Conflict Management and Peace Science, 22:95–111, 2005

ISSN: 0738-8942 print / 1549-9219 online
DOI: 10.1080/07388940590948547

Dangerous Dyads Revisited: Democracies May Not Be That Peaceful After All

HALVARD BUHAUG

Department of Sociology and Political Science
Norwegian University of Science and Technology
Trondheim, Norway

In recent years, the quantitative international relations literature has increasingly paid attention to the potential problem of serially correlated observations in time-series cross-section (TSCS) data. Today, no study using TSCS data is published unless it manages to control for temporal dependence. Using Bremer's (1992) seminal "Dangerous Dyads" article as a starting point, this article has two ambitions. First, it seeks to explore whether the original results are replicable with new and updated data. Second, it aims to uncover whether Bremer's now unconventional choice of statistical model was decisive for his results. The analysis shows that despite clear evidence of serially correlated units, the alternative logit, GEE, and survival models yield relatively similar results. However, Bremer's finding of a monadic democratic peace is not robust to various operationalizations of democracy. His choice of democracy indicator thus seems to have had a significant influence on his conclusions, demonstrating that data selection can be as crucial as the choice of statistical model.

Keywords war, democratic peace, temporal dependence, replication

The article "Dangerous Dyads" by Stuart Bremer (1992) is one of the most frequently cited empirical studies in international relations.[1] This article presented several methodological innovations and provided new insight into the war puzzle. For instance, the analysis demonstrated the very powerful impact of geographic contiguity on the likelihood of interstate conflict. This finding has since been overwhelmingly confirmed by numerous conflict studies and is as such not interesting in this context. Another notable but less-publicized finding in the Bremer article, a monadic democratic peace, has failed to receive the same degree of support in later investigations.[2] Is this finding a consequence of Bremer's unorthodox methodology, or can it be attributed to alternative or obsolete data?

I am grateful to Neal Beck, Kristian S. Gleditsch, Nils Petter Gleditsch, and three anonymous reviewers for valuable comments on an earlier draft. I also thank Steve Chan for providing the original freedom data.

Address correspondence to Halvard Buhaug, Norwegian University of Science and Technology (NTNU), Trondheim NO-7491, Norway. E-mail: halvard.buhaug@svt.ntnu.no

[1] As of 1 September 2004, "Dangerous Dyads" had been cited 228 times by other studies registered in the Social Sciences Citation Index (SSCI). Other popular articles cited in this study are Beck, Katz, and Tucker's (1998) piece on temporal dependence, with 164 citations, and the first article by Oneal, Russett, and colleagues on the liberal peace (Oneal et al., 1996), with 74 citations.

[2] Although the bulk of empirical evidence points to a dyadic democratic peace, there seems to be a growing consensus in the literature of a conditional monadic democratic peace (Russett & Starr, 2000; Russett & Oneal, 2001). In particular, advocates of the revisionist point of view argue that while democracies may be as frequently involved in war as other states, they are less likely to initiate the conflict, they tend to join in coalitions (which inflate the number of warring democracies), they are more successful in winning the war within a short amount of time, and they tend to suffer fewer casualties. Given that Bremer (1992) only treats originators as war participants, a monadic democratic peace seems less likely.

My motivation for posing this question is twofold. On the one side, I aim to explore how updated variables and an extended sample size affect the findings of a prominent but aging study of war. Additionally, the analysis seeks to uncover the empirical consequences of selecting alternative statistical models to handle a specific statistical problem: serially correlated error terms. Bremer argued that in order to study interstate war, one must perform large-*N* multivariate analyses on the appropriate, dyadic level. However, the problem with the customary time-series cross-section (TSCS) data is that the units of observation are not independent, a problem that is particularly pressing when the dependent variable is dichotomous. The probability of war for a dyad at any given time is likely to be affected by the conflict history of that dyad as well as by simultaneous occurrences of war in neighboring dyads (by means of contagion or joining behavior). Recognizing this, Bremer elected to use the rather unconventional Poisson exponential regression model, where the TSCS data structure is converted into a dyad-type setup that lacks a temporal dimension. More recently, scholars have developed several alternative statistical means of dealing with temporally dependent data while at the same time maintaining the TSCS data structure. The central issue here, then, is whether the findings from one of the most influential empirical conflict studies of all times are upheld after a dozen years of data improvement and advances in methodology.

Bremer's article was published long before the data replication requirement was put into effect by international relations journals (see King, 1995, and Bueno de Mesquita et al., 2003), so the original dataset is now unavailable.[3] The replication dataset therefore had to be constructed from scratch, following Bremer's description of the operationalized variables. The first analytical test was to rerun Bremer's Poisson model of war onset and compare the results. Not surprisingly, the analysis was unable to replicate Bremer's exact estimates, even if the discrepancies in most instances were negligible. Above all, Bremer's finding that democracies are more peaceful than other regimes in general was reproduced. Since the original analysis relied on a rather unconventional proxy for democracies, the next model included an alternative democracy indicator based on the much more utilized Polity data. This model showed no indication of the alleged general peacefulness of democratic states, implying that the initial monadic democratic peace finding is indeed conditional on the choice of data. The study then proceeds to present alternative and more conventional statistical models to account for temporally correlated observations. These alternative models largely succeed in reproducing the main findings of the Poisson model, but give little support to the notion of a monadic democratic peace. I conclude that it is as important to pay close attention to the selection of data as to the choice of statistical model. Whereas the different statistical estimation procedures produced generally comparable results, the alternative democracy proxies, updated alliance data, and an expanded sample size caused some variables to behave substantively differently. In the end, Bremer's decision to use the Chan freedom variable in favor of Polity's democracy indicator appears to have had a decisive impact on his finding regarding the democratic peace.

Temporal Dependence

Since the publication of Bremer (1992), there has been a spate of literature about how to correct for a common statistical problem associated with TSCS data, known as temporal dependence (Beck & Katz, 1995, 1997; Beck, Katz, & Tucker, 1998; Benoit, 1996; Heagerty,

[3]The replication standard holds that "sufficient information exists with which to understand, evaluate, and build upon a prior work if a third party could replicate the results without any additional information from the author" (King, 1995, 444). It is, in theory, possible to meet this norm without releasing the data used in the analysis, although that requires a thoughtful and comprehensive description of the collection phase, and even that may not be sufficient if sources for the dataset are updated or otherwise altered after publication of the initial study.

