



## Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable

Nathaniel Beck; Jonathan N. Katz; Richard Tucker

*American Journal of Political Science*, Vol. 42, No. 4 (Oct., 1998), 1260-1288.

Stable URL:

<http://links.jstor.org/sici?&sici=0092-5853%28199810%2942%3A4%3C1260%3ATTSTAW%3E2.0.CO%3B2-4>

*American Journal of Political Science* is currently published by Midwest Political Science Association.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/mpsa.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

# *Taking Time Seriously: Time-Series–Cross-Section Analysis with a Binary Dependent Variable*

Nathaniel Beck, *University of California, San Diego*

Jonathan N. Katz, *University of Chicago*

Richard Tucker, *Harvard University*

Researchers typically analyze time-series—cross-section data with a binary dependent variable (BTSCS) using ordinary logit or probit. However, BTSCS observations are likely to violate the independence assumption of the ordinary logit or probit statistical model. It is well known that if the observations are temporally related that the results of an ordinary logit or probit analysis may be misleading. In this paper, we provide a simple diagnostic for temporal dependence and a simple remedy. Our remedy is based on the idea that BTSCS data are identical to grouped duration data. This remedy does not require the BTSCS analyst to acquire any further methodological skills, and it can be easily implemented in *any* standard statistical software package. While our approach is suitable for any type of BTSCS data, we provide examples and applications from the field of International Relations, where BTSCS data are frequently used. We use our methodology to reassess Oneal and Russett's (1997) findings regarding the relationship between economic interdependence, democracy, and peace. Our analyses show that (1) their finding that economic interdependence is associated with peace is an artifact of their failure to account for temporal dependence yet (2) their finding that democracy inhibits conflict is upheld even taking duration dependence into account.

## **1. INTRODUCTION**

The analysis of time-series—cross-section data with a binary dependent variable (BTSCS data) is becoming more common, particularly in the study of international relations (IR). Moreover, the number of such studies appears to be increasing exponentially.<sup>1</sup> Since it is unlikely that units are statistically unrelated over time, BTSCS observations, like their continuous dependent

We thank John Oneal and Bruce Russett for providing their data and Robert Engle, Gary King, Jonathan Nagler and Glenn Sueyoshi for helpful comments and conversations. A replication data set may be found at <ftp://weber.ucsd.edu:/pub/nbeck>

<sup>1</sup>The vast majority of IR BTSCS analysts study militarized conflict or interstate war; others study alliance and rivalry behavior. A brief list of IR BTSCS analyses using ordinary logit or probit published in the previous eighteen months includes Barbieri (1996), Bennett (1996), Enterline (1996, 1997), Farber and Gowa (1997), Gartzke (Nd.), Gleditsch and Hegre (1997), Henderson (1997), Hermann and Kegley (1996), Huth (1996), Lemke and Reed (1996), Mansfield and Snyder (1996, 1997), Maoz (1996, Nd.), Mousseau (1997), Oneal et al. (1996) and Oneal and Russett (1997). We do not claim that these studies draw incorrect conclusions. However, the possibly faulty (and untested) assumption of temporal independence, inherent in their respective logit/probit analyses, casts some doubt about the validity of their substantive findings.

variable TSCS cousins, are likely to be temporally dependent. It is well known that violations of the assumption of independent observations can result in overly optimistic inferences (underestimates of variability leading to inflated  $t$ -values). Nevertheless, BTSCS data are almost invariably analyzed using ordinary logit or probit analysis, techniques that assume temporal independence.<sup>2</sup> While analysts are certainly aware of the pitfalls of such assumptions, they seem to have overlooked a very simple solution.

Our simple solution is to add a series of dummy variables to the logit specification. These variables mark the number of periods (usually years) since either the start of the sample period or the previous occurrence of an "event" (such as war). A standard statistical test on whether these dummy variables belong in the specification is a test of whether the observations are temporally independent. The addition of these dummy variables to the specification, if the test indicates they are needed, corrects for temporally dependent observations. This simple solution, which can be implemented in *any* software package, allows for accurate estimation of the parameters of temporally dependent BTSCS models.<sup>3</sup>

This simple solution is based on the recognition that BTSCS data are grouped duration data. Note that we do not say "like grouped duration data"; BTSCS data *are* grouped duration data. This recognition permits us to use known and validated event history concepts explicitly designed for temporally dependent data.<sup>4</sup>

In the next section, we briefly discuss the prominence of BTSCS data in international relations and why ordinary logit is inappropriate for BTSCS data in most contexts. The subsequent section illustrates the equivalence of BTSCS and grouped duration data. We also delineate our proposed method for analyzing temporally dependent BTSCS data and discuss its application to the study of conflict/peace. The next section then uses our proposed method to reanalyze one prominent study of conflict (Oneal and Russett 1997).

## 2. BTSCS DATA IN INTERNATIONAL RELATIONS

BTSCS data are most common in international relations.<sup>5</sup> The IR conflict processes literature has favored a theoretical emphasis on dyadic

<sup>2</sup>We freely mix logit and probit analyses here. In the context of this paper they suffer identical flaws which have identical remedies. For simplicity we refer to logit analysis throughout this paper. Those committed to probit analysis should make our recommended changes to the probit specification.

<sup>3</sup>While cross-sectional dependence also causes problems, our goal here is to address the problem of temporal dependence. Our proposed remedy is, however, sufficiently simple that it should be easy to adjoin to any remedy for cross-sectional dependence.

<sup>4</sup>We use the terms event history methods and duration models interchangeably.

<sup>5</sup>Following the pioneering efforts of Berry and Berry (1990, 1992), American state politics researchers also frequently use BTSCS data. Unlike IR researchers, however, they typically begin with

interstate interactions (e.g., Bueno de Mesquita and Lalman 1992; Goertz and Diehl 1993; Vasquez 1993) and an empirical focus on the dyad-year as the unit-of-analysis (e.g. Bremer 1992; Maoz and Russett 1993). Dyad-year data sets typically contain yearly observations on conflict occurrence between pairs of nations (or engagement in some other interstate behavior such as alliance formation or rivalry dissolution). These datasets also include properties of the dyad (which may vary from year to year) to explain the presence or absence of conflict. While our argument generalizes to all BTSCS data, we couch our discussion in terms of IR dyad-year BTSCS data.

BTSCS data shares all the standard characteristics of continuous dependent variable time-series—cross-section data.<sup>6</sup> Formally, a BTSCS model with binary dependent variable,  $y$ , and a vector of independent variables,  $\mathbf{x}$ , has

$$P(y_{i,t} = 1) = f(\mathbf{x}_{i,t}, y_{i,1}, \dots, y_{i,t-1}, \mathbf{x}_{i,1}, \dots, \mathbf{x}_{i,t-1}), \\ i = 1, \dots, N, t = 1, \dots, T \quad (1)$$

where  $f$  is any suitable function that has a range of the unit interval. The inclusion of the lagged values of  $y$  and  $\mathbf{x}$  allows for a very general form of temporal dependence of the observations.<sup>7</sup> We assume the number of time points ( $T$ ) to be reasonably large (say at least 20). This is in contrast to binary panel data, where  $T$  may be as small as two or three. Panel methods are also designed to handle enormous cross-section sample sizes ( $N$ ), ranging into the thousands. While  $N$  is not critical for our interests here, we do not have to solve the problems brought about by large (and asymptotically unbounded)  $N$ 's combined with small (and bounded)  $T$ 's that have plagued

a discrete time event history model, which is then appropriately estimated using BTSCS methods. Meier and McFarland (1992) and Mintrom (1997), for example, correctly allow for duration dependence by adding yearly temporal dummy variables to the logit specification. Typical state politics BTSCS datasets, however, are simpler than their IR counterparts; while IR datasets often contain multiple failures per unit, state datasets typically only have a single failure per unit (i.e., no post-failure observations). The latter datasets also typically track all states for the same time frame. Thus state researchers can use temporal dummy variables that correspond to years. IR researchers, as we will subsequently illustrate, must create temporal dummy variables that track time since the previous event occurrence.

<sup>6</sup>See Beck and Katz (1995, 1996) for a discussion of continuous dependent variable TSCS methods.

<sup>7</sup>Equation 1 is very general. One possible specialization is a latent variable formulation, where temporal dependence is induced by serially correlated errors in the latent variable (Beck and Katz 1997). Equation 1 does *not* imply that one should add a lagged dependent variable to the logit specification. The essential nonlinearity of BTSCS models makes their dynamics much more complex than continuous TSCS models.

panel analysts. This contrast is important, since there are available estimation techniques for interdependent binary panel data (see Diggle, Liang, and Zeger 1994). While some of these techniques *may* prove useful for interdependent BTSCS data, such utility has not yet been demonstrated. But in general, the temporal dimension of BTSCS data are so much richer than its panel counterpart that we would not be overly optimistic about the utility of panel methods for BTSCS data.<sup>8</sup>

Analysts almost invariably simplify Equation 1 to

$$P(y_{i,t} = 1 | \mathbf{x}_{i,t}) = \frac{1}{1 + e^{-\mathbf{x}_{i,t}\beta}} \quad (2)$$

and perform an “ordinary logit” analysis of their data.

