Other programs available for analysis of event history data include: BMDP, EGRET, EPICURE, EPILOG, GAUSS, Sigma Plot, SPIDA, S-PLUS, and Statistica.

REFERENCES

Achen, Christopher H. 1986. *The Statistical Analysis of Quasi-Experiments.* Berkeley: University of California Press.

Allison, Paul D. 1982. "Discrete-time Methods for the Analysis of Event-Histories." In *Sociological Methodology 1982*, ed. S. Leinhardt. Beverly Hills: Sage.

Allison, Paul D. 1984. *Event History Analysis: Regression for Longitudinal Data.* Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07–041. Newbury Park: Sage.

Allison, Paul D. 1995. *Survival Analysis Using the SAS System: A Practical Guide.* Cary: SAS Institute Inc.

Alt, James E., and Gary King. 1994. "Transfers of Governmental Power: The Meaning of Time Dependence." *Comparative Political Studies* 27:190–210.

Alt, James E., and Gary King. 1995. "Aggregation of Binary into Count and Duration Data: An ABCD of Event Process Models." Harvard University. Typescript.

Alt, James E., Gary King, and Curtis Signorino. 1996. "Estimating the Same Quantities from Different Levels of Data: Time Dependence and Aggregation in Event Process Models." Presented at the annual meetings of the Midwest Political Science Association, Chicago.

Altman, Douglas E., and Bianca L. De Stavola. 1994. "Practical Problems in Fitting a Proportional Hazards Model to Data with Updated Measurements of the Covariates." *Statistics in Medicine* 13:301–14.

Alvarez, R. Michael, and Jonathan Nagler. 1995. "Economics, Issues, and the Perot Candidacy: Voter Choice in the 1992 Presidential Election." *American Journal of Political Science* 39:714–44.

Arjas, Elja, and Pekka Kangas. 1992. "A Discrete-Time Method for the Analysis of Event Histories." In *Demographic Applications of Event History Analysis*, ed. James Treussell, Richard Hankinson, and Judith Tilton. Oxford: Clarendon Press.

Bagai, Isha, and B. L. S. Prakasa Rao. 1992. "Analysis of Survival Data with Two Dependent Risks." *Biomedical Journal* 7:801–14.

Bartels, Larry M., and Henry E. Brady. 1993. "The State of Quantitative Political Methodology." In *Political Science: The State of the Discipline II*, ed. Ada W. Finifter. Washington, DC: American Political Science Association.

Beck, Nathaniel. 1995. "Continuous and Discrete Time Methods for Estimating Duration Dependence." Presented at the annual meeting of the Midwest Political Science Association, Chicago.

Beck, Nathaniel. N.d. "Modeling Space and Time: The Event History Approach." In *Research Strategies in the Social Sciences*, ed. Scarbrough and E. Tanenbaum. Oxford University Press. Forthcoming

Beck, Nathaniel, and Richard Tucker. 1996. "Conflict in Space and Time: Time-Series Cross-Section Analysis with Binary Dependent Variables." Presented at the annual meetings of the American Political Science Association, San Francisco.

Berry, Frances Stokes, and William D. Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *American Political Science Review* 84:395–416.

Berry, Frances Stokes, and William D. Berry. 1992. "Tax Innovation in the States: Capitalizing on Political Opportunity." *American Journal of Political Science* 36:715–42.

Berry, Frances Stokes, and William D. Berry. 1994. "The Politics of Tax Increases in the States." *American Journal of Political Science* 38:855–59.

Bienen, Henry, and Nicolas van de Walle. 1989. "Time and Power in Africa." *American Political Science Review* 83:20–34.

Bienen, Henry, and Nicolas van de Walle. 1992. "A Proportional Hazard Model of Leadership Duration." *Journal of Politics* 54:685–719.

Blossfeld, Hans-Peter, and Alfred Hamerle. 1989. "Using Cox Models to Study Multi Episode Processes." *Sociological Methods and Research* 17:432–48.

Blossfeld, Hans-Peter, Alfred Hamerle, and Karl Ulrich Mayer. 1989. *Event History Analysis.* Hillsdale: Lawrence Erlbaum Publishing.

Box-Steffensmeier, Janet M. 1996. "A Dynamic Analysis of the Role of War Chests in Campaign Strategy." *American Journal of Political Science* 40:352–71.

