

## Using Event History Analysis in Historical Research: With Illustrations from a Study of the Passage of Women's Protective Legislation\*

HOLLY J. MCCAMMON

Historians and social scientists often investigate the conditions that influence the occurrence of particular events. For instance, a researcher might be concerned with the causes of revolutionary action in some countries or the forces that unleash racial rioting in major cities. Or perhaps the researcher wishes to examine why industrial workers decide to strike or what prompts policy-makers to pass new legislation.<sup>1</sup> In each of these examples, a qualitative shift occurs, from a circumstance without racial rioting in a particular city, for instance, to one with racial rioting. Event history analysis can aid researchers in uncovering the conditions that lead to such a shift.

Event history analysis is a quantitative method that offers researchers a means of explaining why such events occur. A myriad of types of events can be analyzed using event history analysis. Suitable kinds of events are those marked by a definite and somewhat abrupt transition from one state to another, such as the founding or collapse of an organization or the emergence of a social movement. More gradual transitions from one state to another where there is difficulty pinpointing the moment in time of the transition are usually not amenable to event history analysis.

Event history analysis utilizes event history data which are composed of *event histories* for the nations, organizations, groups, or even individuals examined in the analysis. These event histories are over-time records that reveal when, if at all, the event being studied occurs for each of the cases included in the analysis. In addition to the event histories, additional data for each observation on a variety of factors believed to influence the occurrence of the event are included in the analysis (the specific nature of the data is discussed in greater detail below). Thus, if the event of interest is the transition of a polity from authoritarianism to democracy, not only will the researchers need information on the point in time at which the transition occurred, but they will also need longitudinal (i.e. over-time) data on the factors likely to have facilitated or even hindered this change in government. In short, then, if a researcher is interested in the question of why a historical event occurs for some cases but not for others and if the researcher

\* I am grateful to Larry Griffin, Marcel van der Linden and Karen Campbell for comments on an earlier draft.

1. For a variety of historical studies using event history analysis concerning these and other issues, see the annotated bibliography at the end of this piece.

has longitudinal and quantifiable data on the timing of the events and similar data on the factors likely to have influenced the occurrence of the event, then event history analysis can be a useful tool for the researcher in explaining why such events occur.

Event history analysis is useful because it can explain why such events occur. But its utility also lies in the way in which it allows researchers to explain events. While focusing on a single case permits researchers to gather detailed insights into social dynamics, the case study does not usually provide a *systematic* assessment of the influences necessary for an event to occur.<sup>2</sup> Because event history analysis includes both cases that have and have not experienced the event, a comparison of such cases can be made to determine those conditions that are and are not necessary for the event to occur. In this sense, then, a far more systematic determination of the causes of the event is possible.<sup>3</sup> The “negative” cases – those for which the event did not occur – are not excluded from the analysis (this kind of exclusion occurs almost by definition in most case studies), and the valuable lessons such negative cases offer about the reasons why the event could not occur are incorporated into event history analysis.

Here I provide a discussion of how one can use event history analysis to explain, using a systematic comparison of cases in which the event occurs and cases in which the event does not occur, why the particular historical event happens. I begin with a discussion of the nature of the *data* necessary for event history analysis, then turn to the *statistical technique* used in the analysis and the *interpretation of the results*. Finally, a number of *complexities* associated with event history analysis are explored. For instance, what can be done to analyze events that repeat themselves or multiple kinds of events? Throughout this discussion, the data needs and the method are illustrated with data and an analysis concerning the passage of protective legislation for women in the United States around the turn of the century. These data, drawn from previous research,<sup>4</sup> are particularly suited to event history analysis given that the adoption of new law is a historical event.

2. Susan Olzak, “Analysis of Events in the Study of Collective Action”, *Annual Review of Sociology*, 15 (1989), p. 121. Also, as Stanley Lieberson (“Small N’s and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases”, in Charles C. Ragin and Howard S. Becker (eds), *What Is a Case? Exploring the Foundations of Social Inquiry* (New York, 1992), p. 105) discusses, studies involving a single case are best for revealing that “a given phenomenon exists in some setting” and are perhaps less useful for explaining causal processes.

3. This is true generally of methods that rely on comparative analyses such as Qualitative Comparative Analysis (see Ragin elsewhere in this volume) and Millsean methods of comparison.

4. Holly J. McCammon, “The Politics of Protection: State Minimum Wage and Maximum Hours Laws for Women in the United States, 1870–1930”, *The Sociological Quarterly*, 36 (1995), pp. 217–249; idem, “Protection for Whom? Maximum Hours Laws and Women’s Employment in the United States, 1880–1920”, *Work and Occupations*, 23 (1996), pp. 132–164.

## THE DATA FOR EVENT HISTORY ANALYSIS

A unique feature of event history analysis compared to many other quantitative methods is that it employs data that are simultaneously cross-sectional and longitudinal. Thus the method analyzes both cross-sectional and temporal variation. To put this in more concrete terms, consider data concerning the passage of women's protective legislation. To study the enactment of this legislation using event history analysis, data are needed both over time and across multiple cases. Women's protective laws, enacted – at least ostensibly – to protect women in the workplace, were passed in many states in the US around the turn of the century.<sup>5</sup> The event history measures, then, not only are over time, denoting the year in which a protective law was passed, but also are across observations or, in this case, across US states (i.e. the data are for Alabama, Iowa, New Jersey, etc.).

Figure 1 provides a visual representation of the general structure of the data matrix. For each measure or variable, including both the dependent variable (which indicates the year in which a protective law was enacted in a state – although see the discussion below of the precise nature of the dependent variable) and the various explanatory variables, information is given both across years and across states. The unit of analysis, then, in this study (or each cell in Figure 1) is the “state-year”. The unit of analysis for event history data always designates both a cross-sectional observation and a time unit.

The level of over-time aggregation in event history data (i.e. is decade-level data used? annual data? monthly data?) ideally should be determined by the nature of the research question or by the time frame in which the event of interest occurs. For instance, the state legislatures that enacted women's protective laws met annually (or sometimes biennially) and thus annual-level data are used in this analysis. More frequently, however, the over-time level of aggregation in the data is determined by the nature of the data available to the researcher. Annual data, in particular, are frequently used in quantitative historical research because of their availability from governmental sources.<sup>6</sup> Researchers, however, are sometimes able to construct their own data sets or specific variables from information gleaned

5. Elizabeth Brandeis, “Labor Legislation”, in John R. Commons (ed.), *History of Labor in the United States, 1896–1932*, vol. 4 (New York, 1935), pp. 397–697. In the research presented in this paper, women's protective legislation includes maximum hours laws (that restricted the maximum number of hours women could work), minimum wage laws and laws prohibiting night work among women.