Ward, & Gleditsch, 2002; Raknerud & Hegre, 1997). Temporal dependence implies that the units of observation are correlated over time, so that the value of variable X for unit i at time t is largely dependent on the value of X_i for time $t-1, \ldots, T$. For TSCS data with binary dependent variables (BTSCS), the problem becomes even more apparent. When modeling interstate war for all countries over time, most countries will have series of zeros (i.e., not war) for the entire temporal span. Such sequences of identical values cause substantial autocorrelation. Observations of war may also be serially correlated if they are coded as incidence rather than onset and last over several time periods. Autocorrelation, if substantial, causes underestimated variance and thus indicates stronger correlations (higher t-values) than is true. Beck, Katz, and Tucker, replicating Oneal and Russett (1997), convincingly demonstrate that failing to control for serial dependence may lead to erroneous conclusions. This article seeks to uncover the magnitude of temporal dependence in Bremer's (1992) TSCS data and explore whether more conventional models designed to address this problem manage to reproduce the original findings. Eventually, the analysis will elucidate whether data selection or choice of statistical model can explain the rare finding of a monadic democratic peace in the original study.

Before proceeding to the main content of the article, two other methodological issues that will not be dealt with here should be noted. The first is unit heterogeneity. Although traditionally ignored in empirical IR studies, dyads are likely to have certain characteristics not controlled for in the statistical model that correlate with their propensity to engage in interstate conflict (Bennett & Stam, 2000). The normal procedure to correct for such omitted variable bias is to use a fixed-effects model, which generates a dummy for each unit in the sample. This approach is not undisputed, though, and has serious drawbacks both theoretically and methodologically. See Beck and Katz (2001) for a critique of the application of fixed-effects estimations in BTSCS models.

The second potential problem with Bremer's study relates to selection effects. A simple logit analysis of war onset is unable to capture the underlying strategic interaction of the actors leading to the escalation of a dispute to war (Signorino, 1999). We only observe dyads going to war if they are already involved in a militarized dispute. Factors that influence the risk of dispute may have a different impact on the risk of war. For example, it may well be that democratic states are generally less likely to be involved in disputes, but once they have committed themselves they are more likely to escalate these disputes to war (Fearon, 1994). To solve this problem, Reed (2000) suggests applying a unified model of onset and escalation. Whereas this is a valid approach to address selection bias, expanding the dependent variable to include low-level disputes has theoretical implications that reach well beyond Bremer's focus and is therefore outside the scope of this study.

Research Design

The dataset used by Bremer (1992) includes all dyads—i.e. all pairs of states that were members of the international system—for the period from 1816 to 1965. The dataset used here is as much as possible a replication of Bremer's data, where each variable has been collected and coded in accordance with Bremer's narrative. However, several of the original datasets from which the variables were generated have been modified since the publication of "Dangerous Dyads." In building the replication dataset, I consistently relied on the most recent versions of all variables, which presumably are more accurate and facilitate expanding the sample beyond 1965. See the appendix for details about the operationalization of the variables.

The unit of observation is the nondirected dyad-year. Bremer's dataset included a population of 202,778 units for the 1816 to 1965 time span. The replication dataset includes 203,543 observations for the same period and a total of 533,655 dyad-years for the entire temporal domain, 1816–1993. All independent variables were lagged one time period to avoid endogeneity. This causes some wars to be excluded from the analysis, as not all warring participants were members of the international system in the year prior to the war.[4] All statistical operations were conducted using *Stata* 8.

Table 1 compares the summary statistics for the original data with the replication dataset. Evidently, the two datasets, while being very similar, are not identical. In particular, the distribution of the dependent variable differs. Whereas Bremer's dataset includes 93 dyadic war onsets (first-day participants only) from 1816 to 1965, the replication data contain 100 onsets of war during the same period, and a total of 121 onsets for the expanded period 1816–1993. This deviation stems from different versions of the Correlates of War (COW) data being used; Bremer relied on Small and Singer's (1982) list of interstate wars, while the dependent variable in the replication dataset is based on the most recent version of the COW interstate war dataset (Sarkees, 2000). Since the Small and Singer dataset was released, five wars have been added to the current list of interstate wars before 1980: Guatemala–El Salvador (1876), Egypt–UK (1882), France–Thailand (1893), Poland–Lithuania (1920), and Saudi Arabia–Yemen (1934).[5] However, these five wars were all one-on-one confrontations and comprise only five additional dyads of war onset. The remaining deviation between the dependent variables is due to updated originator status in some multiple-actor conflicts, as is the case with Italy and Mecklenburg in the Seven Weeks War (see the appendix and fn. 15).

Replicating the Poisson Model

In "Dangerous Dyads," Bremer analyzed the data in two stages. First, he explored the bivariate relationships between the exogenous factors and the dependent variable. Based on these findings the seven predictors were dummy-coded, assigning the value of 1 to the expected most war-prone condition in each case (see the appendix). In sum, the seven explanatory variables constitute $2^7 = 128$ possible dyad types. Of these, 118 were present in Bremer's sample, compared to 119 dyad types in the replication data. Each combination, or dyad type, then became the unit of analysis, and the dependent variable measured the number of war onsets that each dyad type experienced during the 1816–1965 temporal domain.[6] In addition to the seven exogenous variables, Bremer controlled for the number of observations for each dyad type by including a logged dyad years count variable. In the replication data, this variable (before transformation) ranges from 2 to 41,813 observations,

[4]Bremer initially had 93 dyad-years of war onset but they were reduced to 85 when time lags were applied to the independent variables. By comparison, in the replication sample, 19 of the 117 observations of war from 1816 to 1980 are excluded when the Chan freedom variable is used, and 15 of the 121 dyadic wars in the period 1816–1993 are dropped when using the Polity democracy dummy. Most of these cases involve German states in the Seven Weeks War that are not recognized as independent regimes by Polity and Chan. Some missing lagged values can also be explained by newly independent states that go to war during their first year. Losing more than ten percent of the war cases certainly raises the question of selection bias, and a quick look at the data is not comforting. Whereas roughly half of the warring dyads with valid lagged democracy values include democracies, only two of the fifteen dyads with missing lagged values include democracies by Bremer's definition (United Kingdom against Egypt in 1882 and Czechoslovakia against Hungary in 1919). One possible solution to maintain the latter cases would be to use current values whenever lagged values are missing.

