Back to the Future:
Modeling Time Dependence in Binary Data

David B. Carter
Curtis S. Signorino*

Department of Political Science
University of Rochester

Work in Progress
Comments Welcome

November 1, 2007

Abstract

Since Beck, Katz, and Tucker (1998), the use of time dummies or splines has become the standard method to model temporal dependence in binary data. We show that there are potential problems with each of these approaches, especially in the case of time dummies. We propose a simpler alternative: using t, t^2, and t^3 to approximate the hazard. This cubic polynomial is trivial to implement and avoids problems with time dummies such as quasi-complete separation and issues with splines such as interpretation or knot selection. It also accommodates non-proportional hazards in a much simpler way than either time dummies or splines. We show via monte carlo analysis that our method performs as well as splines and better than time dummies. We also demonstrate this method with reanalyses of a number of empirical studies such as Oneal and Russett (1997) and Crowley and Skocpol (2001).

*We thank Neal Beck, Giacomo Chiozza, Jocelyn Crowley, Miguel de Figuereido, John Freeman, Hein Goemans, Luke Keele, Eduardo Leoni, Arthur Spirling, and Randy Stone for their comments and suggestions, as well as the participants in the Watson Center for Conflict and Cooperation Seminar Series, the University of Minnesota Political Economy Seminar Series, and the University of California-Berkeley IR Colloquium. We also thank Jocelyn Crowley, Theda Skocpol, Matt Dickinson, and Glenn Palmer for providing their data. We gratefully acknowledge support from the National Science Foundation (SES-0213771, SES-0413381). Emails: `dcarter5@mail.rochester.edu`, `curt.signorino@rochester.edu`.

Contents

1 Introduction

Whether it is the primary focus or not, political scientists are often interested in how the phenomena they study change over time. For example, are democracies more stable the longer they have been established? Are congressional incumbents more likely to survive an election the longer they have been in office? Are two recently warring nations more or less likely to become embroiled in another conflict as time goes on? Ultimately, we would like to understand the causal mechanisms that cause the subject of interest to change (or not) over time. Short of that, we still need some way to account for time in our empirical analyses.

Increasingly, researchers have access to refined (or higher resolution) versions of event history data. In one common form of this data, a binary dependent variable represents whether the event occurred or not during some slice of time. First advocated by Beck, Katz and Tucker (1998), the logit model with time dummies or splines has become the standard method for analyzing this type of data. Researchers who study a wide variety of topics in international relations, American politics, and comparative politics have all adopted the Beck, Katz, and Tucker (1998) (hereafter, BKT) recommendations.

We should be clear in stating from the outset that we completely agree with BKT's main point: scholars should "take time seriously".[1] However, the vast majority of researchers have treated temporal dependence in binary data models more as a statistical nuisance that needs to be "controlled for," rather than as something that is substantively interesting. Indeed, as we will later show, most of those who have followed BKT's advice subsequently ignore time in discussions of empirical results.

Consider Table 1, which lists all published articles we found that follow BKT's advice on using time dummies or splines.[2] We have further classified the citations according to whether the authors interpreted the effect of time or not. Table 1 demonstrates that despite both splines and time dummies being extensively used in every substantive field of political science, virtually no one actually plots and interprets the hazard. In fact, out of 91 studies that utilize splines, only 3 actually plot and interpret a hazard. The track record for time dummies is slightly better, but out of 28 studies that utilize dummies, only

[1] For example, see Alt, King and Signorino (2001).

[2] We compiled this list by locating all published articles that cited BKT in the Social Sciences Citation Index (SSCI) as of July, 2006. We then went through all the articles to determine whether they implemented either splines or time dummies and whether they interpreted the hazard.

4 plot the hazard. In short, the discipline's track record shows that rarely is time taken seriously in a substantial way.

What accounts for researchers taking BKT seriously, but not time? We suspect it stems from the difficulties in either the implementation or interpretation of the two methods they propose. As we will show, serious practical problems can arise in the implementation of time dummies. Researchers can avoid those problems by using splines. Moreover, researchers do not even need to know how to generate the splined time data: Richard Tucker provided a Stata routine (btscs) that researchers can use to include these variables in their regressions. However, most political scientists do not seem to understand what splines are or how to interpret them in their regressions.

In this paper, we propose a simpler alternative that has advantages in terms of both modeling and interpreting time dependence: using t, t^2, and t^3 in one's regression, which serves as a third-order Taylor series approximation to the hazard. In contrast to time dummies, this cubic polynomial approximation is trivial to implement and does not cause the same data problems.[6] The cubic polynomial is similar in many ways to splines, but much easier to interpret. In addition, the cubic polynomial is advantaged relative to splines if researchers are interested in potential non-proportional effects in time among their regressors.

This paper proceeds as follows. In the next section, we briefly discuss the link between duration models and their binary data equivalents. Following that, we examine the implementation and interpretation issues with time dummies and splines. In Section 5, we show via monte carlo analysis that our method generally performs as well as splines and better than time dummies. We then demonstrate this method with empirical replications of studies including Oneal and Russett (1997) and Crowley and Skocpol (2001). Then, we extend the findings of Crowley and Skocpol (2001) using a nonproportional hazards model to demonstrate how the temporal component in many theories can be more richly explored using our method.

[3]The authors do not show the hazard in the text of the article but do offer an appendix with a plot of the hazard upon request.

[4]They do not actually plot the full hazard but do report the hazard for specific years in a table.

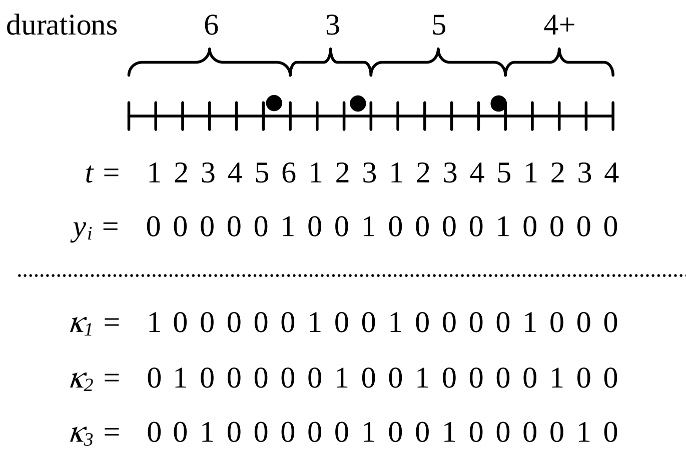
[5]The authors include a table that shows the effect of two covariates on the hazard, but do not show the hazard as a function of time.

[6]We use the term cubic polynomial and t, t^2, and t^3 interchangeably throughout the paper.

Table 1: Use of Splines and Time Dummies

	No Hazard	**Interpret Hazard**
Splines	**88 (96.7%)** Goodcliffe and Hawkins (2006)[3] Hafner-Burton and Montgomery (2006) Dorussen (2006) Caprioli and Trumbore (2006) Rasler and Thompson (2006) Mansfield and Pevehouse (2006) Braithwaite (2005) Buhaug (2005) Benson (2005) Oneal and Russett (2005) Meinke, Staton and Wuhs (2006) Bearce and Omori (2005) Edwards (2005) Melander (2005) Enterline and Greig (2005) Barbieri and Reuveny (2005) Chamberlain and Haider-Markel (2005) Kim and Rousseau (2005) Caprioli and Trumbore (2005) Senese (2005) Besancon (2005) Humphries (2005) Lujala, Gleditsch and Gilmore (2005) Marinov (2005) Caprioli (2005) Sobek (2005) Milner and Kubota (2005) Boehmer and Sobek (2005) Sørli, Gleditsch and Strand (2005) Fritz and Sweeney (2004) Krause (2004) Powers (2004) Mitchell and Prins (2004) Gowa and Mansfield (2005) Sechser (2004) Rasler and Thompson (2004) Goenner (2004) McDonald (2004) Volden and Carruba (2004) Walter (2004) Schamis and Way (2003) Sweeney and Fritz (2003) Cauthen and Peters (2003) Lai (2003*a*) Leblang and Chan (2003) Ray (2003) Reed (2003*b*) Oneal (2003) Gartzke and Li (2003*c*) Gartzke and Li (2003*a*) Sweeney (2003) Choi and James (2003) Mansfield and Reinhardt (2003) Gartzke and Li (2003*b*) Ravlo, Gleditsch and Dorussen (2003) Senese and Vasquez (2003) Bennett and Rupert (2003) Montinola (2003) Lai (2003*b*) Reed (2003*a*) Clark and Regan (2003) Chiozza (2002) Pickering (2002) Kinsella and Russett (2002) Jungblut and Stoll (2002) Davies (2002) Dixon and Senese (2002) Mousseau (2002) Colaresi and Thompson (2002) Peceny, Beer and Sanchez-Terry (2002) Li and Sacko (2002) Zorn (2001) Rasler and Thompson (2001) Drury (2001) Mousseau (2001) Crescenzi and Enterline (2001) Sherman (2001) Henderson and Tucker (2001) Gartzke, Li and Boehmer (2001) Bennett and Stam (2000*a*) Lektzian and Souva (2001) Lemke and Reed (2001) Peet and Simon (2000) Mansfield and Pevehouse (2000) Bennett and Stam (2000*b*) Mousseau (2000) Oneal and Russett (1999*b*) Oneal and Russett (1999*a*)	**3 (3.3%)** Gelpi and Grieco (2001) Simmons (2000) Beck, King and Zeng (2000)
Time Dummies	**24 (85.7%)** Brinks and Coppedge (2006) Volden (2006) Stein (2005) Chang (2005) Krutz (2005) Heath (2005) Lebovic (2004) Stalley (2003) Arceneaux (2003) Lebovic (2003) Gelpi and Feaver (2002) Howard and Roch (2001) Henisz (2002) Dickinson and Tenpas (2002) Ka and Teske (2002) Volden (2002) Ka and Teske (2002) Crowley and Skocpol (2001)[5] Balla (2001) Mooney (2001) Mooney and Lee (2000) Palmer and Whitten (2000) Reed (2000) Thacker (1999) Leblang (1999)	**4 (14.3%)** James (2006) Carpenter and Lewis (2004) Bernard, Reenock and Nordstrom (2003)[4] Clark and Hart (1998)

Figure 1: Binary Representation of Duration Data



2 Temporal Dependence in Binary (Event History) Data

The starting point for BKT is the observation that, increasingly, the binary data we use in political science is a disaggregated (or less aggregated) form of event history data. Although BKT refer to this data as binary time-series cross-section (BTSCS) data, the focus is really on temporal dependence, rather than cross-sectional interdependence. We will similarly focus on time here.

To help make this more concrete, consider the time line of events displayed in Figure 1. (For the time being, ignore the section below the dotted line.) Here, we have three events, denoted by the black dots. The durations, or time between successive events, are shown along the top: 6, 3, and 5. If we were analyzing duration data, our observations would generally correspond to these, along with the last right-censored observation of length 4. Vast literatures exist on duration (or survival) analysis (see, for example, Box-Steffensmeier and Jones (2004)). Rather than reviewing that literature, we simply point out that there are well-known techniques for modeling temporal dependence in duration data. Parametric models like the weibull, log-logistic, or log-normal allow the analyst to estimate whether the hazard is increasing, decreasing, or non-monotonic with time. Alternatively, researchers sometimes opt for the (semi-parametric) Cox model.

Now consider the binary data shown below the timeline and denoted as y_i. As BKT and others have noted, if the data generating process is temporally dependent, then the use of a model such as logit with only a linear $x\beta$ specification is inappropriate, since it implies a constant hazard.[7] The question then becomes one of how to allow for temporal dependence in binary data without being too restrictive concerning the form of that dependence.

As BKT show, derivation of the binary data version of a Cox model is relatively straightforward (Beck, Katz and Tucker, 1998; Meyer, 1990; Narendranathan and Stewart, 1993; Prentice and Gloeckler, 1978; Katz and Sala, 1996). Rather than reproduce their derivation, we refer the reader to BKT (see their appendix). We simply note that if one starts with the assumption of a Cox relative risk model with hazard

$$h(t|x_{i,t}) = h_0(t)\exp(x_{i,t}\beta) \tag{1}$$

then the equivalent binary data model is a complementary log-log (cloglog) model, which can be written as

$$Pr(y_{i,t} = 1) \quad = \quad 1 - \exp(-\exp(x_{i,t}\beta + \kappa_{i,t}\gamma_t)). \tag{2}$$

where t indicates the time since the last event (or the life of the current subject), κ_t is a time dummy for period t, and γ_t is the effect of period t on the probability of an event occurring.[8] It is important to note that the "time dummies" are not just the time counter t. Returning to Figure 1, we have displayed three time dummies below the dotted line. For example, κ_1 will be one whenever $t = 1$ and zero otherwise. Similarly, κ_2 will equal one whenever $t = 2$ and zero otherwise. In general, one needs a κ_t for every t in the data (although more about this later). If the time unit of analysis is the year and the longest duration in the data is 10 years, then there will be ten time dummies. If it is 30, then there will be 30 time dummies. The researcher can include all time dummies and drop the constant or include the constant and drop one time dummy. If time dependence is present in data and its effects are not estimated then omitted variable bias is present, which can bias other coefficients of interest in commonly used binary dependent variable models even when the omitted variables are not correlated with included variables (Yatchew and Griliches, 1985).[9]

[7]See Alt, King and Signorino (2001) for a demonstration of this when durations are distributed gamma.

[8]It is a slight abuse of terminology to refer to the plots of $\Pr(y_{i,t} = 1)$ as plots of the "hazard." However, as the aforementioned proofs demonstrate (e.g., in BKT's appendix), the relationship between $\Pr(y_{i,t} = 1)$ and time is identical to that of the hazard, upon which it is based. An increasing hazard produces a $\Pr(y_{i,t} = 1)$ that increases with time; a decreasing hazard produces a $\Pr(y_{i,t} = 1)$ that decreases with time; and so on for nonmonotonic hazards.