BTSCS data, however, are simply a variant of TSCS data, and we know that TSCS data often shows temporal dependence. Might we not expect BTSCS data to show temporal dependence as well? The probability of dyadic conflict in a given year, for example, is likely to be dependent on the conflict history of that dyad.<sup>9</sup> Remedies for continuous dependent variable TSCS data (Beck and Katz 1996), however, are inapplicable to BTSCS data (Beck and Tucker 1997). It is well known that if the observations are temporally related, results of an ordinary logit or probit analysis may be misleading. Poirier and Ruud demonstrate that probit<sup>10</sup> standard errors are incorrect for time series data with serially correlated errors. These time series results hold for BTSCS data. Simulations reported in Beck and Katz (1997) indicate the severity of these problems, with reported standard errors possibly understating variability by 50 percent or more! While probit analysis of temporally dependent data provides consistent parameter estimates, ignoring this dependence may also lead to severe inefficiency. The incorrect assumption of temporal independence leads to both inaccurate statistical tests and the loss of valuable information in the data.

IR BTSCS analysts routinely acknowledge these problems, but in the absence of better alternatives, continue to ignore temporal dependence and use ordinary logit analysis. Farber and Gowa (1997, 397), for example, agree that “the yearly observations for a dyad cannot be considered to be independent” but they “proceed ignoring this lack of independence. While

<sup>8</sup>IR datasets may have large  $N$ 's; the one we reexamine has an  $N$  of almost 1000. The critical issue is that IR datasets typically have reasonably large  $T$ . Most dyads in our reanalysis, for example, are observed for over twenty years, with more observations per dyad becoming available as more recent data are collected. Our proposed method would not work for datasets with very small  $T$ 's.

<sup>9</sup>As we discussed below, temporal dependence cannot provide a satisfactory explanation by itself, but must, instead, be the consequence of some important, but unobserved, variable.

<sup>10</sup>Their conclusions hold for logit and any other standard binary dependent variable method.

[they] recognize that the power of [their] tests is somewhat overstated as a result, a better solution is not obvious." Oneal and Russett (1997, 283) note that the "greatest danger arises from autocorrelation, but that there are not yet generally accepted means of testing for or correcting this problem in logistic regressions."<sup>11</sup> Some BTSCS analysts have simply given up on logit based methods, opting for less well-known event history methods. Bennett (1997, 12), for example, argues that a "hazard [event history] model is the most appropriate way to analyze alliance durations, and superior to the [ordinary logit] procedure, since hazard models allow corrections for censoring, heterogeneity and duration dependence."

In this paper we will show that the logit, once corrected, is an event history method for BTSCS data. Moreover, we illustrate a simple and easy to implement modification to the logit specification that allows it to handle temporally dependent data. Thus our methodology allows logit oriented BTSCS analysts to continue to use their familiar methods while deriving all the benefits of event history analysis.<sup>12</sup>

### **3. BTSCS DATA IS GROUPED DURATION DATA**

Our solution depends on the recognition that BTSCS data are *identical* to grouped duration data. While we need very little of the specialized language of event history analysis, a few concepts will prove helpful.<sup>13</sup> Event history analysts model the elapsed time until an "event" or "failure," or, equivalently, the length of a non-eventful "spell." In our IR examples an event is conflict, with the duration of spells of peace being modeled. A unit has "survived" or is "at risk" until it fails.<sup>14</sup> The "hazard" rate is, loosely speaking, an indication of how likely failure is to occur at any given time (or more precisely, the rate of failure in any small time interval), provided the unit has survived until that time. If the hazard rate is time invariant, that is, the risk of failure does not depend on how long a unit has survived, the hazard is said to be "duration independent"; if it varies with time, the hazard rate is said to show "duration dependence." Event history analysts model the hazard rate as a function of independent variables, which may or may not be time invariant.

The most common event history methods assume continuous time, so that durations are measured continuously and hazard rates vary continu-

<sup>11</sup>They attempt various *ad hoc* remedies, which we discuss in the reanalysis section.

<sup>12</sup>Our only objection to Bennett's approach of using standard event history methods is that it requires analysts to learn an entirely new methodology.

<sup>13</sup>Introductions to event history methods for political scientists are in Beck (N.d.) and Box-Steffensmeier and Jones (1997).

<sup>14</sup>For simplicity, we initially assume only one possible failure per unit. We relax this assumption below.

ously. But duration data may be "grouped," so that we only know whether a unit has failed in some discrete time interval (with independent variables only measured to the fineness of that interval). This is usually a result of the measurement process, so that instead of recording the exact time of failure, we only record whether a unit failed in some fixed time interval. BTSCS data, as coded, only allows us to know if a conflict occurred sometime during a year.<sup>15</sup>

Annual BTSCS data are equivalent to grouped duration data with an observation interval of one year.<sup>16</sup> The dichotomous dependent variable is one in a given year if there was a failure (for example, conflict) during that year, with the independent variables also being measured yearly.<sup>17</sup> We stress that BTSCS data are, by definition, grouped event history data; no sophisticated mathematical, statistical nor computational argument is required to demonstrate this.

### 3.1 The Grouped Duration Solution

Having noticed the equivalence, we also note that there are standard methods for estimating models with grouped event history data where the observations may be temporally dependent. These methods begin with a continuous time event history model. They are derived under the assumption that observations of this continuous process are only made at discrete intervals,

<sup>15</sup>The beginning and end of conflicts can obviously be determined more accurately, often to the day. Raknerud and Hegre (1997), for example, use the daily dating of wars to convert a BTSCS data set into a continuous time event history data set (since they are interested in the order in which nations join multilateral conflicts). But while it may be possible to more accurately date events, many independent variables are only measured yearly. Our interest is in the use of event history methods to analyze data that has already been coded as BTSCS data.

<sup>16</sup>Some analysts prefer to use the term discrete time duration data rather than grouped duration data. BTSCS data, however, are grouped, not discrete time data. Grouped duration data allows for exits at any time, but we only observe whether an exit has occurred in some time interval. Exits, within the discrete time framework, only occur at discrete time intervals. We do not contend that wars only occur on New Year's Eve! But this distinction has few, if any, practical implications, since discrete time models are analyzed using grouped time concepts.

<sup>17</sup>Note that we are assuming there can be no more than one measured conflict in a year. This may be due to a censoring process, where the only recorded information is whether at least one conflict occurred in a year, or it may be due to something about the conflict process which limits conflicts to one per year. BTSCS data are presented this way. Analysts may have a choice as to whether to use a binary dependent variable or an event count dependent variable; our discussion assumes that either the investigator or some outside data collector has previously decided to only collect information about the binary dependent variable. Alt, King, and Signorino (1997) provide a very interesting treatment of this entire issue. Our point here is much simpler than theirs, since we assume that aggregation decisions have already been made, and so only BTSCS data are available. Event count TSCS models must also take duration dependence into account.

with only one event possible per interval.<sup>18</sup> The most common continuous time duration model is the Cox (1975) proportional hazards model; this model dominates applied work in the social and life sciences.<sup>19</sup>

In this model the instantaneous hazard rate is

$$h(s|\mathbf{x}_{i,s}) = h_0(s)e^{\mathbf{x}_{i,s}\beta} \quad (3)$$

where  $\mathbf{x}_{i,s}$  is the vector of independent variables at (continuously measured) time  $s$ . In this setup the hazard of exit depends both on the independent variables (via the  $e^{\mathbf{x}_{i,s}\beta}$  term) and the length of time the unit has been at risk (via  $h_0(s)$ , the “baseline hazard”). The proportional hazards model is widely used because it allows for estimation of the parameters of interest ( $\beta$ ) in the presence of an unknown, and possibly complicated, time varying baseline hazard.<sup>20</sup> As we shall see, the  $\beta$  in Equation 3 are what logit BTSCS modelers are estimating. Ordinary logit fails because it doesn’t allow for a (nonconstant) baseline hazard. The grouped duration model, although derived from an underlying continuous time Cox proportional hazards model, is easier to estimate, and does not suffer from some problems inherent in the continuous time model.<sup>21</sup> For notational simplicity, let us assume annual data indexed by year  $t$ . The “discrete hazard” in year  $t$  for dyad  $i$  is simply the probability that a dyad will experience conflict sometime during that year. Letting  $y_{i,t}$  be a binary indicator of conflict in dyad  $i$  sometime in year

<sup>18</sup>The grouped duration model was first derived by Prentice and Gloeckler (1978). Readable social science treatments are in Allison (1982), Singer and Willett (1993) and Jenkins (1995). For completeness we lay out the basic argument in an appendix, although it is dependent on some duration results not contained in this paper. Han and Hausman (1990) and Sueyoshi (1995) provide a modern econometric treatment of many of the issues discussed here. Katz and Sala (1996) have applied the grouped duration model to Congressional data.

<sup>19</sup>The assumption of proportional hazards is not innocuous, and surely there are situations where it is a bad assumption. For example, we implicitly assume that there is no heterogeneity in the baseline hazards across units, so that we may pool all observations. But Equation 3 is more general than other common hazard specifications used in event history analysis. The Weibull model is the most common fully parametric event history model. The Weibull model uses a hazard rate which is a special case of Equation 3, with  $h_0(t)$  assumed to follow a specific parametric form. In practice, the proportional hazards model works well, but no one model is perfect for all situations. Since, as we shall see, ordinary logit can be derived from a special case of the proportional hazards model, any criticism of proportional hazards is at least as strong a criticism of ordinary logit.