[5]In Small and Singer's (1982) *Resort to Arms*, the war between the UK and Egypt is coded as an extra-systemic war.

[6]In the replication data, 87 of the 119 dyad types never experienced war, 14 dyad types experienced one war, 5 types were involved in two wars, 4 observations had three wars, another 4 dyad types experienced four wars, 2 types had five wars, 1 dyad type was involved in war six times, 1 type had nine wars, and the last dyad type had sixteen war onsets between 1816 and 1965.

TABLE 1 Descriptive statistics

Variable	Category	Dummy value	Bremer 1816–1965 *N*	%	1816–1965 *N*	%	1816–1993 *N*	%
War onset	No onset	0	202,685	99.95	203,443	99.95	533,534	99.98
	Onset	1	93	0.05	100	0.05	121	0.02
Proximity	Not contiguous	0	189,217	93.3	189,875	93.3	511,570	95.6
	Sea contiguous	1	3,019	1.5	3,132	1.5	5,594	1.1
	Land contiguous		10,542	5.2	10,536	5.2	17,771	3.3
Relative power	Small difference	1	62,055	32.1	64,237	32.4	150,084	29.7
	Medium		56,432	29.2	57,900	29.2	137,757	27.3
	Large	0	74,620	38.6	75,916	38.3	217,239	43.0
Power status	Minor–minor	0	163,604	80.7	163,722	80.7	471,349	88.6
	Major–minor	1	36,907	18.2	36,962	18.2	58,213	10.9
	Major–major		2,267	1.1	2,267	1.1	2,562	0.5
Alliance	No agreement	0	187,424	92.4	188,074	92.4	498,287	93.4
	Defense pact	1	11,176	5.5	11,314	5.7	19,466	3.7
	Neutrality		647	0.3	627	0.3	948	0.2
	Entente		3,531	1.7	3,528	1.7	14,954	2.8
Freedom	Both not free	1	99,580	49.9	96,574	49.0	171,298	51.1
(Chan)	One free	0	78,349	39.3	78,785	40.0	132,717	39.6
	Both free		21,644	10.8	21,754	11.0	31,177	9.3
Democracy	Both not democratic	1	80,801	43.8	85,993	44.7	203,041	43.4
(Polity)	One democratic	0	80,668	43.8	83,178	43.2	206,692	44.1
	Both democratic		22,859	12.4	23,297	12.1	58,495	12.5
Development	Both not developed	0	128,939	65.5	131,762	65.0	313,716	61.4
	One developed		61,823	31.2	63,618	31.4	173,858	34.0
	Both developed	1	7,160	3.6	7,501	3.7	23,820	4.7
Militarization	Both not militarized	0	76,467	39.5	80,071	40.4	220,016	43.1
	One militarized		87,720	45.3	88,655	44.8	226,856	44.4
	Both militarized	1	29,366	15.2	29,395	14.8	63,902	12.5
N			202,778		203,543		533,655	

Note: The values are given for the unlagged variables. The Chan freedom variable is only available for the 1816–1980 period.

with a mean value of 1,696 dyad-years.[7] No information is included on the distribution of the original dyad year measure. Bremer then used the exponential Poisson model to analyze the data. By using the Poisson model and focusing on the number of onsets for each dyad type, Bremer avoided temporal dependence between observations; there is only one observation per unit.[8]

Table 2 compares Bremer's multivariate Poisson model (Model 1) with the primary replication analysis (Model 2) and the alternative replication model where Chan's freedom indicator is substituted with the Polity democracy variable (Model 3). Most of the variables behave similarly in the two models and the substantive effects are almost identical. In particular, Bremer's finding of a monadic democratic peace is reconfirmed. Comparatively

[7]The most war-prone dyad type (one war onset in 17 dyad-years) consists of two contiguous, developed, unfree states, at least one of which is a major power; large power difference; allied; and at least one not militarized.

[8]See Alt, King, and Signorino (2001) for further discussion on the characteristics of the Poisson model and comparison with related methodological applications.

TABLE 2 Poisson regression analysis of dyadic war onset, 1816–1965

Variable	Bremer (1992) Model 1	Replication Model 2	Alternative Model 3
Proximity	1.780***	1.935***	2.535***
	(.362)	(.373)	(.453)
Not large power difference	.619**	.607***	.799***
	(.243)	(.232)	(.252)
At least one major power	.658**	.599**	.843***
	(.263)	(.256)	(.284)
Allied	−.397	−.660**	−.882**
	(.287)	(.299)	(.351)
Both not free (Chan)	1.285***	1.123***	
	(.295)	(.265)	
Both not democratic (Polity)			−.094
			(.228)
Both developed	−1.275**	−1.271***	−1.726***
	(.507)	(.491)	(.555)
Both militarized	.098	.068	−.163
	(.240)	(.241)	(.268)
Log dyad-years	.471***	.569***	.694***
	(.130)	(.129)	(.148)
Intercept	−5.468***	−5.865***	−6.613***
	(1.206)	(1.226)	(1.437)
Chi-square		196.93***	185.28***
N	118	119	119

Note: Standard errors are in parentheses. Model 1 is identical to Table 3 in Bremer (1992, 334).

* $p < .1$; ** $p < .05$; *** $p < .01$, two-tailed tests.

free states *are* more peaceful than other regimes. The most important deviation in the replication is that allied dyads are more peaceful. The original study found a similar effect of alliance only after adding an interaction term between alliances and level of militarization (not tested here). Although unexpected, this finding compares well with other studies (e.g., Oneal et al., 1996) and intuitively makes sense. The deviant behavior of the alliance dummy is a consequence of the different versions of the alliance dataset being used. In terms of marginal effects, proximity exerts the single strongest impact on war propensity in both models. Additionally, the absence of free states, dyadic power parity, at least on the undeveloped state (contrary to Bremer's reported bivariate relationship), and at least one major power in the dyad all substantially increase the probability of war.