[9]Beck, Katz and Tucker (1998) argue that the main problem encountered when time dependence is not dealt with is overly small standard errors. This result is obtained when autocorrelation is present in the data (Beck and Katz, 1997).

The cloglog model in equation 2 has the nice feature that we do not have to assume any particular hazard shape (e.g., increasing, decreasing, or non-monotonic). In principle, this allows consistent β estimates, even if the true shape of the hazard is unknown (Meyer, 1990). However, as BKT demonstrate, the logit model closely approximates the cloglog. Moreover, there is no reason to privilege the cloglog distribution over logit. Therefore, BKT recommend that researchers use logit (or probit), since they are so widely available.[10] BKT advocate the use of natural cubic splines instead of temporal dummies when researchers want a relatively smooth plot of the hazard.

3 Time Dummies for Dummies

Two major problems plague the use of time dummies, complete or quasi-complete data separation and inefficiency. While BKT briefly discuss efficiency issues, they do not discuss the potentially serious data separation problem associated with time dummies. Both problems apply regardless of whether one estimates a logit, cloglog, probit or a number of other models.

3.1 Complete and Quasi-Complete Separation

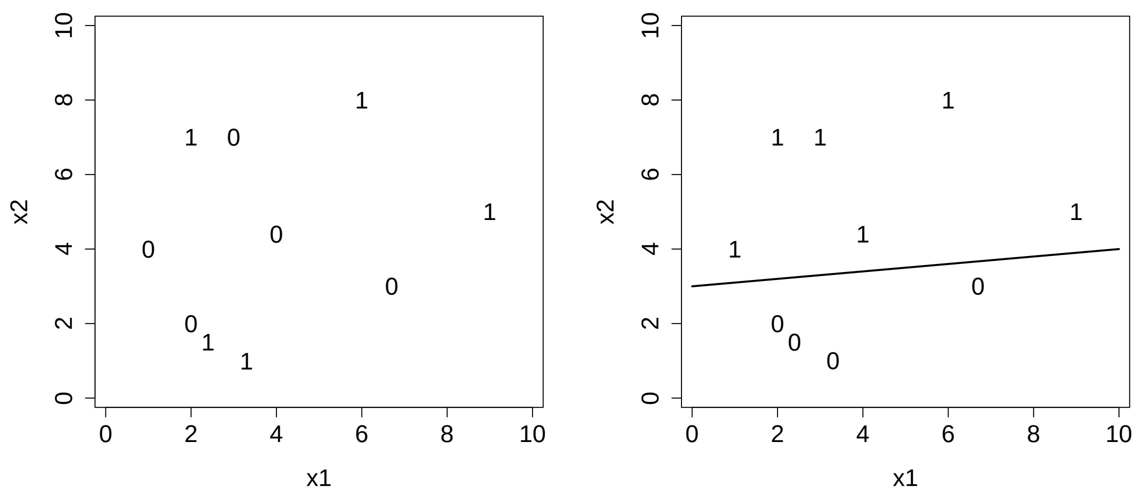
Whenever a binary dependent variable is regressed on a set of independent variables (or one independent variable and a constant) the data either exhibits complete separation, quasi-complete separation, or overlap, as these three categories are mutually exclusive (Albert and Anderson, 1984; Santner and Duffy, 1986).[11] The third category, overlap, is the only one for which a maximum likelihood estimate (MLE) exists.[12] If the data exhibits either complete or quasi-complete separation, no maximum likelihood estimate exists unless the analyst either drops both the offending variables and some observations or utilizes a more complicated estimation method such as penalized maximum likelihood (Firth, 1993; Heinze and Schemper, 2002; Zorn, 2005). In the following we will first briefly explain when data falls into each of these three categories. Second, we will explain how quasi-complete data separation arises as a problem when time dummies are used.

[10]This is not really an issue anymore, as cloglog is widely available in statistical packages such as Stata, SPSS, and R.

[11]Even though Albert and Anderson (1984) only deal with the logit case, cloglog and probit are similarly affected. Data with an ordinal dependent variable can also fall into this classification. See So (1995) for an intuitive and non-technical explanation.

[12]For a formal proof of nonexistence in the logit case, see Albert and Anderson (1984).

Figure 2: Overlap and Perfect Data Separation



(a) Overlap

(b) Perfect Separation

The intuition behind complete and quasi-complete separation are best understood with a graphical illustration. Imagine a regression in which we have a binary dependent variable, y, and two continuous regressors, x_1 and x_2.

Figure 2 demonstrates the cases of overlap and perfect separation for a hypothetical two variable regression. The zeros represent observations where $y = 0$, while the ones represent $y = 1$. In the graph on the left, it is impossible to draw a vector through the points that leaves all ones on one side of the vector and all zeros on the opposite side. Thus, it is not possible to perfectly separate the responses with a vector. The graph on the right in Figure 2 shows a case where perfect separation is possible. It is now possible to draw a vector that perfectly separates the zeros and the ones. More formally stated, data exhibits complete separation if there exists some vector γ, that can correctly separate the data points according to whether $y_i = 0$ or $y_i = 1$. Thus, in this case the following will hold,

$$\begin{cases} \gamma' x_i > 0 \text{ if } y_i = 0 \\ \gamma' x_i < 0 \text{ if } y_i = 1. \end{cases} \tag{3}$$

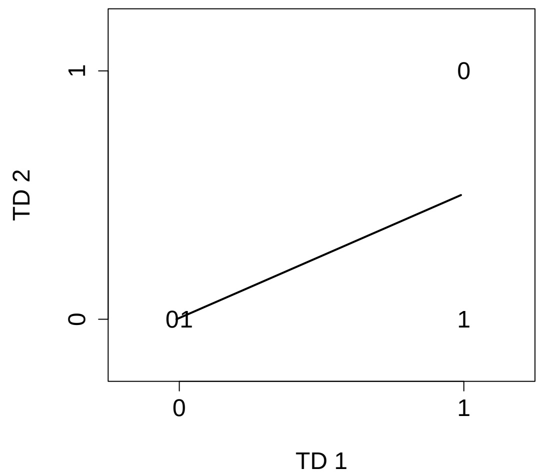
Quasi-complete separation holds when the equalities in the above formulation are weak,

$$\begin{cases} \gamma' x_i \geq 0 \text{ if } y_i = 0 \\ \gamma' x_i \leq 0 \text{ if } y_i = 1. \end{cases} \tag{4}$$

Figure 3: Quasi-Complete Separation in Binary Data with Two Time Dummies

Obs i	Y_i	TD_{1i}	TD_{2i}
1	0	1	0
2	0	0	1
3	1	0	0
4	1	1	0

(a) Data



(b) Graph

Graphically, quasi-complete separation would be present in the graph on the right in Figure 2 if the vector (γ) had to intersect with a zero and/or a one. The third case, overlap, occurs when no vector γ exists for which either of the inequalities in equations 3 or 4 hold. If complete or quasi-complete separation holds for any variable(s), MLE routines will push the value of the estimated $\hat{\beta}$ for the offending regressor(s) to ∞ or $-\infty$ and no maximum likelihood estimate exists. Many preprogrammed routines in commonly used software such as Stata are programmed to look for complete separation in the data and will drop the offending variables and observations stating that the variable(s) "perfectly predicted" y in some number of cases.

In the case with binary regressors such as time dummies, quasi-complete separation is the most likely problem, although complete separation is also possible.[13] To clearly see why quasi-complete separation can arise with time dummies, consider the simple four observation case depicted in Figure 3(a).

To preclude quasi-complete separation, each possible pair $\{Y, TD_{ti}\}$, where $t = 1, 2$, must exhibit all of four possible combinations, $\{0, 0\}$,$\{1, 0\}$,$\{0, 1\}$, and $\{1, 1\}$. In Fig-

[13]Quasi-complete separation is quite unlikely with continuous regressors (So, 1995).

ure 3.1(a), TD_1 fits this requirement, while TD_2 does not. Intuitively, the consequence of this is that $TD_2 = 0$ is a "perfect predictor" of when $Y = 1$. Thus, an MLE procedure will push the parameter estimate for TD_2 towards $-\infty$ and no maximum likelihood estimate exists.

The depiction in Figure 3.1(b) demonstrates informally why the data in Figure 3.1(a) exhibits quasi-complete separation. Notice that the vector shown in Figure 3.1(b) does not completely separate the $Y = 0$ from the $Y = 1$ cases because observations 2 and 3 from Figure 3.1(a) are both at the coordinate $(0, 0)$ in the graph. In this case it not possible to completely separate the responses because any vector must pass through $(0, 0)$, meaning it will "touch" both observations 2 and 3. If we changed the value of TD_2 to 1 for either observation 3 or 4 then there would be no vector that could weakly separate the realizations of the dependent variable. If TD_2 was equal to 1 in observation 2, then we would have complete separation, as $TD_2 = 1$ would be a perfect predictor of when $y = 0$ and $TD_2 = 0$ would be a perfect predictor of when $y = 1$.[14]

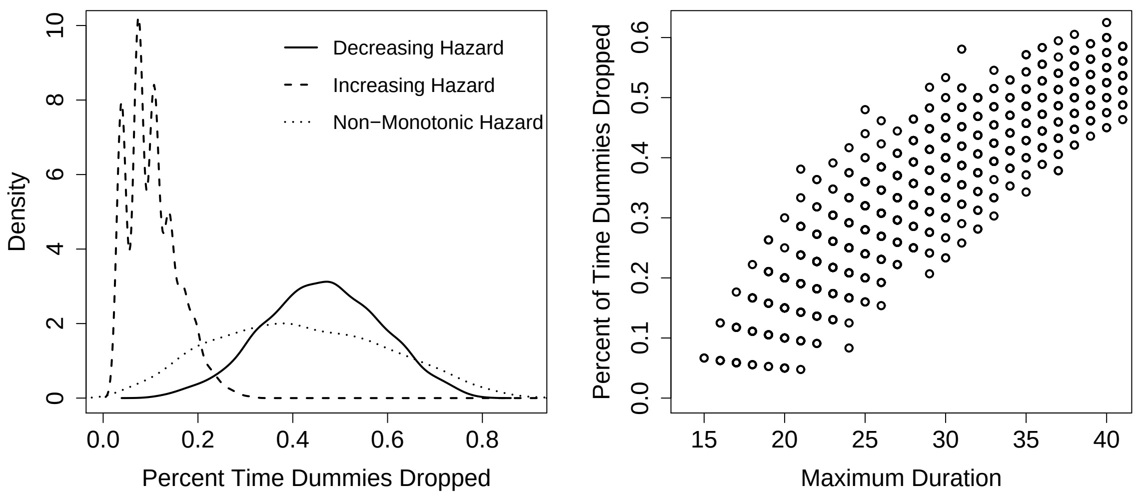
Both monte carlo results and empirical examples demonstrate that quasi-complete separation is a serious problem that should at minimum give pause to researchers before they utilize time dummies, especially if relatively long durations occur in their data. Figure 4 shows two graphs from monte carlo experiments assuming decreasing, increasing, and non-monotonic hazards.[15] Figure 4(a) shows the density of the percent of time dummies dropped under each hazard assumption. In both the decreasing and non-monotonic hazard scenarios, up to 80% of the time dummies are dropped, while up to approximately 25% of time dummies are dropped in the increasing hazard case. Figure 4(b) shows that the percentage of time dummies dropped is an increasing function of the maximum duration in each monte carlo iteration under the non-monotonic hazard assumption.[16] Thus, regardless of hazard shape, separation proves to be a serious problem that suggests that the use of time dummies is problematic, especially if relatively long durations are observed.

[14]This example is put forth for pedagogical reasons, as in reality there would be no "fix" here when complete separation was true because all the observations would have to be dropped unless a more complicated and computationally intensive estimation procedure is used (e.g., penalized MLE).

[15]The design of the monte carlo simulation is described in section 5.

[16]The graphs for the other scenarios look similar and are available upon request from the authors.

Figure 4: Percent of Time Dummies Dropped Due to Quasi-Complete Separation



(a) By Type of Hazard

(b) As a Function of Maximum Duration

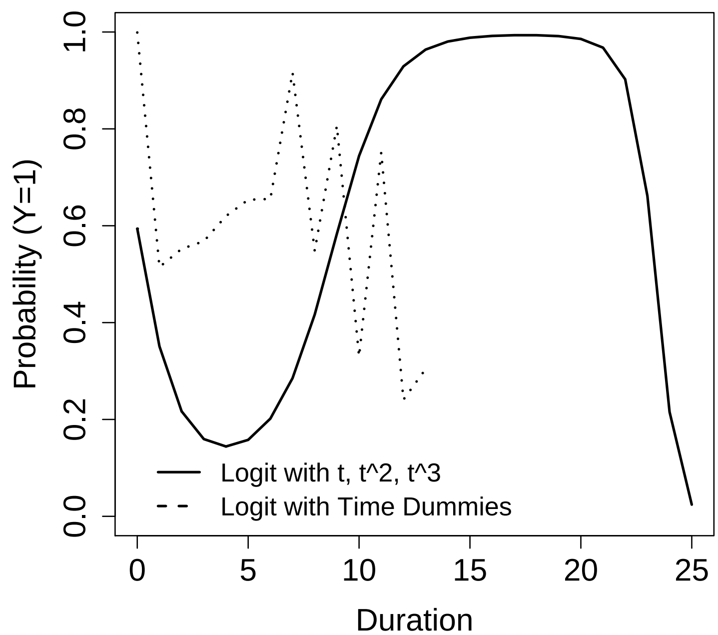
3.2 Empirical Examples of Quasi-Complete Separation

A prominent empirical example where quasi-complete separation exists is in the reanalysis of John Oneal and Bruce Russett's data in BKT, where 3 of 34 time dummies are dropped, along with 916 observations. A very large n of 20990 explains why the percentage dropped, ≈ 0.08, is smaller than what would be implied by the graphs in Figure 4 for a decreasing hazard, as the chance that we observe all of the four combinations of the dependent variable and each of the binary regressors should generally increase with the number of observations. A reanalysis of Palmer, London and Regan (2004) reveals much more severe quasi-complete separation issues, with 19 of 39 dummies and 672 of 2975 observations dropped as a result (See section 6 for a full description).