6. See, for example: US Bureau of the Census, *Historical Statistics of the United States: Colonial Times to 1970* (Washington, DC, 1975); US Bureau of the Census, *Statistical Abstract of the United States* (Washington, DC, various years); President of the United States, *Economic Report of the President* (Washington, DC, various years); US Department of Labor, *Handbook of Labor Statistics* (Washington, DC, various years).

|         |      | Dependent variable                |                   |                      |     |               |
|---------|------|-----------------------------------|-------------------|----------------------|-----|---------------|
|         |      | Explanatory variables             |                   |                      |     |               |
| State   | Year | Passage of protective legislation | Consumers' league | Competitive election | ... | Full suffrage |
| Alabama | 1870 |                                   |                   |                      |     |               |
|         | 1871 |                                   |                   |                      |     |               |
|         | 1872 |                                   |                   |                      |     |               |
|         | .    |                                   |                   |                      |     |               |
| Arizona | 1930 |                                   |                   |                      |     |               |
|         | 1870 |                                   |                   |                      |     |               |
|         | 1871 |                                   |                   |                      |     |               |
|         | 1872 |                                   |                   |                      |     |               |
|         | .    |                                   |                   |                      |     |               |
|         | 1930 |                                   |                   |                      |     |               |
| Wyoming | 1870 |                                   |                   |                      |     |               |
|         | 1871 |                                   |                   |                      |     |               |
|         | 1872 |                                   |                   |                      |     |               |
|         | .    |                                   |                   |                      |     |               |
|         | .    |                                   |                   |                      |     |               |
|         | 1930 |                                   |                   |                      |     |               |

Figure 1. Event history analysis data matrix for a study of the passage of women's protective legislation in the US states, 1870-1930

from archival sources, newspapers, court or legislative documents, organizational reports, or even secondary historical accounts. When data are compiled from such sources, the level of over-time grouping may be more specific than annual-level measures. One word of caution, however. In some cases a higher level of aggregation makes more sense than data indicating the exact timing of the occurrence of the event. Consider the protective legislation data which are annual-level. Daily or even monthly data concerning the dates of passage of such laws would confound the analysis with state-to-state differences in when state legislatures meet, which is not of theoretical interest in the analysis.<sup>7</sup> Thus, annual-level measures for this analysis are desirable.

Given that event history analysis analyzes a shift from one condition to another, the dependent variable is coded as a binary or dichotomous vari-

7. Eliza K. Pavalko, "State Timing of Policy Adoption: Workmen's Compensation in the United States, 1909-1929", *American Journal of Sociology*, 95 (1989), p. 601.

able. That is, the measure is a simple dummy variable taking the value of "0" prior to the occurrence of the event and the value of "1" once the event occurs. The dependent variable for the event history analysis of the passage of women's protective legislation is assigned the value of "0" for all years prior to the passage of a particular state's protective law and "1" for the year in which the law is enacted. The years after the law is enacted and before the end of the period of analysis are coded as missing values on the dependent variable and thus are not included in the analysis.<sup>8</sup>

An example will make this clearer. The period of analysis for the protective legislation study is 1870 to 1930, and California, for instance, passed a protective law in 1911.<sup>9</sup> The dependent variable for this state is assigned the value of "0" from the beginning of the analysis in 1870 until 1910. A value of "1" is assigned to 1911, and missing values are assigned to all years in the 1912–1930 period. The dependent variable is constructed in a similar manner for all states. For states that did not pass a protective law (the negative cases), the value of "0" is assigned for all years from 1870 through 1930.

While the dependent variable is coded as a dichotomous variable, the actual value of the dependent variable for the event history analysis *per se* is the *hazard rate* or, as it is sometimes known, the *transition rate*. After the researcher enters the dichotomous variable into the analysis, the event history statistical software (such as SPSS or SAS) computes the hazard rate. The hazard or transition rate is the probability that the event (i.e. the passage of law) will occur at time "t" (i.e. during a particular year) given that the event has not yet occurred for that case (i.e. for that state) or, in other words, given that the case is still at risk of the event occurring.<sup>10</sup> In short, the hazard rate measures and the event history analyzes the likelihood of the event occurring at a particular point in time.

The hazard rate is, however, reasonably straightforward to compute by hand. For each year, one simply divides the number of events (i.e. the number of laws passed) by the number of observations at risk of the event occurring (i.e. the number of states not yet having passed a protective law). The number of observations at risk of the event occurring is the risk set.<sup>11</sup> Table 1 provides the annual number of laws passed (column 3) and the associated risk set (column 4) and hazard rate (column 5) for some of the years in the protective legislation data (periods when the hazard rate is constant over time are left out to simplify the table). In 1870 when no state had yet enacted protective legislation, the hazard rate is 0 (column 3) divided

8. Note that this also means that data on the explanatory variables are not needed for the years for which the dependent variable is assigned missing values. The explanatory variables can be assigned missing values for these years as well.

9. Below I discuss why the period 1870–1930 was chosen for the analysis.

10. Paul D. Allison, *Event History Analysis: Regression For Longitudinal Event Data* (Newbury Park, 1984), p. 16; Kazuo Yamaguchi, *Event History Analysis* (Newbury Park, 1991), pp. 3, 9–10.

11. Allison, *Event History Analysis*, p. 16.

Table 1. *The risk set and hazard rate*

| (1)<br>Year | (2)<br>Total number of<br>states | (3)<br>Number of laws<br>passed | (4)<br>Risk set or number of<br>states at risk of<br>passing law | (5)<br>Hazard rate |
|-------------|----------------------------------|---------------------------------|--|--------------------|
| 1870        | 38                               | 0                               | 38   | .0                 |
| 1871        | 38                               | 0                               | 38   | .0                 |
| 1872        | 38                               | 0                               | 38   | .0                 |
| 1873        | 38                               | 0                               | 38   | .0                 |
| 1874        | 38                               | 1                               | 38   | .026               |
| 1875        | 38                               | 0                               | 37   | .0                 |
| ⋮           | ⋮                                | ⋮                               | ⋮  | ⋮                  |
| 1884        | 38                               | 0                               | 37   | .0                 |
| 1885        | 38                               | 1                               | 37   | .027               |
| 1886        | 38                               | 1                               | 36   | .028               |
| 1887        | 38                               | 3                               | 35   | .086               |
| 1888        | 38                               | 0                               | 32   | .0                 |
| 1889        | 42                               | 0                               | 36   | .0                 |
| 1890        | 44                               | 1                               | 38   | .026               |
| 1891        | 44                               | 0                               | 37   | .0                 |
| 1892        | 44                               | 1                               | 37   | .027               |
| 1893        | 44                               | 0                               | 36   | .0                 |
| 1894        | 44                               | 0                               | 36   | .0                 |
| 1895        | 44                               | 0                               | 36   | .0                 |
| 1896        | 45                               | 0                               | 37   | .0                 |
| 1897        | 45                               | 1                               | 37   | .027               |
| 1898        | 45                               | 0                               | 36   | .0                 |
| 1899        | 45                               | 3                               | 36   | .083               |
| 1900        | 45                               | 0                               | 33   | .0                 |
| 1901        | 45                               | 1                               | 33   | .030               |
| 1902        | 45                               | 0                               | 32   | .0                 |
| 1903        | 45                               | 2                               | 32   | .063               |
| 1904        | 45                               | 0                               | 30   | .0                 |
| 1905        | 45                               | 0                               | 30   | .0                 |
| 1906        | 45                               | 0                               | 30   | .0                 |
| 1907        | 46                               | 2                               | 31   | .065               |
| 1908        | 46                               | 0                               | 29   | .0                 |
| 1909        | 46                               | 3                               | 29   | .103               |
| 1910        | 46                               | 0                               | 26   | .0                 |
| 1911        | 46                               | 5                               | 26   | .192               |
| 1912        | 48                               | 3                               | 23   | .130               |
| 1913        | 48                               | 6                               | 20   | .300               |
| 1914        | 48                               | 1                               | 14   | .071               |
| 1915        | 48                               | 3                               | 13   | .231               |
| 1916        | 48                               | 0                               | 10   | .0                 |
| 1917        | 48                               | 1                               | 10   | .100               |
| 1918        | 48                               | 0                               | 9  | .0                 |
| 1919        | 48                               | 1                               | 9  | .111               |
| 1920        | 48                               | 0                               | 8  | .0                 |
| ⋮           | ⋮                                | ⋮                               | ⋮  | ⋮                  |