Despite a high correlation between Bremer's two regime type measures, only the Chan freedom variable indicated a significant monadic democratic peace in the bivariate assessment.[9] In order to maintain as many cases as possible, Bremer included only the Chan

[9]The Polity dummy showed joint democracies to be the combination of states least likely to engage in hostilities. Since the Chan freedom dummy also suggested that pairs of democracies are the least war-prone dyads (though mixed dyads were also significantly less belligerent than autocratic dyads), the bivariate analysis actually produced more robust results in favor of a dyadic than a monadic democratic peace. Bremer did not elaborate on this fact.

regime type variable in the multivariate model. Model 3 in Table 2 suggests what would have happened if he had opted for the Polity indicator instead: The "both not democratic" dummy shows a negative but far from statistically significant effect. This finding corresponds well to the work of other scholars who question the notion of a democratic peace at the country level, in particular when the conflict sample is restricted to first-day participants (Bueno de Mesquita et al., 1999; Gleditsch & Hegre, 1997). Evidently, Bremer's finding of a monadic democratic peace is sensitive to the choice of democracy indicator. For the other regressors in Model 3, the estimates remain substantively unchanged from the previous model. Contiguity, followed by lack of development, still represents the largest threat to dyadic harmony.

Alternative Methodological Approaches

Bremer's (1992) study was initially based on a dyadic TSCS dataset with a binary dependent variable. However, for the multivariate analysis, Bremer employed a Poisson regression model where the unit of observation is the dyad type rather than the dyad-year. In doing so, Bremer bypassed the problem of temporally correlated units. Today, several more conventional statistical estimation models are able to correct for temporal dependence. Among these are Beck, Katz, and Tucker's (1998) grouped duration analysis, logit regression with a decaying function of time (Raknerud & Hegre, 1997), general estimation equations (GEEs), and various survival models. The following section briefly presents four such models and empirically evaluates whether selecting a more familiar statistical approach with an expanded sample alters the main lessons from the original analysis. Again, I run parallel models with the Chan and Polity regime variables. The unit of analysis is the nondirected dyad-year. The models that include Polity's democracy measure cover the 1816–1993 period (443,543 valid observations), whereas those that include the Chan freedom dummy are limited to the 1816–1980 period (327,704 valid observations). The dependent variable is a binary indicator of onset of interstate war between pairs of states. Only originator war dyads—that is, dyads in which both parties engaged in hostilities at the very first day of the war—are considered as dyads at war, and only in the first year of conflict.

TSCS data frequently suffer from temporally dependent observations, which may lead to "overly optimistic inferences" (Beck, Katz, & Tucker, 1998, 1261). One way to reduce the duration dependence is to censor consecutive observations of war, and hence measure onset rather than incidence of war, i.e., corresponding to the above procedure. However, since observations of peace are also likely to be correlated over time, it would be equally justifiable to remove successive observations of peace. This solution is obviously not viable, as it would exclude time-varying covariates and thus remove most of the data that might explain why some dyads are more likely to enter a war than others. Following the autoregressive models in adding a lagged dependent variable y_{t-1} to the model could be another solution, although this approach does not solve the problem of duration dependence between observations separated by $t > 1$.

Beck, Katz, and Tucker (1998) present one solution to correct for temporally dependent data. By asserting that TSCS data are identical to grouped duration data, they demonstrate that temporal dependence can be eliminated by including dummy variables that count the number of time periods since either the start of the sample period or the previous "failure" (onset of war). In contrast to most event history methods, which treat time as a continuous event, the grouped duration approach assumes fixed time intervals, allowing for the habitual nation-year and dyad-year data structures. The main distinction between ordinary logit and the grouped duration model is the inclusion of temporal dummy variables, κ_{t-t0}, in the latter model, where κ_{t-t0} contributes to the baseline hazard or probability of failure (war) in each time interval:

$$P(y_{i,t} = 1 \mid x_{i,t} = 0, t0) = \frac{1}{1 + e^{-\kappa_{t-t0}}}.$$

In order to ensure a "smooth," declining probability of war as a function of time (which is desirable but not necessarily achievable by the dummy variables κ_{t-t0}), Beck and colleagues suggest replacing the dummies with natural cubic splines. Using the smoothing splines instead of the temporal dummies also reduces the number of parameters in the model, thus making it more efficient. The splines are tied together at certain knots, the number and location of which are determined by the researcher. A test of duration dependence is then performed simply by running a significance test of the cubic splines.[10]

An alternative and somewhat simpler method for dealing with duration dependence is to add a decaying function of time to the model (Raknerud & Hegre, 1997). This decay variable is computed corresponding to a theoretical expectation, or visual inspection, of the probability, or rate, of war onset as a function of time since previous onset. Figure 1 illustrates frequency and probability of war onset as a function of time since previous onset (or time since the initial year of the dyad in the system) in the replication sample. Clearly, most conflicts follow shortly after previous ones or shortly after the dyads enter the system, and the effect of time on the risk of war appears to decline monotonically until the curve approaches zero.

If we prefer the decaying function variable to the temporal dummies or the cubic splines, we have to determine the half-life parameter that makes the decaying function best fit the data:

$$decay = 2^{-(T/\alpha)},$$

where α is the half-life time (in years) and T is the duration of nonfailure (peace) until the time of observation. By testing different half-life values in the logit model and controlling for the effects of other covariates, the decay variable with a half-life of just one year ($decay = 2^{-T}$) was found to generate the best results.

A more powerful and flexible model to handle temporal autocorrelation is generalized estimating equations (GEE). In contrast to the splines approach, it requires the user to specify the underlying distribution of the dependent variable as well as the within-group correlation structure. In our case with a binary dependent variable, the binomial distribution is selected and the habitual one-year time lag is applied to the autocorrelation structure. See Liang and Zeger (1986) for an introduction to the GEE model and Bennett and Stam (2000) for an empirical assessment of GEE and comparable models.

The final method to correct for temporal dependence considered here is survival analysis. Originally developed for medical purposes to analyze the risk of deaths (hence the term

[10]This software extension to *Stata* is downloadable from www.vanderbilt.edu/~rtucker/programs/btscs. A nice feature of the BTSCS application is that a "number of previous failures" variable can be generated if desired by the researcher. The probability of conflict is not randomly distributed in space and time. It seems fair to expect a dyad with a history of repeated wars to have a higher probability of yet another war than a dyad that has never experienced war (Box-Steffensmeier, Reiter, & Zorn, 2003). The resulting count variable is ordinal, assuming a linear relationship between the baseline hazard and the number of previous wars. To minimize the differences between the models, the conflict history variable is not included here. However, adding the count variable to the grouped duration model improves the overall performance of the model, but the individual effects of the regressors are hardly affected.