While there are no ironclad rules that indicate when time dummies will perform most poorly, monte carlo results indicate that they generally do a very bad job in cases where n is not overly large (< 10000) and the maximum duration is fairly long (≈ 15 or longer). For illustrative purposes, we show two empirical cases in which time dummies perform badly, a replication of the results of Lemke and Reed (1996) originally done by Clark and Hart (1998) and a reanalysis of Dickinson and Tenpas (2002).

Clark and Hart (1998) essentially perform a robustness check for the findings of Lemke

Figure 5: Quasi-Complete Separation in Dickinson and Tenpas (2002)



and Reed (1996) by adding time dummies to the original logit specification. They conclude that Lemke and Reed's original finding, that dyadic "satisfaction" with the international status quo significantly reduces the probability of war, is robust to the inclusion of "controls" for temporal dependence. However, the use of time dummies in this case is quite problematic, as quasi-complete separation arises in all of the models. For instance, in Lemke and Reed's model 5, 12 out of 32 total time dummies (37.5%) are perfect predictors of $Y = 1$.[17] The consequences of this are not trivial, as it is not possible to plot a significant portion of the hazard plot and 497 of 7031 observations are dropped from the analysis. Neither of these two problems would be present if either a cubic polynomial or spline is used.

Another example in which time dummies suffer from complete separation issues is in a study of White House staff retention patterns from 1929–1997 by Dickinson and Tenpas (2002). In this case, 12 of 24 the dummies (50%) are dropped. Figure 5 shows that the time

[17]The same is true of all of the seven models replicated by Clark and Hart (1998), although most of the other models have even more serious separation issues which are beyond the scope of this paper. For more on this see Carter and Signorino (2006).

dummies do a very poor job in estimating the hazard relative to the cubic polynomial. Substantive interpretation of the graph is also easier as the cubic polynomial recovers an intuitive relationship rather than the rather abrupt and puzzling inter-duration year changes recovered by time dummies. The hazard recovered by the cubic polynomial shows that the probability of retention is (unsurprisingly) quite high in the first year of a staffer's tenure and steadily decreases through the fourth year. Staffers that stay on past the fourth year (or one term) then are increasingly likely to stay. If a staffer stays on for more than 8 years (or two terms), then he or she is almost certain to stay for quite a while. As the case shown in Figure 5 is for when the first year of tenure is also the first year in a four year presidential term, this plot nicely captures how staffers that stay on past the first term are typically either "loyalists" or highly competent or effective (and thus not forced to leave). The results shown are also consistent with the theoretical expectations of Dickinson and Tenpas (2002), as their main variable of interest, the percentage of party delegates selected via primary, significantly decreases the retention probabilities for the first four years (for details see Dickinson and Tenpas (2002, 434–438)).

3.3 Efficiency

The time dummies approach is also in many cases very inefficient relative to using either splines or a cubic polynomial. When dealing with data that has a maximum duration of greater than 3 or 4, time dummies use more degrees of freedom than either of the other two approaches. Most data sets in IR that record rare events, such as the data on war used in Oneal and Russett (1997) and reanalyzed in BKT, will have maximum durations much longer than 3 or 4. Indeed, as the maximum duration is 34 in the Oneal and Russett data set, time dummies use 31 more degrees of freedom than t, t^2, and t^3 and also results in a loss of information about both time and certain cases due to quasi-complete separation.

In sum, quasi-complete separation is a serious problem that leads us, in conjunction with efficiency issues, to recommend that researchers avoid the use of time dummies to control for duration dependence in binary data since alternatives that do not suffer from these problems are readily available.

4 Love the Spline, Hate the Spliner

The second approach to time dependence in binary data that BKT advocate is splines. While we do not find splines to be necessarily problematic like time dummies, we still recommend that researchers use t, t^2, and t^3 instead of splines. We base this recommendation on three points, two of which are closely related. First, we point out that splines are largely an unknown entity to most empirical researchers in political science. One consequence of this is that several potentially problematic issues such as knot selection or even the choice of which type of spline to implement are dealt with in a manner that can lead to problems as serious as a biased hazard. Finally, we will demonstrate that a cubic polynomial performs just as well as splines in several Monte Carlo experiments and empirical reanalysis. Taken together, these suggest that empirical researchers might be better served simply by using t, t^2, and t^3.

4.1 What's a Spline?

In theory, splines are a useful and powerful tool that can be utilized to create a smooth function of time (or almost any variable that is approximately continuous). While an enormous technical literature exists on splines in statistics, most political scientists do not have much, if any, training in using splines. Indeed, many authors are quite upfront about viewing splines as an opaque method to control for a statistical nuisance, with one author referring to them as "so-called cubic splines" that are used "to control for any temporal correlation" (Dorussen, 2006, 95). Another set of authors criticize the standard approach as "hiding" important linkages "in the peace-year spline variable" (Goertz, Jones and Diehl, 2005, 747).

There are actually many different kinds of splines (e.g., B-splines, natural cubic splines, quadratic splines, piecewise linear splines), many of which are available as "canned" packages in Stata, SPSS, or R. Without any training in the differences between these different spline packages, researchers will tend (quite understandably) to use whatever their preferred statistical package makes most easy to implement. As a result, researchers using different statistical packages may implement different types of splines. This can make replication problematic, especially if it is not clear what kind of spline was implemented by a researcher. Explaining splines in general is beyond the scope of this paper. However, to help the reader understand our argument concerning splines, it will be helpful to provide some intuition concerning what they are and how they are used in modeling temporal dependence.

For our purposes, we can think of a spline as a smoother — a function that allows us to smooth the relationship between two variables, say a dependent variable y and time t. Most splines allow one to specify points in t where the relationship with y radically changes. Those points are referred to as “knots.” Fewer knots will lead to a smoother relationship, but may miss important changes in the relationship. Specifying more knots allows for more changes to be modeled. On the other hand, the greater the number of knots, the less smooth the relationship. Moreover, the spline may pick up on idiosyncratic changes in the relationship, not general trends.

As we previously mentioned, there are numerous types of splines. However, a key component to all splines is that they generate a set of vectors that are a function of the independent variable (here time t) and associated with the knots. These vectors are referred to as basis vectors and the set of vectors is referred to as a basis matrix.[18] A basis is a set of linearly independent vectors, the linear combination of which can span an entire vector space (i.e., reach any point in that space). These vectors can be included as regressors in our statistical analysis.

To illustrate this, we will briefly explain two simple splines: a piecewise linear spline and a simple cubic spline. Suppose we believe the probability of some event (e.g., war) in observation i is explained by a set of regressors $x_{i,t}$ and some smooth function of time $s(t)$. We can specify the logit equation for this situation as

$$\Pr(y_{i,t} = 1) = \frac{1}{1 + \exp\left[-(x_{i,t}\beta + s(t))\right]}$$

For the piecewise linear spline, $s(t)$ takes the form

$$s(t) = \gamma_1 t_i + \sum_{k=1}^{K} \gamma_{1k}(t_i - \eta_k)_+ \tag{5}$$

where the function $(t_i - \eta_k)_+$ returns the difference between t_i and η_k when it is positive, but equals zero otherwise.[19] The η_k are our k knots. The vector t_i and the k $(t_i - \eta_k)_+$

[18] Suppose we have some vector space S such that it is possible to express every vector in S as a linear combination of k $(x_1, x_2, \ldots, x_k)$ vectors. Then, if the k vectors are linearly independent, the set of k vectors forms a basis for S. See Searle (1982) for more details.

[19] Note that time dummies and the piecewise linear spline in equation 5 have an equivalent basis when the maximum number of knots (η_k) given t are fit. Thus, if we ran a regression including the basis matrix produced by equation 5 with the maximum number of knots the log-likelihood, estimated β coefficients, and predicted probabilities would all be identical.

vectors are the basis vectors that would be included in our logistic regression. The β and γ coefficients would be estimated.

The simple cubic spline is conceptually very similar:

$$s(t) = \gamma_1 t_i + \gamma_2 t_i^2 + \gamma_3 t_i^3 + \sum_{k=1}^{K} \gamma_{1k}\{(t_i - \eta_k)_+\}^3. \tag{6}$$

The only difference here is the inclusion of the t_i polynomials, as well as the k $((t_i - \eta_k)_+)^3$ basis vectors. Again, these would be included as regressors and the β and γ coefficients would be estimated.

In some implementations, the γ spline coefficients are penalized in estimation to constrain the influence of the knots and to ensure a relatively "smooth" fit. Commonly used splines such as the natural cubic spline employed by Beck, Katz and Tucker (1998) are just more complicated versions of equation 6.[20] Another common type of spline is the B-spline, which can have a basis that is equivalent to that of a natural cubic spline, but has the nice properties that the basis vectors will be orthogonal to each other (i.e., no multicollinearity issues) and vary between 0 and 1 (i.e., no numerical instability issues).

Now that we have briefly explained what splines are, we turn to two important aspects of splines: their interpretation and knot selection.

4.2 Interpretation

Interpretation of the role time plays should ultimately be the goal of political scientists who seek to "take time seriously." Rather than simply including the basis vectors produced by splines as regressors and stopping there, researchers should think theoretically about whether and why temporal dependence is present in their data and carefully examine the hazard to interpret the role time plays.

Although BKT provide Stata code for generating splined time, no directions are given concerning how to interpret the results. As we previously noted (see table 1), almost no scholarly work in political science plots and interprets the hazard when splines are implemented. It is not difficult to imagine why this would be the case. We saw in the previous section that applying a spline function to time t generates basis vectors that depend on the knots. It is one thing to insert these into a regression equation and to

[20] A natural cubic spline $s(t)$ is assumed to be linear beyond the two "boundary knots" (i.e., $s''(t) = s'''(t) = 0$). See Hastie and Tibshirani (1990) or Ruppert, Wand and Carroll (2003) for details.

estimate the coefficients (i.e., the γ's) associated with the basis vectors. However, how does one then interpret the results?

It is important to note first that the estimates for the individual spline vectors (e.g., the γ terms in the previous section) are not substantively interesting by themselves. For example, we are not interested in the individual effect of the $((t_i - \eta_2)_+)^3$ vector in Equation 6. Rather, we want to interpret the effect of time t on the estimated $\Pr(y_{i,t} = 1)$. The problem for many substantive researchers, we suspect, is that the spline-based regression equation is in terms of the spline basis vectors, not in terms of t. Although we recommend a simpler approach to modeling and interpreting the effect of time in Section 5, it is important that researchers using splines be able to interpret their results as well. We therefore provide steps for doing so in the Appendix.

4.3 Specifying Knots

One of the most important aspects of implementing splines is appropriately selecting the knots η_k. The number of knots determines the number of basis vectors that are included in subsequent regression analysis. The location of the knots determine where the relationship between the dependent variable and t is allowed to drastically change (here, by censoring the negative values in the difference $(t_i - \eta_k)_+$).

One potential difficulty with splines is that implementation requires the researcher to specify knots. Political science theory tells us little to nothing about where these knots should be placed, which makes their specification rather difficult. For instance, Beck, Katz and Tucker (1998, 1279) choose knots via "a sequence of F-tests". Such an approach is difficult to justify theoretically and arguably amounts to little more than data mining. Given the difficulty in choosing theoretically informed knots, some researchers such as Schultz (2001, 270–271) have simply used the same knots used by Beck, Katz and Tucker (1998), admitting that their choice was ad hoc, while others such as Senese (2005) simply provide no discussion of knot selection at all.

The idea that theory can drive knot selection is in all likelihood not very practical. One option for choosing knots is to simply examine the data, a strategy that can be quite effective when smoothing scatter-plots (Ruppert, Wand and Carroll, 2003, 57–72). However, when we are recovering an underlying hazard via estimation from a regression model, we cannot just examine the data we are smoothing in the same fashion for the simple reason that before specifying and estimating the model we do not know what it looks like. Furthermore, what we eventually see depends upon the knots we choose (see figures 7 and 8). These issues at least partially explains why the current approach in political science is simply to choose the knots that BKT used (i.e., knots at $t = 1, 4$, and 7) or to choose rather arbitrarily three other knots and not discuss or justify the criteria used to do so.

One possible solution to the problem of knot selection is to utilize some criteria to automatically select knots. Existing methods that are computationally feasible (i.e., do not require estimation of all possible models) are generally similar in concept to stepwise regression (Ruppert, Wand and Carroll, 2003, 64). Thus, some model selection criteria is assumed and the set of knots that is optimal given this criteria is chosen. These procedures are quite computationally intensive (Ruppert, Wand and Carroll, 2003, 64–65, 214–222) and would require even more expertise of researchers than implementation of splines that are currently used. Additionally, such computationally intensive techniques are likely to be unnecessary in political science. As noted by Beck and Jackman (1998, 609–610) in their discussion of generalized additive models, relationships implied by political processes are generally thought to be relatively smooth and not overly local in character. We agree with this point and argue that a cubic polynomial will be able to recover just about any hazard one might expect to see in political science. A cubic polynomial does not utilize any knots, rather it assumes that the hazard is a smooth global function of time that can be recovered with the basis vectors t, t^2, and t^3. As we will discuss later, this is not without its limitations either. Indeed, our intention is not to rule out entirely the use of either splines or automatic knot selection techniques. However, the issues discussed in this section, combined with the experimental and empirical evidence we provide in the next section, suggest that a cubic polynomial performs just as well as splines in most substantive settings without the additional complexity.

5 Time Cubed

Having discussed the technical details of time dummies and splines, our recommendation in this section is almost embarrassingly simple: include t, t^2, and t^3 as regressors. To make this concrete, suppose a researcher with regressors $x_{i,t}$ wanted to conduct logistic regression, control for temporal dependence, and interpret the effect of time on $\Pr(y_{i,t} = 1)$.