Box-Steffensmeier, Janet M., Laura W. Arnold, and Christopher J. Zorn. 1997. "Strategic Position Taking and the Timing of Voting Decisions in Congress: A Study of the North American Free Trade Agreement." *American Political Science Review* 91:324–38.

Box-Steffensmeier, Janet M., and Peter Radcliffe. 1996. "The Timing of PAC Contributions: A Split-Population Model." Presented at the annual meeting of the American Political Science Association, San Francisco.

Box-Steffensmeier, Janet M., and Christopher J. Zorn. 1996. "An Examination of the Assumption of Proportionality in Event History Models. " Ohio State University. Typescript.

Brace, Paul. 1984. "Progressive Ambition in the House: A Probabilistic Approach." *Journal of Politics* 46:556–69.

Brace, Paul. 1985. "A Probabilistic Approach to Retirement from the U.S. Congress." *Legislative Studies Quarterly* 10:107–23.

Chung, Ching-Fan, Peter Schmidt, and Ann D. Witte. 1991. "Survival Analysis: A Survey." *Journal of Quantitative Criminology* 7:59–98.

Cioffi-Revilla, Claudio. 1984. "The Political Reliability of Italian Governments: An Exponential Survival Model." *American Political Science Review* 78:318–38.

Cioffi-Revilla, Claudio. 1985. "European Political Cooperation: An Application of Political Reliability Theory to Integration." *International Studies Quarterly* 28:467–92.

Cioffi-Revilla, Claudio, and David Lai. 1995. "War and Politics in Ancient China, 2700 B.C. to 722 B.C.: Measurement and Comparative Analysis." *Journal of Conflict Resolution* 34:467–94.

Collett, D. 1994. *Modeling Survival Data in Medical Research.* London: Chapman and Hall.

Congressional Quarterly. 1992. *Congressional Ethics: History, Facts, and Controversy.* Washington, DC: Congressional Quarterly Press.

Cox, D. R. 1972. "Regression Models and Life Tables." *Journal of the Royal Statistical Society, B* 34:187–220.

Cox, D. R. 1975. "Partial Likelihood." *Biometrika* 62:269–76.

Cox, D. R., and D. Oakes. 1983. *The Analysis of Survival Data.* New York: Chapman and Hall.

Crowder, Martin. 1991. "On the Identifiability Crisis in Competing Risks Analysis." *Scandinavian Journal of Statistics* 18:223–33.

Diermeier, Daniel, and Randall Stevenson. 1994. "A Competing Risks Analysis of Cabinet Duration." Presented at the annual meeting of the Midwest Political Science Association, Chicago.

Farole, Donald, Jeffrey Levine, and William Morgan. 1995. "The Odds Ratio Interpretation and Complete Analysis of Logit Results." Presented at the Midwest Political Science Association, Chicago.

Geddes, Barbara. 1990. "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics." *Political Analysis* 2:131–52.

Gill, Richard, and Martin Schumacher. 1987. "A Simple Test of the Proportional Hazards Assumption." *Biometrika* 74:289–300.

Gourieroux, C., A. Monfort, and A. Trognon. 1984. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica* 52:701–20.

Greene, William. 1993. *Econometric Analysis*, 2nd ed. New York: Macmillan Publishing Company.

Ham, John C., and Robert J. LaLonde. 1996. "The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training." *Econometrica* 64: 175–205.

Han, Aaron, and Jerry A. Hausman. 1990. "Flexible Parametric Estimation of Duration and Competing Risks Models." *Journal of Applied Econometrics* 5:1–28.

Hannan, Michael T. 1989. "Macrosociological Applications of Event History Analysis: State Transitions and Event Recurrence." *Quality and Quantity* 23:351–83.

Harris, C. M., A. R. Kaylan, and M. D. Maltz. 1981. "Refinements in the Statistics of Recidivism Measurement." In *Models in Quantitative Criminology*, ed. J. A. Fox. New York: Academic Press.

Hausman, Jerry A., and Daniel McFadden. 1984. "Specification Tests for the Multinomial Logit Model." *Econometrica* 52:1219–40.

Heckman, James J., and B. Singer. 1984a. "Econometric Duration Analysis." *Journal of Econometrics* 24:63–132.