by 38 (column 4).<sup>12</sup> In 1874, when Massachusetts passed the first protective law, the hazard rate is 1 divided by 38. One can see in Table 1 that at the end of each time period (e.g. each year), the risk set is reduced by the number of observations (e.g. states) experiencing the event during that year. In addition, one can see the variation in the hazard rate over time. For these data, the probability of a state passing a protective law was highest on average in the 1910s. This is not surprising given that this was roughly the progressive era, when many such reforms were adopted.

However, while the hazard rate varies across years, as can be seen in Table 1, an assumption of event history analysis is that the rate is constant across all observations within a given year. Thus, concerning protective legislation, it is assumed that the rate is constant for all states in each year. If a researcher deems this an invalid assumption and can separate and group observations such that the assumption of constancy holds true for the subgroupings, then the researcher can analyze the separate subgroupings. So, for example, for the protective legislation data, perhaps a better assumption might be that the hazard rate for the southern, northeastern, midwestern and western states is different across these regions but relatively constant within regions. If this were the case, these subgroupings of states could be analyzed in separate event history analyses.<sup>13</sup>

Establishing the sample or population of observations to be included in the analysis can aid the researcher in determining the appropriate time period over which the analysis should be conducted.<sup>14</sup> For the analysis of protective legislation, all 48 states are included in the analysis (but see note 12). The earliest

12. There were 38 states in 1870. By the end of the period of analysis there were 48 states (see column 2 of Table 1 for the total number of states in the union). Until a territory becomes a state, it is excluded from the analysis because the territories may have had different processes by which laws were enacted. However, no territory adopted any workplace protective legislation for women.

13. See Beth A. Rubin, "Limits to Institutionalization? A Sectoral Analysis of U.S. Strike Settlements, 1960–1977", *Research in Social Stratification and Mobility*, 11 (1992), pp. 177–202; and Michael T. Hannan and John Freeman, "The Ecology of Organizational Founding: American Labor Unions, 1836–1985", *American Journal of Sociology*, 92 (1987), pp. 910–943 for other examples of subgrouping the data in this manner for event history analysis.

14. Various authors discuss the complexities associated with defining the appropriate population to be included (or sampled) in an analysis. See Larry J. Griffin, Christopher Botsko, Ana-Maria Wahl and Larry W. Isaac, "Theoretical Generality, Case Particularity: Qualitative Comparative Analysis of Trade Union Growth and Decline", *International Journal of Comparative Sociology*, 32 (1991), pp. 115–116; Charles C. Ragin, *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies* (Berkeley, 1987), pp. 41–42; Michael Burawoy, "Two Methods in Search of Science: Skocpol versus Trotsky", *Theory and Society*, 18 (1989), pp. 765–769. As these analysts point out, determining the theoretically relevant population can present difficulties because not all observations experience the event that the researcher is investigating. Such observations are the "negative cases" or the observations experiencing "non-events". Determining the theoretically relevant set of observations that experience non-events can be a challenge, and the researcher will need to rely upon a precise definition of the population at risk of the event occurring, whether or not the event actually took place for cases in the population.

year that a state passed a protective law is 1874 (a maximum hours law in Massachusetts); the latest date is 1923 (a minimum wage law in South Dakota). Noting this, the researcher could choose to analyze the period from 1870 to 1930 and be certain of including all such events in the analysis.<sup>15</sup>

If researchers cannot analyze a period that encompasses the occurrence of all events of interest for the sample or population being studied, researchers confront the problem of *censoring* in event history analysis. Consider a study of the onset of racial rioting in large metropolitan areas. The researcher may include in the analysis a sample of large cities and analyze the incidence of rioting in them. However, the availability of data may necessitate a particular starting date for the analysis that excludes the occurrence of an earlier riot in a particular city. Or data limitations may necessitate an ending date for the analysis that excludes a later riot in another city. For each of these cities, while the dependent variable is coded as “0” for the period of analysis (i.e. no riot occurs in the cities during the period analyzed), in actuality, the variable is censored, and the “0”s to some extent misrepresent reality. The city did experience a riot, but not within the period analyzed. Censoring of the data such as this in event history analysis introduces the possibility of sample selection bias and biased parameter estimates which can distort the conclusions drawn from the analysis. Thus, to the degree possible censored data should be avoided.<sup>16</sup>

Often, though, researchers are not confronted with such difficulties. In the analysis of the passage of protective legislation, eight states did not pass protective laws between 1870 and 1930, but the historical record shows that these eight states also did not pass protective laws in earlier or later years and, thus, the data are not censored. It is possible that a researcher will of necessity analyze a period that is shorter in duration than what may be ideal. The researcher may nonetheless have evidence that the events of interest did not occur outside of this period for those cases not experiencing the events during the period.

## AN EVENT HISTORY ANALYSIS

### *Discrete-time methods*

Two methods of analysis are used in event history analysis: discrete-time methods and continuous-time methods.<sup>17</sup> Researchers must decide which is

15. For detailed discussions of the theoretical and historical importance of choosing beginning and end dates for analyses, see Larry W. Isaac and Larry J. Griffin, “Ahistoricism in Time-Series Analyses of Historical Process: Critique, Redirection and Illustrations from U.S. Labor History”, *American Sociological Review*, 54 (1989), pp. 873–890; Larry J. Griffin and Larry W. Isaac, “Recursive Regression and the Historical Use of ‘Time’ in Time-Series Analysis of Historical Process”, *Historical Methods*, 25 (1992), pp. 166–179; and Isaac *et al.*, “Temporally Recursive Regression and Social Historical Inquiry: An Example of Cross-Movement Militancy Spillover” (this volume).

16. Censoring is a complex issue in event history analysis. For extended treatments of the subject see: Yamaguchi, *Event History Analysis*, pp. 3–9; Nancy Brandon Tuma and Michael T. Hannan, “Approaches to the Censoring Problem in Analysis of Event Histories”, in Karl F. Schuessler (ed.), *Sociological Methodology* (San Francisco, 1979), pp. 209–240.