Using our approach, her logit equation would take the form

$$Pr(y_{i,t} = 1) = \frac{1}{1 + \exp(-(x_{i,t}\beta + \gamma_1 t_i + \gamma_2 t_i^2 + \gamma_3 t_i^3))} \tag{7}$$

where $\gamma_1 t_i + \gamma_2 t_i^2 + \gamma_3 t_i^3$ is a cubic approximation to the hazard. This cubic approximation can accommodate just about any hazard shape (e.g., linear, nonlinear, nonmonotonic) that political scientists typically deal with. In this section, we will first demonstrate with Monte Carlo results that t, t^2, and t^3 does just as well, if not better than either time dummies or splines (as implemented in practice) in a variety of substantively interesting settings. Second, we discuss some potential issues with using a cubic polynomial.

5.1 Monte Carlo Comparison

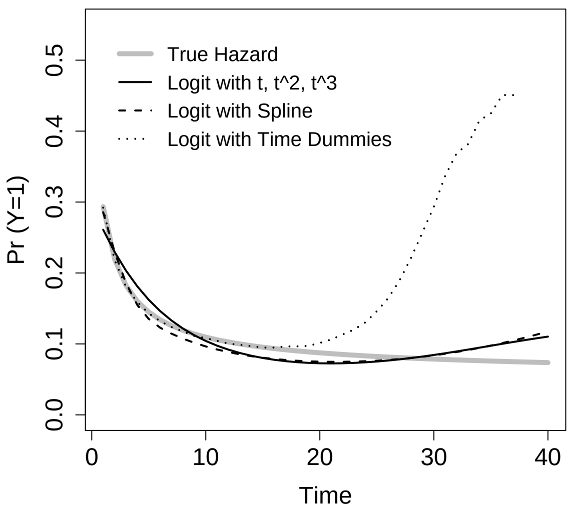
We assume that the data generating process is logit with time dummies.[21] Thus, the link function is

$$Pr(y_{i,t} = 1) = \frac{1}{1 + \exp(-(x_{i,t}\beta + \kappa_{i,t}\gamma))}. \tag{8}$$

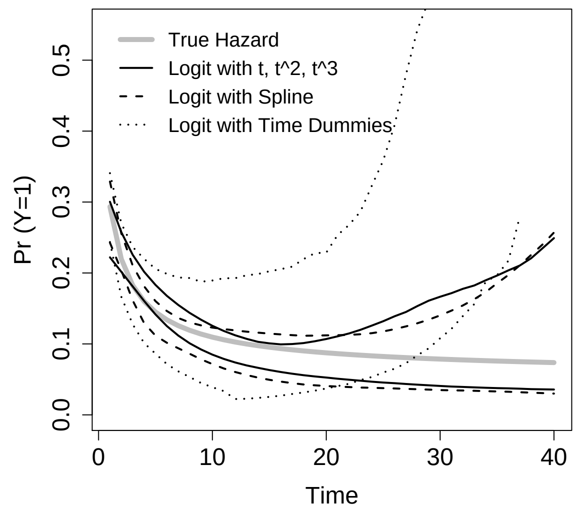
$x_{i,t}$ consists of a constant and a regressor that varies between -2 and 2 drawn from the uniform distribution. The $\kappa_{i,t}$ are the time dummies that depend on the γ's for the shape of the hazard. We conduct Monte Carlo experiments for constant, decreasing, increasing, and two different non-monotonic hazards. We run 10000 monte carlo iterations with $n = 2000$ for each iteration. In each iteration for each hazard shape, we estimate logit with time dummies, cubic B-splines, and a cubic polynomial. In figures 6–9 we provide graphical illustrations that demonstrate how well each method performs on average in recovering the true underlying hazard. To assess how well splines perform as implemented in the discipline we choose knots at 1, 4, and 7 for all spline models.

Figure 6 shows the plotted hazard for the case of a typical decreasing hazard. In both graphs, the thick grey line is the true hazard. 6(a) compares the three methods in terms of how well they recover the true hazard on average, while 6(b) demonstrates efficiency with 95% confidence intervals. The curves produced with spline and the cubic polynomial are quite similar, while the graph produced with time dummies looks considerably worse. The time dummies plot begins increasing sharply after around $t = 20$, which should not happen. Additionally, the confidence intervals are considerably wider around the time dummies plot relative to the spline and t, t^2, and t^3, which demonstrates that we have

[21] We observed virtually identical results when the data generating process is cloglog with time dummies.



(a) Expectation



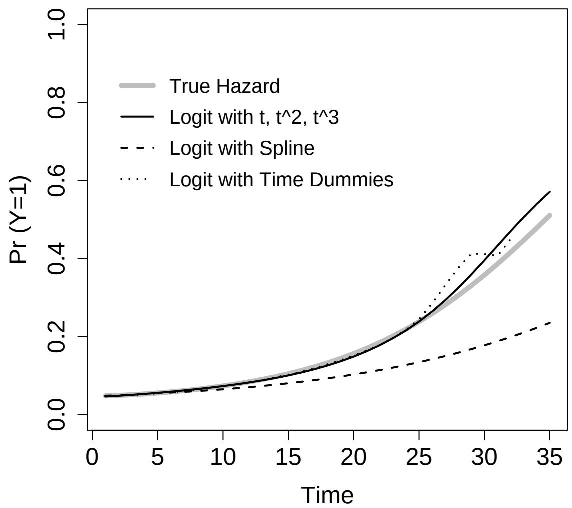
(b) Confidence Intervals

Figure 6: Monte Carlo Comparison: Decreasing Failure Rate

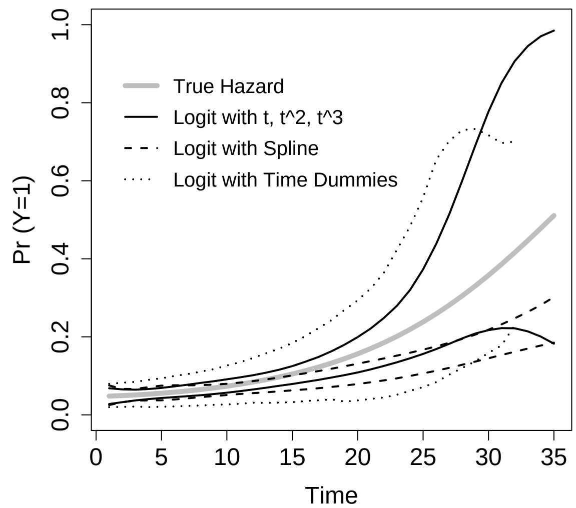
increased uncertainty around these probability estimates. Data separation problems are also evident here, as the time dummies plot is cut off too soon.

Figure 7 shows that splines perform worse than the cubic polynomial in the increasing hazard case while time dummies performs reasonably well. Since the shape of the time dummies plot is dependent upon coefficients estimated for each duration, if there are considerably fewer long durations, time dummies will do much worse than either of the other two methods. Time dummies do well here because the average duration is longer in the increasing hazard case than it was in the decreasing hazard case. The estimated hazard produced by splines shown in 7(a) is biased and diverges considerably from the true hazard after the tenth duration year. 7(b) demonstrates that the 95% confidence bands do not include the true hazard for half of the plot. Thus, in this scenario t, t^2, and t^3 outperforms both of its rivals.

Figure 8 compares how well the three methods recover a non-monotonic hazard that takes the form of an asymmetric parabola. t, t^2, and t^3 outperforms both splines and time dummies in this case, as splines produce a biased estimated hazard and perfect separation plagues time dummies. Figure 8(a) shows that the cubic polynomial essentially does a flawless job of recovering the true hazard while the plot produced by splines is far from



(a) Expectation



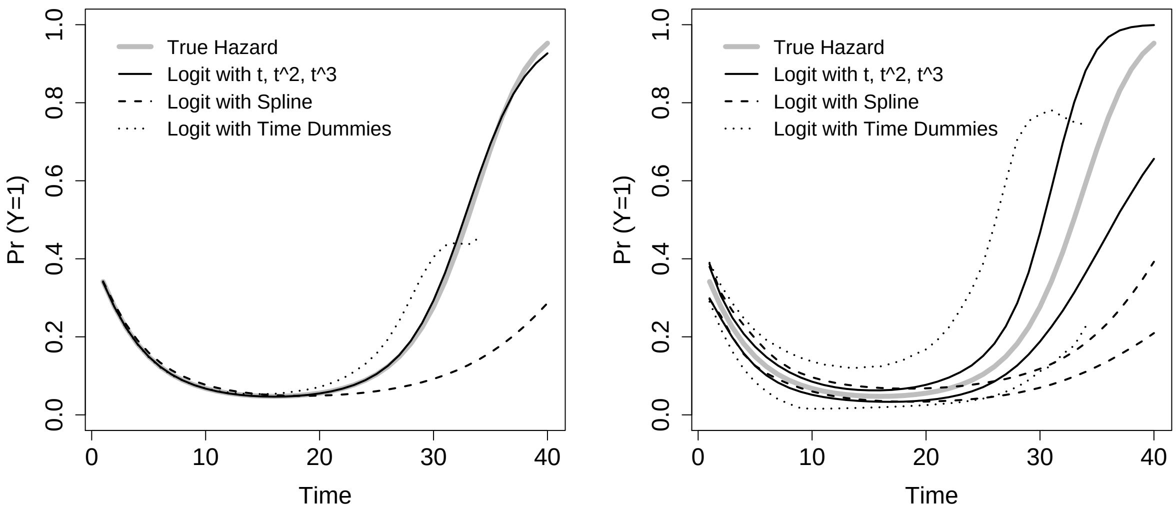
(b) Confidence Intervals

Figure 7: Monte Carlo Comparison: Increasing Failure Rate

the true hazard after t=20. Furthermore, 8(b) shows that the cubic polynomial is more efficient than time dummies, while the 95% confidence bands for the spline hazard are not even close to including the true hazard for about 40% of the plot.

The second non-monotonic hazard we study is produced from the log-logistic distribution and captures situations in which the probability of observing the event of interest within a group initially increases sharply, but then decreases for the remainder of the group's "life". Both the cubic polynomial and splines perform well in this scenario, although time dummies do very poorly. Figure 9(a) shows that the time dummies and splines do slightly better than t, t^2, and t^3 in recovering the true shape of the initial increase in the hazard, although time dummies fails to produce a reasonable estimate after the twelfth duration year and loses over 10 duration years worth of the plot due to separation. In terms of efficiency, the cubic polynomial and splines are comparable while time dummies is relatively inefficient.

In sum, in terms of recovering the hazard, the cubic polynomial on average performs better than splines, as they are implemented in practice, while easily outperforming time dummies. Time dummies perform very poorly even though the data generating process uses time dummies. Even if one chose better knots in implementation of a spline, a cubic



(a) Expectation

(b) Confidence Intervals

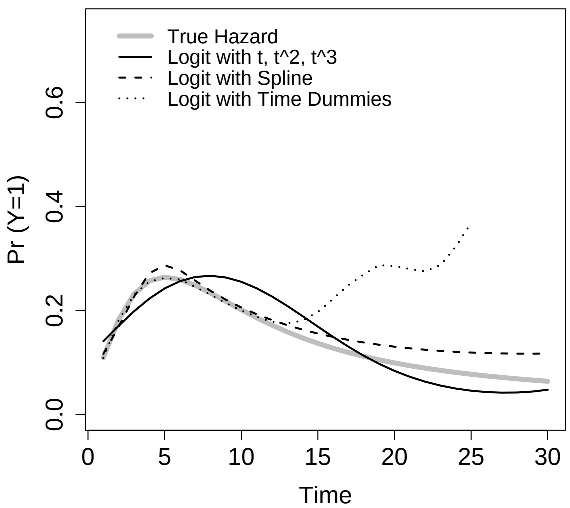
Figure 8: Monte Carlo Comparison: Non-Monotonic Hazard 1

polynomial would perform just as well in any of these scenarios. As the cubic polynomial is much easier to implement, is conceptually simple, and is also quite flexible and simple in terms of modeling non-proportional hazards (see section 6.3.1) we recommend that researchers model time dependence in binary data with t, t^2, and t^3.

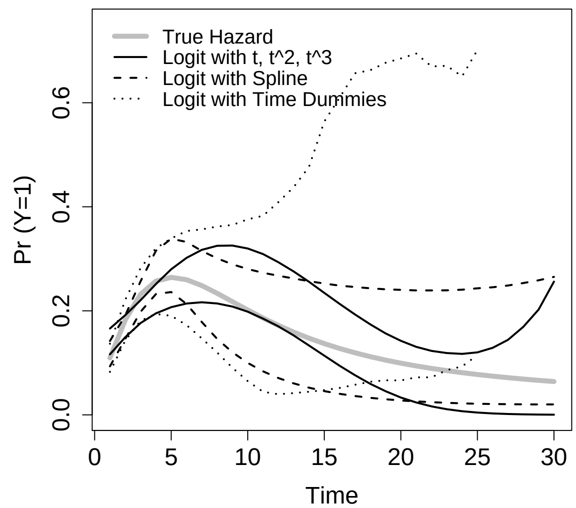
5.2 Potential Issues

Although we find that a cubic polynomial performs at least as splines and much better than time dummies, we are aware of some potential issues with the use of t, t^2, and t^3. We discuss four potential issues here.

The first potential issue pertains to the order of the polynomial. We recommend a cubic polynomial, but in principle any order of polynomial (e.g., quadratic polynomial) could be chosen. Generally, polynomials of odd order are preferable to polynomials of even order. Polynomials of odd order asymptotically have a smaller mean-squared error than polynomials of even order (see Fox (2000) for a simple illustration of this). As far as the exact order to choose, we recommend a third-order cubic polynomial because it will capture any hazard shape (monotonic or otherwise) that is recovered by commonly estimated parametric duration models (e.g., Weibull, log-logistic) and typically seen in



(a) Expectation



(b) Confidence Intervals

Figure 9: Monte Carlo Comparison: Non-Monotonic Hazard 2

semi-parametric models such as the Cox proportional hazards model. A cubic polynomial allows for non-monotonic hazards while avoiding the problems associated with polynomials of higher order (i.e., plots with "kinks" that are sensitive to a small proportion of observations).

