Modeling Duration Dependence

Christopher J. W. Zorn
Department of Political Science, Emory University, Atlanta, GA 30322
e-mail: czorn@emory.edu., http://www.emory.edu/POLS/zorn/

As applications of duration analysis have burgeoned in political science, scholars have become increasingly aware of the potential substantive importance of *duration dependence*: the extent to which the conditional hazards of the events of interest are rising or falling over time. Here I discuss the issue of duration dependence, focusing on the distinction between "spurious" dependence due to unobserved heterogeneity and "true" duration dependence due to state dependence in the process of interest. I present a simple extension of a commonly used parametric duration model—the Weibull model—which allows researchers to assess the influence of causal variables on the nature and extent of duration dependence in their data. I then illustrate the application of this "generalized Weibull" model using data on the duration of international alliances.

1 Introduction

IN RECENT YEARS, models for durations have become widely used in political science. This has been especially true of parametric models, particularly the Weibull model, which has come to dominate duration analyses in political science (e.g., Bennett 1997, 1998; Bennett and Stam 1996; Bueno de Mesquita and Siverson 1995; McCarty and Razaghian 1999; Vuchinich and Teachman 1993; Werner 1999). One of the reasons for the widespread use of the Weibull has been its capacity for dealing with *duration dependence:* the extent to which the conditional hazard of the event of interest occurring is increasing or decreasing over time.[1] While it has been considered in a range of contexts, among political scientists the issue of duration dependence has been most clearly addressed in the debate over such dependence in cabinet durations (e.g., Alt and King 1994; Warwick 1992, 1994; Warwick and Easton 1992).

Recently, Bennett (1999) has argued for the use of parametric models, such as the Weibull, in analyzing political data where the nature of the duration dependence is of substantive interest. His primary assertion regarding duration dependence is that, in many if not most cases, duration dependence is of substantive as well as statistical importance, and that (parametric) models may be chosen to allow us to examine hypotheses regarding that dependence. Here, I address the issue of duration dependence, focusing on its causes and its consequences for model estimation. In particular, I consider the separate issues of dependence due to

Author's note: Thanks go to the John M. Olin Foundation for their support of this research and to Jan Box-Steffensmeier, Dan Reiter, and the anonymous referees for helpful comments. All remaining errors are my own.

[1] An alternative approach to duration dependence is that suggested by Beck et al. (1998); see also Laird and Oliver (1981).

unobserved heterogeneity and that due to state dependence, and outline the strategies which have been developed to address such dependence. I then present a simple extension of the Weibull model which allows the extent of duration dependence to vary according to a specified set of covariates. Such models, which allow duration dependence to vary according to the characteristics of the data, have the potential for widespread application in political science research. Theories of international relations, for example, suggest that the phenomenon of the "democratic peace" may be due to democratic leaders' inability to sustain a losing effort in the face of domestic political pressure (e.g., Reiter and Stam 1998; Russett 1993). If this is the case, a direct implication is that wars in which democracies take part will have greater (positive) duration dependence than those involving autocracies. That is, the hazards of war termination should, *ceteris paribus*, increase over time for democratic states, but may not do so for autocracies.[2] Finally, I illustrate this "generalized" model through an application to the question of international alliance termination, showing how such a model can allow researchers to gain new substantive insights about the process under study.

2 Duration Dependence and Heterogeneity in Event History Models

In models of event histories, duration dependence may arise for two conceptually distinct reasons. The first is *unobserved heterogeneity* in the data. It has long been understood that, if some observations are more likely to experience the event of interest than others, and if the factors contributing to this propensity are not accounted for in the systematic portion of the model, then negative duration dependence will be observed (see, e.g., Heckman 1991; Lancaster 1979; Omori and Johnson 1993). To see why this is the case, consider a population in which the hazards of the event of interest are constant over time, but consist of a mixture of high- and low-hazard individuals. As time passes, those individuals who are more likely to experience the event of interest will do so and be removed from the data. As a result, over time the data increasingly come to consist of low-hazard individuals, and estimates of the average hazard which fail to account for the distinction among subjects will appear to decline over time.[3] In economics, this type of duration dependence is known as "spurious" duration dependence (e.g., Elbers and Ridder 1982); this cause of duration dependence has received some attention among political scientists as well (e.g., Beck 1998; Bennett 1997, 1998).

Several aspects of spurious duration dependence deserve mention here. First, it is important to recognize that the effect of heterogeneity will *always* be to make the observed duration dependence more negative (i.e., to cause the conditional hazard to decrease more, or increase less, over time).[4] In the simple case of mixtures of exponential (i.e., time-constant)

[2]In this light, Bennett and Stam's (1996) finding of no significant duration dependence in the length of interstate wars may be due to their failure to distinguish between the shapes of the conditional hazards of wars involving autocracies [which may be essentially constant, or even declining (e.g., Goemans 1999)] from those of democracies, which may be rising over time (see also Bennett and Stam 1998).

[3]This characterization of duration dependence goes back at least to Silcock's (1954, p. 435) study of job turnover, in which he notes that "entrants with a high value [on the heterogeneity parameter]..., comprising largely the drifters of the industrial world, the misfits and the dissatisfied, will tend to leave, both by resignation and dismissal, more quickly than the remainder of the entrants. Those who remain in employment gradually become 'self-selected' in the sense that the proportion with low values...increases with length of service and in consequence the rate of wastage steadily diminishes."

[4]In fact, unobserved heterogeneity has serious consequences which go far beyond its effect on observed duration dependence. Vermunt (1997, p. 189) notes that

> . . . unobserved heterogeneity biases the duration dependence downward, even if it is not correlated with the observed covariates. If the unobserved risk factors are correlated with the time-constant covariates included in the model, in other words, if there is selection bias, not only are the model parameters biased, but there will also be spurious time-covariate interactions. If there are

or decreasing hazards, the observed mixture will also be declining. For mixtures of increasing hazards, however, the observed hazard may be either decreasing or nonmonotonic, first rising and then falling (Omori and Johnson 1993). Second, note that this bias occurs even if the unobserved heterogeneity is uncorrelated with observed variables, and such heterogeneity may arise whether or not the omitted variables vary over time. In the example given above, omission of the time-constant covariate distinguishing high- and low-risk individuals is all that is required for spurious negative dependence to result.

