"Temporal Dependence and International Conflict"

October 27, 2000

Dan Reiter
Department of Political Science
Emory University
Atlanta, GA 30322
dreiter@emory.edu
(404) 727-0111

Presented at the annual meeting of the Peace Science Society, New Haven, CT. Thanks to Christopher Zorn, and for data to Bruce Russett, and John Oneal. All remaining errors are mine.

"Temporal Dependence and International Conflict"

Event history analysis, the examination of how the amount of time spent in a single state as well as external factors affect the transition to a second state, is seeing increasing usage in political science. This is good news, as it indicates a general sensitivity to the necessity of using econometric techniques appropriate to the nature of both the data and the hypotheses being tested. Further good news is that political scientists are becoming increasingly sophisticated in their use of event history techniques.

Like all powerful tools, however, event history analysis illuminates best when used appropriately. This paper makes two points about potential pitfalls in the application of event history techniques. First, scholars need to be sensitive to issues of left-censoring when conducting event history analysis. Left-censoring occurs when the beginning of the phase under study is unknown; for example, if the topic is the tenure of a government, left-censoring occurs if the point at which the government took power is unknown or omitted from the data. Left-censoring is a real problem in the application of event history techniques, and how one treats (or does not treat) left-censoring can have important effects on empirical results. Second, scholars need to think about how to handle the issue of multiple transitions to this second stage, or, more briefly, multiple failures. Often, the phase under examination will have multiple failures, an example being a peace treaty suffering multiple violations from the time it is put into effect. There are multiple ways of treating multiple failures, and one's decision as to how to treat them can have important effects on results.

The rest of this paper proceeds in three parts. The first presents event history analysis briefly, before laying out the left-censoring and multiple failures issues themselves. The second part presents an empirical application to the question of the causes of international conflict. In this section, I explain why left-censoring and multiple failures are both matters of concern regarding the application of event history analysis to the study of international relations. I then reexamine a data set of international conflict from 1950-1985 which has been discussed in two previous papers (Oneal and Russett, 1997; Beck, Katz, and Tucker, 1998). Using two different methods of event history analysis, one discrete and one continuous, I demonstrate that different treatments of the left-censoring and multiple failures issues can generate substantively significant differences in results. Specifically, I find that different treatments of these two issues can generate different conclusions as to whether economic prosperity, contiguity, common alliance membership, and international trade inhibit international conflict. The third section offers conclusions and suggestions as to how to treat these two issues.

I. Event History Analysis, Left-Censoring, and Multiple Failures.

Some political science theories call for examining what causes a phenomenon to shift from one phase to another, such as when peace terminates in war, when leadership's hold of power terminates in ouster, when alliances are formally dissolved, and so forth. For these types of questions, event history analysis (or hazard, survival, or duration analysis) is appropriate. The basic approach is to model as a function the lifespan of the phase of interest, or put differently, the incremental risk of a particular event happening at

each point in time in the life of a subject; the occurrence of the event is called failure (see Box-Steffensmeier and Jones, 1997; Bennett, 1999). Event history analysis has been applied to understand a number of important questions in political science, including cabinet stability and the duration of governments (King et al, 1990; Alt and King, 1994; Warwick and Easton, 1992), the adoption of hate crime laws (Grattet et al, 1998), the adoption of education reform (Mintrom, 1997), and the adoption of state lotteries (Berry and Berry, 1990). Event history analysis has also been applied to questions in international relations, such as the duration of alliances (Gaubatz, 1996; Reed, 1997; Bennett, 1997), wars (Bennett and Stam, 1996, 1998), enduring rivalries (Bennett, 1996, 1998b), peace following third party intervention into conflicts (Diehl, Reifschneider, and Hensel, 1996), the international sources of democracy (Reiter 2001), and the tenure of leaderships after war (Bueno de Mesquita and Siverson, 1995).

There are a variety of types of hazard analysis, many of which differ primarily on the basis of differing assumptions about the nature of duration dependence, that is, the relationship between the time elapsed since the beginning of the phase and the likelihood of failure at any given point. Another important distinction is between continuous time models, which assume that failure can occur at any point in time, and discrete time models, which assume that failure can occur only at certain intervals. My aim in this paper is not to describe the array of different approaches. However, there is one event history approach which deserves brief summary because of its recent application to international conflict, the empirical area examined in the next section of this paper. This approach was introduced recently by Beck, Katz, and Tucker (1998, hereafter BKT). It uses a conventional logit model, but includes temporal dummy variables to account for a

particular case being located at a different point in the lifespan of a phase. A series of dummy variables are included to represent all the possible lengths of time since the beginning of the dyad or the last failure. A variation of this approach is to create a cubic spline function which represents all of the time dummies with a smaller number of variables, thereby saving degrees of freedom. This solution is a form of an event history approach, which "depends on the recognition that BTSCS [binary time series cross sectional data] are *identical* to grouped duration data" (BKT, 1264, italics in original).[1] A number of scholars have begun to use the BKT approach in application to international conflict (for example, Bennett and Stam, 1998a; Enterline, 1998; Dassel and Reinhardt, 1999).

Event history analysis works most straightforwardly in the analysis of data in which the beginning of the phase is known, the end is known, and there is only one failure. An example would be the study of human mortality, when a person's dates of birth and death are known, and he or she can only die once. In the study of politics, however, data are often not this simple.

First, the beginning and end of the phase in question are often not known. Consider a study with data ranging from t(0) to t(k), with subjects A, B, C, and D, as displayed in Figure 1 (see Blossfeld and Rohwer 1995: 34-5).