A second potential issue involves the interpretation of the estimated β's of the polynomial. The short answer is that when estimating a cubic polynomial (or any polynomial of order greater than one or two) the individual coefficients are not interpretable in an obvious way. However, this is really not an issue since the hazard is the quantity of interest, not the coefficients.

The last two potential issues involve possible multicollinearity and numerical instability. Given that the t, t^2, and t^3 terms are highly correlated, multicollinearity is a potential issue with such an approach. Having said this, we have found no evidence that it presents any problems in numerous Monte Carlo experiments and empirical reanalysis. Another potential problem involves potential numerical instability in a MLE routine. Instability could result from large differences in magnitude between t, t^2, and t^3 and other regressors given that t^2 and t^3 can get quite large depending on the maximum value of t.[22] Given

[22]For example, if the maximum duration is $t = 25$ then t^3 varies from 0 to 15625.

that this is always a potential concern when running MLE routines, the solution here is no different than what common practice should already be. Simply examine the range of all the data and t, t^2, and t^3 and rescale variables as necessary via division by some multiple of 10. We have found utilizing $\frac{t}{100}$ and its square and cube generally works quite well. Or, using t, t^2, and $\frac{t^3}{1000}$ also works well. In either case, users should remember to use the appropriate scaling when plotting $Pr(y_{i,t} = 1|t)$.

6 Empirical Examples

In this section we replicate and extend the empirical analysis of Oneal and Russett (1997), Palmer, London and Regan (2004), and Crowley and Skocpol (2001) using a cubic polynomial. In all three cases, t, t^2, and t^3 yields results that are nearly identical substantively to those reached with spline. Additionally, we show that a cubic polynomial outperforms time dummies in the reanalysis of Oneal and Russett (1997) and Palmer, London and Regan (2004) and allows for a much richer investigation of temporal effects present in Crowley and Skocpol (2001). Finally, we use non-nested comparative model testing to assess whether logit with a cubic polynomial or logit with spline or time dummies is more appropriate in a maximum likelihood framework.

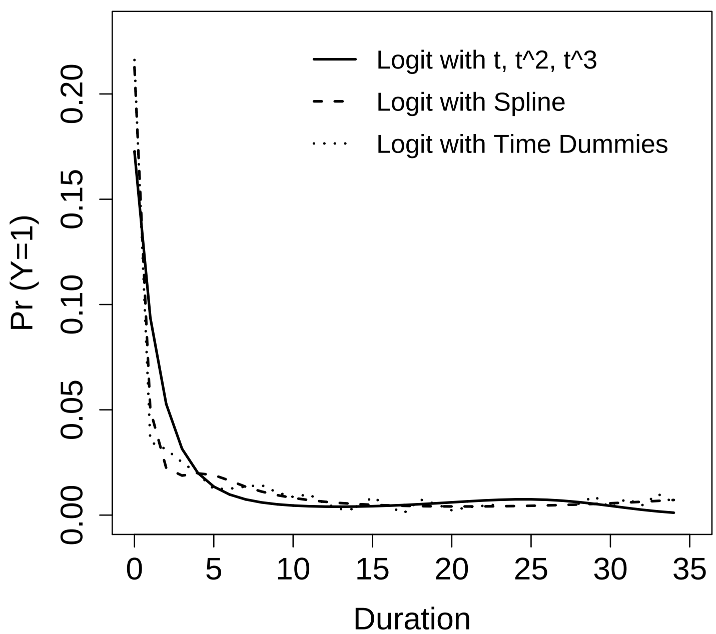
6.1 Trade and International Conflict

In an influential article, Oneal and Russett (1997) find that the probability of war is significantly lowered both when two states are democracies and when two states have high levels of trade. Beck, Katz and Tucker (1998) show that the latter relationship no longer significantly lowers the probability of conflict when time dependence is taken into account. Thus, "political liberalism" seems to find support even when the effects of time are taken into account, while "economic liberalism" is not robust. We replicate the results found in both Oneal and Russett (1997) and Beck, Katz and Tucker (1998) and demonstrate that logit with t, t^2, and t^3 obtains the same results.

Table 2 contains the results of using logit with time dummies, spline, and a cubic polynomial.[23] Columns 1 and 3 are replications of the results found in Beck, Katz and Tucker (1998), while columns 2 and 4 show logit with B-spline and t, t^2, and t^3 respectively. Notice that the results across the four different specifications are remarkably similar. The

[23] We also replicated Oneal and Russett's original findings using logit, but do not present them in Table 2.

Figure 10: Hazards for Oneal and Russett Replication



main finding, that trade does not significantly affect the probability of conflict in the period under study, holds up when using logit with t, t^2, and t^3.

In order to more fully assess whether the results of logit with time dummies or spline are substantively no different than of logit with t, t^2, and t^3, we also plot the hazard in Figure 10. The curves are quite similar, although the logit with t, t^2, and t^3 probability estimate for the year immediately following a conflict (or in some cases the first year included in the data) is slightly lower, ≈ 0.17, than either of the other two estimates, which are ≈ 0.22. The t, t^2, and t^3 curve also decreases less drastically after the first year, but the overall differences are slight in terms of substantive interpretation.

In order to assess whether a cubic polynomial or the methods prescribed by Beck, Katz and Tucker (1998) are the most appropriate in a maximum likelihood framework, we conduct non-nested comparative model tests.[24] The distribution-free test developed by Clarke (2003) compares the individual log-likelihood values from two models. For every

[24]Note that a cubic polynomial is nested within a cubic spline. For Clarke's test, the only consequence of this is inefficiency.

observation $i = 1, \ldots, n$ in which the first model has a higher log-likelihood than the second model an $n \times 1$ positive difference vector is given a one in the ith spot, and a zero otherwise. The null hypothesis states that half of the elements in the positive difference vector are one, while the other half are zero. To employ the test statistic, we just sum the n elements of the positive difference matrix and use an exact binomial test with size n and null probability 0.5.[25] The Vuong test determines whether the average value of the log-likelihood ratio is significantly different than zero Vuong (1989). Thus, while Clarke's test examines the individual log-likelihoods, Vuong's test deals with the average of the ratio of the summed log-likelihoods.

We compare logit with t, t^2, and t^3 and logit with natural cubic spline, but are unable to compare our method to logit with time dummies because we dropped 916 observations due to quasi-complete separation.[26] The Clarke test provides support for the hypothesis that logit with a cubic polynomial is significantly better than logit with natural cubic spline. The sum of the positive difference vector is 12752 out of 20990 observations, which means that our method outperforms BKT's method 60.8% of the time. The probability of seeing this result if the models are equally good (i.e., the sum of the positive difference vector divided by n is $\frac{1}{2}$.) is arbitrarily close to zero. The Vuong test statistic is unable to reject the null hypothesis that the two models are equally good, as the p-value is 0.499. Thus, depending on the test chosen, t, t^2, and t^3 is at least as good, if not better than natural cubic spline in the context originally analyzed by Oneal and Russett (1997).

6.2 Parliamentary Democracies and the Initiation of Militarized Disputes

Palmer, London and Regan (2004) examine how heterogeneity among parliamentary democracies affects conflict initiation and escalation decisions. The authors push beyond the standard approach in the democratic peace literature and challenge the idea that all democracies are a homogenous set. In particular, they examine how the following factors affect the initiation and escalation of militarized inter-state disputes: whether or not there is a single ruling party or multiple parties in a coalition, the "left-right" orientation of the ruling party or coalition, the size of the ruling coalition, and whether the government

[25] We can do this because we are dealing with the sum of n Bernoulli trials, which is distributed binomial.

[26] Even if we were able to run the tests, since logit with time dummies estimates 38 parameters compared to 10 for logit with t, t^2, and t^3, the Akaike or Schwartz correction penalizes the former model enough to make the tests rather uninteresting.

is a minority or majority government. Analysis is conducted on eighteen parliamentary democracies during the period 1949–1992 with separate logit models for initiation and escalation.[27] While the investigation of how the behavior of parliamentary democracies is affected by their past involvement in conflicts would also be of interest, Palmer, London and Regan (2004) do not pursue it for one of the same reasons we criticize time dummies in section 3. As mentioned above, time dummies bring about severe separation issues in their data that result in the loss of 672 out of 2975 observations ($\approx$ 22.6%) as well as 19 of 39 time dummies ($\approx$ 48.7%). Given that almost half of the time dummies had to be dropped, the authors quite reasonably concluded that it was "too high a price to pay (21)", especially given that almost half of the information about time is lost anyhow.

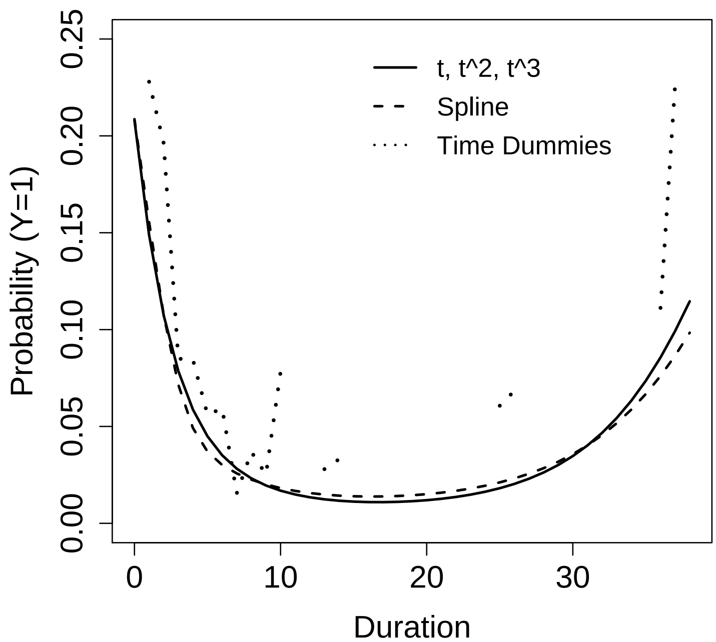
We replicate their results regarding parliamentary democracies' decisions to initiate militarized interstate disputes (MIDs) and reanalyze their data accounting for time with a cubic polynomial, time dummies, and cubic B-spline. The results of the four regressions are shown in table 3.[28] Note that when the effects of time are estimated, none of the variables that pertain to the type of government are statistically significant. The only statistically significant variable is the state's relative military power, which is not a particularly striking finding. Thus, once the effects of time are accounted for, heterogeneity among parliamentary democracies is no longer a strong predictor of conflict initiation.

We also estimate the hazard for logit with a cubic polynomial, spline, and time dummies. The resulting hazard plots are shown in figure 11. Notice that the estimated hazard plots for the cubic polynomial and spline are essentially identical, while the time dummies plot is jagged and missing almost half of the plot due to quasi-complete separation. The plot shows that conflict is most likely in years immediately following conflict, becoming quite unlikely after approximately five years have passed without a militarized conflict. The probability of a parliamentary democracy initiating a MID starts to increase after around two decades of peace, increasing through the maximum duration of 38 years. While full analysis of this hazard via a thorough examination of cases is beyond the scope of this paper, this counterintuitive upward trend in conflict initiation among democracies is a trend worthy of further investigation.[29]

Again, we run comparative model tests to assess which model is the "best" model in an MLE framework. The distribution free test provides support for our proposed method

[27] They test for selection effects and find none.
[28] Likelihood ratio tests show that any of the methods for estimating the effects of time are warranted.
[29] 442 observations reach a duration of greater than 20, so this is not a negligible portion of the sample.

Figure 11: Quasi-Complete Separation in Palmer, London and Regan (2004)



when tested against spline. The sum of the positive difference matrix when testing logit with a cubic polynomial against logit with spline is 2325 out of 2975 (78.2%). The probability of seeing this result if the models are equally good is arbitrarily close to 0. Thus, the distribution free test provides strong support for our proposed method. As an additional check, we also ran the Vuong test. The Vuong test returns a p-value of 0.500 and therefore does not provide the same support for our method as Clarke's test does. Thus, while our method generally does better when we examine all individual observations, outperforming BKT's method in over 78% of the observations, when looking at the average log-likelihood ratio neither model outperforms the other.

6.3 Associational Formation in the United States

Crowley and Skocpol (2001) use grouped duration analysis to assess the factors influencing associational formation in the United States from 1860 through the 1920s. Using a remarkable data set that tracks the founding of state-level units for twenty-one organizations in the forty-eight continental states, they use logit with time dummies to determine whether factors associated with modernization and/or the after-effects of the American Civil War were key determinants of associational formation, finding support for the latter notion but little for the former.

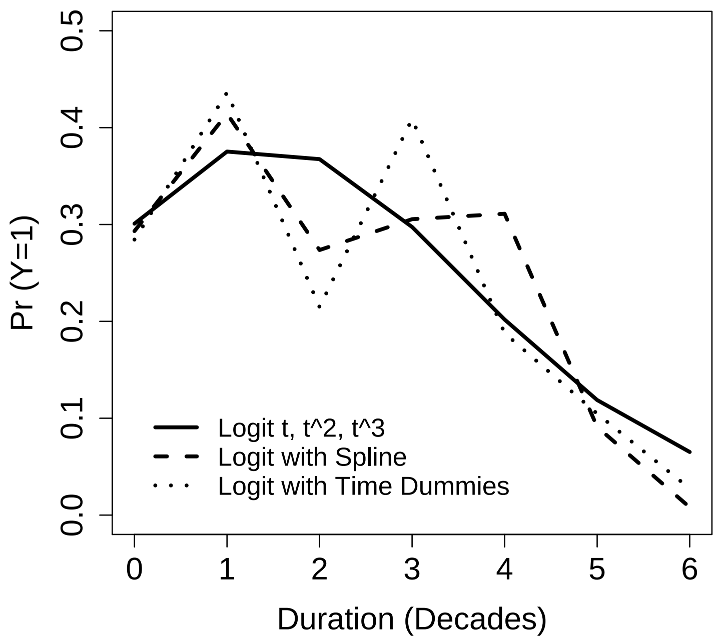
Crowley and Skocpol's dependent variable measures whether a state-level branch of a membership federation was created in a given decade. Examination of Table 4 indicates that the use of a cubic polynomial does not alter the authors' substantive results in any substantial way.[30] The graph in Figure 12 shows the predicted hazard rates as a function of time for logit with a cubic polynomial, spline, and time dummies. Although Crowley and Skocpol did not include a hazard graph in their analysis, it is essential in order to substantively assess the effect of time on associational formation. The hazard plot further damages the modernization thesis, as it shows that the probability a state-level branch of a federation formed is greatest about a decade after the organization was founded. The hazard subsequently decreases and there are slight bumps around the third or fourth decade in the plots produced by time dummies and B-spline. Thus, for all organizations under study except four, this peak came between 1870–1900, which is consistent with civil war effects explanation favored by Crowley and Skocpol.