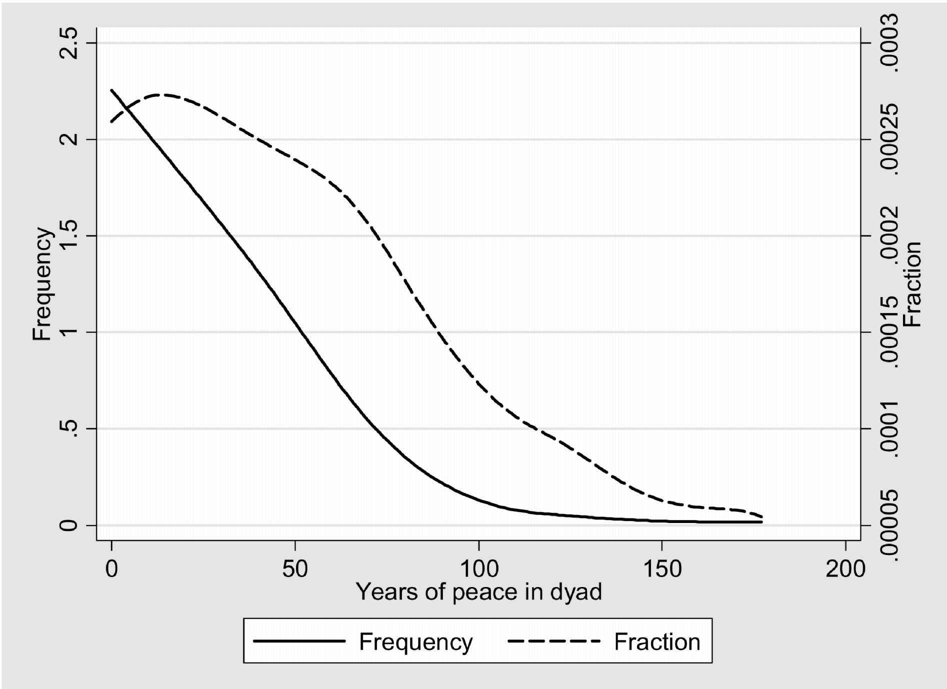

FIGURE 1 Dyadic war onset as a function of years since last onset, 1816–1993.

"survival"), this method allows for right-hand censoring as well as time-varying variables. The least restrictive of the survival models is the semiparametric Cox proportional hazard model (Cleves, Gould, & Gutierrez, 2004). In contrast to the TSCS model that assumes grouped time duration, the proportional hazard model is a continuous duration model that allows for a varying baseline hazard, making it more robust than fully parametric models such as the Weibull model. The hazard function is essentially the probability of failure (war) within a small time interval $(t, t + \Delta t)$, given that the country has survived until time t:

$$h(t) = \lim_{t \to 0} \frac{p(t + \Delta t > T \geq t \mid T \geq t)}{\Delta t}.$$

In order to perform a Cox regression analysis, the data have to be converted to survival-time format, with information on the dyads at the time of transition in the system only. This reduces the overall size of the dataset and corrects for temporal dependence between the peace years as well.

Table 3 reveals the results from the alternative models. Focusing first on the two spline models (Models 4 and 5), the general picture is similar to the earlier results. Once again, the regression output questions the notion of a monadic democratic peace. Model 4 suggests that pairs of comparatively unfree states are neither more nor less war-prone than mixed and free pairs of states. Model 5 actually suggests a dyadic autocratic peace. This may well be consistent with some previous findings that dissimilar regime types are the most hazardous combination (Gleditsch & Hegre, 1997), although the model would need an additional variable (joint democracy) to capture such a parabolic relationship. The lack of support for a monadic democratic peace may also be attributable to Bremer's operationalization of the democracy dummy, which is quite inclusive by today's standards. I will return to this issue below.

TABLE 3 Logit, GEE, and Cox regression analysis of dyadic war onset, 1816–1993

	Logit w/cubic splines		Logit w/ decaying function		GEE		Cox proportional hazard	
Variable	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Proximity	3.547***	3.799***	3.507***	3.744***	3.518***	3.813***	3.841***	3.997***
	(.271)	(.260)	(.265)	(.252)	(.318)	(.309)	(.345)	(.327)
Not large	.768***	.858***	.737***	.833***	.741***	.877***	1.404***	1,407***
power difference	(.253)	(.248)	(.252)	(.250)	(.272)	(.274)	(.338)	(.356)
At least one	1.375***	1.491***	1.286***	1.369***	1.276***	1.358***	2.228***	2.190***
major power	(.281)	(.261)	(.283)	(.264)	(.323)	(.313)	(.370)	(.370)
Allied	−.544*	−.441	−.575**	−.473*	−.599**	−.558**	−.660**	−.652**
	(.278)	(.282)	(.274)	(.274)	(.274)	(.276)	(.322)	(.316)
Both not free	.166		.414*		.430		.235	
(Chan)	(.251)		(.236)		(.303)		(.338)	
Both not democratic		−.480**		−.284		−.276		−.129
(Polity)		(.219)		(.208)		(.260)		(.277)
Both developed	−1.382**	−1.697***	−1.409***	−1.878***	−1.418***	−1.964***	−1.272**	−1.498***
	(.535)	(.527)	(.540)	(.535)	(.536)	(.540)	(.567)	(.562)
Both militarized	.415*	.544**	.421*	.555**	.413*	.557**	.387	.406
	(.243)	(.230)	(.248)	(.234)	(.249)	(.242)	(.349)	(.341)
Peace years	−1.330***	−2.575***						
	(.492)	(.547)						
Decay function			.922	2.542***				
($\alpha = 1$)			(.703)	(.460)				
Intercept	−8.720***	−7.450***	−10.275***	−10.342***	−10.251***	−10.281***		
	(.608)	(.522)	(.285)	(.256)	(.328)	(.294)		
Wald chi-square	577.03***	800.42***	545.61***	726.73***	473.66***	487.03***	541.07***	484.98***
N	328,725	443.543	328,725	443,543	328,713	442,841	157,005	175,287

Note: Robust standard errors are in parentheses. Estimates for three cubic splines in Models 4 and 5 are not reported.
* $p < .1$; ** $p < .05$; *** $p < .01$, two-tailed tests.