Finally, the existence of spurious duration dependence suggests that a more accurate picture of the true shape of the conditional hazard may be obtained through better model specification. Lancaster (1979, p. 948; see also Lancaster and Nickell 1980) was among the first to note that the inclusion of additional covariates mitigated negative duration dependence in his study of unemployment duration. More recently, Bennett (1999, p. 262) points out the close relationship between duration dependence and model specification, noting that "if the causal factor driving duration dependence is measured and included as an independent variable, then unexplained duration dependence in the form of a significant p may disappear, even if the hazard rate *is* changing consistently over time." As an example, he cites Bennett and Stam's (1996) study, in which a more fulsome specification of a model of the duration of interstate wars yields no evidence of duration dependence, where previous work had uncovered negative dependence.

A second cause of duration dependence is positive or negative persistence, or what is often termed "true" duration dependence or "state dependence" in econometrics (e.g., Elbers and Ridder 1982; Heckman 1991). Such duration dependence arises when the value of the hazard at any point in time depends on the amount of time which has already elapsed. The potential importance of such "true" duration dependence has long been recognized across a range of fields. In economics, duration dependence has been a central question in studies of business cycles (Mudambi and Taylor 1991; Sichel 1991), speculative bubbles in markets (e.g., McQueen and Thorley 1994), sales [e.g., Zuehlke's (1987) study of the Florida housing market], and other areas. Sociologists and demographers have addressed the substantive importance of duration dependence (e.g., Felmlee et al. 1990), and political scientists have also tended to imbue duration dependence with substantive importance (e.g., Clark and Hart 1998; McCarty and Razaghian 1999; Vuchinich and Teachman 1993; Werner 1999).

In contrast to spurious dependence, the importance of true duration dependence lies in what it reveals about the nature of the process under study, particularly as it relates to the propensity of a state toward self-perpetuation. For example, recent studies of international rivalries suggest that the hazard of such a rivalry ending in any particular year increases over time (e.g., Cioffi-Revilla 1998; Bennett 1998), a phenomenon Bennett (1999, p. 259) has attributed to the tendency for rivals "to wear themselves out and become more willing to settle their disputes over time." Similarly, Sichel (1991, p. 259) finds significant positive duration dependence in postwar economic contractions, but not in expansions, leading him to conclude that "expansions do not seem to burn themselves out, while contractions do."

> unobserved risk factors which also have an effect on changes in the values of particular endogenous time-dependent covariates, the effects of these covariates will be, at least partially, spurious. In addition, unobserved common risk factors may lead to dependent competing risks. And, finally, unobserved heterogeneity may invalidate the assumption of conditional independence in models for repeatable events or other types of clustered observations and may lead to spurious effects of time-varying covariates.

Lancaster (1990, pp. 299–300) shows that neglected heterogeneity biases estimated hazards toward zero as well; see also Omori and Johnson (1993). For good recent discussions of the broader issue of heterogeneity in event history models, see Vermunt (1997) and Box-Steffensmeier and Zorn (1999).

In both instances, the fact of negative state dependence (i.e., that longer periods spent in that state make exit from it more likely) holds important substantive insights about which researchers are concerned. Conversely, true negative duration dependence corresponds to positive state dependence, and suggests a process which becomes self-reinforcing over time. Vuchinich and Teachman's (1993, p. 561) analyses of wars and strikes exhibit negative duration dependence, leading those authors to conclude that "strikes and wars tend to become entrenched, and are less likely to end, the longer they last." Clearly, whether a process exhibits characteristics consistent with the former or the latter will often be of significance to investigators.

From this discussion of the causes of duration dependence, we may draw at least two important points. First is the need to be aware of, and account for, both mechanisms by which duration dependence may arise. This is particularly the case if "true" duration dependence is of substantive concern, since even in instances where one expects and finds rising hazards, the presence of heterogeneity may still have a deleterious effect on both the estimates of that dependence as well as on the structural parameters of the model. As a practical matter, this means that models which allow for both flexible estimation of the shape of the conditional hazard and the incorporation of heterogeneity effects (or other corrections for heterogeneity) are necessary to separate out the two types of dependence.[5] A range of such models is available, and some have begun to receive use in political science applications (e.g., Vuchinich and Teachman 1993; Bennett 1997).[6]

Second, if we admit to the possibility of substantive interest in the shape of the duration dependence, then it stands to reason that scholars may be further interested in the extent to which that dependence varies across different observations in their data. That is, we may want to know whether the hazard of the event of interest is more persistent (i.e., declining faster or rising more slowly) for some observations than it is for others. A number of analyses point to the potential interest in finding answers to such a question. Sichel's (1991) study of business cycles concludes by noting that changes in macroeconomic policy after World War II may be responsible for the fact that contractions, but not expansions, exhibit positive duration dependence, suggesting that the shape of the hazard for cycle reversals may be tied to policy decisions. Similarly, Werner's examination of the failure of peace settlements considers, among other variables, the influence of imposed regime changes on the durability of peace. She notes that "the terms of an imposed settlement also frequently reduce the loser's ability to unilaterally defect by requiring that the loser disarm upon surrender" (1999, p. 916). While one result of such a settlement is certainly to lengthen the peace itself, we might also predict that the degree to which the hazard of a recurrent

[5] A central question in the econometric literature has been the ability to distinguish these effects from one another, and numerous studies on the identifiability of a model with both state dependence and heterogeneity exist (e.g., Elbers and Ridder 1982; Heckman 1991; Heckman and Singer 1984a, b; Honoré 1990; Ridder 1990). In general, the identifying restrictions necessary to estimate such a model (e.g., the inclusion of at least one covariate and the assumption that the first moment of the mixing distribution is not infinite) are relatively innocuous.