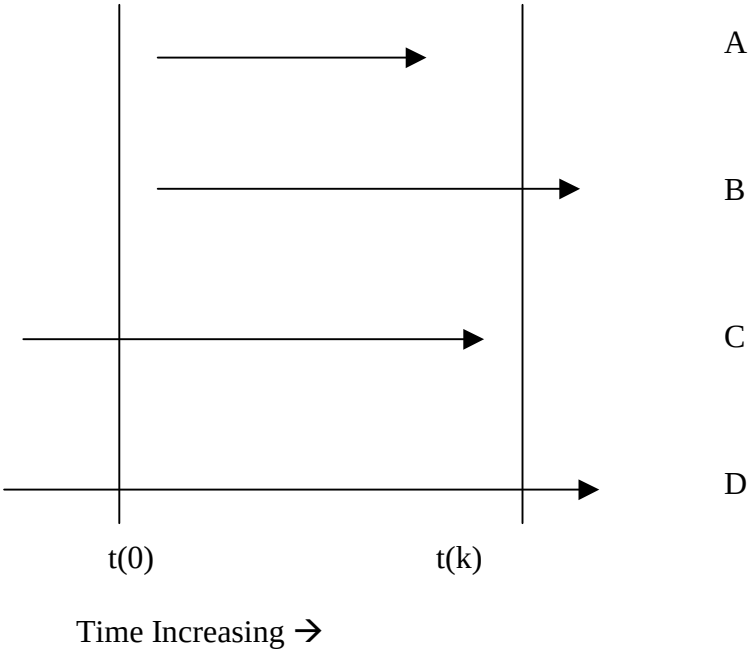

Figure 1: Left- and Right-Censoring in Event History Analysis

Subject A is unproblematic: both the beginning and end of the phase are known, so there is a complete set of information about its appearance, non-failure, and eventual failure. The beginning of subject B is known, but its failure is not, as information has not been collected after t(k). This is known as right censoring. While it is not known when there is failure, the rest of the information can be used because it is known that there was no failure from the beginning of the phase up to t(k).

For subject C, the point of failure is known, but the point of beginning is not known; this is known as left-censoring.[2] This presents more severe information problems, as it is unknown exactly where in the hazard function--the beginning, middle, or end--this panel should begin. Again, event history analysis attempts to build an understanding of when in the lifespan of a phase failure is likely to occur. When there is left-censoring, it cannot be known exactly where in the lifespan the observed data rests,

making it quite difficult to make use of the data which is available. For example, if one observed an individual living seven years before dying of a stroke, and one did not know the age of the subject, then drawing conclusions as to the shape of the hazard function would be quite difficult, as the individual might have died at the age of 17, 47, or 87. Left-censoring is "practically intractable" in the sense that bias is essentially unavoidable, unless the hazard rate is constant, the start times are assumed to be constant, or left-censored cases are omitted (Guo, 1993: 217; Blossfeld and Rohwer, 1994: 34). The nature of the bias depends on whether the left-censored case is completely omitted (which introduces sample bias), or if it is assumed to begin at point t(0) when observation begins (systematic measurement error). Lastly, Subject D has both right and left censoring.

Aside from left-censoring is the issue of "multiple failure," that is, analyzing the course of a phase in which there can be multiple failures, or multiple occurrences of a single event. Put differently, allowing for multiple failures means that the occurrence of an event does not necessarily require that the subject exit the data set (on multiple failures, see Blossfeld and Hamerle, 1989). An example might be examining how long it takes for a elections to be called against a governing party; if the governing party wins these elections and remains in office, then it would be inappropriate to stop collecting data after the first election.

There are basically three ways to deal with this matter. The first is to allow multiple failures within a single subject, that is, to model the hazard function on the basis of possible multiple failures. Within this approach, one can assume that past failures affect the likelihood of future failures—for example, unsuccessful elections make future

calls for election more likely—or not. A second approach is to allow only one failure per subject, but then to assume that a new subject starts after each failure. In other words, a new subject is restarted after each failure, so time to failure is measured (and modeled) since the last failure. Third, one could model the time to only the first failure. This is of course appropriate when the question asked is focused on the first failure rather than failure in general, such as a medical concern with first heart attacks rather than heart attacks in general, or individuals' first elections to public office.

The choice of approach to multiple failures should be driven by theory. The first approach, allowing for multiple failures, is more appropriate when the theory predicts that past experience is not washed away by each failure. In the elections example, it may be that unsuccessful elections heighten partisanship, making future challenges more likely. The second approach is more appropriate when the theory predicts that past events have no effect on the likelihood of failure. For example, if one assumes that there is no difference between an election that overthrows the governing party and one which does not, then it might be more appropriate to examine each post-election phase as a separate case. I will discuss in greater detail theoretical issues involving multiple failures and international conflict in the next section.

II. Left Censoring, Multiple Failures, and International Conflict.

The occurrence of war is easy to observe, but how does one observe the occurrence of peace? The first modern scholars satisfied themselves with collecting systematic data merely on the occurrence of war, an approach which of course precludes

the unbiased testing of factors which lead to peace or war (for example, Small and Singer, 1976). In the 1980s, scholars took a step forward by collecting data on crises as well as wars (for example, Gochman and Maoz, 1984). The approach was to collect data on all crises, which would then permit testing hypotheses which predict when crises escalate to war and when they do not. Focusing on the crisis-war phase is of course an improvement over the earlier approach, as it includes data of states not at war (though in crisis) and states at war. Remaining, however, was the question of how states got into crisis in the first place, which potentially introduces selection bias problems (Fearon, 1994; Achen, 1986). Starting in the late 1980s, scholars began to build data sets which included all states (or all pairs of states, or dyads) at peace (see, for example, Maoz and Abdolali, 1989). This permitted analysis of the peace-crisis phase, and permitted peace-war rather than crisis-war analysis, thereby eliminating that source of selection bias. This approach of analyzing all states for some time period, or all dyads, or all politically relevant dyads, took off in the 1990s (see, for example, Maoz and Russett, 1993; Bremer, 1993; Oneal and Russett, 1997; Bennett and Stam, 1998a).