The magnitude of the estimates for the jointly not free and jointly not democratic dummies also differs from the respective Poisson models above. This may in part be a consequence of using different statistical estimation procedures, but it is also caused by the extended temporal domain in the latter models. If we run the logit model only for the initial 1816–1965 period (results not shown here), the Chan freedom variable once again suggests a monadic democratic peace significant at the 1% level of uncertainty. Similarly, the estimate for the Polity measure is no longer significantly different from zero. The militarized dummy behaves likewise; in the full sample (Models 3 and 4), it indicates that more militarized dyads are more frequently involved in wars, but for the period until 1965 we find no evidence of such a connection, supporting Bremer's original finding. The effect of alliances is consistently negative but weak.

The comparably large impact of geographic proximity and major power status suggests that the sample could have been restricted to so-called politically relevant dyads (proximate states and dyads consisting of at least one major power) without losing many observations of war. This would eliminate a substantial number of "irrelevant" dyads—that is, dyads that are very unlikely ever to engage in a war (e.g., Mali–Paraguay)—and simultaneously reduce the problem of rare events.[11] The estimated effects of the other regressors are not substantively affected by the choice of statistical model. Finally, the peace-years count variable and the three splines all strongly improve the fitness of the logit model, thus demonstrating that the units suffer from considerable autocorrelation.[12]

The results from the models with a decaying function of time (Models 6 and 7) are largely similar to the splines approach. The only substantial variation pertains to the regime type proxies. In line with the replication Poisson models, we find unfree dyads to be more belligerent than the reference category whereas war involvement of autocratic pairs of states is indistinguishable from the random dyad. In line with previous findings but contrasting with Bremer's results, alliances reduce the likelihood of war onset. Militarization is again positively associated with the general war proneness of the dyad, but the impact of militarization on war propensity weakens considerably when we limit the model to the initial sample period. Quite surprisingly, we see that the decay variable has a negligible effect on Model 6, even though a Wooldridge test suggests that the decay successfully captures the serial dependence of the observations. In the full sample (Model 7), the decaying function of time adds significant power to the model.

The regression estimates from the GEE models are almost identical to those of the decaying approach. Allied and highly developed minor powers that are separated by water or third parties constitute the least war-prone combination of states, regardless of political system involved. The final, alternative method to Bremer's approach is the survival model, here represented by the semiparametric Cox model (Models 10 and 11). These models generate results that are even closer to the original Poisson model; the relative militarization dummy now has no impact on the estimated hazard of war. As with the previous models, neither of the regime proxies is able to explain any variation on the dependent variable.

In general, the alternative models that correct for duration dependence yielded almost identical estimates, for the most part similar to the Poisson findings. Regardless of model, proximity was found to have the largest positive impact on the dependent variable, whereas at least one major power, power parity, and at least one country being less developed also

[11]Following King and Zeng (1999), the logit models were also tested by the *relogit* command (*Stata*), which is designed to correct biased coefficients due to very rare events. The corrected models did not differ substantively from the results reported here.

[12]Another way to estimate the level of autocorrelation in panel-data models is to run the Wooldridge test for serial correlation, *Stata* command *xtserial* (Wooldridge, 2002). This test, too, reveals significantly correlated errors in the logit model before adding splines.

TABLE 4 Cross-tabulation of Chan and Polity democracy indicators, 1816–1980

	No war		War	
	Both not democratic	At least one democracy	Both not democratic	At least one democracy
Both not free	136,246	30,089	53	17
At least one free	15,606	146,807	0	26

Note: All numbers are frequencies.

contribute substantially to increase the risk of war. The effects of alliances and militarization were found to be less robust, but generally suggest that allied pairs of states less often take opposite sides in a war, while relatively militarized dyads are more belligerent than less militarized dyads. The decaying function and the cubic splines added significant strength to three of the logit models, demonstrating substantial duration dependence between units over time. Hence, Bremer made a good choice in selecting a statistical model that would not be biased by temporally correlated residuals.

Arguably the most interesting discovery, the analysis revealed that Bremer's claim of a monadic democratic peace was heavily influenced by his choice of democracy proxy. Moreover, although statistically insignificant in most models, the Chan and Polity democracy variables consistently indicated opposite effects; the former points towards a monadic democratic peace, while the later suggests a weak autocratic peace. How do we explain this different behavior? Table 4 offers a simple, yet fruitful assessment of degree of correspondence between the alternative democracy measures. In both "no war" and "war" samples, a majority of the cases (86% and 82%, respectively) have similar values on the Chan and Polity variables, which explains the high correlation between the measures. Still, 17 warring dyads are inconsistently coded, containing at least one unfree but democratic state. A closer examination of these cases reveals that no less than 12 of the 17 dyads with seeming democratic (but unfree) states, i.e. having a Polity democracy score of at least 5, would not be classified as democratic had we rather employed the now conventional democ–autoc index with a similar value threshold.[13] It is likely that some of the 26 war dyads with at least one free and one democratic state (it need not be the same state) would also be reclassified as not democratic by this procedure. Accordingly, the use of raw democracy scores rather than modified polity values also has some impact on the estimated effect of democracy and amplifies the differences between the alternative democracy measures.[14]

Concluding Remarks

Bremer's (1992) "Dangerous Dyads" was published before the replication standard became established in international relations. Nevertheless, by reconstructing his dataset, it was possible to replicate his findings and to extend them with new data and more sophisticated

[13]The following states at war are coded as democratic but unfree (democracy score with democ–autoc score in parenthesis): Denmark 1864, democ=5 (democ–autoc=2); Japan 1894, 5 (1); Greece 1897, 10 (10); Japan 1900, 5 (1); Japan 1904, 5 (1); Spain 1909, 6 (6); Greece 1912, 10 (10); Yugoslavia 1912, 5 (4); Greece 1913, 10 (10); Yugoslavia 1913, 5 (4); Yugoslavia 1914, 5 (4); Poland 1920, 8 (8); Lithuania 1920, 7 (4); Japan 1937, 5 (1); Japan 1938, 5 (1); Japan 1939, 5 (1).