Again, we run both Clarke's test and the Vuong test to assess which model is most appropriate for Crowley and Skocpol's data in an MLE framework. The distribution free test provides support for our proposed method when tested against both time dummies and spline. The sum of the positive difference matrix when testing logit with a cubic polynomial against logit with time dummies is 1762 out of 2529 (69.7%).[31] The probability of seeing this result if the models are equally good is arbitrarily close to 0. Thus, the distribution free test provides strong support for our proposed method. As an additional check, we also ran the Vuong test. The Vuong test returns a p-value of 0.498 and therefore does not provide the same support for our method as Clarke's test does. Thus, while our method generally does better when we examine all individual observations, outperforming BKT's

[30]Readers may notice that our estimated coefficient on the constant differs slightly from what Crowley and Skocpol reported. Crowley and Skocpol utilized Stata for their analysis and included all the time dummies as well as a constant term instead of dropping either the constant or the first time dummy. Stata arbitrarily chooses the sixth time dummy to drop due to collinearity problems (we tried this using Stata 8). We chose to estimate a constant and to drop the first time dummy rather than to drop the sixth time dummy. We also divided several of their variables by a multiple of 10 to make the regressors all of similar relative magnitudes. This accounts for the larger coefficients on several of the variables (e.g., Urban Growth). One can recover Crowley and Skocpol's coefficients by dividing our coefficients by the same number we divided a given variable by. We also ran logit with B-spline and found results that were substantively no different.

[31]Since data separation is not a problem in Crowley and Skocpol (2001) we are able to directly compare the use of a cubic polynomial versus time dummies.

Figure 12: Hazards for Crowley and Skocpol Replication



method in nearly 70% of the observations, when looking at the average log-likelihood ratio neither model outperforms the other.

We also ran both the distribution free test and the Vuong test to compare logit with t, t^2, and t^3 and logit with spline. The results were essentially the same as for logit with time dummies, with Clarke's test returning a positive difference matrix sum of 1467 (58%) and a p-value arbitrarily close to 0, while the p-value for the Vuong test was 0.500.

6.3.1 Accounting For Non-Proportional Hazards

An additional advantage of using t, t^2, and t^3 is that it is relatively easy to account for non-proportional hazards.[32] As Box-Steffensmeier and Zorn (2001) note, nearly all duration models that are commonly used by political scientists assume that the hazards are proportional relative to each other over time. However, there are many instances where, substantively, we would expect these hazards to be non-proportional. The proportional

[32]Thanks are due to Hein Goemans for suggestions on this point.

hazards assumption may not only prevent researchers from examining important empirical phenomena, but also result in bias and incorrect standard errors if it is not warranted. For regressors that exhibit non-proportionality, Box-Steffensmeier and Zorn (2001) propose the interaction of the regressors with $ln(Time)$. When using logit or cloglog with a cubic polynomial, we can simply interact the regressors that exhibit non-proportional effects on the hazard with t, t^2, and t^3.

We hasten to point out that estimating and interpreting the effects of time is not just an econometric issue, but a theoretical one as well. Even though Crowley and Skocpol utilize logit with time dummies as prescribed by BKT, beyond the inclusion of the dummies as "controls" they do not really investigate the effect of time, even though (as we demonstrate below) their theoretical framework has an important temporal component.

Crowley and Skocpol (2001, 814) explicitly state that they are interested in *when* and *how* membership associations in the United States developed and grew. They note that despite the popularity of the idea among historians that socio-economic modernization and high levels of immigration were the main catalysts of associational development, an historical-institutionalist perspective also has a lot of explanatory power. In particular, they focus on the pivotal role that the U.S. Civil War (1861–1865) plays in the historical-institutionalist account of associational development. Since the U.S. government lacked a large standing army or well-developed bureaucracy prior to the Civil War, voluntary federations were assembled across the states to aid in raising an army. After Northern victory, civic associational structures built to support the war effort remained in place or were replicated by the founders of new civic associations.[33] The historical institutionalist account just briefly summarized implies that associational development was propelled by the after-effects of associational building and development during the U.S. Civil War.

Importantly, while Crowley and Skocpol have nicely analyzed *how* associational development was influenced by the Civil War, they have very little to say about *when* these effects materialized across time. Despite including time dummies as "temporal controls", they do not estimate or plot the hazard across the seven decades included in their analysis. This is unfortunate as the historical-institutionalist account that they discuss implies that time plays a significant role in associational development in the postbellum U.S. If the historical-institutionalist account is correct, we would expect growth in associations to be strongest in the immediate aftermath of the Civil War, which is exactly what we find in

[33] For a complete account, see Crowley and Skocpol (2001, 814–816) and the sources they cite.

figure 12.

Given that the key variable that captures the effect of the Civil War varies across states, we would expect the temporal dynamics of associational development to vary depending on how much mobilization for and fighting in the Civil War was experienced in each state. Recall that two of Crowley and Skocpol's primary variables measure the effects of the Civil War on associational formation in the postbellum United States. These two variables measure the percent of the population in Union armies and the pension dollars per pensioner respectively. We have strong reason to suspect that the effects of the Civil War exhibit non-proportionality over the 7 decade postbellum era, as the individuals and organizational structures that were built and/or gained experience during the war effort would have had their effect within a couple decades or atrophied (in the case of organizational structures) or even died (in the case of individuals). Thus, we should expect the Civil War effect to taper off considerably as we observe durations that imply an association being formed in the 1900's. The pension dollars variable was the most substantively significant of the two (see Crowley and Skocpol (2001, 825)[34]) and exhibited non-proportionality. We used a likelihood ratio test where the restricted model did not have the time interactions with the variable of theoretical interest and found that the restrictions were rejected at well below the 0.05 level.[35]

The graph in Figure 13b plots the probability that a federated voluntary association forms as we vary both time and the pension dollars per pensioner in each state. When a federated association is created (duration=0), which is 1860 for 17 of the 21 associations, the effect of pension dollars per pensioner is dramatic, as it increases the probability from about 0.15 to nearly 0.50. As the time between the start of the association and the emergence of a state-level branch increases, the effect of pension dollars per pensioner decreases markedly. When pension dollars per pension is very high (i.e., \$200–\$250), the probability drops dramatically in a nearly linear fashion. However, in states where pension dollars per pensioner is relatively low, the effects of time are non-linear. The hazard increases for approximately two decades after the founding of the federation and subsequently decreases, although the decrease is less dramatic than in states where per

[34]We replicated Crowley and Skocpol's table 5 using logit with a cubic polynomial and obtained very similar results.

[35]Box-Steffensmeier and Zorn (2001) advocate examination of the Schoenfield residuals plotted against the variable(s) of interest when using a continuous time Cox proportional hazards model. When using logit or cloglog, a likelihood ratio test will pick up whether the restriction of proportionality is correct and be much easier to implement. See King (1998, 84–85) and Greene (2003, 349) for details.

pensioner pension dollars is high. This relationship indicates that organizations took off very fast where the Union armies had a large postbellum footprint and were generally followed several decades later by states where the organizational influence of the Union army was negligible. Contrast this with the graph in Figure 13a that depicts the equivalent plot under the proportional hazards assumption. Here the effect of time is constant across different levels of pension dollars; thus, the hazard shape is identical to that in Figure 12, although it shifts upward as pension dollars increase. Thus, while Crowley and Skocpol were correct about the positive impact of Union armies, they did not fully explore the temporal dimension of the relationship, which actually provides even stronger support for their theoretical perspective. In sum, when time is modeled in a theoretically informed way we are able to produce richer empirical analysis and to actually "take time seriously".

7 Conclusion

BKT made an important methodological and substantive contribution in demonstrating that BTSCS data was simply grouped duration data. This simple observation, coupled with recommendations for how to deal with temporal dependence has generally improved the quality of empirical research that analyzes BTSCS data. This being said, we take issue with BKT's recommendations that researchers utilize logit with time dummies or splines to analyze BTSCS data that exhibits time dependence. We show that time dummies suffer from complete and quasi-complete separation issues and inefficiency issues, while splines suffer from problems of interpretation and knot selection. Additionally, there is more to "taking time seriously" than just adding regressors as "controls." We argue that researchers should also plot and interpret the hazard rather than treat temporal dependence as a nuisance. Failing to do so not only neglects substantively interesting and important information, but can also lead to bias in other coefficients of interest due to omitted variable bias.

We show that our simpler alternative outperforms time dummies and performs as well as splines in monte carlos and empirical applications: using t, t^2, and t^3, which serves as a Taylor series approximation to just about any shape of hazard. t, t^2, and t^3 is trivial to implement and avoids the problems associated with time dummies and splines. t, t^2, and t^3 is also quite flexible in that it allows researchers to model substantively interesting non-proportionality of the hazard in regressors. Thus, while we agree with BKT in that researchers need to take time seriously, we recommend that this be done with t, t^2, and

t^3.

A Procedure for Plotting the Estimated Effect of Time in Spline-Based Models

The following assumes the researcher has already conducted logistic regression and included the spline basis vectors as regressors. To plot the estimated $\Pr(y_{i,t} = 1)$ as a function of time, the researcher should follow these steps:

1. Construct a new time vector $\tilde{t} = \{1, 2, 3, \ldots, \max(t)\}$. This will be used in the next step and will serve as the x-axis for the plot.

2. Apply the spline function that was used for the data analysis *in exactly the same way* (i.e., same number and location of knots) to the $\tilde{t}$ vector. This will provide basis vectors for $\tilde{t}$ that correspond to those used in the regression. The ordering of the observations will also correspond to the ordering of time in $\tilde{t}$.

3. Using the logistic regression's parameter estimates, calculate $\Pr(y_{i,t} = 1)$ for each row in the newly generated ($\tilde{t}$) basis vectors, substituting those basis vectors into their corresponding locations in the regression equation, while holding all other (non-spline) variables constant at some value (e.g., their means or medians).

4. The estimated probabilities will be ordered according to the ordering in $\tilde{t}$. The researcher then needs only plot the estimated $\Pr(y_{i,t} = 1)$ with $\tilde{t}$ along the x-axis.

References

Albert, A. and J.A. Anderson. 1984. "On the Existence of Maximum Likelihood Estimates in Logistic Regression Models." *Biometrika* 71(1):1–10.

Alt, James E., Gary King and Curtis S. Signorino. 2001. "Aggregation Among Binary, Count, and Duration Models: Estimating the Same Quantities from Different Levels of Data." *Political Analysis* 9(1):1–24.

Arceneaux, Kevin. 2003. "The Conditional Impact of Blame Attribution on the Relationship Between Economic Adversity and Turnout." *Political Research Quarterly* 56(1):67–75.

Balla, Steven J. 2001. "Interstate Professional Associations and the Diffusion of Policy Innovations." *American Politics Research* 29(3):221–245.

Barbieri, Katherine and Rafael Reuveny. 2005. "Economic Globalization and Civil War." *Journal of Politics* 67(4):1228–1247.

Bearce, David H. and Sawa Omori. 2005. "How Do Commercial Institutions Promote Peace?" *Journal of Peace Research* 42(6):659–678.

Beck, Nathaniel and Jonathan N. Katz. 1997. "The Analysis of Binary Time-Series-Cross-Section Data and/or The Democratic Peace.".

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(4):1260–1288.

Beck, Nathaniel and Simon Jackman. 1998. "Beyond Linearity by Default: Generalized Additive Models." *American Journal of Political Science* 42(2):596–627.

Beck, Neal, Gary King and Langche C. Zeng. 2000. "Improving Quantitative Studies of International Conflict: A Conjecture." *American Political Science Review* 94(1):21–35.

Bennett, D. Scott and Allan C. Stam. 2000*a*. "Research Design and Estimator Choices in the Analysis of Interstate Dyads - When Decisions Matter." *Journal of Conflict Resolution* 44(5):653–685.

Bennett, D. Scott and Allan C. Stam. 2000*b*. "A Universal Test of an Expected Utility Theory of War." *International Studies Quarterly* 44(3):451–480.

Bennett, D. Scott and Matthew C. Rupert. 2003. "Comparing Measures of Political Similarity - An Empirical Comparison of S Versus Tau-b in the Study of International Conflict." *Journal of Conflict Resolution* 47(3):367–393.

Benson, Michelle. 2005. "The Relevance of Politically Relevant Dyads in the Study of Interdependence and Dyadic Disputes." *Conflict Management and Peace Science* 22(2):113–133.

Bernard, Michael, Christopher Reenock and Timothy Nordstrom. 2003. "Economic Performance and Survival in New Democracies - Is There a Honeymoon Effect?" *Comparative Political Studies* 36(4):404–431.

Besancon, Marie. 2005. "Relative Resources: Inequality in Ethnic Wars, Revolutions, and Genocides." *Journal of Peace Research* 42(4):393–415.

Boehmer, Charles R. and David Sobek. 2005. "Violent Adolescence: State Development and the Propensity for Militarized Interstate Conflict." *Journal of Peace Research* 42(1):5–26.

Box-Steffensmeier, Janet and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. Cambridge, UK: Cambridge University Press.

Box-Steffensmeier, Janet and Christopher J.W. Zorn. 2001. "Duration Models and Proportional Hazards in Political Science." *American Journal of Political Science* 45(4):972–988.

Braithwaite, Alex. 2005. "Location, Location, Location ... Identifying Hot Spots of International Conflict." *International Interactions* 31(3):251–273.