Significantly, looking at all dyads over a certain time period means examining data with time series or panel characteristics. However, many of these studies used logit or probit techniques, because of the dichotomous nature of the dependent variable (peace/crisis, peace/war, etc.), but did not include corrections for potential bias introduced by time dependence within panels. Notably, some scholars were aware of the problems of using time series techniques with logit or probit, and argued that such an approach was preferred to the alternatives of generalized least squares or event history analysis (Huth, 1996: 265-7). Most recently, scholars have begun to use event history

analysis as a means of dealing with panel data with a dichotomous dependent variable. BKT applied their temporally adjusted logit approach to understanding the causes of conflict among dyads. Other scholars have also applied more conventional hazard techniques such as Cox and Weibull models, generating some interesting empirical results (Raknerud and Hegre, 1997; Werner, 1998).

The application of event history analysis to the study of the causes of international conflict is likely to confront issues of left-censoring and multiple failures. Regarding left-censoring, many studies will endeavor to analyze the behavior of dyads (or individual states) across time but will not have data going back to the birth of the dyad. Many studies will use Correlates of War (Singer and Diehl, 1991), MID (Gochman and Maoz, 1984), and Polity III data (Jaggers and Gurr 1995), which go back to 1816 for the first two and 1800 for Polity III. Many states entered the nation-state system after 1816, reducing the issue of left censoring to just a handful of states and dyads. However, the states which did enter the nation state system before 1816 (and hence would experience left censoring) are especially interesting because they are major powers and especially prone to conflict; they include the United States, Britain, France, Spain, Germany (Prussia), Italy (Sardinia), and Russia, to name a few. Some of these dyads go back far before 1816; for example, the Britain-Spain dyad extends back arguably to the eleventh century.

Unfortunately, the left-censoring problem is bigger than just a handful of dyads for most analyses. There are many important data sets which begin later than 1816, meaning that left censoring will affect an even larger fraction of dyads. Several crisis data sets begin in the twentieth century: the Interstate Crisis Bargaining data set begins in

1918 (see Brecher, Wilkenfeld, and Moser, 1988); the SHERFACS data set begins in 1945 (Sherman, 1994); the COPDAB data set begins in 1953 (Azar, 1980); Blechman and Kaplan (1978) begin in the postwar period as well. Limits on trade data availability force studies on the causes of trade to begin in 1905 for great powers and even later for more inclusive studies (Gowa and Mansfield, 1993; Bliss and Russett, 1998; Morrow, Siverson, and Tabares, 1998); such limits also force studies on the connections between trade and war to start in 1950 (see Oneal and Russett, 1997, 1999a, b; Barbieri, 1996 uses data from 1870-1938). United Nations voting patterns as indicators of national preferences are of course limited to the postwar period (Gartzke, 1998). Limitations on the availability of data on general economic performance, internal political disorder, and public opinion routinely limit studies of diversionary war to the postwar period (see, for example, Gelpi, 1997; Meernik and Waterman, 1996; Miller, 1995).[3] Such left-censoring will of course also affect new data sets which cover limited spans of time.

In the context of studying international conflict, the problem with left-censoring is that it risks introducing systematic selection bias; dyads which have had long terms of peace before the beginning of the data set "lose" those long stretches of peace. In other words, it makes peaceful dyads look much less peaceful. Consider the US-Canada dyad in 1951. If the data set began in 1950, then the thirty-one years of peace between the US and Canada going back to Canada's emergence as a nation state in 1920 are not accounted for in the analysis, and the dyad instead is coded as having one year of peace.

The simplest approach to left-censoring is to truncate the data on the left, and impose the artificial assumption that all dyads begin their lives at the start of the observed data set (point t(0) in Figure 1). BKT take this approach; their data run from 1950-1985,

and for dyads which began before 1950 they start the temporal dummy clock in 1950, so that for example the US and Canada in 1951 are assumed to have gone one year since the last dispute.

Two points are worth making about left-truncation. First, if all the dyads began at the same time in the (unobserved, pre-data set) past, then this avoids the problem, in the sense that estimating time-varying covariates (such as what the capability ratio between two states was in a given year) can proceed for the data which is available (Jenkins, 1995). Unfortunately, this assumption does not hold for this type of problem, as dyads are born in different years (a dyad's beginning being conventionally defined as when both members of the dyad are recognized as members of the international system). Second, even if dyads began in different years, then one could assume that duration begins at the time of the data set if the beginning of the data set coincides with a beginning of a new era or international order (BKT, 1272). Hence, if one assumes that the end of World War II founds a new security order, then perhaps all dyads can be assumed to begin in 1946. This assumption is interesting and potentially valid, but it requires theoretical exposition as it may apply to some hypotheses and not others. Specifically, this assumption may be more defensible for system-level hypotheses than for dyad-level hypotheses. While the end of World War II may herald a new era for the effects of polarity on conflict, it is unlikely to herald a new era for the effects of, for example, dyad-level trade on the likelihood of dyad-level conflict.

Multiple failures are also an issue for the study of international conflict. Dyads which are conflictual (such as enduring rivals) have more than one dispute in their history. Choice of treatment of multiple failure needs to be theoretically driven. If one is

interested in the duration of peace after conflict, then it may be appropriate to restart each conflict after each dispute, if one proposes for example that each dispute washes away progress towards the improvement of relations established by the last spell of peace. However, if one instead hypothesizes that a history of past disputes makes future disputes more likely, then the potential for bias emerges. Specifically, the single failure approach would conflate dissimilar cases, for example a young dyad (for example, two newly decolonized African countries in 1962) and a dyad with long-standing rivalry soon after its latest dispute (for example, US-Soviet Union in 1964). Hence, if both dyads experience conflict then using the single failure approach this would be seen as evidence favoring the "past disputes make future disputes more likely" hypothesis, though in the African case the experience of a dispute is in fact evidence against the hypothesis as the dyad has no history of past disputes.