[14]In the original data, the Polity democracy classification is more inclusive than Chan's freedom dummy. By substituting the Polity democracy measure with a similar dummy made from a democ–autoc index, we get fewer democratic dyads and the number of inconsistent (unfree but democratic) regimes decreases. Still, the regression estimate of a modified democ–autoc dummy remains negative (and insignificant).

methods. The fact that the initial replication model produced almost identical results to the original analysis is a major credit to Stuart Bremer and his thorough description of the data and operationalization procedures. As in the original analysis, I consistently find geographical proximity to be the strongest predictor of dyadic war onset. Major power status, development level, and relative power difference were also found to affect the probability of war across all models, even if their marginal impacts varied somewhat. Alliances, having no apparent effect in Bremer's original model, had a considerable pacifying effect in the replication models. This deviation can only be explained by differences between the versions of the alliance data being used. The effect of militarization was less robust to model selection and the extension of the temporal domain.

Bremer's claim of a monadic democratic peace is very much dependent on his choice of regime type indicator. The alternative models tested here give little support to the alleged general peacefulness of democratic regimes. While this may partly be attributable to the very inclusive operationalization of the Polity democracy dummy, it corresponds well to the majority of empirical works on the democratic peace. In order to uncover any regime-related variation with respect to war involvement, a dummy of joint democratic states would have been a more appropriate indicator.

Ultimately, alternative operationalizations, data updates, and an extended temporal domain appear to have had a greater influence on Bremer's original findings than the selection of estimator. While recent methodological changes have improved greatly our ability to deal with specific statistical problems, we should not ignore the task of evaluating the reliability, validity, and robustness of our empirical measures. This sentiment was indeed shared by Stuart Bremer, who noted that "the empirical study of world politics runs the risk of stagnation unless we devote attention to generating the data required to adequately make use of our theoretical and methodological refinements" (Bremer, Regan, & Clark, 2003, 11).

References

Alt, E., G. King, and C. Signorino. 2001. Aggregation among binary, count, and duration models: Estimating the same quantities from different levels of data. *Political Analysis* 9(1): 21–44.

Beck, N., and J. N. Katz. 1995. What to do (and not to do) with time-series-cross-section data. *American Political Science Review* 89(3): 634–647.

Beck, N., and J. N. Katz. 1997. The analysis of binary time-series-cross-section data and/or the democratic peace. Paper presented at the annual meeting of the Political Methodology Group, Columbus, OH.

Beck, N., and J. N. Katz. 2001. Throwing out the baby with the bath water: A comment on Green, Kim, and Yoon. *International Organization* 55(2): 487–495.

Beck, N., J. N. Katz, and R. Tucker. 1998. Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science* 42(4): 1260–1288.

Bennett, D. S., and A. C. Stam. 2000. Research design and estimator choices in the analysis of interstate dyads: When decisions matter. *Journal of Conflict Resolution* 44(5): 653–685.

Benoit, K. 1996. Democracies really are more pacific (in general): Reexamining regime type and war involvement. *Journal of Conflict Resolution* 40(4): 646–657.

Box-Steffensmeier, J. M., D. Reiter, and C. J. W. Zorn. 2003. Temporal dynamics and heterogeneity in the quantitative study of international conflict. In *Economic interdependence and international conflict: New perspectives on an enduring debate*, eds. E. Mansfield and B. Pollins (273–288). Ann Arbor, MI: University of Michigan Press.

Bremer, S. A. 1992. Dangerous dyads. *Journal of Conflict Resolution* 36(2): 309–341.

Bremer, S. A., P. M. Regan, and D. H. Clark. 2003. Building a science of world politics: Emerging methodologies and the study of conflict. *Journal of Conflict Resolution* 47(1): 3–12.

Bueno de Mesquita, B., J. D. Morrow, R. M. Siverson, and A. Smith. 1999. An institutional explanation of the democratic peace. *American Political Science Review* 93(4): 791–807.

Bueno de Mesquita, B., N. P. Gleditsch, P. James, G. King, C. Metelits, J. L. Ray, B. Russett, H. Strand, and B. Valeriano. 2003. Symposium on replication in international studies research. *International Studies Perspectives* 4(1): 72–107.

Chan, S. 1984. Mirror, mirror on the wall . . . are the freer countries more pacific? *Journal of Conflict Resolution* 28(4): 617–648.

Cleves, M. A., W. W. Gould, and R. G. Gutierrez. 2004. *An introduction to survival analysis using Stata* (revised edition). College Station, TX: Stata Corp.

Fearon, J. D. 1994. Domestic political audiences and the escalation of international disputes. *American Political Science Review* 88(3): 577–599.

Gibler, D. M., and M. Sarkees. 2004. Measuring alliances: The correlates of war formal interstate alliance dataset, 1816–2000. *Journal of Peace Research* 41(2): 211–222.

Gleditsch, N. P., and H. Hegre. 1997. Peace and democracy: Three levels of analysis. *Journal of Conflict Resolution* 41(2): 283–310.

Gurr, T. R., K. Jaggers, and W. H. Moore. 1989. Polity II. *DDIR Update* 3(4): 1–7.

Heagerty, P., M. D. Ward, and K. S. Gleditsch. 2002. Windows of opportunity: Window subseries empirical variance estimators in international relations. *Political Analyst* 10(3): 304–317.

King, G. 1995. Replication, replication. *PS: Political Science and Politics* 28(3): 444–452.

King, G., and L. Zeng. 1999. *Logistic regression in rare events data*. Department of Government, Harvard University, available from http://gking.harvard.edu

Liang, K.-Y., and S. L. Zeger. 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73: 13–22.

Oneal, J. R., and B. Russett. 1997. The classical liberals were right: Democracy, interdependence, and conflict, 1950–1985. *International Studies Quarterly* 41: 267–294.

Oneal, J. R., F. H. Oneal, Z. Maoz, and B. Russett. 1996. The liberal peace: Interdependence, democracy, and international conflict, 1950–1985. *Journal of Peace Research* 33(1): 11–28.

Raknerud, A., and H. Hegre. 1997. The hazard of war. Reassessing the evidence for the democratic peace. *Journal of Peace Research* 34(4):385–404.

Ray, J. L. 1995. *Democracy and international conflict*. Columbus, SC: University of South Carolina Press.

Reed, W. 2000. A unified statistical model of conflict onset and escalation. *American Journal of Political Science* 44(1): 84–93.