[6] It is important to note that, contrary to Bennett's (1999) suggestion, semiparametric models such as that of Cox (1972) also account for duration dependence and, in many instances, do so more flexibly and with fewer potentially restrictive distributional assumptions than do parametric approaches. Moreover, statistical tests for changes in the baseline hazards (e.g., regression of smoothed baseline hazards on time) do exist for the Cox model, and in many instances visual examination of baseline hazard plots is actually more informative about the nature of the duration dependence than are significance tests of parameter estimates. This fact, combined with its overall flexibility, useful mathematical properties (Fleming and Harrington 1991), and the ease with which it handles heterogeneity and multiple events, contributes to the Cox model's canonical status in biostatistics and medicine. Because of its centrality in political science research, however, I focus hereinafter on the Weibull model.

war decreases over time will be greater for nations subject to such an imposed regime change. In the discussion below, I introduce such a model and demonstrate its application by replicating a recent study in which questions of both heterogeneity and state dependence are of considerable importance.

3 Modeling Duration Dependence

I next consider the general issue of duration dependence in the more specific context of Weibull models for survival times. Following the notation of Lancaster (1990), we may write the hazard for the Weibull model as

$$h(t) = \alpha \lambda_i^{\alpha} t^{\alpha - 1} \tag{1}$$

where we typically define $\lambda_i = \exp(X_i \beta)$ and α is the "shape" parameter of the Weibull distribution. The latter, which is usually estimated along with the β's, determines the shape of the hazard function: $\alpha > 1$ denotes rising hazards over time, while $\alpha < 1$ denotes falling hazards, with the special case of $\alpha = 1$ corresponding to an exponential model with a constant hazard rate over time. The Weibull model is well developed (e.g., Kalbfleisch and Prentice 1980) and, as noted above, has been used in a host of political science applications.

In the presence of unobserved, individual-specific heterogeneity, the most widely used modification to the Weibull has been a random-effects (or "frailty") approach in which the individual-specific effects are assumed to follow a gamma distribution (e.g., Butler and Worrall 1989; Lancaster 1979, 1985; Lancaster and Nickell 1980; McDonald and Butler 1987; see also Vaupel et al. 1979; Larsen and Vaupel 1993). If, for example, we assume that the random effects are gamma-distributed,[7] with a mean of 1.0 and variance θ, the observed distribution of hazards takes the form

$$h(t) = \frac{\alpha \lambda_i^{\alpha} t^{\alpha - 1}}{1 + \theta(\lambda_i t)^{\alpha}} \tag{2}$$

(Blossfeld and Rohwer 1995, p. 252). Intuitively, the standard Weibull hazard is "rescaled" by a factor proportional to the level of heterogeneity in the data. In the case of no individual-specific heterogeneity (i.e., $\theta = 0$), this model reduces to the standard Weibull; higher levels of heterogeneity have the effect of reducing the overall variability of the hazard. Figure 1 plots both homogenous and heterogenous Weibull hazards, varying the values of α and θ. If hazards are constant or decreasing (i.e., $\alpha \leq 1$), the introduction of heterogeneity decreases both the level of the hazard and the extent of duration dependence. For $\alpha > 1$, the introduction of gamma heterogeneity yields a nonmonotonic hazard which first rises and then falls (cf. Omori and Johnson 1993).

Models incorporating heterogeneity in this fashion have been shown to do a good job of recovering the structural parameters of interest when random heterogeneity is present (e.g., Petersen 1993), including the presence of "true" duration dependence. They cannot,

[7]The primary advantage of assuming gamma-distributed heterogeneity is that it leads to computationally tractable expressions for the conditional hazard (e.g., Lancaster 1990, pp. 65–70); alternative parametric distributions for specifying heterogeneity are discussed by Hougaard (1986). Heckman and Singer (1984a, b) show that these models can be highly sensitive to the distributional assumptions made about the random effects and propose a nonparametric alternative for the distribution of frailties [but see Manton et al. (1986) for an alternative viewpoint]. Other approaches include a fixed-effects specification, though for a range of reasons (e.g., Yamaguchi 1986; Gönül and Srinivasan 1993) such models have not seen widespread application.

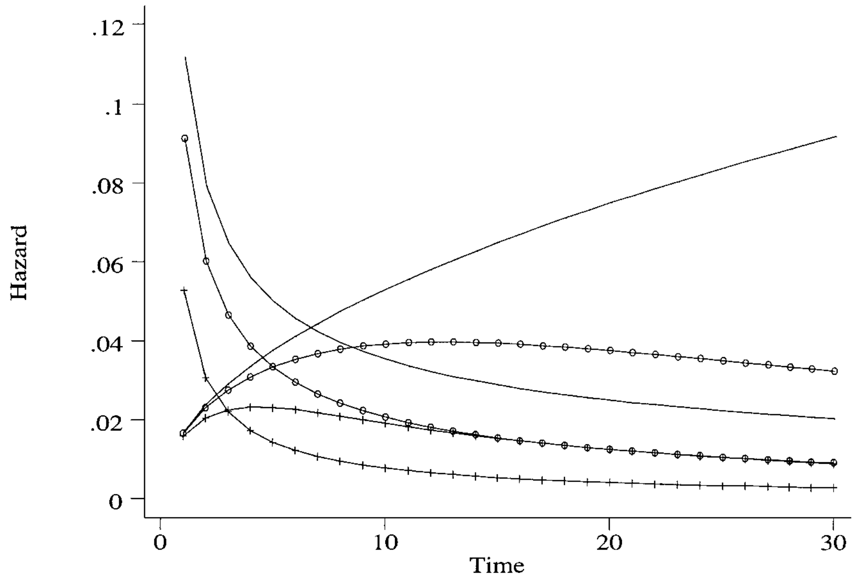

Fig. 1 Weibull hazards with and without gamma heterogeneity. All plots are based on a Weibull distribution with $\lambda = 0.05$; rising hazards have $\alpha = 1.5$, and declining hazards have $\alpha = 0.5$. Smooth lines are hazards for $\theta = 0$ (i.e., no heterogeneity), circled lines for $\theta = 1.0$, and crossed lines for $\theta = 5.0$. See text for details.