If one wishes to assume that past behavior and events have a persevering effect on the likelihood of recurring disputes, then a multiple failure approach may be preferable. Similarly, if one has a theory about how ongoing interaction between states generates certain patterns of behavior, then one would want to track the evolution of the dyad across time allowing for the multiple inclusion of disputes. One might assume that there is memory only of massive disputes, such as interstate or systemic wars. One might assume that there is a memory function which decays as time extends back; so, a dyad might get a hostility score which is raised for each past dispute, adjusted for how long ago it occurred. Of course, this approach begins to look like more conventional time series models which assume there is autoregression in the error term or simply include a lagged dependent variable on the right hand side. Significantly, current statistical

packages allow for probit or similar models with panel data.[4] Again, choice of approach should be driven by theory; how past disputes affect behavior is itself a complex theoretical question which has attracted much attention, and it is not limited to a particular theoretical perspective (see, for example, Jervis, 1976; Levy, 1994; Reiter, 1996).

Different approaches to left-censoring and multiple failures can generate different results. I provide an example of this kind of variation in results of an analysis of a data set of all politically relevant dyads from 1950-1985 initially generated by Oneal and Russett (1997) and used by BKT.[5] Using this data facilitates comparison with methods used by BKT. I summarize the data set very briefly, as it is discussed in greater detail in these two sources and elsewhere. Each case is a dyad year, for all dyads which are politically relevant, that is, when one is a major power or when the dyad is geographically contiguous. The dependent variable is 1 if there was a MID between the two states during the year in question, 0 otherwise. Oneal and Russett (1997) provide extensive description of the independent variables; briefly, they include:

Democracy: Each state gets a score of –10 to 10 for how democratic it is, higher numbers being more democratic, and the dyadic score is the lower of the two scores.

Growth: Each state gets a value for its average growth in per capita gross domestic product (GDP) over three years, and the dyadic score is the lower of the two scores.

Alliance: This equals 1 if both countries are formally allied, or if both are allied with the United States; 0 otherwise.

Contiguity: This equals 1 if the two states are contiguous, 0 otherwise.

Capability Ratio: This expresses the asymmetry between the two states, and it gets higher when one state becomes more powerful than the other, and is 1 when both states are equal. Correlates of War data measure capabilities.

Trade: Each state gets a score for its imports and exports from and to the other state in the dyad divided by its GDP, and the dyadic score is the lower of the two scores.

Peace Years: Number of years since the members of the dyad were on opposite sides of a MID. If the two states have never been on opposite side of a MID, then it is a count of the number of years since both states emerged as recognized members of the international system.

Democracy, Growth, Alliance, Capability Ratio, and Trade are hypothesized to be negatively related with the dependent variable; Contiguity is hypothesized to be positively related to the dependent variable.

I explore the effects of choosing different approaches to left-censoring and multiple failures by examining two types of event history models: the BKT temporally adjusted logit model, and a standard Cox model. There are important differences between these two models; the BKT model has a logit functional form, while the Cox model estimates the hazard rate of failure. Also, the BKT model is an example of a discrete time model, as it assumes that failure can only happen at discrete points in time. The Cox model on the other hand is a continuous time model, meaning that it allows for failure at any point in time. A semi-parametric model like the Cox model leaves the shape of the hazard rate relatively unspecified before estimation, an advantage if we do not have a theoretical reason to specify one pattern of time-dependence over another and

are interested primarily in the effects of covariates (Blossfeld and Rohwer, 1994: 212; Box-Steffensmeier and Jones, 1997: 1432).

I was able to replicate the results of Oneal and Russett (1997) and BKT.[6] In Tables 1 and 1a, I present various approaches to the two different problems using BKT. In Model 1, I first present the results of Oneal and Russett's (1997) unadjusted logit analysis. Model 2 uses the BKT model and produces essentially similar results as BKT. Two substantively significant differences between the two models are that neither Growth nor Trade is statistically significant in Model 2; the effects of trade on conflict have in particular been a source of scholarly dispute (see Barbieri, 1996; Oneal and Russett, 1999a, b).

[Table 1 about here]

Model 3 retains the single failure assumptions, but does not truncate the data. Specifically, for dyads which begin before 1950, I assume that the dyad begins at the first pre-1950 MID, or at the beginning of the dyad itself if the dyad experienced no MIDs at all. This approach provides fuller information on dyads beginning before 1950, as it counts the actual number of years of peace rather than assuming that peace begins in 1950. If this change is made, the trade variable becomes negatively and significantly correlated with the likelihood of conflict. These results are in contrast to the BKT (1272-3) expectation that "results should be relatively insensitive to a few differences in judgment on this issue [of left censoring]."[7]

In Model 4, we present a second solution to left-censoring: omission of all left-censored cases (that is, all dyads which began in 1950 or before). This approach has the advantage of eliminating the left-censoring bias with no additional data requirements.

This approach produces similar results as Model 2, that used by BKT, except that Growth is significant and negative. However, there is a tremendous loss of information, as nearly two-thirds of the cases are dropped. Additionally, this reduction is systematic, as the age of dyads is not randomly distributed. Two thirds of the remaining sample of 7227 dyads contain two African countries (due to the wave of decolonization in the 1950s and 1960s), and there are virtually no European-European or major power-major power dyads in the smaller sample. In short, the omission solution to left-censoring creates even bigger problems of information loss and sample bias.

Table 1a presents further results for different solutions to left-censoring and multiple failure data. The left most column is of Model 2 taken from Table 1: I include it for easy reference. Model 5 returns to truncated data—all dyads begin in 1950 at the earliest. Differing from Model 3, however, this model presents one approach which allows for multiple failure in a single dyad. Specifically, rather than restarting a new dyad after each dispute, I record the age of dyad, again where no dyad can begin before 1950. Note that Peace Years is included as a control variable, which is the number of years since the last dispute. This dispute helps distinguish between older dyads with peaceful histories and older dyads with more conflictual histories. In Model 5, both Growth and Trade are significant and negative. Model 6 combines both techniques, leaving the data untruncated and not restarting the count after each dispute. In this model, Trade is negative and significant, though Growth and Contiguity are insignificant.