<sup>20</sup>The grouped model could easily be adapted to fully parametric duration models. Given the dominance of the semi-parametric Cox approach in applied work, we see no reason to pursue the fully parametric approach here. Alt, King, and Signorino (1997) derive the grouped model for a continuous time gamma duration model.

<sup>21</sup>In particular, the continuous time model has problems if there are many units that exit at the same time.

$t$ , the discrete hazard is just  $P(y_{i,t} = 1)$ . This is the probability estimated by logit analysis. But the logit probability (Equation 2) is *not* the same as the discrete hazard constructed by aggregating the continuous hazard rate of Equation 3. The discrete hazard rate corresponding to Equation 3 is (as shown in the Appendix)

$$P(y_{i,t} = 1 | \mathbf{x}_{i,t}) = h(t | \mathbf{x}_{i,t}) = 1 - \exp(-e^{\mathbf{x}_{i,t}\beta + \kappa_{t-t_0}}) \quad (4)$$

where  $\mathbf{x}_{i,t}$  now represents the observed value of the independent variable for the entire year  $t$ .  $\kappa_{t-t_0}$  is a dummy variable marking the length of the sequence of zeros that precede the current observation; for first events,  $t_0 = 0$ . We use  $t - t_0$  instead of the simpler  $t$  subscript because the notation must allow for multiple events; in that case  $t_0$  marks the time of the previous event and  $t - t_0$  is the length of the spell of peace from  $t_0$  until  $t$ .<sup>22</sup> We use the more complicated notation even when  $t_0 = 0$  to remind us that the temporal dummies mark the length of prior spells of peace, which will not always be the current year index,  $t$ .<sup>23</sup>

### 3.2 The Logit Solution

The grouped duration model differs from ordinary logit in two ways. First, it is a binary-dependent variable model using what is known as a “complementary log-log (cloglog) link” instead of the more familiar logit (or probit) link.<sup>24</sup> Second, the specification contains the temporal dummy variables,  $\kappa_{t-t_0}$ . The distinction between the cloglog and logit links is trivial; the inclusion of the temporal dummy variables is not. Let us eliminate the trivia first.

<sup>22</sup>To be concrete we show one particular assignment of the  $\kappa$ .

$t$	1	2	3	4	5	6	7	8	9
$y$	1	0	0	0	1	0	1	1	0
$\kappa$	$\kappa_1$	$\kappa_2$	$\kappa_3$	$\kappa_4$	$\kappa_1$	$\kappa_2$	$\kappa_1$	$\kappa_1$	$\kappa_2$

As with any saturated set of dummy variables we must either not estimate a constant term or drop one dummy variable. For notational simplicity we assume the former, though most statistical packages do the latter. This should cause no problems.

<sup>23</sup>We should note that in constructing the discrete hazard rate in Equation 4 we have implicitly assumed that the unit has “survived” to period  $t$ . That is, we should condition on the event that  $y_{i,t-1} = 0$  in the definition of the discrete hazard. Since we will allow for multiple events, we drop this added notation. This assumption, however, is important. In particular, for the case of multiple failures (conflicts) this means that we are treating the second and subsequent events as if they were new units in the data.

<sup>24</sup>This terminology is from the “generalized linear model” (GLM approach [McCullagh and Nelder 1989]). A link function specifies the relationship between a linear predictor ( $\mathbf{x}_{i,t}\beta$ ) and the dependent variable. The logit and cloglog links are two common links for binary dependent variable models.

*The cloglog vs. logit link*

The two links transform probabilities by

$$\text{cloglog}(P) = \log(-\log(1 - P)) \text{ and} \quad (5)$$

$$\text{logit}(P) = \log\left(\frac{P}{1 - P}\right) \quad (6)$$

These are the inverses of the transforms used in Equations 4 and 7 and plotted in Figure 1.

We see that the two links are almost identical when the probability of an event is less than 25 percent, and are extremely similar so long as the probability of an event does not exceed 50 percent. If the probabilities of an event are small, either a logit or a cloglog link can be used. For typical event history data (especially in IR) the probability of an event in any given time period will be small. The two links will differ only in the unlikely case (for event history data) that many observations have a probability of failure exceeding 50 percent. And even here, who is to say which is the best model to use? While the cloglog model is the exact grouped duration analogue of the most widely used Cox proportional hazards model, it may not be appropriate in every instance. The logit link corresponds to a (complicated) continuous time duration model (Sueyoshi 1995). While the Cox proportional hazards model is computationally convenient, there is no reason to assume that the data were generated in order to make computations simple! There appears to be little if any cost then to use the more familiar logit link for typical BTSCS data. There are clear benefits to using the logit link. It is well understood by researchers, is estimable with any software package, does not require learning new methods (generalized linear models), and most importantly, can be extended easily in a variety of interesting ways.<sup>25</sup> We, therefore, recommend that researchers use

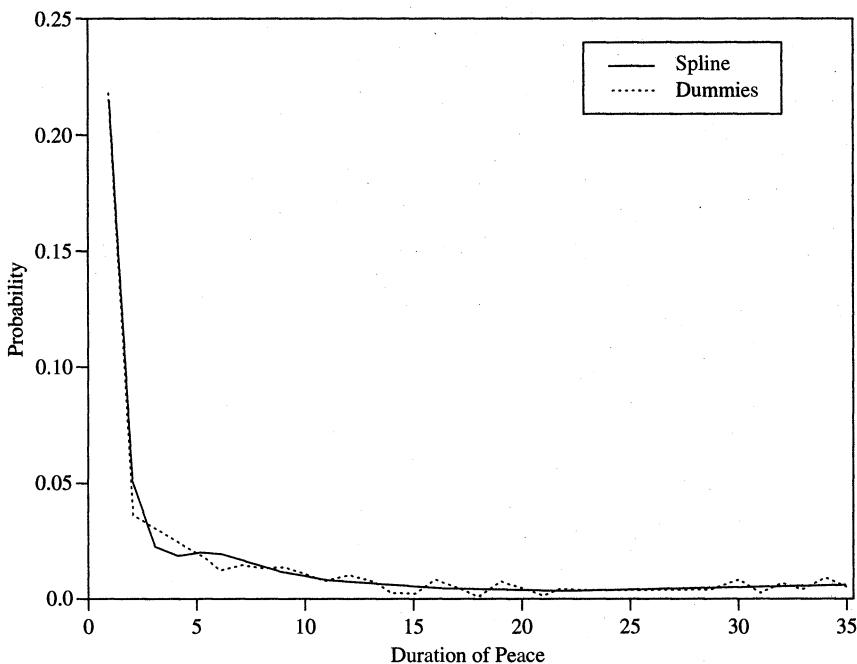
$$P(y_{i,t} = 1 | \mathbf{x}_{i,t}) = h(t | \mathbf{x}_{i,t}) = \frac{1}{1 + e^{-(\mathbf{x}_{i,t}\beta + \kappa_{t-t_0})}} \quad (7)$$

which is the logistic analogue of Equation 4.

*Temporal dummy variables*

Using the logit rather than the cloglog link allows us to focus on the second way that Equation 7 differs from ordinary logit: the inclusion of the

<sup>25</sup>Allison (1982, 87–90) and Farhmier and Wagenpfeil (1996) show how the logit can be extended to the multinomial logit to handle multiple types of failures. Thus we could allow for data where  $y_{i,t}$  denotes a series of unordered outcomes, so long as the outcomes satisfy the independent risks assumption underlying the “competing risks” model. Any remedies which allow logit to deal with cross-sectional dependence will also be easy to combine with the logit link.

**Figure 1. Comparison of cloglog and logit Transforms**

temporal dummies,  $\kappa_{t-t_0}$ . These are the grouped duration analogue of the continuous time baseline hazard function,  $h_0$ . Omitting these dummies is equivalent to assuming that the baseline hazard is constant, so that the model shows duration independence. While such a situation can occur, event history analysts typically allow for duration dependence, at least initially, and then test whether the model can be simplified by imposing duration independence. The costs of incorrectly imposing duration dependence are, at a minimum, inefficiency and incorrect standard errors, and in some complicated cases may even lead to inconsistent parameter estimates. It is exactly these problems that the Cox proportional hazard model avoids. It is simple enough to include the temporal dummies in the logit specification. Before doing so, however, one should determine whether they are required. Temporal dummies should *not* be included in the specification if the observations are already temporally independent, since the temporal dummies might then introduce unnecessary multicollinearity. The test of whether the temporal dummies should be included is a standard likelihood ratio test of the hypothesis that all the  $\kappa_{t-t_0} = 0$ . If the null hypothesis of temporal independence is rejected, then all the  $\kappa_{t-t_0}$  should be included in the logit specification.

Thus Equation 7 is the generalization of ordinary logit that allows for temporally interdependent observations. As we have just seen, it is easy to both test and correct the logit for temporally dependent observations.