Brinks, Daniel and Michael Coppedge. 2006. "Diffusion is No Illusion - Neighbor Emulation in the Third Wave of Democracy." *Comparative Political Studies* 39(4):463–489.

Buhaug, Halvard. 2005. "Dangerous Dyads Revisited: Democracies May Not Be That Peaceful After All." *Conflict Management and Peace Science* 22(2):95–111.

Caprioli, Mary. 2005. "Primed for Violence: The Role of Gender Inequality in Predicting Internal Conflict." *International Studies Quarterly* 49(2):161–178.

Caprioli, Mary and Peter F. Trumbore. 2005. "Rhetoric Versus Reality - Rogue States in Interstate Conflict." *Journal of Conflict Resolution* 49(5):770–791.

Caprioli, Mary and Peter F. Trumbore. 2006. "Human Rights Rogues in Interstate Disputes, 1980–2001." *Journal of Peace Research* 43(2):131–148.

Carpenter, Daniel P. and David E. Lewis. 2004. "Political Learning From Rare Events: Poisson Inference, Fiscal Constraints, and the Lifetime of Bureaus." *Political Analysis* 12(3):201–232.

Carter, David B. and Curtis S. Signorino. 2006. "Data Separation and the Statistical Exploration of Wars." *Unpublished Manuscript, University of Rochester* .

Cauthen, James N.G. and C. Scott Peters. 2003. "Courting Constituents: District Elections and Judicial Behavior on the Louisiana Supreme Court." *Justice System Journal* 24(3):265–281.

Chamberlain, Robert and Donald P. Haider-Markel. 2005. ""Lien On Me": State Policy Innovation in Response to Paper Terrorism." *Political Research Quarterly* 58(3):449–460.

Chang, Eric C.C. 2005. "Electoral Incentives for Political Corruption Under Open-List Proportional Representation." *Journal of Politics* 67(3):716–730.

Chiozza, Giacomo. 2002. "Is There a Clash of Civilizations? Evidence From Patterns of International Conflict Involvement, 1946–1997." *Journal of Peace Research* 39(6):711–734.

Choi, Seung-Whan and Patrick James. 2003. "No Professional Soldiers, No Militarized Interstate Disputes? A New Question for Neo-Kantianism." *Journal of Conflict Resolution* 47(6):796–816.

Clark, David H. and Patrick M. Regan. 2003. "Opportunities to Fight - A Statistical Technique for Modeling Unobservable Phenomena." *Journal of Conflict Resolution* 47(1):94–115.

Clark, David H. and Robert A. Hart. 1998. "Controlling Duration Dependence in Conflict Analyses: A Replication and Extension of "Regime Types and Status Quo Evaluations"." *American Journal of Political Science* 42(4):1335–1342.

Clarke, Kevin A. 2003. "Nonparametric Model Discrimination in International Relations." *Journal of Conflict Resolution* 47(1):72–93.

Colaresi, Michael and William R. Thompson. 2002. "Strategic Rivalries, Protracted Conflict, and Crisis Escalation." *Journal of Peace Research* 39(3):263–287.

Crescenzi, Mark J.C. and Andrew J. Enterline. 2001. "Time Remembered: A Dynamic Model of Interstate Interaction." *International Studies Quarterly* 45(3):409–431.

Crowley, Jocelyn Elise and Theda Skocpol. 2001. "The Rush to Organize: Explaining Associational Formation in the United States, 1860s–1920s." *American Journal of Political Science* 45(4):813–829.

Davies, Graeme A.M. 2002. "Domestic Strife and the Initiation of International Conflicts - A Directed Dyad Analysis, 1950–1982." *Journal of Conflict Resolution* 46(5):672–692.

Dickinson, Matthew J. and Kathryn Dunn Tenpas. 2002. "Explaining Increasing Turnover Rates Among Presidential Advisors, 1929–1997." *Journal of Politics* 64(2):434–448.

Dixon, William J. and Paul D. Senese. 2002. "Democracy, Disputes, and Negotiated Settlements." *Journal of Conflict Resolution* 46(4):547–571.

Dorussen, Han. 2006. "Heterogeneous Trade Interests and Conflict - What You Trade Matters." *Journal of Conflict Resolution* 50(1):87–107.

Drury, A. Cooper. 2001. "Sanctions as Coercive Diplomacy: The U.S. President's Decision to Initiate Economic Sanctions." *Political Research Quarterly* 54(3):485–508.

Edwards, Martin S. 2005. "Investor Responses to IMF Program Suspensions: Is Noncompliance Costly?" *Social Science Quarterly* 86(4):857–873.

Enterline, Andrew J. and J. Michael Greig. 2005. "Beacons of Hope? The Impact of Imposed Democracy on Regional Peace, Democracy, and Prosperity." *Journal of Politics* 67(4):1075–1098.

Firth, David. 1993. "Bias Reduction of Maximum Likelihood Estimates." *Biometrika* 80:27–38.

Fox, John. 2000. *Nonparametric Simple Regression: Smoothing Scatterplots*. Thousand Oaks, CA: Sage.

Fritz, Paul and Kevin Sweeney. 2004. "The (De)limitations of Balance of Power Theory." *International Interactions* 30(4):285–308.

Gartzke, Eric and Quan Li. 2003*a*. "All's Well That Ends Well: A Reply to Oneal, Barbieri & Peters." *Journal of Peace Research* 40(6):727–732.

Gartzke, Eric and Quan Li. 2003*b*. "Measure for Measure: Concept Operationalization and the Trade Interdependence-Conflict Debate." *Journal of Peace Research* 40(5):553–571.

Gartzke, Eric and Quan Li. 2003*c*. "War, Peace, and the Invisible Hand: Positive Political Externalities of Economic Globalization." *International Studies Quarterly* 47(4):561–586.

Gartzke, Eric, Quan Li and Charles R. Boehmer. 2001. "Investing in the Peace: Economic Interdependence and International Conflict." *International Organization* 55(2):391–438.

Gelpi, Christopher and Joseph M. Grieco. 2001. "Attracting Trouble - Democracy, Leadership Tenure, and the Targeting of Militarized Challenges, 1919–1992." *Journal of Conflict Resolution* 45(6):794–817.

Gelpi, Christopher and Peter D. Feaver. 2002. "Speak Softly and Carry a Big Stick? Veterans in the Political Elite and the American Use of Force." *American Political Science Review* 96(4):779–793.

Goenner, Cullen F. 2004. "Uncertainty of the Liberal Peace." *Journal of Peace Research* 41(5):589–605.

Goertz, Gary, Byran D. Jones and Paul F. Diehl. 2005. "Maintenance Processes in International Rivalries." *Journal of Conflict Resolution* 49(5):742–769.

Goodcliffe, Jay and Darren G. Hawkins. 2006. "Explaining Commitment: States and the Convention Against Torture." *Journal of Politics* 68(2):358–371.

Gowa, Joanne and Edward D. Mansfield. 2005. "Alliances, Imperfect Markets, and Major-Power Trade." *International Organization* 58(4):775–805.

Greene, William H. 2003. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Hafner-Burton, Emilie M. and Alexander H. Montgomery. 2006. "International Organizations, Social Networks, and Conflict." *Journal of Conflict Resolution* 50(1):3–27.

Hastie, T.J. and R.J. Tibshirani. 1990. *Generalized Additive Models*. New York, NY: Chapman & Hall.

Heath, et al. 2005. "Women on the Sidelines: Women's Representation on Committees in Latin American Legislatures." *American Journal of Political Science* 49(2):420–436.

Heinze, Georg and Michael Schemper. 2002. "A Solution to the Problem of Separation in Logistic Regression." *Statistics in Medicine* 21:2409–2419.

Henderson, Errol A. and Richard Tucker. 2001. "Clear and Present Strangers: The Clash of Civilizations and International Conflict." *International Studies Quarterly* 45(2):317–338.

Henisz, Witold J. 2002. "The Institutional Environment for Infrastructure Investment." *Industrial and Corporate Change* 11(2):355–389.

Howard, Robert M. and Christine H. Roch. 2001. "Policy Change and the State Courts: The Case of Education Finance Reform." *Justice System Journal* 22(2):137–153.

Humphries, Macartin. 2005. "Natural Resources, Conflict, and Conflict Resolution - Uncovering the Mechanisms." *Journal of Conflict Resolution* 49(4):508–537.

James, Scott C. 2006. "Patronage Regimes and American Party Development From "The Age of Jackson" to the Progressive Era." *British Journal of Political Science* 36:39–60.

Jungblut, Bernadette and Richard J. Stoll. 2002. "The Liberal Peace and Conflictive Interactions: The Onset of Militarized Interstate Disputes, 1950–1978." *Journal of Peace Research* 39(5):527–546.

Ka, Sangjoon and Paul Teske. 2002. "Ideology and Professionalism - Electricity Regulation and Deregulation Over Time in the American States." *American Politics Research* 30(3):323–343.

Katz, Jonathan N. and Brian R. Sala. 1996. "Careerism, Committee Assignments, and the Electoral Connection." *American Political Science Review* 90(1):21–33.

Kim, Hyung Min and David L. Rousseau. 2005. "The Classical Liberals Were Half Right (Or Half Wrong): New Tests of the "Liberal Peace"." *Journal of Peace Research* 42(5):523–543.

King, Gary. 1998. *Unifying Political Methodology.* Ann Arbor, MI: The University of Michigan Press.

Kinsella, David and Bruce Russett. 2002. "Conflict Emergence and Escalation in Interactive International Dyads." *Journal of Politics* 64(4):1045–1068.

Krause, Volker. 2004. "Hazardous Weapons? Effects of Arms Transfers and Defense Pacts on Militarized Disputes, 1950–1995." *International Interactions* 30(4):349–371.

Krutz, Glen S. 2005. "Issues and Institutions: "Winnowing" in the U.S. Congress." *American Journal of Political Science* 49(2):313–326.

Lai, Brian. 2003*a*. "The Effects of Different Types of Military Mobilization on the Outcome of International Crises." *Journal of Conflict Resolution* 48(2):211–229.

Lai, Brian. 2003*b*. "Examining the Goals of U.S. Foreign Assistance in the Post-Cold War Period, 1991–1996." *Journal of Peace Research* 40(1):103–128.

Leblang, David A. 1999. "Domestic Political Institutions and Exchange Rate Commitments in the Developing World." *International Studies Quarterly* 43(4):599–620.

Leblang, David and Steve Chan. 2003. "Explaining Wars Fought by Established Democracies: Do Institutional Constraints Matter?" *Political Research Quarterly* 56(4):385–400.

Lebovic, James H. 2003. "The Limits of Reciprocity: Tolerance Thresholds in Superpower Conflict." *Journal of Peace Research* 40(2):139–158.

Lebovic, James H. 2004. "Uniting for Peace? Democracies and United Nations Peace Operations After the Cold War." *Journal of Conflict Resolution* 48(6):910–936.

Lektzian, David and Mark Souva. 2001. "Institutions and International Cooperation - An Event History Analysis of the Effects of Economic Sanctions." *Journal of Conflict Resolution* 45(1):61–79.

Lemke, Douglas and William Reed. 1996. "Regime Types and Status Quo Evaluations: Power Transition Theory and the Democratic Peace." *International Interactions* 22:143–164.

Lemke, Douglas and William Reed. 2001. "War and Rivalry Among Great Powers." *American Journal of Political Science* 45(2):457–469.

Li, Quan and David Sacko. 2002. "The (Ir)relevance of Militarized Interstate Disputes for International Trade." *International Studies Quarterly* 46(1):11–43.

Lujala, Päivi, Nils Petter Gleditsch and Elisabeth Gilmore. 2005. "A Diamond Curse? Civil War and a Lootable Resource." *Journal of Conflict Resolution* 49(4):538–562.

Mansfield, Edward D. and Eric Reinhardt. 2003. "Multilateral Determinants of Regionalism: The Effects of GATT/WTO on the Formation of Preferential Trading Arrangements." *International Organization* 57(4):829–862.

Mansfield, Edward D. and Jon C. Pevehouse. 2000. "Trade Blocs, Trade Flows, and International Conflict." *International Organization* 54(4):775–808.

Mansfield, Edward and Jon Pevehouse. 2006. "Democratization and International Organizations." *International Organization* 60(1):137–167.

Marinov, Nikolay. 2005. "Do Economic Sanctions Destabilize Country Leaders?" *American Journal of Political Science* 49(3):564–576.

McDonald, Patrick J. 2004. "Peace Through Trade or Free Trade?" *Journal of Conflict Resolution* 48(4):547–572.

Meinke, Scott R., Jeffrey K. Staton and Steven T. Wuhs. 2006. "State Delegate Selection Rules for Presidential Nominations, 1972–2000." *Journal of Politics* 68(1):180–193.

Melander, Erik. 2005. "Gender Equality and Intrastate Armed Conflict." *International Studies Quarterly* 46(4):695–714.

Meyer, Bruce D. 1990. "Unemployment Insurance and Unemployment Spells." *Econometrica* 58(4):757–782.

Milner, Helen V. and K. Kubota. 2005. "Why the Move to Free Trade? Democracy and Trade Policy in the Developing Countries." *International Organization* 59(1):107–143.

Mitchell, Sara McLaughlin and Brandon C. Prins. 2004. "Rivalry and Diversionary Uses of Forces." *Journal of Conflict Resolution* 48(6):937–961.

Montinola, Gabriella R. 2003. "Who Recovers First? Banking Crises Resolution in Developing Countries." *Comparative Political Studies* 36(5):541–574.

Mooney, Christopher Z. 2001. "Modeling Regional Effects on State Policy Diffusion." *Political Research Quarterly* 54(1):103–124.

Mooney, Christopher Z. and Mei-Hsien Lee. 2000. "The Influence of Values on Consensus and Contentious Morality Policy: U.S. Death Penalty Reform, 1956–1982." *Journal of Politics* 62(1):223–239.

Mousseau, Demet Yalcin. 2001. "Democratizing With Ethnic Divisions: A Source of Conflict?" *Journal of Peace Research* 38(5):547–567.