Heckman, James J., and B. Singer. 1984b. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52:251–320.

Heckman, James J., and B. Singer. 1985. "Social Science Duration Analysis." In *Longitudinal Analysis of Labor Market Data*, ed. James J. Heckman and B. S. Singer. New York: Cambridge University Press.

Heckman, James J., and Robert J. Willis. 1977. "A Beta-logistic Model for the Analysis of Sequential Labor Force Participation by Married Women." *Journal of Political Economy* 85(1): 27–58.

Hibbing, John R. 1982a. *Choosing to Leave.* Washington, DC: University Press of American.

Hibbing, John R. 1982b. "Voluntary Retirement from the U.S. House: Who Quits?" *American Journal of Political Science* 26:467–84.

Hill, Daniel H., William G. Axinn, and Arland Thornton. 1993. "Competing Hazards with Shared Unmeasured Risk Factors." *Sociological Methodology* 23:245–77.

Hoem, J. M. 1990. "Limitations of Heterogeneity Techniques." In *Convergent Issues in Genetics and Demography*, ed. J. Adams, et al. Oxford: Oxford University Press.

Hougaard, Philip. 1991. "Modelling Heterogeneity in Survival Data." *Journal of Applied Probability* 28:695–701.

Inter-university Consortium for Political and Social Research. 1993. *Roster of United States Congressional Officeholders, 1788-1988.* University of Michigan.

Jacobson, Gary C. 1992. *The Politics of Congressional Elections.* 3rd ed. New York: Harper Collins.

Jacobson, Gary C., and Samuel Kernell. 1983. *Strategy and Choice in Congressional Elections.* New Haven: Yale University Press.

Jones, Bradford S. 1994. "*A Longitudinal Perspective on Congressional Elections.*" Ph.D. diss. State University of New York at Stony Brook.

Kalbfleisch, John, and Ross Prentice. 1980. *The Statistical Analysis of Failure Time Data.* New York: Wiley.

Katz, Jonathan N., and Brian R. Sala. 1996. "Careerism, Committee Assignments, and the Electoral Connection." *American Political Science Review* 90:21–33.

Kiefer, Nicholas. 1988. "Economic Duration Data and Hazard Functions." *Journal of Economic Literature* 26:646–79.

King, Gary C., James E. Alt, Nancy E. Burns, and Michael Laver. 1990. "A Unified Model of Cabinet Dissolution in Parliamentary Democracies." *American Journal of Political Science* 38: 846–71.

King, Gary C., Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton: Princeton University Press.

Kuk, Anthony Y. C. 1992. "A Semiparametric Mixture Model for the Analysis of Competing Risks Data." *Australian Journal of Statistics* 34:169–80.

Lancaster, T. 1979. "Econometric Methods for the Duration of Unemployment." *Econometrica* 47:939–56.

Mintrom, Michael. 1994. "*Policy Entrepreneurship in Theory and Practice: A Comparative State Analysis of the Rise of School Choice as a Policy Idea.*" Ph.D. diss. State University of New York at Stony Brook.

Namboodiri, N. Krishnan. 1996. "Event History Class Handout." Ohio State University. Typescript.

Namboodiri, N. Krishnan, and C. W. Suchindran. 1987. *Life Table Techniques and Their Applications.* Orlando: Academic Press.

Olzak, Susan. 1987. "Causes of Ethnic Conflict and Protest in Urban America, 1877–1889." *Social Science Research* 16:185–210.

Olzak, Susan. 1989. "Labor Unrest, Immigration, and Ethnic Conflict: Urban America 1880–1915." *American Journal of Sociology* 94:1303–33.

Petersen, Trond. 1985. "Fitting Parametric Survival Models with Time-Dependent Covariates." *Applied Statistics* 35:281–8.

Petersen, Trond. 1991. "The Statistical Analysis of Event Histories." *Sociological Methods and Research* 19:270–323.

Prentice, R. L., and V. T Farewell. 1986. "Relative Risk and Odds Ratio Regression." *Annual Review of Public Health* 7:335–38.

Proschan, F. 1963. "Contribution to the Discussion of Paper by D. R. Cox." *Journal of the Royal Statistical Society, Serial B* 34:472–5.