17. Allison, *Event History Analysis*, p. 14.



the more appropriate for their data. The decision depends largely (although not entirely) on the level of over-time aggregation in the data. Continuous-time methods should be utilized when the exact moment of the event is known and measured in the data, that is, when the data are continuous. This is often the case with individual-level data that are routinely used in biological research (e.g. in mortality research) and some sociological research (e.g. concerning marriage and divorce patterns). Discrete-time methods, on the other hand, are used when the data are aggregated or grouped over larger time units, that is, over discrete time units, such as months, years, or even decades. With quantitative historical data, given the typical nature of the available data, as discussed above, this broader level of aggregation is common. For this reason, the analyses and discussion below consider discrete-time methods.<sup>18</sup>

Yamaguchi lists two additional circumstances that can influence a researcher's decision to utilize discrete- or continuous-time methods.<sup>19</sup> First, if the data contain ties (multiple observations experiencing the event at the same point in time), then discrete methods are appropriate. Ties in the data can introduce bias in the parameter estimates when continuous-time methods are used, and ties are not at all uncommon in historical data. For example, six states passed protective laws in 1913. Second, under certain circumstances, discrete-time methods offer a reasonable approximation of continuous-time methods. This is the case when the hazard rate or the probability of an event occurring over the various time units is generally small, where small is defined as a hazard rate of less than 0.10. However, as Yamaguchi points out, this rule of thumb is somewhat flexible.<sup>20</sup> Analyses show that even where a moderate number of the probabilities are greater than 0.10, the discrete-time method continues to offer a reasonable approximation of the continuous-time method.

The analysis below provides an illustration of event history analysis using discrete-time methods, a technique that is appropriate to use, most importantly, when the data are over discrete time units, that is, for instance, when the data are aggregated over months or years as opposed to data that pin-points the specific date of event occurrence. Such aggregated time units are commonly used in quantitative historical research because researchers seldom are able to design and administer a survey for data collection that allows them to collect data on the exact timing of events.

### *Logistic regression*

Given that the dependent variable in event history analysis is a probability, discrete-time event history data can be modeled using logistic

18. See Nancy Brandon Tuma and Michael T. Hannan, *Social Dynamics: Models and Methods* (Orlando, 1984), *passim*; Allison, *Event History Analysis*, pp. 22–33; Yamaguchi, *Event History Analysis*, pp. 101–160; and David R. Cox and D. Oakes, *The Analysis of Survival Data* (New York, 1983), *passim* for detailed treatments of continuous-time methods.

19. Yamaguchi, *Event History Analysis*, pp. 16–17.

20. *Ibid.*, p. 42.

regression.<sup>21</sup> Regression allows the researcher to estimate the impact of various explanatory variables on a dependent variable. However, because the dependent variable for event history analysis is a probability, it cannot be estimated to be less than 0 or greater than 1. Such an estimate would be statistically impossible because a probability cannot be less than 0 or greater than 1. A logit transformation of the dependent variable or, more specifically, of the hazard rate remedies this difficulty.<sup>22</sup> If  $P(t)$  is the probability of the event occurring at time “ $t$ ” (that is,  $P(t)$  is the hazard rate), then a logit transformation of  $P(t)$  is:

$$\log (P(t) / (1 - P(t))),$$

or simply the log odds where the odds are the ratio of the probabilities for two mutually exclusive states,  $P(t)$  and  $1 - P(t)$ . The natural base of the logarithm is used in the equation. Thus, while  $P(t)$  varies between 0 and 1, the log odds (or the logit) varies between negative and positive infinity. Simply put, the log odds can take on any value generated by the regression estimate. It need not be bounded by 0 and 1. Any standard statistical package with logistic regression capabilities (e.g. SPSS, SAS) will compute the log odds of the dependent variable.

The regression equation then takes the form:

$$\log (P(t) / (1 - P(t))) = a + b_1X_1 + b_2X_2,$$

where the dependent variable (on the left-hand side of the equation) is the logit or the log odds and  $b_1$  and  $b_2$  (on the right-hand side) are the parameter estimates of the impact of the explanatory variables,  $X_1$  and  $X_2$ , on the dependent variable. The intercept is represented by “ $a$ ” on the right-hand side. The coefficients,  $b_1$  and  $b_2$ , generated by the logistic regression indicate the change in the log odds for a one unit increment in  $X_1$  and  $X_2$  respectively.

### *An application*

Table 2, column 1 presents parameter estimates from an event history analysis of various conditions that influence the likelihood of a state’s passage of a protective workplace law for women, where such a protective law can be a maximum hours law, a minimum wage law, or a restriction on night work for women.<sup>23</sup> A maximum likelihood estimation technique is the best

21. *Ibid.*, p. 15; Allison, *Event History Analysis*, p. 17–18.

22. Eric A. Hanushek and John E. Jackson, *Statistical Methods for Social Scientists* (New York, 1977), pp. 187–189.

23. McCammon, “The Politics of Protection”, pp. 217–249 presents a more detailed analysis of similar data.

Table 2. Discrete-time event history analysis estimates of the effect of selected variables on the likelihood of a state adopting women's protective legislation, 1870-1930 (standard errors in parentheses)

|  | (1)                 | (2)                | (3)                | (4)                | (5)                |
|--|---------------------|--------------------|--------------------|--------------------|--------------------|
| Consumers' League  | 1.73***<br>(.383)   | 1.59***<br>(.494)  | 1.24***<br>(.324)  | 1.01***<br>(.378)  | 1.03<br>(.573)     |
| Competitive gubernatorial election                             | 3.23***<br>(1.22)   | 3.04***<br>(1.21)  | 2.42**<br>(1.02)   | 2.62**<br>(1.21)   | .87<br>(1.76)      |
| Proportion voting for Progressive Party presidential candidate | 8.52***<br>(1.87)   | 3.78<br>(2.21)     | 7.13***<br>(1.41)  | 6.77***<br>(1.75)  | 7.50***<br>(1.60)  |
| Women's suffrage   | 1.91***<br>(.510)   | 1.50***<br>(.552)  | 1.93***<br>(.404)  | 1.38***<br>(.490)  | 1.48***<br>(.510)  |
| Percentage of workers female                                   | 8.54***<br>(3.15)   | 7.20**<br>(3.33)   | 6.18**<br>(2.72)   | 4.83<br>(3.00)     | 1.45<br>(5.14)     |
| Legislative session  | 3.10***<br>(.748)   | 3.20***<br>(.752)  | 2.93***<br>(.583)  | 3.19***<br>(.770)  | 2.03***<br>(.769)  |
| 1870s  | — <sup>a</sup>      | -3.22***<br>(1.10) | —                  | —                  | —                  |
| 1880s  | —                   | -1.58***<br>(.645) | —                  | —                  | —                  |
| 1890s  | —                   | -1.36**<br>(.598)  | —                  | —                  | —                  |
| 1900s  | —                   | -1.82***<br>(.553) | —                  | —                  | —                  |
| 1920s  | —                   | -2.16**<br>(1.09)  | —                  | —                  | —                  |
| Number of past events  | —                   | —                  | —                  | 11.63<br>(22.21)   | —                  |
| Time elapsed since last event                                  | —                   | —                  | —                  | -0.05<br>(2.26)    | —                  |
| Constant   | -11.10***<br>(1.61) | -9.19***<br>(1.69) | -9.65***<br>(1.30) | -9.94***<br>(1.55) | -8.64***<br>(2.06) |
| N  | 1,585               | 1,585              | 1,747              | 1,747              | 2,384              |
| Likelihood-ratio chi-squared                                   | 103.8***            | 124.3***           | 145.2***           | 214.6***           | 52.4***            |
| Degrees of freedom   | 6                   | 11                 | 6                  | 8                  | 6                  |