however, address the issue of variability in that dependence across observations, that is, the degree to which state dependence differs across observations. A simple way of doing so is to allow the α parameter in the Weibull model to vary according to some function of a set of covariates. Because of the restriction $\alpha > 0$, an obvious choice for that function is the exponential:

$$\alpha_i = \exp(Z_i \gamma) \tag{3}$$

where Z includes a constant term and the use of the exponential link ensures that the estimate of α is strictly positive. Substitution of (3) into (1) or (2) yields a "generalized" Weibull model,[8] in which the vector of parameters γ may then be estimated along with the β's, typically via maximum likelihood. In this respect, the model presented here is analogous to heteroscedastic models developed in the context of OLS (Harvey 1976), probit (Alvarez and Brehm 1995), and event count (King 1989) models.[9]

The result is a model that allows the magnitude of the duration dependence in the data to vary in substantively interesting ways. The parameter estimates γ indicate the extent to which the variables in Z increase or decrease the degree of duration dependence. Thus, for example, variables with positive coefficients would correspond to observations with more positively sloped hazards, i.e., observations where the hazards are either declining at a slower rate or increasing at a faster one. Moreover, because it nests the standard Weibull

[8]The language is analogous to that used in discussions of variance function models (e.g., Davidian and Carroll 1987). Also note that this extension of the Weibull is not novel (see, e.g., Blossfeld and Rohwer 1995), though to this point it has not been used in political science.

[9]See, more generally, Jorgensen (1997). Moreover, it should be noted that this extension need not be limited to the Weibull model; Blossfeld and Rohwer give illustrations for similar parameterizations of the Gompertz, lognormal, log-logistic, and sickle models (1995, Chap. 7).

model, this model is amenable to standard likelihood-based tests (e.g., Wald or likelihood ratio) for inference about the effect of those variables on the Weibull shape parameter. For example, twice the difference of the log-likelihoods between the generalized Weibull and a standard (i.e., constant α) Weibull model is a chi-square test for the null hypothesis of no joint effect of the variables in Z on the duration dependence. Such a model allows one to test substantively interesting hypotheses about differences in the duration dependence of different subsets of his or her data. Finally, the generalized Weibull may also allow for gamma heterogeneity, thus allowing researchers to separate out the effects of heterogeneity and state dependence on the estimated duration dependence.

4 Duration Dependence and the Termination of International Alliances

I illustrate the application of the generalized Weibull with an example from recent research on international alliances. A prominent theme in recent work on international alliances has been the extent to which such alliances become institutionalized. Bennett (1997, p. 855) states this case succinctly: "Alliances often develop formal institutions in the form of intergovernmental organizations to facilitate cooperation among member states. As organizations, these institutions will seek to perpetuate themselves once they have become established; the better established they become, the more likely they will be to extend themselves." The crux of the argument is that, to the extent that alliances become self-perpetuating institutions, the likelihood of their dissolution ought to decrease over time. Such a theory is not, however, supported by the evidence; Bennett finds that hazards of rivalry termination are in fact rising over time and suggests that institutionalization may actually lead to more fragile, rather than more durable, alliances. He notes, however, that "it is unclear . . . how to directly measure alliance institutionalization or the underlying problems demanding institutional solutions" (1997, pp. 855–856).

But while measurement and operationalization problems certainly plague any attempt at formal quantification of alliance institutionalization, the methods addressed herein suggest an alternative approach. By estimating a single duration dependence parameter, Bennett's approach implicitly imposes the restriction that the extent of institutionalization is constant across the whole range of alliances in his data. But if the institutionalization hypothesis is correct, then to the extent that certain kinds of alliances are more or less institutionalized than others, those alliances ought to exhibit less or more duration dependence, respectively. Moreover, current theories of alliance termination suggest a number of the dimensions on which this variability ought to occur. Three such theories are the influence of security concerns, the size of the alliance, and the purpose for which the alliance was formed.

Capability-aggregation theories of alliances (e.g., Walt 1987) emphasize the central role of security concerns in the creation and maintenance of international alliances. That is, changes in the security status of allied nations are held to be directly related to variation in alliance duration: improvements in security due to conditions outside the alliance ought to increase the odds of its dissolution, while such changes due to the alliance itself make the alliance more attractive and thus more durable. Alternatively, theories based on the notion of security–autonomy tradeoffs (e.g., Morrow 1991) suggest that "asymmetric" alliances between states with widely differing military capabilities will be more durable than those among relatively balanced states. Both theories, however, hold implications for questions of alliance institutionalization as well. One might expect, for example, that alliances which yield large security gains for the nations in question, or those which occur in a hostile international climate, would also be more likely to become more formalized over time. Conversely, to the extent that asymmetric alliances are "unusually attractive"

(Morrow 1991, p. 916), states may have greater incentives to formalize such arrangements as compared to those among more evenly matched allies. As a result, we might expect asymmetric alliances to be more state-dependent, and thus exhibit more negative (or less positive) duration dependence than their symmetric counterparts.

Size and purpose offer more direct expectations for the extent of alliance persistence. Bennett (1997) asserts that, due to fewer commitments and vagueness of purpose, large alliances may last longer than smaller ones, and his analysis bears this out. One might also speculate, however, that larger alliances are also more likely to require formal mechanisms for their administration, including ensuring that commitments are met and preventing "free riding," and that this greater level of institutionalization will lead to larger alliances being more self-perpetuating and less positively duration-dependent than smaller ones. Similarly, if "institutions are created to help states solve problems" (Bennett 1997, p. 855), then alliances which are created for a single purpose (for example, winning a specific war) ought not require the level of institutional development necessary for those of a more general nature. Thus, we might expect that, e.g., wartime alliances will be less self-perpetuating, and so exhibit more positive duration dependence, than those occurring under other conditions.