Robinson, D. T., and L. Smith-Lovin. 1990. "Timing of Interruption in Group Discussions." In *Advances in Group Processes: Theory and Research,* Vol. 2, ed. E. J. Lawler, B. Markovsky, C. Ridgeway, and H. Walker. Greenwich: JAI Press.

Rhode, David W. 1979. "Risk-Bearing and Progressive Ambition: The Case of Members of the United States House of Representatives." *American Journal of Political Science* 23:1–26.

Schmidt, Peter, and Ann Dryden Witte. 1988. *Predicting Recidivism Using Survival Models.* New York: Springer-Verlag.

Schmidt, Peter, and Ann Dryden Witte. 1989. "Predicting Criminal Recidivism Using 'Split Population' Survival Time Models." *Journal of Econometrics* 40:141–59.

Singer, B., and S. Spilerman. 1976a. "The Representation of Social Processes by Markov Models." *American Journal of Sociology* 82:1–54.

Singer, B., and S. Spilerman. 1976b. "Some Methodological Issues in the Analysis of Longitudinal Surveys." *Annals of Economic and Social Measurement* 5:447–74.

Singer, Judith D., and John B. Willett. 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.

Small, Kenneth A., and Cheng Hsiao. 1985. "Multinomial Logit Specification Tests." *International Economic Review* 26:619–27.

Small, Melvin, and J. David Singer. 1982. *Resort to Arms: International and Civil Wars*. 2nd ed. Beverly Hills: Sage.

Stimson, James A. 1985. "Regression in Space and Time: A Statistical Essay." *American Journal of Political Science* 29:914–47.

Sueyoshi, Glenn T. 1992. "Semiparametric Proportional Hazards Estimation of Competing Risks Models with Time-Varying Covariates." *Journal of Econometrics* 51:25–58.

Sueyoshi, Glenn T. 1995. "A Class of Binary Response Models for Grouped Duration Data." *Journal of Econometrics*. 40:141–59.

Swaim, Paul, and Michael Podgursky. 1994. "Female Labor Supply Following Displacement: A Split-Population Model of Labor Force Participation and Job Search." *Journal of Labor Economics* 12: 640–56.

Teachman, Jay D., and Mark D. Hayward. 1993. "Interpreting Hazard Rate Models." *Sociological Methods and Research* 21:340–71.

Tillema, Herbert K. 1991. *International Armed Conflict Since 1945: A Bibliographic Handbook of Wars and Military Interventions*. Boulder: Westview Press.

Tillema, Herbert K. 1994. "Cold War Alliances and Overt Military Intervention, 1945–1991." *International Interactions* 20:249–78.

Trussel, J., and T. Richards. 1985. "Correcting for Unmeasured Heterogeneity Using the Heckman-Singer Procedure." In *Sociological Methodology*, ed. Nancy B. Tuma. San Francisco: Jossey-Bass.

Tuma, Nancy Brandon, and Michael T. Hannan. 1979. "Approaches to the Censoring Problem in Analysis of Event Histories." In *Sociological Methodology 1979,* ed. K. F. Schuessler, San Francisco: Jossey-Bass.

Tuma, Nancy Brandon, and Michael T. Hannan. 1984. *Social Dynamics: Models and Methods*. Orlando: Academic Press.

Tuma, Nancy Brandon, Michael T. Hannan, and L. Groeneveld. 1979. "Dynamic Analysis of Event Histories." *American Journal of Sociology* 84:820–54.

Vaupel, James W., and Anatoli I. Yashin. 1985. "Heterogeneity's Ruses: Some Surprising Effects of Selection of Population Dynamics." *The American Statistician* 39:176–85.

Warwick, Paul. 1992. "Economic Trends and Government Survival in West European Parliamentary Democracies." *American Political Science Review* 86:875–86.

Yamaguchi, Kazuo. 1990. "Logit and Multinomial Logit Models for Discrete-Time Event-History Analysis: A Causal Analysis of Interdependent Discrete-State Processes." *Quality and Quantity* 24:323–41.

Yamaguchi, Kazuo. 1991. *Event History Analysis*. Newbury Park: Sage.

Zhang, Junsen, and Saul D. Hoffman. 1993. "Discrete-Choice Logit Models: Testing the IIA Property." *Sociological Methods & Research* 22:193–213.