### *Cubic splines*

Equation 7 requires the estimation of the coefficients of many dummy variables. Unless  $N$  is large, estimates will not be precise. While this is not a problem if our interest is in estimating  $\beta$ , we may have some interest in the  $\kappa$  themselves. Note that the  $\kappa_{t-t_0}$  are easily interpretable as “baseline” probabilities (or hazards) in that

$$P(y_{i,t} = 1 | \mathbf{x}_{i,t} = 0, t_0) = \frac{1}{1 + e^{-\kappa_{t-t_0}}} \quad (8)$$

These baseline hazards give the probability of failure in each time interval when all the independent variables are zero. If the independent variables are measured so that these zeros are substantively meaningful, then the baseline hazards are of substantive interest.

While the path traced out by the  $\kappa_{t-t_0}$  is easily interpretable, the imprecision with which the  $\kappa$  are estimated may give a false impression that the baseline hazard is jagged. We would expect it to be smooth, that is, baseline hazard rates should change relatively slowly over time, rather than jumping around from year to year.

One solution to this problem is to replace the dummy variables in Equation 7 with a smooth function of  $t - t_0$  (we cannot directly use  $t - t_0$  since there is no reason to assume that the baseline hazard is a linear function of time). In earlier work we recommended “cubic smoothing splines” (Beck and Jackman 1997; Beck and Turner 1997). But while these work very nicely, they do require software (such as S-Plus) that often is not readily accessible. One can obtain almost the same degree of smoothness with “natural cubic splines” (Eubank 1988), which are easy to implement with widely available software packages (such as Stata). Natural cubic splines fit cubic polynomials to a predetermined number of subintervals of a variable. These polynomials are joined at “knots,” with the number and placement of the knots specified by the analyst. Smoothness is imposed by forcing the splines, and their first and second derivatives, to agree at each of the knots. Thus each knot only uses up one degree of freedom, so that we can flexibly fit a cubic spline using up only a very few degrees of freedom. The estimated spline coefficients can then be used to trace out the path of duration dependence.

One advantage of the spline is that it facilitates a test of the hypothesis of duration dependence. With many temporal dummy variables, the likelihood ratio test for whether they are all zero may have poor finite sample

properties. The equivalent test on the spline formulation requires testing only whether a small number of spline coefficients are zero.

Analysts can choose either the dummy variable or the spline formulation; neither will have significant consequences for the estimation of  $\beta$ . We have a slight preference for the spline formulation. Users hesitant to deal with natural splines can use the simpler dummy variable specification with little loss if they are primarily interested in examining the effects of the substantive independent variables. We use both approaches in our replication, though we rely primarily on the spline approach.

Since the logit with temporal dummy (or spline) variables is more general than ordinary logit, and since we can easily test the null hypothesis of duration independence, there is no reason not to undertake logit analysis of BTSCS data, adding the temporal variables when they are required. This is not to say that there might not be better methods for estimating some models. While the Cox proportional hazards model is widely used and works well in practice, no model can be expected to be optimal for all problems. We expect that logit analysis with temporal dummy or spline variables will work well for most BTSCS data sets, and undoubtedly this approach is superior to ordinary logit.

### 3.3 Complications

Before turning to our reanalysis, several complications must be discussed. These complications would not arise if the data were independent, but they are inherent if we are unwilling to make that assumption. The event history approach simply makes these problems (and possible solutions) clearer.

#### *Multiple failures*

The first problem is that BTSCS data allows for multiple failures per unit. Many event history analyses simply model time until the first (or only) failure, but the nature of BTSCS data allows for more than one failure per unit.<sup>26</sup> Ordinary logit avoids this problem by assuming that the probability of failure in any year is the same as in any other year (conditional only on the independent variables), so that second and subsequent failures are assumed to be generated identically to first failures. In our construction of the  $\kappa$ 's we have also used this assumption, since the only relevant information in the  $\kappa$  is time since the most recent event. However implausible the assumption that second and subsequent events are independent of the number and timing of previous events, this assumption is weaker than the ordinary logit assumption that all observations are independent.

<sup>26</sup>If only one failure per unit were possible, we should discard all data after the first failure. But in BTSCS data we have observations through a fixed time  $T$ .

Since the assumption that second spells are independent of first spells is questionable, one solution might be to limit the analysis to the initial event. While losing data on second events is inefficient, it does allow for consistent estimation of  $\beta$  without having to model the dependence of second and later events on earlier events. Of course it would be better to correctly model repeated events. One easy way to do this is to include in the specification a variable which counts the number of previous events. This approach, while primitive, is better than ignoring the problem. A related issue common to IR studies is that events may appear to take place over the course of several years. If conflicts really are multi-year, we should simply drop all but the first year of the conflict from the analysis. If we have a theory about the duration of peace, we should not include spells of conflict in testing that theory. However, since we can observe different conflicts in consecutive years, this would be tantamount to discarding new, but very short, spells of peace. A decision on how to proceed should be made on theoretical grounds. But if we observe multi-year spells of conflict, it is difficult to maintain the assumption that yearly observations are independent of each other. Duration dependence may manifest itself in the finding that conflicts are more likely to follow other conflicts.

#### *Left censoring*

The second concern has to do with what event history analysts call "left censoring." Spells are left censored if we do not know when they began.<sup>27</sup> For example, if our first dyadic observation is 1951, we do not know if a spell of peace began in 1951, 1950, or before. This may not be a large problem in IR, since we can often begin analyses at the start of a new international order or security regime (the Congress of Vienna or the beginning of the Cold War). Our proposed method allows left censoring so long as all observations are equally left censored. For example, if the Cold War began in 1947, but our data starts in 1954, left censoring causes literally no problems for our proposed method.<sup>28</sup> All that is required is that the  $\kappa$  for any given year reflect the same length of prior peace spell length for all units.

This could cause problems, for example, for dyads that enter the data set after the starting year. In our reanalysis, for example, some dyads enter the data set after one of the members became independent. Suppose the data set

<sup>27</sup> Spells are "right censored" if we do not know when they ended. Units that are right censored simply contribute a string of zeros, with no final one, to the logit likelihood. These are not a problem for grouped duration logit analysis.

<sup>28</sup> All we lose are estimates of the nuisance yearly dummies from 1948 through 1953. See Jenkins (1995) for a formal proof and a good discussion of the interaction of grouped duration analysis with various sampling designs. Jenkins shows that it is irrelevant, for the estimation of  $\beta$ , if the sample period begins with  $t_2$  but that spells actually began at  $t_1 < t_2$ .

begins in 1951 but a dyad enters the data set in 1962. Should the dummy variable for that observation be  $\kappa_1$  or  $\kappa_{12}$ ? If our example (and in our re-analysis) it seems reasonable to use  $\kappa_1$  here. But analysts will have to make judgments before beginning their own data analysis. Results should be relatively insensitive to a few differences in judgment on this issue.

#### *Variables that are fixed across units*

The third potential problem with our method is that it does not allow researchers to use independent variables that vary by time but not across units. In IR such variables are measured at the system level. Some examples of systemic level variables are the concentration of power or the number of nation-states in the world at any given time. These variables will be highly collinear with the  $\kappa$ .<sup>29</sup> Inclusion of the  $\kappa$  in the specification makes it unlikely that the coefficients of these systemic variables will remain statistically significant. This will cause problems for some, but not all, research agendas.<sup>30</sup> Systemic level variables are, for example, rare in dyad-year studies of conflict.

If system level variables account for most of the duration dependence, then our test for it will indicate that we cannot reject the hypothesis of duration independence. At that point researchers can confidently use ordinary logit analysis, including system-level variables. This is the optimal situation, since the system-level variables theoretically explain duration dependence. However, we fear that this situation is rare.

There may be other situations that remain problematic. If the system level variables are important, we might choose to ignore duration dependence if it is not serious (as indicated by a baseline hazard function that looks fairly flat). Sometimes the cure may be worse than the disease! Some situations will occur where the researcher is faced with a choice between two evils. The analysis of data is an art, not a science. No one method will ever solve all possible problems. But a test for whether the temporal variables belong in the logit specification, even with the system level variables included, at least alerts the researcher to the existence of potential problems caused by temporally dependent observations.

#### *Missing data*

The fourth problem is that missing data becomes more troublesome in the presence of duration dependence. The assumption of independence

<sup>29</sup>They are not perfectly collinear if there are multiple events per unit, since the  $\kappa$  then no longer simply mark  $t$ . They will also not be perfectly collinear with the temporal spline. But they might be highly collinear.

<sup>30</sup>The problem is identical to that associated with fixed unit effects in models with independent variables that are constant within units. However, this problem has not caused researchers to abandon fixed effects modeling.

allows the analyst to omit all observations with missing data—subject, of course, to the usual caveats about missing data (Little and Rubin 1987).<sup>31</sup> Our method also allows for the elimination of observations with missing data so long as the correct time dummy variable is retained. Thus we cannot allow missing observations on the dependent variable (or we must assume that there were no missing years of conflict). In practice we will encounter relatively little, if any, missing data in the conflict variable, since IR researchers have gone to great lengths to code this data. But missing data on the dependent variable could be a potential problem for other types of BTSCS analyses. Keeping this in mind, we now turn to a reanalysis of one prominent BTSCS study.