To address these possibilities, I reexamine the institutionalization hypothesis using Bennett's (1997) time-varying data on 207 alliances between 1816 and 1984. These data were also used in earlier studies by Morrow (1991) and Gaubatz (1996).[10] For comparison purposes, I estimate both standard and generalized Weibull models of alliance durations, and for maximum generality I include all of Bennett's variables in both parts of the generalized model.[11] Bennett also considers the possibility of unobserved heterogeneity, though his final integrated model finds only insignificant evidence of such heterogeneity. Accordingly, I also estimate both models with gamma heterogeneity and discuss those results here; doing so also provides a means to separate out heterogeneity and covariate effects on the shape of the hazard itself.[12]

Table 1 presents estimates from standard and generalized Weibull models with gamma heterogeneity for Bennett's (1997) alliance termination data. Results for the general model are broadly consistent with Bennett's analysis, including the finding of no significant heterogeneity for the full model ($p = 0.17$). A likelihood-ratio test indicates that the generalized Weibull provides a significantly better fit than the standard model [$\chi^2(11) = 27.76$, $p = 0.004$], and a number of covariates have a significance influence on both the level and the shape of the hazard. Also, it is interesting to note that, in contrast to the standard model, the generalized model uncovers significant levels of heterogeneity. One possible explanation for this finding is that covariate effects on α are confounded with the individual-specific effects; once covariate effects on the shape of the hazard are taken into consideration, the individual-specific effects become apparent.

The results for specific variables in Table 1 support a number of the hypotheses outlined above. Somewhat paradoxically, and as in the study by Bennett (1997), the effect of an external improvement in security conditions has the effect of both extending the duration

[10] Details of variable sources and codings are available in Bennett's paper; summary statistics for all variables, as well as data and commands for replicating the analyses herein, are available at the *Political Analysis* website (http://www.westviewpress.com/politicalanalysis/index.html).

[11] All models were estimated using Blossfeld and Rohwer's (1995) *TDA* program; other statistical packages which can estimate the generalized Weibull include LIMDEP 7.0 (Greene 1997). In addition, any user-written routine for estimating a Weibull model may be easily modified to include the parameterization indicated in Eq. (3)

[12] Estimates for the alliances models which do not incorporate heterogeneity vary little from those presented in Table 1 and are reproduced in an appendix available at the *Political Analysis* website (http://www.westviewpress.com/politicalanalysis/index.html).

Table 1 Weibull models with gamma heterogeneity: Alliance terminations[a]

Variable	*Standard Weibull*	*Generalized Weibull* λ	α
(Constant)	−4.931**	−3.189**	1.267**
	(0.400)	(0.611)	(0.489)
Change in security	−0.147**	−0.175**	0.078
	(0.033)	(0.053)	(0.041)
Alliance security improvement	−0.012*	−0.017	−0.015*
	(0.006)	(0.009)	(0.007)
Mutual threat	−1.000	−1.257*	1.256
	(0.901)	(0.522)	(0.985)
Capability change	53.536**	36.881**	−1.871
	(10.303)	(9.629)	(7.345)
Symmetry	0.125	−0.217	−0.350
	(0.233)	(0.222)	(0.222)
Capability concentration	−1.310**	−1.666**	−0.158
	(0.391)	(0.373)	(0.391)
Liberal alliance	−0.655*	−1.215**	−0.211
	(0.279)	(0.296)	(0.254)
Polity change	0.067	−0.169	0.457
	(0.311)	(0.311)	(0.400)
Number of states	−0.122**	−0.368**	−0.082**
	(0.043)	(0.077)	(0.027)
Wartime alliance	1.551**	1.041**	0.718
	(0.402)	(0.254)	(0.383)
War Termination	0.419*	0.811**	0.877**
	(0.186)	(0.263)	(0.253)
α	1.649**	—	
	(0.385)		
θ	1.714	8.218**	
	(1.307)	(2.252)	
lnL	−771.45	−757.57	

[a] $N = 207$ (113 terminations and 94 censored cases). Cell entries are coefficient estimates; numbers in parentheses are standard errors. See text for details.
* $p < 0.05$ (two-tailed).
** $p < 0.01$ (two-tailed).

and increasing the extent of duration dependence, though the latter effect is of only marginal statistical and substantive significance. The security effects of the alliance itself, in contrast, have exactly the expected relationship: improvements in security due to the alliance both extend its duration and decrease the slope of the hazard. The finding that hazards for alliances which provide dramatic security benefits increase more slowly is consistent with the theory that such alliances are more likely to be self-reinforcing over time. In contrast, wartime alliances, as well as the ending of wars in which alliance members are engaged, have the opposite effect: such alliances are both shorter and more duration-dependent (i.e., their hazards are rising at a steeper rate). These results are also consistent with the notion that such alliances lack both the underlying motives and the institutional mechanisms for self-perpetuation. The same holds true for the effect of war terminations by member states: such events not only increase the hazard that alliances will end, but also reduce the degree of positive state dependence, and thus persistence, in that alliance.

It is also interesting to note that a number of other variables in the model exert a significant influence on the hazard itself but have little or no effect on the shape of that hazard. The presence of a mutual threat and changes in allies' capabilities, for example, both have their predicted effects on the odds of an alliance termination but exert effectively zero influence on the extent to which those alliances are more or less persistent over time. Perhaps most interesting is the case of liberal or democratic alliances: if, as suggested by Gaubatz (1996) and others, democracies are more stable and less able to change their positions rapidly than autocratic states, then one might expect alliances composed of democracies to exhibit greater institutionalization and, therefore, lower duration dependence than others. Such a hypothesis is not borne out by the data, however; liberal alliances appear to be no more or less persistent and institutionalized than those among other states. Interestingly, however, accounting for the possible covariate effects on the shape of the hazard has the effect of magnifying the size of this variable's effect on the mean hazard substantially; these results thus work to confirm the findings of Gaubatz and others who show that alliances among democracies are longer-lived than those among other kinds of regimes.