[Table 1a about here]

Model 7 presents a second approach to dealing with multiple failures. This model restarts the count after each dispute, but it also records the number of disputes in the age

of the dyad.[8] This approach allow the examination of two different dimensions of time dependence, the frequency of past conflict and the length of time since the last conflict. More specifically, it facilitates the more precise separation of different kinds of dyads. Return to our comparison between a 1962 African dyad which has seen no conflict since its birth in 1960, and the 1964 US-Soviet dyad which has seen no conflict since 1962, but experienced substantial conflict before that. A similarity between the two dyads is that neither dyad has had an opportunity to build up a trusting relationship over a long period of peace. Such long periods of peace, according to the security communities literature, contribute to the kinds of social learning necessary to build stable, peaceful international relationships (Adler and Barnett, 1998). However, the US-Soviet dyad suffers the additional problem of a conflictual history in comparison to the *tabula rosa* of the African dyad. The presence of a string of disputes, aside from the length of time since the last dispute, can make further disputes especially likely, as it may indicate the presence of an enduring rivalry (Goertz and Diehl, 1992) or it may have entrenched conflictual modes of behavior or hostile images of the adversary (Jervis, 1976). Including both Number of Disputes and Peace Years allows the efficient identification of both this similarity and this difference. In Model 7, Trade is not significant; in Model 8, which does not truncate the data at 1950, neither Trade or Allies are significant.

A couple last points about Models 7 and 8. Though the added variable is simply the count of the Number of Disputes, there are several ways to measure how conflictual the relationship is. One could assume that more recent disputes are more significant than older disputes, so one could discount a past dispute the farther back in time it is.[9] Or, one could assume there is generational memory, that is, disputes which have occurred within

the last generation (defined as 20 or 25 years, for example) are counted, but anything older is not counted (on generational memory and learning, see Jervis, 1976). Or, one could record the percentage of years in the history of the dyad in which there has been a dispute (Werner, 1998).

Table 2 presents analysis from a Cox event history model, a continuous model as opposed to the BKT model which is discrete. The Cox model does not include the number of Peace Years as an independent variable, as the correlation between Peace Years and the length of duration are quite high. Four models are tested, varying whether the data is left-truncated at 1950 or is allowed to extend the first pre-1950 dispute, and whether or not multiple failures are allowed for within a single dyad. Regarding the latter point, single failure means that a new dyad is considered to begin after a dispute has occurred; allowing for multiple failures with a Cox model is just that, allowing for a single dyad to fail more than once (see Blossfeld and Hamerle, 1989).

[Table 2 about here]

As seen in Table 1, the results change as the specification changes. Growth is significant in all models except for single failure, untruncated. The Allies variable is significant in all models except for multiple failures, truncated. The Trade variable is not significant in any of the models.

There are important differences in specification and results across all of these models. Specifically, the standard error estimates for Trade, Growth, Contiguity, and Allies are inconsistent across model specifications. These changes indicate that different techniques for handling temporal dependence can generate substantively different empirical results. Within the BKT framework, we varied the truncation of the data,

whether the temporal count was restarted at the beginning of the dyad, and whether a control variable for the number of past disputes was included. Within the Cox framework, we varied the truncation of the data and whether the temporal count was restarted at the beginning of the dyad. In short, variation in treatment of methodological issues such as left-censoring and multiple failures has substantively important consequences which scholars must be sensitive to.

III. Conclusions.

This paper has demonstrated two potential pitfalls in the application of event history analysis. First, scholars should be aware of potential problems introduced by left-censoring in data. Second, there are different ways to treat data with multiple failures, each of which incorporates different theoretical assumptions. Left-censoring and multiple failures are especially likely to appear in the study of international conflict, given the nature of available data.

Empirical analysis demonstrated the importance of sensitivity to these two issues. I applied two event history techniques to existing data on the causes of conflict among politically relevant states from 1950-1985. The first technique—temporally adjusted logit—is a discrete-time event history model; the second—a Cox model—is a continuous-time model. In both cases, changes in treatment of these two issues changed results, especially regarding the effects of Trade, Growth, Allies, and Contiguity on the likelihood of conflict.

Which approaches should scholars choose? Scholars should select techniques which match their theoretical needs. Left truncation can be appropriate if all the subjects begin at the same time prior to truncation, or if they can assume a constant hazard rate. Omission of left-censored cases is one option, but this entails information loss and possibly the introduction of systematic sample bias. Given that the conditions justifying left truncation are unlikely and that omission may introduce more problems than it may solve, scholars may be best off leaving data untruncated. Choice of single or multiple failure models should be theoretically determined, as how past events affect the likelihood of future failures should be determined by learning or other similar theories.

For the study of international conflict, scholars should consider three aspects of model building to minimize bias and best match theories. First, the results here reinforce the importance of accounting for temporal dependence in the study of conflict. Across all of the models considered here, temporal controls were statistically significant. Scholars should not be shy about exploring means of accounting for time, as it is readily available in common statistical packages. Besides the BKT and Cox models described here, other options include probit models which account for panel data and Generalized Estimating Equations (see Zeger and Liang, 1986).[10] Second, left-truncation ought to be avoided where possible. This clearly introduces bias for the study of dyad-level conflict, and measuring the number of years since the last dispute pre-1950 is relatively easy to do; indeed, that data can be easily generated using the EUGENE statistical package, which is publicly available (see Bennett and Stam, 1998a). Third, a single failure model which accounts for previous disputes may be the most preferred approach. This allows assessing two different dimensions of time dependence: the time since last dispute, and

the degree to which the past relationship has been conflictual. Additionally, it allows the efficient distinction between different types of dyads which might otherwise be conflated together, further reducing bias and increasing our confidence in the results.