#### 4. A REASSESSMENT OF THE LIBERAL PEACE

Russett and his colleagues (Russet 1990, 1993; Maoz and Russett 1992; Maoz and Russett 1993; Oneal et al. 1996) have pioneered one of the most important current research projects about the causes of militarized conflict.<sup>32</sup> Their work on the “Liberal Peace” in particular has captivated IR researchers. To date, there have been two components of this liberal peace: a political one (democracies are less likely to fight with other democracies) and an economic one (trading partners are less likely to engage in militarized conflict). In fact, one of the most celebrated propositions in the IR/IPE literature is that democracies do not wage war on one another.<sup>33</sup> Moreover, the classical economic liberal argument that economic interdependence inhibits war has received extensive empirical support for almost two decades.<sup>34</sup> Russett and his colleagues (Oneal et al. 1996) strengthened the confidence in these findings by showing that the effects of economic interdependence and democracy are inversely related to the onset of military hostilities, even when controlling for several important confounding factors. Oneal and Russett (1997) claim to have improved upon the Oneal et al. (1996) specification to further connect these two major strands of research on the causes of conflict. Oneal and Russett, in exploring the interrelationship between liberalism (po-

<sup>31</sup>Only intra-unit missing data is a particular problem for our method. Obviously our proposed method, like any method, is sensitive to the choice of which units to analyze. This choice is often controversial in international relations (see, for example, Bremer 1992; Maoz and Russett 1993; Lemke 1995).

<sup>32</sup>More than seventy-five articles have been published or presented at conferences in the last five years that have relied on case selection criteria, variable measurement, or substantive foci originally developed or pursued by Russett and the members of his group.

<sup>33</sup>This has been confirmed in myriad empirical studies and has prompted Levy (1988) to imply that this is the only law-like generalization in IR. For recent overviews of the democratic peace literature see Chan (1997) and Ray (1997).

<sup>34</sup>The first prominent statistical study was conducted by Polacheck (1980) and the most recent analysis can be found in Gartzke (1978). For a recent overview of the economic interdependence literature see McMillan (1997).

litical and economic) and militarized conflict, found that, during the Cold War era, higher levels of democracy, as well as trade, lowered the probability of hostilities between pairs of nations. These results, seemingly the most robust of their genre, appear to have solidified conventional wisdom regarding the relationships between economic interdependence, democracy, and war. They therefore conclude that the classical liberal prescription for peace, trade, and democracy, is sound.

The Oneal and Russett results are the Russett's research group's most recently published and, arguably, most rigorous, empirical support for the liberal economic and political peace research program. Their observations, however, were presumed to be independent. That is, Oneal and Russett (hereinafter O/R) performed ordinary logit analyses on BTSCS data without accounting for temporal dependence. We use our proposed method to reanalyze their data to see whether their findings survive more appropriate statistical tests.<sup>35</sup>

The dataset we use, generously provided by O/R, contains 20990 dyad-years, comprised of 827 "politically relevant dyads" observed annually from 1951 through 1985.<sup>36</sup> Some dyads are observed for all thirty-five years, while others are observed for a shorter subperiod. The median observation length is twenty-two years.<sup>37</sup> The dependent variable, militarized conflict, is whether or not a dyad engaged in a militarized interstate dispute in a given year. While earlier researchers typically used interstate war as a dependent variable, recent research has frequently examined militarized interstate disputes. Interstate wars are a small subset of militarized interstate disputes. The latter include any event involving the threat or actual use of military force, while the former require a substantial number of battle deaths.<sup>38</sup>

<sup>35</sup>Oneal and Russett propose, in a footnote, a series of methodological solutions. Initially, they regress a variable that is the number of prior years of dyadic peace on the trade variable and then add the residuals from the regression to the logit specification. However, this does not correct for temporal dependence. The temporal variables added to the logit specification to correct for duration dependence may not be arbitrarily changed without undoing the correction for temporally dependent observations. They also claim that a modified Cochrane-Orcutt correction for temporal dependence did not change their results. We know of no way, however, to modify the Cochrane-Orcutt procedure to handle BTSCS data. Finally, O/R report that bootstrapped standard errors differed only slightly from their reported standard errors. Standard bootstrapping, however, does not work with interdependent observations (Freedman and Peters 1984).

<sup>36</sup>A dyad is "politically relevant" if the nations are geographically proximate or if one state is a major power. The analysis of politically relevant dyad-years is a prominent IR BTSCS design. The limitation of the dataset to politically relevant dyads, or a particular conception of relevancy, is not without criticism, but is irrelevant to the methodological issues we are concerned with.

<sup>37</sup>Their data set has gaps in some dyadic observations. Although we did *not* attempt to fill these in, we did correct the temporal variables for these gaps.

<sup>38</sup>We initially maintain O/R's coding decision to count every dispute year as a separate conflict, even when many of these were merely a continuation of the same event. As we shall see, this decision turns out to have been crucial.

The two key independent variables, democracy and trade, are both dyadic measures. The dyadic democracy variable is constructed by creating democracy scores (using Polity III data) for each member of the dyad and taking the dyadic score as the lesser of the two (Oneal and Russett refer to this as the "weak link" assumption). We rescaled democracy to run from -1 to 1. The trade variable measures the importance of dyadic trade to the less trade-oriented of the two partners. The importance of trade is measured by the ratio (in percent) of dyadic trade to the GDP of each partner. Following O/R, trade is lagged one year so that low trade does not proxy a current dispute.

O/R also use a series of control variables. Alliance is a dummy variable measuring whether the dyad partners were allied (or both were allied with the United States). Contiguous is a dummy variable indicating the geographical contiguity of both states. Capability Ratio measures the dyadic balance of power. Using the Correlates of War material capabilities index, it is the ratio (in percent) of the stronger nation's score to the weaker nation's. Finally, economic growth measures the lesser of the rates of economic growth (as a percent) of the partners. Detailed discussion of the O/R data set and research design is contained in their original paper.<sup>39</sup> The analyses which correct for duration dependence either use a natural cubic spline in a variable we call peace years or the set of dummy variables created from peace years. Peace years counts the length of the spell of peace preceding the current observation. For observations with no previous dyadic disputes, this variable is simply  $t - 1$ , since the time index starts at zero; subsequent to a dispute, this variable is  $t - t_0$  (where  $t_0$  is the time index of the most recent dispute). The variable peace years ranges from zero to 34.

Column I of Table 1 shows the original O/R results. We were able to replicate Oneal and Russett's (1997) original estimates exactly.<sup>40</sup> We limited our reanalysis to the temporal dependence issues discussed in our paper. Examining only their specification 1, we will not present alternative substantive models of conflict.<sup>41</sup> Results indicate that both democracy and trade lower the probability of a militarized dispute; they appear to be both statistically and substantively significant. The control variables, as O/R predicted, also exhibit substantively important effects.

A different picture emerges, however, when we correct for temporally dependent observations using grouped duration methods. (See Columns II

<sup>39</sup> Since our interest is in examining the consequences of temporal dependence, we do not consider issues of operationalization or case-selection, nor do we consider specifications outside the O/R framework.

<sup>40</sup> All of our analyses were done with Stata, Version 5. Note that some variables were rescaled to simplify the reading of the tables.

<sup>41</sup> Their other specifications are similar to that examined here. We have applied our method to their specifications two through six and obtained similar results.

**Table 1. Comparison of Ordinary Logit and Grouped Duration Analyses**

Variable	Ordinary Logit	Grouped Duration		
	I	Logit Dummy <sup>a</sup> II	Logit Spline III	Cloglog Dummy <sup>b</sup> IV
Democracy	-0.50 (0.07)	-0.55 (0.08)	-0.54 (0.08)	-0.49 (0.07)
Economic Growth	-2.23 (0.85)	-1.15 (0.92)	-1.15 (0.92)	-0.81 (0.76)
Alliance	-0.82 (0.08)	-0.47 (0.09)	-0.47 (0.09)	-0.43 (0.08)
Contiguous	1.31 (0.08)	0.70 (0.09)	0.69 (0.09)	0.55 (0.08)
Capability Ratio	-0.31 (0.04)	-0.30 (0.04)	-0.30 (0.04)	-0.30 (0.04)
Trade	-66.13 (13.44)	-12.67 (10.50)	-12.88 (10.51)	-12.50 9.96
Constant	-3.29 (0.08)	-0.94 (0.09)	-0.96 (0.09)	-1.11 (0.08)
Peace Years			-1.82 (0.11)	
Spline(1) <sup>c</sup>			-.24 (0.03)	
Spline(2) <sup>c</sup>			-.08 (0.01)	
Spline(3) <sup>c</sup>			-.01 (0.003)	
Log Likelihood	-3477.6	-2554.7	-2582.9	-2554.1
df	20983	20036	20979	20949
N=20990				

Standard errors in parentheses

<sup>a</sup>31 temporal dummy variables in specification not shown<sup>b</sup>3 dummy variables and 916 observations dropped due to outcomes being perfectly predicted<sup>b</sup>34 temporal dummy variables in specification not shown<sup>c</sup>Coefficients of Peace Years cubic spline segments

and III for the logit link and Column IV for the cloglog link results.) A test for whether the temporal dummies (a likelihood ratio test of I vs. II), or the temporal splines (I vs. III), are required indicate strong duration

dependence.<sup>42</sup> Likelihood ratio tests of specification II versus I yielded  $\chi^2$  statistics of 1778 with 31 degrees of freedom; the test of III versus I yields a statistic of 1789 with 4 degrees of freedom. The test of I versus II drops 916 perfectly predicted observations from both logits so that log likelihoods are comparable. The probability of obtaining either result by chance is, to computer precision, zero.<sup>43</sup> Thus the O/R logits clearly show duration dependence.