Turning finally to the issue of alliance size, we find that the number of states involved in an alliance has a similar effect to that of security issues: as the number of allied nations increases, both the hazards of termination and the (positive) slope of those hazards decrease. This effect is illustrated graphically in Fig. 2, which plots the estimated mean hazards for the two models in Table 1 for two hypothetical alliances: one consisting of two nations (the lowest value possible) and one containing eight allies (roughly the 90th percentile of that variable). In Fig. 2, the smooth lines correspond to hazards estimated by the standard model (with heterogeneity), while the circled lines plot those from the generalized model. In both instances, we observe the nonmonotonic hazards associated with the Weibull/gamma model: the hazard of termination first rises, then falls over the course of the alliance's life span. The effect on the shape of the hazard is also apparent; among two-nation alliances,

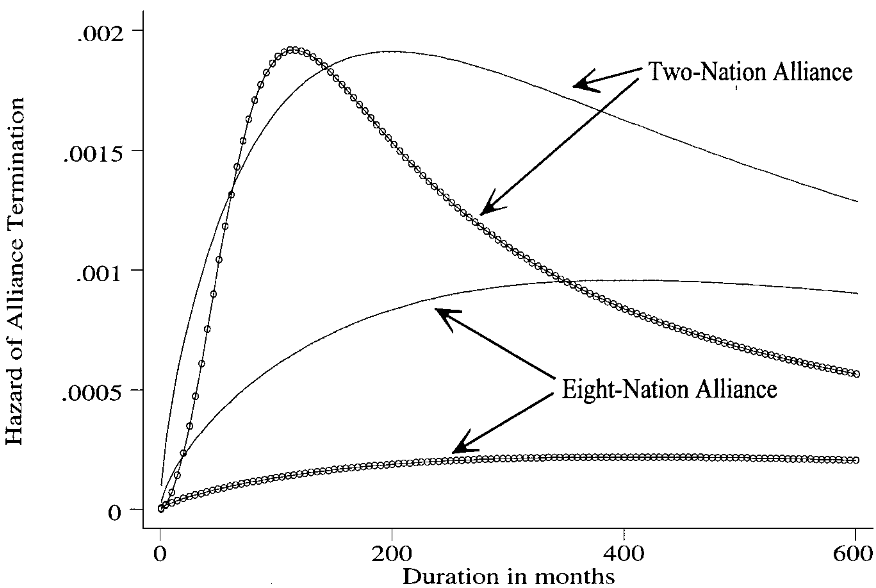

Fig. 2 Estimated hazards, standard and generalized Weibull models with heterogeneity, by number of states. Estimated hazards are for a mean case (i.e., one with all other covariates at their mean values). Smooth lines are estimates for the standard model, and circled lines for the generalized Weibull, both with gamma heterogeneity. See text for details.

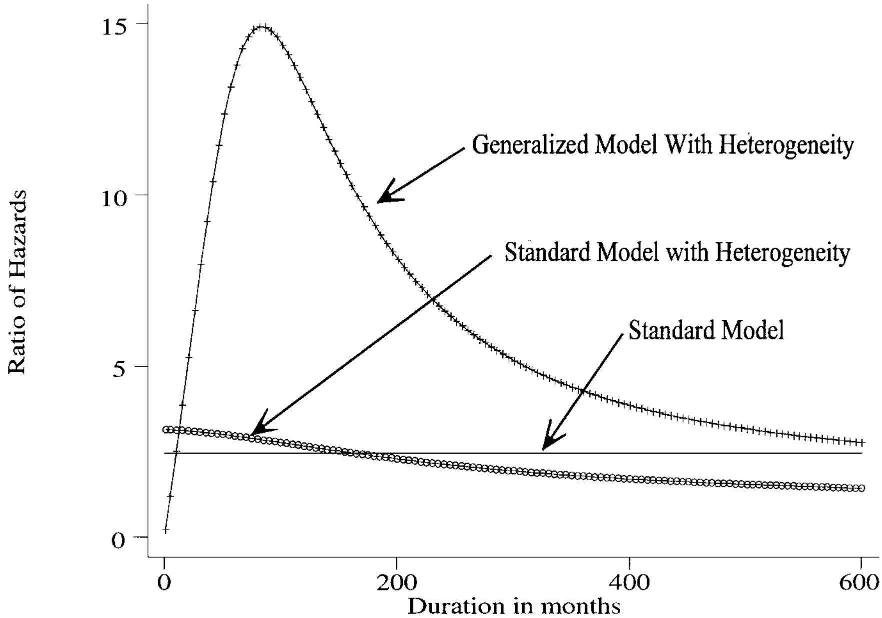

Fig. 3 The ratios of the predicted hazards for two-nation alliances to those for eight-nation alliances presented in Fig. 2 over time for standard, heterogeneous, and generalized heterogeneous Weibull models. See text for details.

this rise is abrupt and large, while that for larger alliances is more gradual. Perhaps most striking is the differences across the two types of models: because of its effect on both the α and the λ parameters, the influence of alliance size is much more pronounced in the generalized model than in the standard Weibull analysis.