\*\*  $p \leq .05$

\*\*\*  $p \leq .01$

<sup>a</sup> Variable not in the equation.

method for generating the parameter estimates and is used here.<sup>24</sup> The dependent variable in this analysis is coded as "0" for every year that a state did not enact a protective law and as "1" for the year, in which the

24. Paul D. Allison, "Discrete-Time Methods for the Analysis of Event Histories", in Samuel Leinhardt (ed.), *Sociological Methodology* (San Francisco, 1982), pp. 73-76.

state legislated any one of these types of law. After the passage of such a law, data for the following years are coded as missing data and, therefore, these years are not included in the analysis. As discussed above, the dichotomous dependent variable is transformed into a hazard rate, and the log odds of the hazard rate is modeled with logistic regression.

As can be seen in Table 2, all of the explanatory variables entered into the equation have a significant and positive influence on the probability of a state passing a protective law. The standard errors of the coefficients are provided in parentheses, and the significance of the coefficients is determined with a Wald statistic, which is the squared ratio of the coefficient to its standard error and has a chi-squared distribution. The standard error, the Wald statistic, and its probability will usually be supplied by the event history analysis statistical package.

The results in Table 2 show that state legislatures around the turn of the century were more likely to pass protective laws for women when a Consumers' League was organized in the state. This was a middle- and upper-class women's group that lobbied for workplace protection for women. Also, states were more willing to enact protective laws when certain political conditions prevailed, for instance when a recent gubernatorial election was competitive; perhaps this made those in office unsure of their political constituencies so that they searched for ways of insuring their support in the future. Protective laws were also more likely to be adopted when there was significant support for Progressive Party presidential candidates in the state and when women possessed full voting rights in the state. In addition, the greater the percentage of the labor force that was female, the more likely law-makers were to pass protective legislation, as if a more pressing need for the law were perceived. Finally, a finding that is, of course, obvious: legislatures were more likely to pass these laws when they were in session. This measure must be included in the analysis to control for the fact that only during these years could legislatures pass laws.

One can determine the "fit" of the model or whether the model containing the explanatory variables in addition to the intercept is a significant improvement over a model with only the intercept by comparing the log likelihoods for the two equations. The log likelihoods are compared with the test statistic, the likelihood-ratio chi-square ( $L^2$ ), which appears at the bottom of the table and is computed as two times the positive difference between the log likelihood for the model containing both the explanatory variables and intercept (or the "tested model") and the log likelihood for the model containing only the intercept (or the "constant rate model").<sup>25</sup> Typically, statistical programs compute  $L^2$  and its significance level for the researcher (both SPSS and SAS do this) or, at a minimum, they compute  $-2$  times the log likelihood for both the tested and constant rate models

25. Yamaguchi, *Event History Analysis*, p. 20.

which can then be used to compute  $L^2$  by simply finding their positive difference.

The model in column 1 of Table 2 has a likelihood-ratio chi-square of 103.8 with 6 degrees of freedom. The number of degrees of freedom is determined by the difference in the number of regressors in the model with the explanatory variables and in the model without them. An  $L^2$  of 103.8 is greater than the critical value for a .01 level of significance with 6 degrees of freedom, and thus a model including these explanatory variables is a significant improvement over a model containing only the intercept. This is not surprising given that all of the coefficients in the model are significant.<sup>26</sup>

The number of cases included in the analysis ( $N = 1,585$ ) is also reported at the bottom of the table. Note that the  $N$  for column 1 is substantially larger than the number of states included in the analysis because the unit of analysis is the state-year. Multiple observations are included in the data for each state over the period of analysis, 1870 to 1930. However, the number of cases is not 48 times 61 (or the number of states in the analysis times the number of years in the period of analysis, which would be 2,928) because state-years are excluded from the analysis after a state passes a protective law for women.<sup>27</sup> Forty states passed protective laws and thus have truncated data in the analysis.

## SOME MATTERS OF FURTHER INTEREST IN EVENT HISTORY ANALYSIS

### *Over-time variation in the hazard rate*

An assumption of the model in column 1 of Table 2 is that the hazard rate is constant over time. As was seen in Table 1, this is not the case. The hazard rate is higher during some years – especially in the 1910s – because a substantial number of states passed protective laws during these years. To test the validity of the assumption concerning a constant hazard rate, a set of dummy variables can be added to the analysis, in this case, each representing a decade-long period.<sup>28</sup> Five such dummy variables are added

26. See Hanushek and Jackson, *Statistical Methods*, pp. 65–68 for a general discussion of statistical inference and Hubert M. Blalock, *Social Statistics* (New York, 1979), pp. 280–292 for a discussion of the chi-square statistic.

27. In addition, state-years are excluded from the analysis for years prior to a territory becoming a state (see note 12). Also, there are a few data points missing for Georgia given the unavailability of data for gubernatorial elections in that state for 1884, 1885, 1910–1917, 1920, 1921, 1928–1930. Losing data for these years, however, does not censor the data. Georgia did not pass a protective law.

28. The decade variables take the value of “0” for all years except the particular decade the variable represents. For these years, the variable equals “1”. So for the 1870 measure, the variable equals “0” for all years except 1870 through 1879; for these years the value equals “1”. Individual year dummies could be used instead. For this analysis, however, which contains data over 61 years, this

to the model in column 2 of Table 2. The omitted decade is the 1910s, the decade in which the greatest number of states passed protective laws. The inclusion of these measures in the model allows one to assess whether the likelihood of a state passing a protective law is significantly different in the 1910s compared to the other decades.

Indeed, the results in column 2 show that the effect for all of the dummy variables is significantly smaller than that for the 1910s, meaning that the hazard rate is significantly lower in these other decades than it was during the 1910s. States were significantly more likely to pass women's protective laws during the 1910s than during the other decades, even after the other explanatory factors (listed in the table) are statistically controlled. It does appear, however, comparing the magnitude of the coefficients for the dummy variables over time, that the likelihood of a state passing a protective law increased over time roughly until the 1910s (although the coefficient for the 1900s does not fit this trend perfectly) but then decreased substantially in the 1920s, probably due to a saturation effect given that the bulk of states had passed such laws by the 1920s.

One can also use the model log likelihoods to determine if the model including the decade dummy variables is a statistically significant improvement over the model without the dummies. Two times the positive difference between the log likelihoods for the models in columns 1 and 2 is 20.5 (this is also the simple difference between  $-2$  log likelihood) with 5 degrees of freedom (the number of dummies added to the second model). The critical chi-squared value is 15.1 (for a .01 probability level). The computed difference is greater than this. Thus, the equation including the decade dummies significantly improves the fit of the model. (Note that the likelihood-ratio in column 2 does not represent this comparison but rather a comparison of the tested and constant rate models.)