Table 1: Logit Analysis of Dyadic Conflict, 1950-1985

Variables	Model 1: Oneal/Russett, unadjusted logit	Model 2: truncated, restart peace years after dispute	Model 3: untruncated, restart peace years after dispute	Model 4: delete all left-censored dyads
Democracy	-.0497 (.0074)*	-.0548 (.00786)*	-.0430 (.00806)*	-.0544 (.0195)*
Growth	-.0224 (.0085)*	-.0140 (.00902)	-.0140 (.00913)	-.0479 (.0160)*
Allies	-.821 (.0800)*	-.486 (.0881)*	-.256 (.0915)*	-.579 (.172)*
Contiguity	1.31 (.0796)*	.672 (.0875)*	.277 (.0900)*	1.51 (.263)*
Capability Ratio	-.00307 (.000417)*	-.00307 (.000416)*	-.00338 (.000421)*	-.00583 (.00165)*
Trade	-66.1 (13.4)*	-13.0 (10.6)	-29.8 (12.9)*	1.81 (14.6)
Peace Years	---	-.844 (.0391)*	.0601 (.0115)*	-1.58 (.224)*
Cubic Spline (1)	---	-.0153 (.00138)*	.0156 (.000905)*	-.222 (.0526)*
Cubic Spline (2)	---	.00920 (.00115)*	-.0134 (.000819)*	.0728 (.0225)*
Cubic Spline (3)	---	-.00291 (.000569)*	.00346 (.000242)*	-.0102 (.00595)*
Constant	-3.29 (.0792)*	-1.136 (.0912)*	-1.20 (.0911)*	-2.30 (.297)*
Prob>chi2	.0000	.0000	.0000	.0000
n	20990	20990	20990	7227
Log likelihood	-3477.5602	-2634.3131	-2241.3309	-650.65404

*significant at the .05 level, one-tailed test

Table 1a: Logit Analysis of Dyadic Conflict, 1950-1985: Multiple Failure Models

Variables	Model 2: truncated, restart peace years after dispute	Model 5: truncated, no restart of peace years after dispute	Model 6: untruncated, no restart of peace years after dispute	Model 7: truncated, restart peace years after dispute	Model 8: untruncated, no restart peace years after dispute
Democracy	-.0548 (.00786)*	-.0385 (.00815)*	-.0544 (.00965)*	-.0420 (.00829)*	-.0388 (.00837)*
Growth	-.0140 (.00902)	-.0212 (.00919)*	-.00507 (.0104)	-.0227 (.00956)*	-.0212 (.00972)*
Allies	-.486 (.0881)*	-.633 (.0890)*	-.527 (.117)*	-.253 (.0930)*	-.147 (.0943)
Contiguity	.672 (.0875)*	.687 (.0870)*	.121 (.106)	.641 (.0926)*	.558 (.0937)*
Capability Ratio	-.00307 (.000416)*	-.00241 (.000395)*	-.00236 (.000404)*	-.00206 (.000376)*	-.00190 (.000363)*
Trade	-13.0 (10.6)	-29.5 (12.7)*	-29.1 (14.4)*	-8.95 (10.2)	-17.2 (11.0)
Number of Disputes	---	---	---	.184 (.00954)*	.126 (.00972)*
Peace Years	-.844 (.0391)*	-.283 (.00970)*	-.696 (.0256)*	-.771 (.040)*	-.486 (.0227)*
Cubic Spline (1)	-.0153 (.00138)*	-.00310 (.000353)*	-.000517 (.0000737)*	-.0145 (.00142)*	-.00265 (.000232)
Cubic Spline (2)	.00920 (.00115)*	.00315 (.000494)*	.000464 (.0000802)*	.00874 (.00117)*	.00139 (.000163)*
Cubic Spline (3)	-.00291 (.000569)*	-.00136 (.000317)*	-.0000982 (.0000245)*	-.00275 (.000573)*	-.000125 (.0000372)*
Constant	-1.136 (.0912)*	-2.23 (.0983)*	-.290 (.125)*	-1.76 (.102)*	-.832 (.119)*
Prob>chi2	.0000	.0000	.0000	.0000	.0000
Log likelihood	-2634.3131	-2652.1828	-1535.0114	-2434.3421	-2444.109

N=20990

*significant at the .05 level, one-tailed test

Table 2: Cox Model Analysis of International Conflict, 1950-1985

Variables	Model 9: single failure, truncated	Model 10: single failures, untrunc.	Model 11: multiple failure, truncated	Model 12: multiple failures, untrunc.
Democracy	-.0389 (.00665)*	-.0384 (.00668)*	-.0425 (.0125)*	-.0465 (.0127)*
Growth	-.0130 (.00636)*	-.00926 (.00658)	-.0341 (.0117)*	-.0248 (.0124)*
Allies	-.195 (.0665)*	-.201 (.0667)*	-.265 (.162)	-.317 (.167)*
Contiguity	.412 (.0668)*	.417 (.067)*	.662 (.164)*	.701 (.170)*
Capability Ratio	-.00215 (.000407)*	-.00230 (.000432)*	-.00249 (.000961)*	-.00285 (.000110)*
Trade	-7.06 (9.41)	-4.35 (9.15)	-21.4 (18.0)	-17.7 (17.7)
Number of Prev. Disputes	.0755 (.00477)*	.078 (.00494)*	.193 (.020)*	.213 (.202)*
Prob>chi2	.0000	.0000	.0000	.0000
Log Likelihood	-6561.733	-6486.2885	-5400.9253	-5074.4707

Cell entries are coefficient estimates. Standard errors are robust.
*Significant at the .05 level. All tests are one-tailed.

References

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

Adler, Emanuel and Michael Barnett, eds. 1998. *Security Communities*. Cambridge: Cambridge University Press.

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

Azar, Edward. 1980. "The Conflict and Peace Data Bank (COPDAB) Project." *Journal of Conflict Resolution* 24 (1): 143-152.

Barbieri, Katherine. 1996. "Economic Interdependence: A Path to Peace or a Source of Interstate Conflict" *Journal of Peace Research* 33 (February): 29-49.

Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42 (October): 1260-1288.

Bennett, D. Scott. 1996. "Security, Bargaining, and the End of Interstate Rivalry." *International Studies Quarterly* 40 (June): 157-84.

Bennett, D. Scott. 1997. "Testing Alternative Models of Alliance Duration, 1816-1984." *Journal of Politics* 41 (July): 846-878.

Bennett, D. Scott. 1998. "Integrating and Testing Models of Rivalry Duration." *American Journal of Political Science* 42 (October): 1200-1232.

Bennett, D. Scott. 1999. "Parametric Models, Duration Dependence, and Time-Varying Data Revisited." *American Journal of Political Science* 43 (January): 256-270.

Bennett, D. Scott and Allan C. Stam III. 1996. "The Duration of Interstate Wars, 1816-1985." *American Political Science Review* 90 (June): 239-257.

Bennett, D. Scott and Allan C. Stam III. 1998a. "Comparative Theory Testing across Multiple Levels of Analysis." Unpublished ms, November 4.

Bennett, D. Scott and Allan C. Stam III. 1998b. "The Declining Advantages of Democracy: A Combined Model of War Outcomes and Duration." *Journal of Conflict Resolution* 42 (June): 344-366.

Berry, Frances Stokes and William D. Berry. 1990. "State Lottery Adoptions As Policy Innovations: An Event History Analysis." *American Political Science Review* 84 (June): 395-415.

Blechman, Barry M. and Stephen S. Kaplan. 1978. *Force Without War: U.S. Armed Forces as a Political Instrument*. Washington: Brookings Institution.

Bliss, Harry and Bruce Russett. 1998. "Democratic Trading Partners: The Liberal Connection." *Journal of Politics* 60 (November): 1126-47.

Blossfeld, Hans-Peter and Alfred Hamerle. 1989. "Using Cox Models to Study Multiepisode Processes." *Sociological Methods and Research* 17 (May): 432-448.

Blossfeld, Hans-Peter and Gotz Rohwer. 1995. *Techniques of Event History Modeling: New Approaches to Causal Analysis*. Mahwah, NJ: Lawrence Erlbaum.

Box-Steffensmeier, Janet M. and Bradford S. Jones. 1997. "Time is of the Essence: Event History Models in Political Science." *American Journal of Political Science* 41 (October): 1414-1461.

Brecher, Michael, Jonathan Wilkenfeld, and Sheila Moser. 1988. *Handbook of International Crises*. New York: Pergamon Press.

Bremer, Stuart A. 1993. "Democracy and Militarized Interstate Conflict, 1816-1965." *International Interactions* 18: 231-249.

Bueno de Mesquita, Bruce and Randolph M. Siverson. 1995. "War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability." *American Political Science Review* 89 (December): 841-855.

Dassel, Kurt and Eric Reinhardt. 1999. "Domestic Strife and the Initiation of Violence at Home and Abroad." *American Journal of Political Science* 43 (January): 56-85.

Diehl, Paul F., Jennifer Reifschneider, and Paul R. Hensel. 1996. "United Nations Intervention and Recurring Conflict." *International Organization* 50 (Autumn): 683-700.

Enterline, Andrew. 1998. "Regime Changes and Interstate Conflict, 1816-1992." *Political Research Quarterly* 51 (June): 385-410.

Fearon, James D. 1994. "Signaling versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model." *Journal of Conflict Resolution* 38 (June): 236-269.

Gartzke, Erik. 1998. "Kant We All Just Get Along? Opportunity, Willingness, and the Origins of the Democratic Peace." *American Journal of Political Science* 42 (January): 1-27.

Gaubatz, Kurt Taylor. 1996. "Democratic States and Commitment in International Relations." *International Organization* 50 (Winter): 109-139.

Gelpi, Christopher. 1997. "Democratic Diversons: Governmental Structure and the Externalization of Domestic Conflict." *Journal of Conflict Resolution* 41 (April): 255-282.

Gochman, Charles S. and Zeev Maoz. 1984. "Militarized Interstate Disputes, 1816-1976: Procedures, Patterns, and Insights." *Journal of Conflict Resolution* 28 (December): 585-615.

Goertz, Gary and Paul F. Diehl. 1992. "The Empirical Importance of Enduring Rivalries." *International Interactions* 18:1-11.

Gowa, Joanne. 1998. "Politics at the Water's Edge: Parties, Voters, and the Use of Force Abroad." *International Organization* 52 (Spring): 307-324.

Gowa, Joanne and Edward D. Mansfield. 1993. "Power Politics and International Trade." *American Political Science Review* 87 (June): 408-420.

Grattet, Ryken, Valerie Jenness, and Theodore R. Curry. 1998. "The Homogenization and Differentiation of Hate Crime Law in the United States, 1978 to 1995: Innovation and Diffusion in the Criminalization of Bigotry." *American Sociological Review* 63 (April): 286-307.

Guo, Guang. 1993. "Event History Analysis for Left-Truncated Data." *Sociological Methodology* 23: 217-243.

Huth, Paul K. 1988. *Extended Deterrence and the Prevention of War*. New Haven: Yale University Press.

Huth, Paul K. 1996. *Standing Your Ground: Territorial Disputes and International Conflict*. Ann Arbor: University of Michigan.

Jaggers, Keith and Ted Robert Gurr. 1995. "Tracking Democracy's Third Wave with the Polity III Data." *Journal of Peace Research* 32 (November): 469-482.

Jenkins, Stephen P. 1995. "Easy Estimation Methods for Discrete-Time Duration Models." *Oxford Bulletin of Economics and Statistics* 57 (1): 129-138.

Jervis, Robert. 1976. *Perception and Misperception in International Relations*. Princeton: Princeton University Press.

King, Gary, James E. Alt, Nancy Elizabeth Burns, and Michael Laver. 1990. "A Unified Model of Cabinet Dissolution of Parliamentary Democracies." *American Journal Of Political Science* 34 (August 1990): 846-871.