Estimation that accounts for duration dependence has dramatic consequences for the O/R finding. The coefficient of trade, in particular, is reduced by a factor of five (and becomes statistically insignificant). These results provide no evidence for a liberal economic peace. Not all coefficients, however, are affected by controlling for duration dependence. Our reanalysis leaves the democracy coefficient and standard error basically unchanged. Thus the Oneal and Russett evidence in support of the liberal economic peace is an artifact of their incorrect assumption of temporal independence; their findings about the liberal political peace, however, are upheld.<sup>44</sup> The consequence of controlling for duration dependence on any variable is difficult to predict in advance. But, as we see here, the consequences of failing to correctly account for duration dependence in typical logit estimation may be enormous.<sup>45</sup>

#### **4.1 Links and Splines**

A closer examination of Table 1 reveals that it appears to make little difference whether we use the logit or cloglog link. The estimates for the two different links are even more similar than Table 1 (Columns II and IV) indicates, since the transformation of the independent variables into probabilities differs slightly between the two links. The mean difference in predicted probabilities between the two models is 0.007 percent, with only about 2 percent of all dyad-years having predicted probabilities of a dispute differing by more than 1 percent. Thus, as we recommended, subsequent analyses use only the logit link.

<sup>42</sup>This can also be seen by looking at the *t*-ratios of the four terms that comprise the cubic spline in peace years: 14, 9, 7, and 4.

<sup>43</sup>Tests on specifications in subsequent tables reveal similar results and are not shown here.

<sup>44</sup>The control variables are also differentially impacted. Although the coefficient of capability ratio is almost unchanged, in magnitude and statistical significance, the coefficients of the three other control variables are cut by half (with the economic growth coefficient even becoming statistically insignificant).

<sup>45</sup>We have applied these methods to their specifications 2 through 6 and obtained similar results. For example, Oneal and Russett (1997, 283) state that they “re-estimated [their] equation (6) with indicator variables for all years but one [with] results consistent with those [they] report.” We reestimated their Equation 6 with temporal dummies and found that the coefficient on trade dropped by a factor of four, (becoming statistically insignificant), and that the coefficient on the trend of trade dropped by a factor of five (also becoming statistically insignificant). While the coefficients on the two democracy variables declined by 30 percent, they remained strongly statistically significant.

Results using a natural cubic spline in peace years appear in Column III. A comparison of Columns II and III shows that it makes no difference in terms of estimating  $\beta$  whether we use temporal dummy variables or a cubic spline in peace years. Since we prefer the spline setup, all subsequent analyses are performed using the natural cubic spline in the length of prior spells of peace.<sup>46</sup>

#### 4.2 Why Duration Dependence Affects the Findings on Economic Interdependence

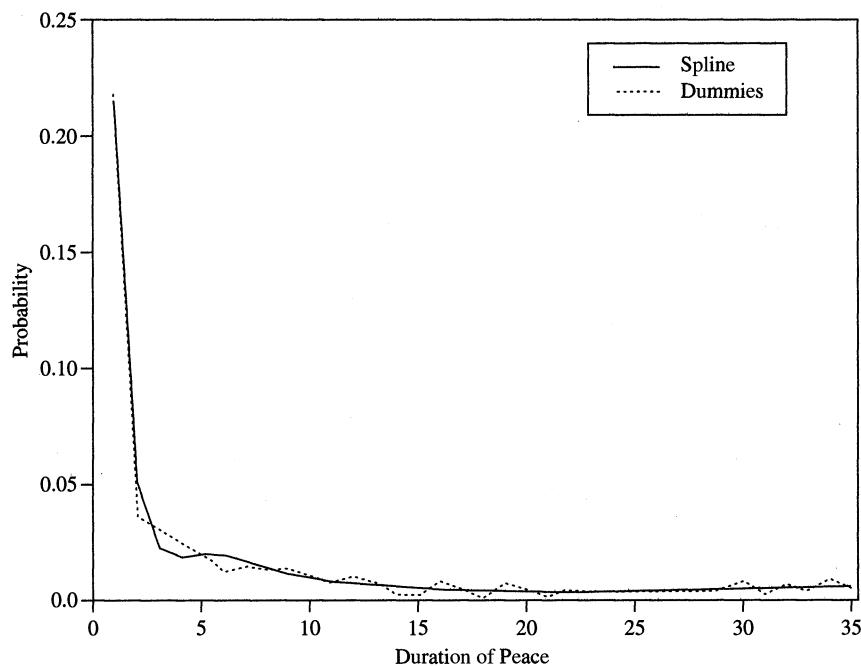
Temporal dependence clearly has dramatic effects on Oneal and Russett's finding that economic interdependence decreases conflict. Oneal and Russett (1997, 283) claim that they theoretically expect a high correlation between trade and length of spells of peace and hence conclude that trade really does lessen conflict. But the problem in this explanation is that it does not take into account the correlation between trade and lengths of spells of conflict, particularly when combined with a higher than average probability of a subsequent conflict immediately following the initial onset.

We can better understand why accounting for duration dependence so strongly affects the O/R finding on trade by using some basic event history techniques. We begin with an examination of the estimated hazard function which is computed for the logit analyses by setting all independent variables at their means (except for the two dummy variables which are set to their modal value of zero). The estimated hazard function, plotted against the length of peace spell, peace years, is shown in Figure 2.

The probability of a dispute immediately following a prior dispute is almost 25 percent. It immediately falls to about 5 percent the next year and to about 2 percent the third year, where it remains for the rest of the spell of peace. Thus, much of what the duration dependent logit highlights is the dependence of the probability of a dispute on an immediately preceding dispute. Counting the latter years of multi-year disputes as new disputes, and failing to correct for dependence between these disputes, is what leads to the Oneal and Russett finding that trade lowers the probability of the onset of a dispute.

It appears that economic interdependence does not dampen the probability of a dispute, but it does diminish the duration of a dispute once it occurs. Remember that trade is lagged one year so that the previous year's

<sup>46</sup>All splines allowed for three knots, placed at 1, 4, and 7 years of peace. The number of knots was chosen by a sequence of  $F$ -tests; a variety of knot placements were tried to ascertain the one with the best performance. Small changes in the number or placement of the knots had no effect on the results. The natural cubic spline estimated here is similar to the smoothing splines shown in Beck and Jackman (1997) and Beck and Tucker (1997). We also reran the analyses for subsequent tables using temporal dummy variables, obtaining almost identical results.

**Figure 2. Discrete Hazard of Dispute**

trade predicts the current probability of a dispute. Trade averages 0.22 percent of GDP prior to one year disputes. This is only slightly lower than the 0.23 percent of GDP that trade averages prior to a year of peace. But, in the last year of peace prior to a multi-year dispute, trade averages only 0.15 percent of GDP. Thus trade is not a good predictor of whether a dispute will occur, but if one does, it is a good predictor of whether it will be lengthy. Low trade may prolong conflicts, but it does not appear to cause them.

#### 4.3 The Effect of Multiple Disputes

##### *The elimination of ongoing dispute years*

We can further examine the contaminating effects of long spells of disputes by eliminating ongoing years of a dispute from the analysis. Five hundred forty-two dyad-years with a dispute are thus dropped.<sup>47</sup> Results of this analysis are in Table 2.

<sup>47</sup>All disputes that continue for more than one year are dropped, even if disputes in subsequent years have different identification codes.

**Table 2. Grouped Duration Analyses: No Continuing Dispute Years**

Variable	I. Logit Ordinary		II. Logit Group Dur.	
	$\hat{\beta}$	se	$\hat{\beta}$	se
Democracy	-0.40	0.10	-0.39	0.10
Economic Growth	-3.43	1.25	-4.01	1.25
Alliance	-0.48	0.11	-0.37	0.11
Contiguous	1.35	0.12	0.99	0.12
Capability Ratio	-0.20	0.05	-0.22	0.05
Trade	-21.08	11.30	-3.81	9.68
Constant	-4.33	0.11	-3.57	0.17
Peace Years			0.39	0.16
Spline(1) <sup>a</sup>			0.09	0.03
Spline(2) <sup>a</sup>			-0.03	0.01
Spline(3) <sup>a</sup>			0.003	0.003
Log Likelihood		-1846.9		-1751.4
df		20441		20437
N = 20448				

<sup>a</sup>Coefficients of Peace Years cubic spline

Dropping the latter years of a dispute, even without accounting for duration dependence, reduces the trade coefficient by a factor of three, leaving it barely statistically significant. A likelihood ratio test, however, clearly shows remaining duration dependence. When we account for this (Column II), the effect of dyadic trade is again greatly reduced and is now not even close to being statistically significant. The elimination of ongoing dispute years, even accounting for duration dependence, has little effect on the democracy coefficient.<sup>48</sup>

#### *Time until first failure*

We can also examine the contaminating effects of disputes on later disputes by confining our analysis to first disputes (eliminating observations on 3999 dyad years which followed an initial dispute). This analysis avoids any problems associated with the need to model the conditional probability of second and later disputes. Results appear in Table 3, Column I.