Figure 2 also illustrates another characteristic of the generalized Weibull model: its relaxation of the assumption of proportionality in covariate effects. Hazards in the standard model are, by construction, proportional; that is, the ratio of the estimated hazards for (e.g.) rivalries in which at least one nation is democratic to those without a democratic member is constant over time (Lancaster 1990, p. 36). This property is relaxed in the case of models which incorporate heterogeneity; in such models, however, the extent of departure from proportionality depends on the degree of heterogeneity present in the data. In contrast, a model which allows α to vary according to a function of the covariates imposes no such restriction on the hazards. Intuitively, this is because allowing for variable effects in the shape parameter permits an independent variable to exert both a mean shift in the hazard (through its effect on the hazard rate) and something akin to a "slope change" on the shape of the hazard as well. Put differently, incorporation of covariate effects into α permits the effects of independent variables on the hazard to vary over time, in a manner similar to explicit time $\times$ covariate interactions (e.g., Schemper 1992). This added flexibility is a significant advantage of the generalized Weibull, since violations of the proportional hazards assumption render interpretation of covariate effects problematic at best and impossible at worst.[13]

The generalized model's relaxation of the proportional hazards restriction is illustrated in Fig. 3, which plots the ratio of the estimated hazards in Fig. 2 for each of three models: the standard models with and without gamma heterogeneity and the generalized model

[13]On the more general issue of the proportional hazards assumption in the most widely used duration models, see Box-Steffensmeier and Zorn (1998).

with heterogeneity. The figure graphs the hazard for two-nation alliances divided by that for one consisting of eight members, over time. Note that, for the standard model without heterogeneity, the assumption of proportional hazards means that the ratio is constant over time and takes on a value of $\exp(-0.149(2-8)) = 2.445$. The nonproportionality of the heterogeneous model is clear in the circled line, where we see that the effect of alliance size on the hazard decreases in magnitude over time, albeit slightly. Much more striking is the degree of nonproportionality in the variable's effect in the generalized model. There, differences in the hazard due to alliance size first increase rapidly, achieving their maximum effects approximately 6 years into the alliance, and then decline; at longer durations, the effects of such differences are roughly similar across the three models. Such a finding is interesting, in that it suggests that smaller alliances which manage to overcome their initial (greater) hazard of dissolution may become as institutionalized, and thus as stable, as those involving large numbers of states.

5 Conclusion

The increased use of duration models in political science reflects a greater concern on the part of empirical scholars over the use of models appropriate to the data being examined, as well as the realization that such models offer the potential for additional insights into the political processes under scrutiny. Among the latter, political scientists have increasingly come to realize the potential substantive importance of considering duration dependence in event history analyses, and that realization has sparked widespread use of the Weibull model as a means of doing so. It is important to understand, however, the processes by which such dependence may arise and, in particular, the potential effects of model specification, heterogeneity, and state dependence on the observed hazards. Moreover, if duration dependence is of substantive interest, then it behooves us to examine whether and how that dependence may vary across different observations, according to one or more variables suggested by our theory. That is, we ought to go beyond simply *accounting for* such dependence and instead attempt to *explain* it in a theoretically meaningful way.

Here I have offered a simple extension to parametric duration models which allows a more explicit account to be taken of duration dependence in political science data. For simplicity, and because of its wide use in political science, I have focused on the parameterization of that dependence in the Weibull model, showing how such a generalized model can both confirm previous findings (when no systematic influence on duration dependence is present) and offer new insights into well-trod fields of inquiry. Moreover, taken in tandem with models which allow for incorporation of unobserved heterogeneity through random effects, this development offers the potential for separating such substantively interesting dependence from that due to unobserved, individual-specific factors outside the specified model.

References

Alt, James, and Gary King. 1994. "Transfers of Governmental Power: The Meaning of Time Dependence." *Comparative Political Studies* 27(2):190–210.

Alvarez, R. Michael, and John Brehm. 1995. "American Ambivalence Towards Abortion Policy: Development of a Heteroskedastic Probit Model of Competing Values." *American Journal of Political Science* 39(Nov.): 1055–1082.

Beck, Nathaniel. 1998. "Modelling Space and Time: The Event History Approach." In *Research Strategies in the Social Sciences*, eds. E. Scarbrough and E. Tanenbaum. London: Oxford University Press.

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(Oct.):1260–1288.

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

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

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

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

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

Blossfeld, Hans-Peter, and Götz Rohwer. 1995. *Techniques of Event History Modeling*. Mahwah, NJ: Lawrence Erlbaum.

Box-Steffensmeier, Janet M., and Christopher J. W. Zorn. 1998. "Duration Models and Proportional Hazards in Political Science." Paper presented at the Annual Meeting of the Midwest Political Science Association, April 23–25, Chicago, IL.

Box-Steffensmeier, Janet M., and Christopher J. W. Zorn. 1999. "Modeling Heterogeneity in Duration Analysis." Paper presented at the 16th Annual Meeting of the Society for Political Methodology, July 15–18, College Station, TX.

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

Butler, Richard J., and John D. Worrall. 1991. "Gamma Duration Models with Heterogeneity." *The Review of Economics and Statistics* 73:161–166.

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

Cox, D. R. 1972. "Regression Models and Life Tables." *Journal of the Royal Statistical Society, Series B* 34(2): 187–220.

Davidian, M., and R. J. Carroll. 1987. "Variance Function Estimation." *Journal of the American Statistical Association* 82:1079–1091.

Elbers, Chris, and Geert Ridder. 1982. "True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model." *Review of Economic Studies* 49:403–409.

Felmlee, Diane, Susan Sprecher, and Edward Bassin. 1990. "The Dissolution of Intimate Relationships: A Hazard Model." *Social Psychology Quarterly* 53(1):13–30.

Fleming, Thomas R., and David P. Harrington. 1991. *Counting Processes and Survival Analysis*. New York: Wiley.

Gaubatz, Kurt Taylor. 1996. "Democratic States and Commitment in International Relations." *International Organization* 50(Jan.):109–130.

Goemans, Hein. 1999. "*The Termination of the First World War*." Manuscript. Durham, NC: Duke University.

Gönül, Füsan, and Kannan Srinivasan. 1993. "Consumer Purchase Behavior in a Frequently Bought Product Category: Estimation Issues and Managerial Insights from a Hazard Function Model with Heterogeneity." *Journal of the American Statistical Association* 88(Dec.):1219–1227.

Greene, William. 1997. *LIMDEP Version 7.0: User's Manual*. Bellport, NY: Econometric Software, Inc.

Harvey, A. C. 1976. "Estimating Regression Models with Multiplicative Heteroscedasticity." *Econometrica* 44(May):461–465.