Allowing the hazard rate to vary over time in column 2 also renders the proportion voting for Progressive Party presidential candidates insignificant. This suggests that concerning the impact of progressive period dynamics on the likelihood of states passing protective laws, the more important explanatory factor is over-time variation rather than cross-sectional variation. The over-time dummy variables had a significant effect on the passage of protec-

would be a burdensome number of variables to include in the analysis. Thus, the decade measures are used. See Allison, *Event History Analysis*, pp. 19–20 for a further description of this method. In addition, see Larry Isaac *et al.*, "Temporally Recursive Regression" and Isaac and Griffin, "Ahistoricism in Time-Series Analysis of Historical Process", pp. 873–890 for a discussion of recursive regression, another method of detecting over-time variation in quantitative analyses, specifically in the effects of the explanatory variables in time-series analysis. See Beth A. Rubin and Brian T. Smith, "Forged Ties: Cooperation and Conflict in the Metals Industries", *Social Science Research*, 21 (1992), pp. 115–133 for an application of recursive regression to event history analysis.

tive legislation in column 2 in that they capture the (over-time) difference between the progressive years (encompassed in the 1910 decade) and the other decades. On the other hand, the measure of the proportion voting for Progressive candidates for the most part is a measure of cross-state differences in Progressive Party voting in the 1912 presidential election, and it is no longer significant in the second model.<sup>29</sup>

### *Repeated events*

The analyses in columns 1 and 2 of Table 2 ignore some potentially important information. The dependent variable examined in these analyses measures only a state's first passage of a protective law. While 42 states adopted some form of protective legislation, 20 of these states enacted additional protective laws concerning women's employment. That is, some states passed two or even all three types of protective law (six states, in fact, enacted all three types of law – maximum hours, minimum wage and prohibitions on night work laws). That some states passed more than one type raises the issue of *repeated events*. Many events examined with event history analysis are events that repeat themselves. For instance, workers' decision to strike and the outbreak of rioting can both occur more than once for the same workplace or the same city respectively. In the analysis of protective legislation, states can adopt more than one form of law and thus can experience the repeated event of the passage of protective legislation.

To include information on repeated events in an analysis is a reasonably straightforward procedure. The researcher must code the dependent variable so that the multiple instances of the event are indicated by the measure. For time units after the occurrence of the first event, if a later event occurs, rather than the variable taking on missing values, the variable is assigned the value "0" for the time units before the occurrence of the second event and "1" for the time unit in which the second event occurs. If a third or more events occur, this coding procedure is repeated. Thus, for California, which passed a maximum hours law in 1911, a minimum wage law in 1913, and a restriction on night work in 1918, the dependent variable is coded as follows:

29. Researchers could also include regional dummy variables (or the equivalent cross-sectional measure) in an analysis to determine if the hazard rate varied across regions. But, as discussed earlier, analysts could also examine separate models for the different regions. Separate models would be particularly warranted if one expects the effects of the explanatory factors to vary across regions. But also, if one expects the effects to differ across regions, interaction terms between the regional dummies and the appropriate explanatory variable could be included as an alternative to the separate regional equations. In addition to the possibility of variation in the effects over time and across regions, the pattern of results may also vary for the different types of law (i.e. maximum hours laws, minimum wage laws and night work laws). This possibility is discussed below in the section entitled "Estimating multiple types of events".

for 1870–1910 the value is “0”,  
 for 1911 the value is “1”,  
 for 1912 the value is “0”,  
 for 1913 the value is “1”,  
 for 1914–1917 the value is “0”,  
 for 1918 the value is “1”,  
 for 1919–1930 the value is missing because the state had passed all three measures.

A state, then, can contribute more than one passage of law to the analysis. By including all repeated events in the protective legislation data, 162 observations are added to the analysis. This information is ignored in the analysis in columns 1 and 2 of Table 2 where repeated events are not included and in which, as can now be discerned, only the adoption of a “first” protective law is analyzed.

The results of an event history analysis where all instances in which a state passed a protective law are included are shown in column 3 of Table 2. As can be seen, including the later passages of law leaves the results virtually unchanged (compare columns 1 and 3; for simplicity the decade dummy variables are left out of the analysis). All variables remain significant and positive. The standard errors for the coefficients are slightly smaller in column 3 than in column 1 given that the N for the analysis in column 3 is somewhat larger. The results suggest that a similar set of causal dynamics governs both the passage of a state’s first protective law and the passage of later laws.

There are some complexities, however, associated with the inclusion of repeated events in event history analysis. The first issue that arises is the potential for a lack of independence among such events.<sup>30</sup> An analysis that includes repeated events must be free of dependence among such events. For the protective legislation example, this means that the passage of a second law in a state must not be dependent upon the passage of a first law (in the next section this possibility is examined further). Or, if the passage of early and later laws are correlated in some way, this relationship must be accounted for by the explanatory variables included in the analysis. If such independence between events cannot be insured, the standard errors for the coefficients may be biased.<sup>31</sup>

In order to minimize dependence among repeated events, the researcher can include two measures as explanatory variables in the analysis that control for the influence of a case’s past event history on the present event.<sup>32</sup> The

30. See Daniel J. Myers, “Racial Rioting in the 1960s: An Event History Analysis of Local Conditions”, *American Sociological Review*, 62 (1997), p. 101 for a discussion of this problem.

31. Allison, *Event History Analysis*, p. 54.

32. Trond Petersen, “The Statistical Analysis of Event Histories”, *Sociological Methods and Research*, 19 (1991), p. 299; Allison, *Event History Analysis*, p. 54; idem, “Discrete-Time Methods”, p. 93.



first is a measure of the number of past events that have occurred at the time of a repeated event. The second is a measure of the time elapsed since the occurrence of the last event (which is set equal to zero if no previous event has occurred). These two measures are included in the model in column 4 of Table 2. Here the dependent variable is the same as that in column 3 in that it indicates the passage of all protective laws in a state. However, the inclusion of the number of past events and time elapsed does not substantially alter the findings (although the impact of the percentage of the labor force that is female becomes insignificant), and neither measure is significant. This suggests that there is little interdependence among the different types of protective legislation or that the passage of one type of protective law is not likely to affect the passage of other types.

The second complexity that can occur with the inclusion of repeated events is that it is possible (and there may be sound theoretical reasons for believing) that the conditions leading to the occurrence of a first event are not the same as those producing later events.<sup>33</sup> In fact, the results in column 3 may be misleading if this is the case. This can easily be discerned by comparing separate analyses, one for first events (i.e. a state's first passage of a protective law) and one for later events (i.e. passage of second and even third protective laws). If the associated sets of predictors are similar across two such models, then the researcher can reasonably assume that the same causal processes govern both earlier and later events. The results in column 5 can be compared with those in column 1 to make this judgement. While column 1 provides an analysis of the passage of a state's first protective law, column 5 provides an analysis of only the passage of later laws. Some of the variables remain statistically significant across the two models (votes for Progressive candidates, women's suffrage and legislative session); others, however, are no longer significant (the presence of a Consumers' League, competitive elections and female workers). These differences suggest that the researcher would have to consider that the passage of later laws occurs under substantially different circumstances than the passage of a first protective law for women.