Levy, Jack S. 1994. "Learning and Foreign Policy: Sweeping a Conceptual Minefield." *International Organization* 48 (Spring): 217-312

Maoz, Zeev and Nasrin Abdolali. 1989. "Regime Types and International Conflict, 1817-1976." *Journal of Conflict Resolution* 33 (March): 3-35.

Maoz, Zeev and Bruce Russett. 1993. "Normative and Structural Causes of Democratic Peace, 1946-1986." *American Political Science Review* 87 (September): 624-638.

Meernink, James and Peter Waterman. 1996. "The Myth of the Diversionary Use of Force by American Presidents." *Political Research Quarterly* 49 (September): 573-590.

Miller, Ross A. 1995. "Domestic Structures and the Diversionary Use of Force." *American Journal of Political Science* 39 (August): 760-785.

Mintrom, Michael. 1997. "Policy Entrepreneurs and the Diffusion of Innovation." *American Journal of Political Science* 41 (July): 738-770.

Morrow, James D., Randolph M. Siverson, and Tressa E. Tabares. 1998. "The Political Determinants of International Trade: The Major Powers, 1907-1990." *American Political Science Review* 92 (September): 649-661.

Oneal, John R. and Bruce M. Russett. 1997. "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950-1985." *International Studies Quarterly* 41 (June): 267-294.

Oneal, John R. and Bruce Russett. 1999a. "Assessing the Liberal Peace with Alternative Specifications: Trade Still Reduces Conflict." *Journal of Peace Research* 4 (36): 423-442.

Oneal, John R. and Bruce Russett. 1999b. "The Kantian Peace: The Pacific Benefits of Democracy, Interdependence, and International Organization, 1885-1992." *World Politics* 52(October): 1-37.

Raknerud, Arvid and Håvard Hegre. 1997. "The Hazard of War: Reassessing the Evidence for the Democratic Peace." *Journal of Peace Research* 34 (4): 385-404.

Reed, William. 1997. "Alliance Duration and Democracy: An Extension and Cross-Validation of 'Democratic States and Commitment in International Relations.'" *American Journal of Political Science* 41 (July): 1072-1078.

Reiter, Dan. 1996. *Crucible of Beliefs: Learning, Alliances, and World Wars*. Ithaca: Cornell University Press.

Reiter, Dan. 2001. "Does Peace Nurture Democracy?" *Journal of Politics* (forthcoming).

Sherman, Frank L. 1994. "SHERFACS: A Cross Paradigm, Hierarchical and Contextually Sensitive, Conflict Management Dataset." *International Interactions* 20: 79-100.

Singer, J. David and Paul F. Diehl. 1991. *Measuring the Correlates of War*. Ann Arbor: University of Michigan Press.

Small, Melvin and J. David Singer. 1976. "The War-Proneness of Democratic Regimes, 1816-1965." *Jerusalem Journal of International Relations* 1 (Summer 1976): 50-69.

Warwick, Paul and Stephen T. Easton. 1992. "The Cabinet Stability Controvery: New Perspectives on a Classic Problem." *American Journal of Political Science* 36 (February): 122-46.

Werner, Suzanne. 1998. "Political Similarity and the Onset of Militarized Disputes, 1816-1985." Unpublished ms., Emory University, Atlanta, Georgia.

Zeger, Scott L. and Kung-Yee Liang. 1986. "Longitudinal Data Analysis for Discrete and Continous Outcomes." *Biometrics* 42 (1): 121-130.

Notes

[1] BKT remark (1264n): "Our only objection to Bennett's approach of using standard event history methods is that it requires analysts to learn an entirely new methodology." I group hazard analysis and the BKT logit technique under the general rubric of "event history" for rhetorical convenience, and because the BKT approach has many of the same strengths and potential weaknesses of standard event history, as discussed in the next section.

[2] Guo (1993) defines left-censoring slightly differently. For him, if a subject enters and exits before data is collected, it is left-censored; if it enters before data is collected but its exit is observed, it is left-truncated.

[3] Gowa (1998:311) is an exception, though she could begin her study of US behavior no earlier than 1870 because of limitations in the availability of US GNP data.

[4] The latest versions of STATA and LIMDEP permit probit analysis of data with panel characteristics.

[5] The BKT data is available at http://www.fas.harvard.edu/~rtucker/papers/grouped/grouped3.html. Thanks to David Davis for providing me the version of the data which contains country codes and years. The time since last dispute data was generated from Eugene version 1.14 (Bennett and

Stam, 1998a), which can be downloaded from http://www.personal.psu.edu/faculty/d/s/dsb10/.

[6] I replicated exactly Oneal and Russett though not BKT. This is probably due to slight differences in the generation of cubic splines. To generate the cubic splines, I use the spbase program publicly available from STATA, specifying three knots (for details, see Enterline, 1998).

[7] Going back to the beginning of each dyad would be problematic, as there would be missing data for pre-1950 cases.

[8] This is easy to do in STATA; one defines the data as survival data using stset, then one creates the failure count variable using stgen failcount=nfailures(). Partial correlation between the Number of Disputes and Peace Years variables is moderate; -.22 for untruncated data and -.24 for truncated data. The absolute value of no partial correlation was higher than .33.

[9] For example, each dispute could be multiplied by: $(.9)^{y-1}$, where y=number of years between past dispute and current case, so there is a steady decay as one goes back farther into the past.

[10] GEE is easy to execute in STATA using the xtgee command. Using a GEE model with a probit link and an AR-1 error term, Democracy, Growth, Allies, Contiguity, and Capability Ration all have the expected signs are significant at the .05 level, one-tailed test. Trade is signed correctly and significant at the .10 level, one-tailed test. Inclusion of Number of Failures as an independent variable does not change the above results; Number of Failures is itself significant at the .05 level. All significance tests were conducted using robust standard errors.