Heckman, James J. 1991. "Identifying the Hand of Past: Distinguishing State Dependence from Heterogeneity." *American Economic Review* 81(May):75–79.

Heckman, James J. and Burton Singer. 1984a. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52(Mar.):271–320.

Heckman, James J. and Burton Singer. 1984b. "The Identifiability of the Proportional Hazards Model." *Review of Economic Studies* 51:231–241.

Honoré, Bo E. 1990. "Simple Estimation of a Duration Model with Unobserved Heterogeneity." *Econometrica* 58(Mar.):453–473.

Hougaard, Philip. 1986. "Survival Models for Heterogeneous Populations Derived from Stable Distributions." *Biometrika* 73(Aug.):387–396.

Jorgensen, Bent. 1997. *The Theory of Dispersion Models*. London: Chapman and Hall.

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

King, Gary. 1989. "Event Count Models for International Relations: Generalizations and Applications." *International Studies Quarterly* 33:123–147.

Laird, Nan, and Donald Oliver. 1981. "Covariance Analysis of Censored Survival Data Using Log-Linear Analysis Techniques." *Journal of the American Statistical Association* 96(June):231–240.

Lancaster, Tony. 1979. "Econometric Methods for Duration of Unemployment." *Econometrica* 47:939–955.

Lancaster, Tony. 1985. "Generalized Residuals and Heterogenous Duration Models, with Application to the Weibull Model." *Journal of Econometrics* 28:155–169.

Lancaster, Tony. 1990. *The Econometric Analysis of Transition Data*. New York: Cambridge University Press.

Lancaster, Tony, and Stephen J. Nickell. 1980. "The Analysis of Re-Employment Probabilities for the Unemployed (with Discussion)." *Journal of the Royal Statistical Society, Series A* 143(2):141–165.

Larsen, Ulla, and James W. Vaupel. 1993. "Hutterite Fecundability by Age and Parity: Strategies for Frailty Modeling of Event Histories." *Demography* 30(Feb.):81–102.

Manton, Kenneth G., Eric Stallard, and James W. Vaupel. 1986. "Alternative Models for the Heterogeneity of Mortality Risks Among the Aged." *Journal of the American Statistical Association* 81(Sept.):635–644.

McCarty, Nolan, and Rose Razaghian. 1999. "Advice and Consent: Senate Responses to Executive Branch Nominations." *American Journal of Political Science* 43(Oct.):1122–1143.

McDonald, James B., and Richard J. Butler. 1987. "Some Generalized Mixture Distributions with an Application to Unemployment Duration." *The Review of Economics and Statistics* 69(May):232–240.

McQueen, Grant, and Steven Thorley. 1994. "Bubbles, Stock Returns, and Duration Dependence." *Journal of Financial and Quantitative Analysis* 29(Sept.):379–401.

Morrow, James D. 1991. "Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances." *American Journal of Political Science* 35(Nov.):904–933.

Mudambi, Ram, and Larry W. Taylor. 1991. "A Nonparametric Investigation of Duration Dependence in the American Business Cycle: A Note." *Journal of Political Economy* 99(3):654–656.

Omori, Yasuhiro, and Richard A. Johnson. 1993. "The Influence of Random Effects on the Unconditional Hazard Rate and Survival Functions." *Biometrika* 80(4):910–914.

Petersen, Trond. 1993. "Recent Advances in Longitudinal Methodology." *Annual Review of Sociology* 19:425–454.

Reiter, Dan, and Alan C. Stam III. 1998. "Democracy, War Initiation and Victory." *American Political Science Review* 92(June):377–389.

Ridder, Geert. 1990. "The Nonparametric Identification of Generalized Accelerated Failure-Time Models." *Review of Economic Studies* 57:167–182.

Russett, Bruce. 1993. *Grasping the Democratic Peace: Principles for a Post Cold War World*. Princeton, NJ: Princeton University Press.

Schemper, Michael. 1992. "Cox Analysis of Survival Data with Non-Proportional Hazard Functions." *The Statistician* 41:455–465.

Sichel, Daniel E. 1991. "Business Cycle Duration Dependence: A Parametric Approach." *The Review of Economics and Statistics* 73(May):254–260.

Silcock, Herbert. 1954. "The Phenomenon of Labor Turnover." *Journal of the Royal Statistical Society, Series* A 117(4):429–440.

Vaupel, James W., Kenneth G. Manton, and Eric Stallard. 1979. "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." *Demography* 16(Aug.):439–454.

Vermunt, Jeroen K. 1997. *Log-Linear Models for Event Histories*. Thousand Oaks, CA: Sage.

Vuchinich, Samuel, and Jay Teachman. 1993. "Influences on the Duration of Wars, Strikes, Riots and Family Arguments." *Journal of Conflict Resolution* 37(Sept.):544–568.

Walt, Stephen M. 1987. *The Origins of Alliances*. Ithaca, NY: Cornell University Press.

Warwick, Paul. 1992. "Rising Hazards: An Underlying Dynamic of Parliamentary Government." *American Journal of Political Science* 36(Nov.):857–876.

Warwick, Paul. 1994. *Government Survival in Parliamentary Democracies*. Cambridge: Cambridge University Press.

Warwick, Paul, and Stephen T. Easton. 1992. "The Cabinet Stability Controversy: New Perspectives on a Classic Problem." *American Journal of Political Science* 36(Feb.):122–146.

Werner, Suzanne. 1999. "The Precarious Nature of Peace: Resolving the Issues, Enforcing the Settlement, and Renegotiating the Terms." *American Journal of Political Science* 43(July):912–934.

Yamaguchi, Kazuo. 1986. "Alternative Approaches to Unobserved Heterogeneity in the Analysis of Repeatable Events." In *Sociological Methodology 1986*, ed. Nancy Brandon Tuma. Washington, DC: American Sociological Association.

Zuehlke, Thomas W. 1987. "Duration Dependence in the Housing Market." *The Review of Economics and Statistics* 69:701–709.