### *Estimating multiple types of events*

Probably a naive assumption of the models in Table 2 is that the causal processes influencing the passage of the three types of protective legislation that are included in the dependent variable are the same. It may be, rather, that the conditions giving rise to laws restricting the number of hours that women could work, for instance, differ from the forces resulting in a state

33. Michael T. Hannan and Glenn R. Carroll, "Dynamics of Formal Political Structure: An Event-History Analysis", *American Sociological Review*, 46 (1981), pp. 19–35 discuss this possibility in their work on changes in political structures from military regimes to multiparty structures.

legislature adopting a minimum wage policy for women, and these dynamics, in turn, differ from those producing restrictions on night work for women.

To examine whether these different events have different causal patterns requires the construction of distinct dependent variables for each of the types of protective legislation and the estimation of separate models for each of the types of law. In estimating these separate models, one can also explicitly test the hypothesis that a state's adoption of one type of protective legislation increases the likelihood that in time a state will enact another type of protective law. Once an initial law is passed, a "policy feedback" effect may make it likelier that a state passes additional reforms.<sup>34</sup>

To examine these possibilities, additional models are estimated and the results are presented in Table 3.<sup>35</sup> Column 1 provides the results for a model estimating the effects of the various explanatory variables on the passage of maximum hours laws; column 2, the results for minimum wage laws; and column 3, the results for laws restricting night work for women. As is readily apparent, the model works well in explaining the circumstances leading to the adoption of an hours law (column 1). All coefficients continue to have a positive effect. The model is less useful in explaining the passage of minimum wage laws (column 2). Only the proportion voting for Progressive Party candidates and full suffrage for women increase the likelihood that a state will enact a minimum wage law for women. Similarly, only the presence of a Consumers' League and years in which legislatures are in session significantly increase the chances of a state passing a law regulating night work (column 3). The model does not work well for this last type of law either.

In addition, while the results in column 1 are robust for maximum hours laws, further analyses (not shown) show that the findings in columns 2 and 3 are highly sensitive to the presence or absence in the model of specific explanatory variables, suggesting that adequate specification of the processes giving rise to minimum wage laws and restrictions on night work need considerably more historical study.<sup>36</sup>

Many states enacted hours laws before minimum wage laws and restrictions on night work. Thus it is possible that the passage of an hours law may have eased the way for the enactment of the other two types of law. To test this idea, the minimum wage and night work models are re-estimated in columns 4 and 5 respectively, and the maximum hours law variable is

34. Theda Skocpol, *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States* (Cambridge, 1992), pp. 57–60.

35. For a lengthier treatment of event history analysis involving multiple events, see Allison, *Event History Analysis*, pp. 42–50.

36. See McCammon, "The Politics of Protection", pp. 234–238 for an analysis of the conditions under which legislatures adopted minimum wage laws.

Table 3. *Discrete-time event history analysis estimates of the effect of selected variables on the likelihood of a state adopting women's maximum hours, minimum wage and night work legislation, 1870-1930 (standard errors in parentheses)*

| Dependent variable   | (1)<br>Max. hours   | (2)<br>Min. wage   | (3)<br>Night work  | (4)<br>Min. wage  | (5)<br>Night work  |
|--|---------------------|--------------------|--------------------|-------------------|--------------------|
| Consumers' League  | 1.38***<br>(.379)   | -.18<br>(.832)     | 1.54***<br>(.596)  | -.22<br>(.829)    | 1.22**<br>(.612)   |
| Competitive gubernatorial election                             | 2.87***<br>(1.17)   | .92<br>(2.38)      | -.78<br>(1.27)     | .55<br>(2.44)     | -1.24<br>(1.33)    |
| Proportion voting for Progressive Party presidential candidate | 8.58***<br>(1.80)   | 10.41***<br>(1.98) | 1.54<br>(2.87)     | 9.91***<br>(2.04) | .90<br>(2.90)      |
| Women's suffrage   | 1.43***<br>(.474)   | 2.21***<br>(.684)  | .51<br>(.619)      | 2.01***<br>(.719) | .08<br>(.649)      |
| Percentage of female workers                                   | 8.18***<br>(3.11)   | 1.91<br>(3.33)     | 2.50<br>(4.97)     | -.46<br>(8.04)    | .46<br>(5.41)      |
| Legislative session  | 3.04***<br>(.742)   | 8.61<br>(19.3)     | 1.59**<br>(.771)   | 8.62<br>(19.3)    | 1.63**<br>(.772)   |
| Passage of maximum hours law lagged 1 yr                       | -a                  | -                  | -                  | .67<br>(.811)     | 1.21<br>(.714)     |
| Constant   | -10.62***<br>(1.55) | -16.01<br>(19.5)   | -6.70***<br>(1.65) | -15.54<br>(19.5)  | -6.41***<br>(1.68) |
| N  | 1,618               | 2,425              | 2,331              | 2,425             | 2,331              |
| Likelihood-ratio chi-squared                                   | 93.5***             | 62.9***            | 16.8**             | 63.6***           | 19.8***            |
| Degrees of freedom   | 6                   | 6                  | 6                  | 7                 | 7                  |

\*\*  $p \leq .05$

\*\*\*  $p \leq .01$

<sup>a</sup> Variable not in the equation.

included as an explanatory variable in both. The variable is lagged by one year because in some states hours laws were passed in the same years that the other two types of law were passed. Lagging the hours variable insures that the analyses examine only cases where an hours law was adopted prior to the enactment of either a minimum wage or night work law.<sup>37</sup> These

37. The same hours law variable used as a dependent variable in column 1 of Table 3 cannot be used as an explanatory variable in the analyses in columns 2 and 3 because the variable used in

findings indicate that the earlier passage of a maximum hours law did not significantly increase the likelihood that a state would adopt a minimum wage law or restrictions on night work for women, confirming the conclusions of the last section that the various types of law are not unduly dependent on one another.

## CONCLUSION

Event history analysis is a quantitative method that allows researchers to investigate the causes of the occurrence or non-occurrence of historical events. This discussion has presented various instructions on how to conduct an event history analysis along with a discussion of some of the complexities associated with the method. An advantage of the method over the case study is that event history analysis allows one to compare cases in which the event occurs with cases in which the event does not occur, thereby more systematically uncovering the causal dynamics that produce or do not produce the events of interest. As is often the case in historical research, the analyst chooses a “positive” case to unravel the causal dynamics producing a particular historical outcome – that is, a case in which the historical event occurs. Event history analysis, on the other hand, is a method that requires the inclusion not only of positive cases, but negative cases as well. The method necessitates that the researcher compare cases in which the event occurs with those in which the event did not occur to more fully uncover the reasons underlying the event’s occurrence.