Limiting our analyses to the onset of the first dispute eliminates about 20 percent of the data and results in an increase in all the standard errors. The pacific effect of democracy remains almost unaffected by this limitation

<sup>48</sup>The estimated pacific impact of economic growth dramatically increases when ongoing dispute years are omitted.

**Table 3. Grouped Duration Analyses: Second Disputes Differ**

Variable	I. First Disputes		II. Prior Disputes	
	$\hat{\beta}$	se	$\hat{\beta}$	se
Democracy	-0.46	0.13	-0.41	0.08
Economic Growth	-2.29	1.78	-2.09	0.97
Alliance	-0.42	0.16	-0.25	0.09
Contiguous	1.11	0.17	0.69	0.09
Capability Ratio	-0.19	0.06	-0.20	0.04
Trade	-3.55	11.73	-9.39	10.19
Prior Disputes (#)			0.17	0.01
Constant	-3.21	0.21	-1.60	0.10
Peace Years	-1.08	0.24	-1.67	0.11
Spline(1) <sup>a</sup>	-0.18	0.05	-0.22	0.03
Spline(2) <sup>a</sup>	-0.07	0.02	0.07	0.01
Spline(3) <sup>a</sup>	-0.01	0.05	-0.01	0.003
Log Likelihood		-964.1		-2393.0
df		16980		20978
N		16691		20990

<sup>a</sup>Coefficients of Peace Years cubic spline segments

(despite a slight increase in its standard error). But once again, the estimated impact of economic interdependence drastically decreases. Increased dyadic trade does not reduce the likelihood of an initial dyadic dispute onset; dyadic democracy does.

A less drastic way to allow for differing conditional probabilities of a dispute given the number of prior disputes is to add to the logit a counter measuring the number of prior dyadic disputes. These results are in Table 3, Column II. While the results are not as dramatic as the limitation to first disputes only, they clearly show the pacific effect of democracy but not of trade. Accounting for temporal dependence clearly has dramatic effects on O/R's finding that trade decreases conflict. When the O/R estimation is corrected for that dependence, the finding that trade reduces conflict simply disappears (although trade may reduce the length of conflicts once they occur). Our reassessment of the O/R finding, however, leaves intact their conclusion about the pacific effects of democracy.

## 5. CONCLUSION

The analysis of binary dependent variable time-series—cross-section data are becoming more common, particularly in the study of international conflict. Virtually all analyses of this type of data use ordinary logit, ignoring issues of temporal interdependence of the data. We have shown that al-

lowing for temporal dependence in logit analysis is easy once we recognize that BTSCS data are grouped duration data. Such data can then be analyzed by adding temporal dummy variables (or a temporal spline) to the logit specification. This can be done using any statistical software package.

This remedy has advantages over other attempts to correct for temporal dependence in BTSCS data. Because it uses standard logit routines, it can be combined with remedies that address other problems. In particular, it is simple to combine our method with Huber (1967) standard errors, which solve other problems inherent in BTSCS data. One can also allow for heterogeneity using our method (Jenkins 1995). Solutions to one problem should be applicable to others; real data are seldom subject to only one problem. A related advantage is that logit is well-known and well understood by researchers.

Our proposed method forces logit analysts to think about some problems that naturally occur to the event history analyst (which do not naturally occur to the logit analyst). In particular, our approach requires analysts to think about whether they are modeling spells of peace, spells of conflict, or both. It also requires pondering the modeling of second and subsequent events for the same unit. These considerations are critical to the modeling process.

There are clearly other possible ways to allow for temporal dependence in BTSCS data. Ideally we would model that dependence as a function of other variables. Our proposed method simply treats serial correlation as a nuisance which impedes estimation of the  $\beta$ . Nothing is explained by noting that hazard rates change with time.

Our treatment of duration dependence parallels the earlier methods for estimating time series models with serially correlated errors that treats the dependence as an estimation nuisance. While we agree with more modern treatments that advocate directly modeling the dynamics instead of treating it as a nuisance, we note that simply ignoring the nuisance can lead to severely incorrect inferences. A theoretically based specification of this duration dependence would be best, but it is easier to give this advice than to implement it. In the meantime we surely do not wish to continue the current practice of ignoring potentially serious duration dependence.

The analogy to the estimation of time series models with serially correlated errors is also helpful in understanding when duration dependence might cause serious problems for BTSCS data. As with serially correlated errors in a linear regression, a small amount of duration dependence, even if statistically significant, will only cause a small amount of harm. Thus researchers must assess not only whether they can reject the null hypothesis of duration independence, but also the severity of duration dependence (by examining the estimated baseline hazards). Also, as with serially correlated

errors, the degree of harm is related to the temporal structure of the independent variables. Estimation problems caused by ignoring duration dependence increase as the independent variables themselves trend or otherwise show time variation. Duration dependence, then, will not always have enormous consequence. For example, duration dependence should be less of a problem in data where the units are observed less frequently, so that dyad-year data will generally show more duration dependence than will dyad-decade data. But, as with serially correlated errors, researchers cannot ignore the potential problems that might be caused by duration dependence. Fortunately our proposed test and correction are easy to implement, so there is no reason for research to ignore these potential problems.

We have applied our methodological remedy to Oneal and Russett's (1997) analysis of militarized conflict during the Cold War period, which found that both political and economic liberalism inhibit conflict. Our re-analyses show that democracy clearly inhibits conflict, but that trade (at least as Oneal and Russett measure it) does not. But while trade may not inhibit conflict, it does appear to shorten spells of conflict. The differences between the original analysis and our reanalysis are considerable. Temporal dependence in BTSCS models is not a minor problem that can be ignored at the cost of a small error. And there is no reason to commit these errors. The inclusion of temporal variables in the specification is a simple solution, available to all researchers, providing a low cost cure to the problem of temporally dependent BTSCS data.

*Manuscript submitted 3 September 1997.*

*Final manuscript received 12 December 1997.*

---

## APPENDIX

### The Simple Math of Grouped Durations

---

This appendix derives the grouped duration model. We present it here for completeness and because many standard event history texts do not present this result. This appendix assumes familiarity with basic duration concepts.

Start with a continuous time Cox proportional hazards model, so

$$h_i(t) = h_0(t)e^{\mathbf{x}_{i,t}\beta} \quad (9)$$

where  $i$  refers to units,  $t$  refers to continuous time,  $\mathbf{x}_{i,t}$  is a vector of independent variables, and  $h_0(t)$  is the unspecified baseline hazard.

Letting  $S(t)$  be the probability of surviving beyond  $t$ , we use the basic identity that

$$S(t) = \exp\left(-\int_0^t h(\tau)d\tau\right) \quad (10)$$

We only observe whether or not an event occurred between time  $t_k - 1$  and  $t_k$  (assuming annual data) and are interested in the probability of this event,  $P(y_i, t_k = 1)$ . This probability is one minus the probability of surviving beyond  $t_k$  given survival up to  $t_k - 1$ . Assuming no prior events (so  $t0 = 0$ ) and using Equation 10, we then get

$$P(y_{i,t_k} = 1) = 1 - \exp\left(-\int_{t_{k-1}}^{t_k} h_i(\tau)d\tau\right) \quad (11a)$$

$$= 1 - \exp\left(-\int_{t_{k-1}}^{t_k} e^{\mathbf{x}_{i,t_k}\beta} h_0(\tau)d\tau\right) \quad (11b)$$

$$= 1 - \exp\left(-e^{\mathbf{x}_{i,t_k}\beta} \int_{t_{k-1}}^{t_k} h_0(\tau)d\tau\right) \quad (11c)$$

(Note that  $\mathbf{x}$  is indexed by  $t_k$  not  $\tau$  because we assume that the independent variables are only measured for an entire interval and not for every instant in the interval  $t_k - 1$  to  $t_k$ .) Since the baseline hazard is unspecified, we can just treat the integral of the baseline hazard as an unknown constant. Defining

$$\alpha_{t_k} = \int_{t_{k-1}}^{t_k} h_0(\tau)d\tau \text{ and} \quad (12)$$

$$\kappa_{t_k} = \log(\alpha_{t_k}) \quad (13)$$

we then have

$$P(y_{i,t_k} = 1) = 1 - \exp\left(-e^{\mathbf{x}_{i,t_k}\beta}\alpha_{t_k}\right) \quad (14a)$$

$$= 1 - \exp\left(-e^{\mathbf{x}_{i,t_k}\beta + \kappa_{t_k}}\right) \quad (14b)$$

This is exactly a binary dependent variable model with a cloglog link

---

## REFERENCES

- Allison, Paul D. 1982. "Discrete-Time Methods for the Analysis of Event Histories." In *Sociological Methodology*, ed. Samuel Leinhardt. San Francisco: Jossey-Bass. pp. 61–98.
- Alt, James E., Gary King, and Curtis Signorino. 1997. Estimating the Same Quantities from Different Levels of Data: Time Dependence and Aggregation in Event Process Models. Technical report. Department of Government, Harvard University (<http://GKing.Harvard.Edu/preprints.shtml>).
- Barbieri, Katherine. 1996. "Economic Interdependence: A Path to Peace or a Source of Interstate Conflict?" *Journal of Peace Research* 33:29–50.
- Beck, Nathaniel. N.d. "Modelling Space and Time: The Event History Approach." In *Research Strategies in Social Science*, ed. Elinor Scarbrough and Eric Tanenbaum. Oxford: Oxford University Press.
- Beck, Nathaniel, and Jonathan N. Katz. 1995. "What To Do (and Not To Do) with Times-Series Cross-Section Data." *American Political Science Review* 89:634–47.