Mousseau, Michael. 2000. "Market Prosperity, Democratic Consolidation, and Democratic Peace." *Journal of Conflict Resolution* 44(4):472–507.

Mousseau, Michael. 2002. "An Economic Limitation to the Zone of Democratic Peace and Cooperation." *International Interactions* 28(2):137–164.

Narendranathan, W. and M.B. Stewart. 1993. "How Does the Benefit Effect Vary as Unemployment Spells Lengthen." *Journal of Applied Econometrics* 8(4):361–381.

Oneal, John R. 2003. "Measuring Interdependence and Its Pacific Effects: A Reply to Gartzke and Li." *Journal of Peace Research* 40(6):721–725.

Oneal, John R. and Bruce M Russett. 1997. "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950–1985." *International Studies Quarterly* 41(2):267–293.

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

Oneal, John R. and Bruce Russett. 1999*b*. "Is the Liberal Peace Just an Artefact of Cold War Interests? Assessing Recent Critiques." *International Interactions* 25(3):213–241.

Oneal, John R. and Bruce Russett. 2005. "Rule of Three, Let It Be? When More Really Is Better." *Conflict Management and Peace Science* 22(4):293–310.

Palmer, Glenn, Tamar R. London and Patrick M. Regan. 2004. "What's Stopping You?: The Sources of Political Constraints on International Conflict Behavior in Parliamentary Democracies." *International Interactions* 30(1):1–24.

Palmer, Harvey D. and Guy D. Whitten. 2000. "Government Competence, Economic Performance and Endogenous Election Dates." *Electoral Studies* 19(2–3):413–426.

Peceny, Mark, Caroline Beer and Shannon Sanchez-Terry. 2002. "Dictatorial Peace?" *American Political Science Review* 96(1):15–26.

Peet, Curtis and Mark V. Simon. 2000. "Social Purpose and Militarized Disputes." *Journal of Peace Research* 37(6):659–677.

Pickering, Jeffrey. 2002. "Give Me Shelter: Reexamining Military Intervention and the Monadic Democratic Peace." *International Interactions* 28(4):293–324.

Powers, Kathy L. 2004. "Regional Trade Agreements as Military Alliances." *International Interactions* 30(4):373–395.

Prentice, R.L. and L.A. Gloeckler. 1978. "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data." *Biometrics* 34(1):57–67.

Rasler, Karen A. and William R. Thompson. 2001. "Rivalries and the Democratic Peace in the Major Power Subsystem." *Journal of Peace Research* 38(6):659–683.

Rasler, Karen A. and William R. Thompson. 2004. "The Democratic Peace and a Sequential, Reciprocal, Causal Arrow Hypothesis." *Comparative Political Studies* 37(8):879–908.

Rasler, Karen A. and William R. Thompson. 2006. "Contested Territory, Strategic Rivalries, and Conflict Escalation." *International Studies Quarterly* 50(1):145–167.

Ravlo, Hilde, Nils Petter Gleditsch and Han Dorussen. 2003. "Colonial War and the Democratic Peace." *Journal of Conflict Resolution* 47(4):520–548.

Ray, James Lee. 2003. "Explaining Interstate Conflict and War: What Should Be Controlled For?" *Conflict Management and Peace Science* 20(2):1–31.

Reed, William. 2000. "A Unified Statistical Model of Conflict Onset and Escalation." *American Journal of Political Science* 44(1):84–93.

Reed, William. 2003*a*. "Information and Economic Interdependence." *Journal of Conflict Resolution* 47(1):54–71.

Reed, William. 2003*b*. "Information, Power, and War." *American Political Science Review* 97(4):633–641.

Ruppert, David, M.P. Wand and R.J. Carroll. 2003. *Semiparametric Regression*. Cambridge, UK: Cambridge University Press.

Santner, Thomas J. and Diane E. Duffy. 1986. "A Note on A. Albert and J.A. Anderson's Conditions for the Existence of Maximum Likelihood Estimates in Logistic Regression Models." *Biometrika* 73(3):755–758.

Schamis, Hector E. and Christopher Way. 2003. "Political Cycles and Exchange Rate-Based Stabilization." *World Politics* 56(1):43–.

Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy.* Cambridge, UK: Cambridge University Press.

Searle, Shayle R. 1982. *Matrix Algebra Useful for Statistics.* New York, NY: John Wiley & Sons.

Sechser, Todd S. 2004. "Are Soldiers Less War-Prone Than Statesmen?" *Journal of Conflict Resolution* 48(5):746–774.

Senese, Paul D. 2005. "Territory, Contiguity, and International Conflict: Assessing a New Joint Explanation." *American Journal of Political Science* 49(4):769–779.

Senese, Paul D. and John A. Vasquez. 2003. "A Unified Explanation of Territorial Conflict: Testing the Impact of Sampling Bias, 1919–1992." *International Studies Quarterly* 47(2):275–298.

Sherman, Richard. 2001. "Democracy and Trade Conflict." *International Interactions* 27(1):1–28.

Simmons, Beth A. 2000. "International Law and State Behavior: Commitment and Compliance in International Monetary Affairs." *American Political Science Review* 94(4):819–835.

So, Ying. 1995. "A Tutorial on Logistic Regression." *SUGI Proceedings* .

Sobek, David. 2005. "Machiavelli's Legacy: Domestic Politics and International Conflict." *International Studies Quarterly* 49(2):179–204.

Sørli, Mirjam E., Nils Petter Gleditsch and Håvard Strand. 2005. "Why Is There So Much Conflict in the Middle East?" *Journal of Conflict Resolution* 49(1):141–165.

Stalley, Phillip. 2003. "Environmental Scarcity and International Conflict." *Conflict Management and Peace Science* 20(2):33–58.

Stein, Jana Von. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance." *American Political Science Review* 99(4):611–622.

Sweeney, Kevin J. 2003. "The Severity of Interstate Disputes - Are Dyadic Capability Preponderances Really More Pacific?" *Journal of Conflict Resolution* 47(6):728–750.

Sweeney, Kevin and Paul Fritz. 2003. "Jumping on the Bandwagon: An Interest-Based Explanation for Great Power Alliances." *Journal of Politics* 66(2):428–449.

Thacker, Strom Cronan. 1999. "The High Politics of IMF Lending." *World Politics* 52(1):38–.

Volden, Craig. 2002. "The Politics of Competitive Federalism: A Race to the Bottom In Welfare Benefits?" *American Journal of Political Science* 46(2):352–363.

Volden, Craig. 2006. "States as Policy Laboratories: Emulating Success in the Children's Health Insurance Program." *American Journal of Political Science* 50(2):294–312.

Volden, Craig and Clifford J. Carruba. 2004. "The Formation of Oversized Coalitions in Parliamentary Democracies." *American Journal of Political Science* 48(3):521–537.

Vuong, Quang. 1989. "Likelihood Ratio Tests for Model Selection and Nonnested Hypotheses." *Econometrica* 57:307–333.

Walter, Barbara F. 2004. "Does Conflict Beget Conflict? Explaining Recurring Civil War." *Journal of Peace Research* 41(3):371–388.

Yatchew, Adonis and Zvi Griliches. 1985. "Specification Error in Probit Models." *The Review of Economics and Statistics* 67(1):134–139.

Zorn, Christopher. 2001. "Estimating Between and Within Cluster Covariate Effects, With an Application to Models of International Disputes." *International Interactions* 27(4):433–445.

Zorn, Christopher. 2005. "A Solution to Separation in Binary Response Models." *Political Analysis* 13:157–170.

Table 2: Oneal and Russett Logit Replication

	Time Dummies	B-Spline	Natural Cubic Spline	t, t^2, t^3
Constant	-0.943	-0.966	-0.965	-1.209
	(0.093)	(0.093)	(0.093)	(0.090)
Democracy	-0.547	-0.546	-0.546	-0.537
	(0.080)	(0.080)	(0.080)	(0.078)
Economic Growth	-0.115	-0.115	-0.115	-0.155
	(0.092)	(0.092)	(0.092)	(0.090)
Alliance	-0.471	-0.470	-0.470	-0.489
	(0.090)	(0.090)	(0.090)	(0.087)
Contiguous	0.699	0.694	0.694	0.667
	(0.089)	(0.089)	(0.089)	(0.087)
Capability Ratio	-0.303	-0.304	-0.304	-0.308
	(0.042)	(0.042)	(0.042)	(0.042)
Trade	-12.675	-12.884	-12.889	-14.078
	(10.499)	(10.505)	(10.505)	(10.650)
t			-1.820	-7.457
			(0.111)	(0.340)
t^2				4.422
				(0.345)
t^3				-0.788
				(0.088)
Spline1		-2.331	-245.712	
		(0.156)	(26.123)	
Spline2		-3.637	79.705	
		(0.277)	(10.951)	
Spline3		-6.711	-11.028	
		(0.246)	(2.759)	
Spline4		-2.294		
		(0.361)		
Log-Likelihood	-2554.723	-2582.876	-2582.877	-2658.931
N =	20074	20990	20990	20990

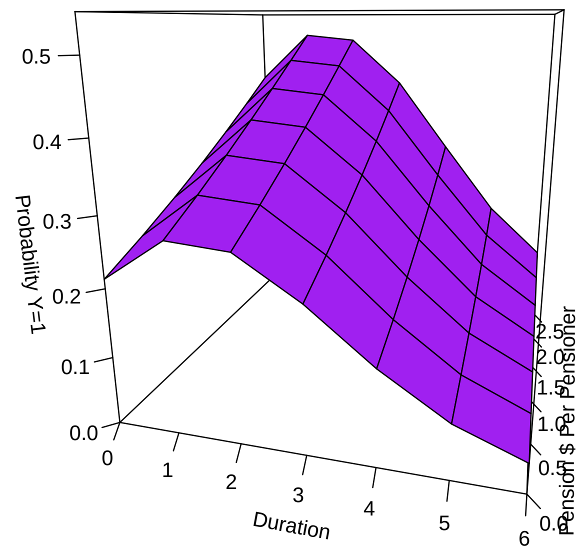
Table 3: Palmer, London and Regan (2004) Logit Reanalysis

	Original	Time Dummies	B-Spline	t, t^2, t^3
Constant	-4.474	-2.313	-2.328	-2.350
	(0.748)	(0.860)	(0.853)	(0.852)
Coalition Score	0.142	0.072	0.070	0.072
	(0.067)	(0.067)	(0.066)	(0.066)
One Pivotal Party?	-0.015	-0.111	0.009	0.025
	(0.489)	(0.528)	(0.591)	(0.595)
Multiple Pivotal Parties?	-0.636	-0.357	-0.267	-0.267
	(0.390)	(0.090)	(0.522)	(0.527)
% of Seats held by Govt.	0.018	0.006	0.005	0.005
	(0.007)	(0.007)	(0.007)	(0.007)
Minority Govt.?	-0.627	-0.638	-0.555	-0.560
	(0.421)	(0.587)	(0.582)	(0.587)
Single-Party Govt.?	-0.691	-0.444	-0.352	-0.354
	(0.383)	(0.530)	(0.525)	(0.530)
Military Power	0.450	0.244	0.245	0.248
	(0.039)	(0.042)	(0.042)	(0.042)
t				-4.231
				(0.686)
t^2				1.658
				(0.576)
t^3				-0.156
				(0.119)
Spline1			-1.806	
			(0.377)	
Spline2			-3.721	
			(0.762)	
Spline3			-2.820	
			(0.496)	
Spline4			-0.633	
			(0.600)	
Log-Likelihood	-737.075	-659.093	-677.723	-678.276
N =	2975	2303	2975	2975

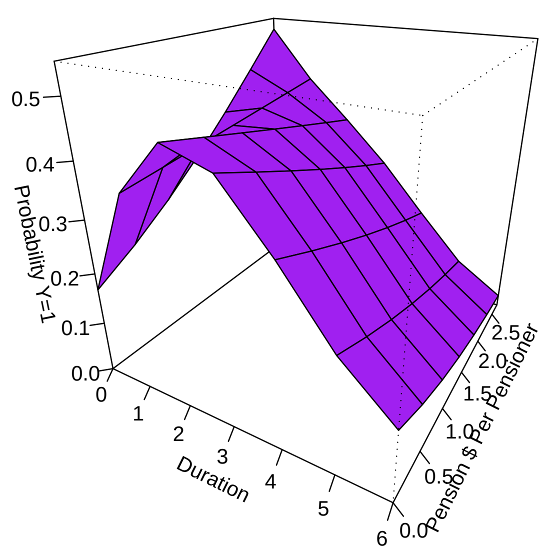
Table 4: Crowley and Skocpol Logit Replication

	Time Dummies	t, t^2, t^3
Constant	-2.443	-2.191
	(0.434)	(0.426)
Urban Growth	0.112	0.095
	(0.061)	(0.059)
Manufacturing per capita	-0.056	-0.145
	(0.295)	(0.290)
Railroad Mi. per capita	-18.834	-18.481
	(10.641)	(10.523)
Teachers per capita	-9.496	-7.128
	(9.531)	(8.684)
Percent literate	0.055	0.040
	(0.053)	(0.053
Percent in Union Armies	0.029	0.031
	(0.010)	(0.010)
Pension $ per pensioner	0.475	0.378
	(0.077)	(0.075)
Electoral Competitiveness	0.045	0.043
	(0.016)	(0.016)
Foreign Born Growth	0.028	0.004
	(0.089)	(0.082)
Population Growth	-0.002	-0.001
	(0.001)	(0.001)
Odd Fellows per capita	-3.660	0.350
	(6.690)	(6.560)
Percent Protestant	0.056	0.072
	(0.029)	(0.028)
Neighbor Effects	0.315	0.283
	(0.046)	(0.044)
t		0.545
		(0.175)
t^2		-0.226
		(0.097)
t^3		0.014
		(0.014)
Log-Likelihood	-1508.480	-1544.439
N =	2529	2529

Figure 13: Proportional and Non-Proportional Hazards for Crowley and Skocpol (2001) Replication



(a) Proportional Hazard



(b) Non-Proportional Hazard