While the data requirements are rather rigorous – both over-time and cross-sectional quantitative data are needed – the method does allow a level of analysis that is unavailable to the researcher who examines one or just a few cases. Event history analysis allows a thorough and rigorous assessment of *many* observations and the characteristics associated with those observations to gauge why the historical event of interest occurs in some of those cases and why it does not occur in others. Possibly the entire population of social entities at risk of the event occurring can be included in the analysis. In the study presented here, all of the existing US states were incorporated to examine the historical circumstances that led them to adopt or not adopt women’s protective legislation. An examination of just one or even a handful

column 1 is assigned missing values for all years after a state passed an hours law. The hours law variable used as an explanatory variable in columns 2 and 3, rather, must be assigned the value of “0” for all years prior to the state’s passage of an hours law and “1” for the year in which an hours law is passed *and* all later years as well. Note also that lagging variables with time series data pooled across states such as that found here presents a difficulty. The lagging procedure assigns the value for the year 1930 for the preceding state to the year 1870 (see Figure 1). This, however, can be corrected manually. No state passed an hours law in 1870, so “0” (rather than the 1930 value) is assigned to 1870.

of states could not offer the same scope of investigation that event history was able to offer in this case.

Event history analysis, however, is not without its disadvantages. Given the rigorous data requirements, it is possible that the researcher will not have the needed data for all observations or over the entire time period. And the work involved in collecting such information even when it is available can present a task that itself can be all-consuming. As Tilly points out concerning quantitative methods generally:

the scale and complexity of such an investigation produce important periods when the researchers are so preoccupied with problems of coding, file construction, statistical procedure, computer techniques, and coordination of the whole effort that they practically lose contact with the people, events, places, and times they are studying.<sup>38</sup>

And certainly this too can be true at times of event history analysis. In the end, however, it is the researcher who must decide if enough quantifiable information exists to make the investigation viable and it is the researcher who controls whether coding and statistical technique dominate the analysis or whether event history analysis is simply a useful tool that can be utilized to reveal insights about historical change.

Event history analysis certainly is not useful in all cases of historical inquiry, but if the historical question posed by the researcher is one that seeks to uncover why particular events occur and if the appropriate data are available, even if data gathering is an arduous task, then event history analysis can be a valuable and worthy research tool.

#### ANNOTATED BIBLIOGRAPHY

For the most accessible discussions of discrete-time methods in event history analysis, see:

Allison, Paul D., "Discrete-time Methods for the Analysis of Event Histories", in Samuel Leinhardt (ed.), *Sociological Methodology* (San Francisco, 1982), pp. 61–98.

Allison, Paul D., *Event History Analysis: Regression For Longitudinal Event Data* (Newbury Park, 1984).

Yamaguchi, Kazuo, *Event History Analysis* (Newbury Park, 1991).

There are various recent historical studies that employ discrete-time methods. See, for example:

38. Charles Tilly, *As Sociology Meets History* (Orlando, 1981), p. 61. Elsewhere in this discussion, Tilly extols the virtues of quantitative methods, as well as being somewhat realistic about their shortcomings.

- Kim, Chulsoo, "Determinants of the Timing of Social Policy Adoption", *Journal of Sociology and Social Welfare*, 23 (1996), pp. 5–29.
- McCammon, Holly J., "The Politics of Protection: State Minimum Wage and Maximum Hours Laws for Women in the United States, 1870–1930", *The Sociological Quarterly*, 36 (1995), pp. 217–249.
- Minkoff, Debra C., "The Organization of Survival: Women's and Racial-Ethnic Voluntarist and Activist Organizations, 1955–1985", *Social Forces*, 71 (1993), pp. 887–908.
- Mizruchi, Mark S. and Linda Brewster Stearns, "A Longitudinal Study of the Formation of Interlocking Directorates", *Administrative Science Quarterly*, 33 (1988), pp. 194–210.
- Olzak, Susan and Elizabeth West, "Ethnic Conflict and the Rise and Fall of Ethnic Newspapers", *American Sociological Review*, 56 (1991), pp. 458–474.
- Pavalko, Eliza K., "State Timing of Policy Adoption: Workmen's Compensation in the United States, 1909–1929", *American Journal of Sociology*, 95 (1989), pp. 592–615.
- Western, Bruce, "A Comparative Study of Working-class Disorganization: Union Decline in Eighteen Advanced Capitalist Countries", *American Sociological Review*, 60 (1995), pp. 179–201.

For detailed treatments of continuous-time methods, see:

- Allison, Paul D., *Event History Analysis: Regression For Longitudinal Event Data* (Newbury Park, 1984).
- Cox, David R. and D. Oakes, *The Analysis of Survival Data* (New York, 1983).
- Petersen, Trond, "The Statistical Analysis of Event Histories", *Sociological Methods and Research*, 19 (1991), pp. 270–323.
- Petersen, Trond, "Time-Aggregation Bias in Continuous-Time Hazard-Rate Models", in Peter V. Marsden (ed.), *Sociological Methodology* (Cambridge, 1991), pp. 263–290.
- Tuma, Nancy Brandon and Michael T. Hannan, "Dynamic Analysis of Event Histories", *American Journal of Sociology*, 84 (1979), pp. 820–854.
- Tuma, Nancy Brandon and Michael T. Hannan, *Social Dynamics: Models and Methods* (Orlando, 1984).
- Yamaguchi, Kazuo, *Event History Analysis* (Newbury Park, 1991).

Some examples of historical research utilizing continuous-time methods are:

- Carroll, Glenn R. and Yangchung Paul Huo, "Organizational and Electoral Paradoxes of the Knights of Labor", in Glenn R. Carroll (ed.), *Ecological Models of Organizations* (Cambridge, 1988), pp. 175–193.
- Hannan, Michael T. and Glenn R. Carroll, "Dynamics of Formal Political

- Structure: An Event-History Analysis”, *American Sociological Review*, 46 (1981), pp. 19–35.
- Hannan, Michael T. and John Freeman, “The Ecology of Organizational Founding: American Labor Unions, 1836–1985”, *American Journal of Sociology*, 92 (1987), pp. 910–943.
- McCarthy, John D., Mark Wolfson, David P. Baker and Elaine Mosakowski, “The Founding of Social Movement Organizations: Local Citizens’ Groups Opposing Drunken Driving”, in Glenn R. Carroll (ed.), *Ecological Models of Organizations* (Cambridge, 1988), pp. 71–84.
- Myers, Daniel J., “Racial Rioting in the 1960s: An Event History Analysis of Local Conditions”, *American Sociological Review*, 62 (1997), pp. 94–112.
- Olzak, Susan, “The Political Context of Competition: Lynching and Urban Racial Violence, 1882–1914”, *Social Forces*, 69 (1990), pp. 395–421.
- Rubin, Beth A., “Limits to Institutionalization? A Sectoral Analysis of U.S. Strike Settlements, 1960–1977”, *Research in Social Stratification and Mobility*, 11 (1992), pp. 177–202.

Various sources offer treatments of censoring in event history analysis. See, for example:

- Tuma, Nancy Brandon and Michael T. Hannan, “Approaches to the Censoring Problem in Analysis of Event Histories”, in Karl F. Schuessler (ed.), *Sociological Methodology* (San Francisco, 1979), pp. 209–240.
- Yamaguchi, Kazuo, *Event History Analysis* (Newbury Park, 1991).