- Beck, Nathaniel, and Jonathan N. Katz. 1996. "Nuisance vs. Substance: Specifying and Estimating Time-Series-Cross-Section Models." *Political Analysis* 6:1–36.
- Beck, Nathaniel, and Jonathan N. Katz. 1997. "The Analysis of Binary Time-Series-Cross-Section Data and/or The Democratic Peace." Presented at the annual meeting of the Political Methodology Group, Columbus, OH.
- Beck, Nathaniel, and Richard Tucker. 1997. Conflict in Time and Space. Center for International Affairs Working Paper, No. 97–8. Harvard University (<http://www1.columbia.edu/sec/dlc/ciao/wpsfrm.html>).
- Beck, Nathaniel, and Simon Jackman. 1997. Getting the Mean Right is a Good Thing: Generalized Additive Models. Working paper. Political Methodology WWW Site ([http://wizard.ucr.edu/polmeth/working\\_papers97](http://wizard.ucr.edu/polmeth/working_papers97)).
- Bennett, D. Scott. 1996. "Security, Bargaining, and the End of Interstate Rivalry." *International Studies Quarterly* 40:157–84.
- Bennett, D. Scott. 1997. "Testing Alternative Models of Alliance Duration, 1816–1984." *American Journal of Political Science* 41:846–78.
- Berry, Frances Stokes, and William D. Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *American Political Science Review* 84:395–415.
- Berry, Frances S., and William D. Berry. 1992. "Tax Innovation in the States: Capitalizing on Political Opportunity." *American Journal of Political Science* 36:715–42.
- 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* 41:1414–61.
- Bremer, Stuart A. 1992. "Dangerous Dyads: Conditions Affecting the Likelihood of Interstate War, 1816–1965." *Journal of Conflict Resolution* 36:309–41.
- Bueno de Mesquita, Bruce, and David Lalman. 1992. *War and Reason: Domestic and International Imperatives*. First ed. New Haven: Yale University Press.
- Chan, Steve. 1997. "In Search of Democratic Peace: Problems and Promise." *Mershon International Studies Review* 41:59–91.
- Cox, D. R. 1975. "Partial Likelihood." *Biometrika* 62:269–76.
- Diggle, Peter J., Kung-Yee Liang, and Scott L. Zeger. 1994. *Analysis of Longitudinal Data*. Oxford: Oxford University Press.
- Enterline, Andrew. 1996. "Driving While Democratizing." *International Security* 20:183–196.
- Enterline, Andrew. 1997. "Fledgling Regimes: Is the Case of Inter-War Germany Generalizable?" *International Interactions* 22:245–277.
- Eubank, Randall L. 1988. *Spline smoothing and nonparametric regression*. New York: Marcel Dekker.
- Farber, Henry S., and Joanne Gowa. 1997. "Common Interests or Common Polities: Reinterpreting the Democratic Peace." *Journal of Politics* 59:393–417.
- Farhniér, Ludwig, and Stefan Wagenpfeil. 1996. "Smoothing Hazard Functions and Time-Varying Effects in Discrete Duration and Competing Risks Models." *Journal of the American Statistical Association* 91:1584–94.
- Freedman, David, and Stephen Peters. 1984. "Bootstrapping a Regression Equation: Some Empirical Results." *Journal of the American Statistical Association* 79:97–106.
- Gartzke, Erik. 1998. "Kant We All Just Get Along?: Opportunity, Willingness and the Origins of the Democratic Peace." *American Journal of Political Science* 42:1–27.
- Gleditsch, Nils Petter, and Håvard Hegre. 1997. "Peace and Democracy: Three Levels of Analysis." *Journal of Conflict Resolution* 41:283–310.
- Goertz, Gary, and Paul Diehl. 1993. "Enduring Rivalries: Theoretical Constructs and Empirical Patterns." *International Studies Quarterly* 37:147–71.
- Han, Aaron, and Jerry A. Hausman. 1990. "Flexible Parametric Estimation of Duration and Competing Risk Models." *Journal of Applied Econometrics* 5:1–28.

- Henderson, Errol A. 1997. "Ethnic Conflict, The Similarity of States, and the Onset of War, 1820–1989." *Journal of Conflict Resolution* 4:649–68.
- Hermann, Margaret G., and Charles W. Kegley. 1996. "Ballots, a Barrier against the Use of Bullets and Bombs: Democratization and Military Intervention." *Journal of Conflict Resolution* 40:436–60.
- Huber, Peter J. 1967. The Behavior of Maximum Likelihood Estimates Under Non-Standard Conditions. In *Proceedings of the Fifth Annual Berkeley Symposium on Mathematical Statistics and Probability*, ed. Lucien M. LeCam and Jerzy Neyman. Vol. I. Berkeley: University of California Press.
- Huth, Paul. 1996. *Standing Your Ground: Territorial Disputes and International Conflict*. Ann Arbor: University of Michigan Press.
- Jenkins, Stephen P. 1995. "Easy Estimation Methods for Discrete-Time Duration Models." *Oxford Bulletin of Economics and Statistics* 57:129–38.
- Katz, Jonathan N., and Brian Sala. 1996. "Careerism, Committee Assignments and the Electoral Connection." *American Political Science Review* 90:21–33.
- Lemke, Douglas. 1995. "The Tyranny of Distance: Redefining Relevant Dyads." *International Interactions* 21:23–38.
- Lemke, Douglas, and William Reed. 1996. "Regime Types and Status Quo Evaluations: Power Transition Theory and the Democratic Peace." *International Interactions* 22:143–164.
- Levy, Jack. 1988. "Domestic Politics and War." *Journal of Interdisciplinary History* 18:653–73.
- Little, Roderick J. A., and Donald B. Rubin. 1987. *Statistical Analysis with Missing Data*. New York: Wiley.
- Mansfield, Edward, and Jack Snyder. 1996. "The Effects of Democratization on War." *International Security* 20:196–207.
- Mansfield, Edward, and Jack Snyder. 1997. "A Response to Thompson and Tucker." *Journal of Conflict Resolution* 41:457–461.
- Maoz, Zeev. 1996. *Domestic Sources of Global Change*. Ann Arbor: University of Michigan Press.
- Maoz, Zeev. N.d. "Realist and Cultural Critiques of the Democratic Peace: A Theoretical and Empirical Re-Assessment." *International Interactions*. Forthcoming.
- Maoz, Zeev, and Bruce Russett. 1992. "Alliance, Contiguity, Wealth, and Political Stability: Is the Lack of Conflict Among Democracies A Statistical Artifact?" *International Interactions* 17:245–267.
- Maoz, Zeev, and Bruce Russett. 1993. "Normative and Structural Causes of Democratic Peace, 1946–1986." *American Political Science Review* 87:639–56.
- McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*. 2nd ed. London: Chapman and Hall.
- McMillan, Susan M. 1997. "Interdependence and Conflict." *Mershon International Studies Review* 41:33–58.
- Meier, Kenneth J., and Deborah R. McFarlane. 1992. "State Policies on Funding of Abortions: A Pooled Time Series Analysis." *Social Science Quarterly* 73:690–8.
- Mintrom, Michael. 1997. "Policy Entrepreneurs and the Diffusion of Innovation." *American Journal of Political Science* 41:738–70.
- Mousseau, Michael. 1997. "Democracy and Militarized Interstate Collaboration." *Journal of Peace Research* 34:73–87.
- Oneal, John R., and Bruce Russett. 1997. "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950–1985." *International Studies Quarterly* 41:267–94.
- Oneal, John R., Francis H. Oneal, Zeev Maoz, and Bruce Russett. 1996. "The Liberal Peace: Interdependence, Democracy and International Conflict." *Journal of Peace Research* 33:11–29.
- Poirier, Dale J., and Paul A. Ruud. 1988. "Probit with Dependent Observations." *Review of Economic Studies* 55:593–614.

- Polacheck, Solomon. 1980. "Conflict and Trade." *Journal of Conflict Resolution*. 24:55–78.
- Prentice, R., and L. Gloeckler. 1978. "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data." *Biometrics* 34:57–67.
- Raknerud, Arvid, and Håvard Hegre. 1997. "The Hazard Of War: Reassessing Evidence of the Democratic Peace." *Journal of Peace Research* 34:385–404.
- Ray, James Lee. 1997. "The Democratic Path to Peace." *Journal of Democracy*. 8:49–64.
- Russett, Bruce. 1990. "Economic Decline, Electoral Pressure, and the Initiation of International Conflict." In *The Prisoners of War*, ed. Charles Gochman and Ned Allan Sabrosky. Lexington, Mass.: Lexington Book.
- Russett, Bruce. 1993. *Grasping the Democratic Peace*. Princeton: Princeton University Press.
- Singer, Judith D., and John B. Willett. 1993. "It's About Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events." *Journal of Educational Statistics* 18:155–95.
- Sueyoshi, Glenn T. 1995. "A Class of Binary Response Models for Grouped Duration Data." *Journal of Applied Econometrics* 10:411–31.
- Vasquez, John A. 1993. *The War Puzzle*. Cambridge: Cambridge University Press.