Russett, B., and J. R. Oneal. 2001. *Triangulating peace: Democracy, interdependence, and international organizations*. New York: Norton.

Russett, B., and H. Starr. 2000. From democratic peace to Kantian peace: Democracy and conflict in the international system. In *Handbook of war studies II*, ed. M. I. Midlarsky (93–128). Ann Arbor, MI: University of Michigan Press.

Sarkees, M. R. 2000. The correlates of war data on war: An update to 1997. *Conflict Management and Peace Science* 18(1): 123–144.

Signorino, C. S. 1999. Strategic interaction and the statistical analysis of international conflict. *American Political Science Review* 93(2): 279–297.

Singer, J. D., S. Bremer, and J. Stuckey. 1972. Capability distribution, uncertainty, and major power war, 1820–1965. In *Peace, war, and numbers*, ed. B. Russett (19–48). Beverly Hills, CA: Sage.

Small, M., and J. D. Singer. 1982. *Resort to arms: International and civil wars*, 1816–1980. Beverly Hills, CA: Sage.

Stinnett, D. M., J. Tir, P. Schafer, P. F. Diehl, and C. Gochman. 2002. The correlates of war project direct contiguity data, Version 3. *Conflict Management and Peace Science* 19(2): 58–66.

Wooldridge, J. M. 2002. *Econometric analysis of cross section and panel data.* Cambridge, MA: MIT Press.

Appendix

Dependent Variable

Interstate War Onset

The focus of Bremer's study is dyadic war onset. To this purpose, he used the Correlates of War's interstate war dataset from 1982 (Small & Singer, 1982) with a few but unspecified modifications. Since the Exponential Poisson model was chosen as the method of analysis, the unit of analysis was dyad type and the dependent variable was a discrete measure of the number of war onsets for each dyad type. This article additionally employs logit, GEE, and Cox models where the dyad year is the unit of analysis and the dependent variable is a dichotomous indicator of dyadic war onset, as identified in version 3.0 of the COW project's interstate war dataset (Sarkees, 2000). Measuring onset rather than incidence, thus censoring consecutive observations of war, reduces the serial correlation between cases (Ray, 1995). However, using onset does not reduce the dependence over time between observations of peace, so observations on the dependent variable may still be correlated over time.

Only dyads in which both states were original combatants were considered as cases of war, given that the "question of why wars begin is fundamentally different from the questions of why wars grow in size, duration, or severity" (Bremer, 1992, 320). By definition, a state is an original participant if it is involved on the first day of the war. Hence, Germany vs. Poland is the only dyad involved in the onset of World War 2. By comparison, ten states were involved on the first day of the Seven Weeks War (1866), constituting 16 dyad-years (8×2) of war onset.[15] Although Bremer mentions this peculiarity, which is likely to have had a substantial impact on the model, he does not return to the subject when interpreting the results of the statistical analysis.

Independent Variables

Geographical Proximity

To explore the impact of proximity on the probability of war, Bremer relied on the COW Direct Contiguity dataset. In this dataset, the states in a dyad are either land contiguous, sea contiguous (separated by a maximum of 150 miles of water), or not contiguous. Bremer reports that 10,542 of the 202,778 cases were made up of land contiguous pairs of states, while 3,019 dyad-years were sea contiguous. The replication data, using the recent version 3.0 of the COW contiguity data (Stinnett et al., 2002), yield nearly exactly the same figures: 10,536 and 3,132 of the 203,543 observations are land and sea contiguous, respectively. Prior to the multivariate analysis, Bremer recoded all exogenous variables into dummies, reflecting the most war prone condition as suggested by the bivariate analyses. Consequently, the proximity dummy was assigned a value of 1 if the states in the dyad were contiguous by either land or sea.

Relative Power

Using the COW National Material Capability dataset (Singer, Bremer, & Stuckey, 1972), Bremer constructed a Composite Index of National Capability (CINC), an index that gives the joint average score for military, economic, and demographic capability.[16] The relative capability ratios for all dyad-years were calculated, and the dyads were classified into three groups. A capability ratio of less than or equal to 3 corresponded to a small power

[15] Due to missing freedom and democracy scores for several of the German states, only five pairs of states involved in the Seven Weeks War dyads are actually included in the analysis. In contrast, Bremer (1992, 321) reports that the Seven Weeks War is represented by twice as many cases. The deviating number of cases is caused by different versions of the COW data being used. In the 1982 version, the Seven Weeks War is coded as a war of Prussia and Italy against Austria-Hungary and eight German states, where all actors were involved on the first day (May 15), hence constituting a 9×2 contingency matrix. Since the Chan dataset lacks information on four of the German states, Bremer's analysis is limited to ten (5×2) dyads. By comparison, version 3.0 of the COW dataset lists Italy as entering the war at a later stage (May 20) and is thus no longer considered an originator. Further, Mecklenburg-Schwerin is now coded on the side of Prussia (and Italy), hence the originator matrix consists of sixteen (8×2) dyads. Since Mecklenburg-Schwerin is among the four German states with missing regime type data, the replication analysis only includes five dyads (5×1) from the Seven Weeks War.

[16] Cf. Bremer (1992) for further details on how to calculate the CINC scores.

difference; if the ratio was higher than 10, the relative power was considered to be large, whereas a ratio between 3 and 10 implied a medium power difference. Bremer's reported variable distribution compares well with the replication data (in parentheses), which use the revised version 2.1: 62,066 (64,237) units are characterized by a small power difference, 56,432 (57,900) by a medium power difference, and 74,620 (75,916) dyad-years exhibit a large power difference. The dichotomized relative power variable, representing the expected most war prone condition, receives the value of 1 if there is a small or medium (i.e., not large) power difference, given that the bivariate analysis found power parity to be more hazardous than power preponderance.

Power Status

Next, a power status variable was generated, giving the number of major powers in the dyad, as defined by the COW System Membership dataset. During the 1816–1965 temporal domain 2,267 (2,267) of the observations consisted of two major powers, 36,907 (36,962) units contained one major power, while there were 163,604 (163,722) minor–minor dyad-years. The power status dummy variable equals 1 if there was at least one major power in the dyad.

Alliance

In the Correlates of War Formal Alliances dataset (Gibler & Sarkees, 2004), dyads are classified as belonging to one of four groups: (a) defense pact: 11,176 (11,314) observations, (b) neutrality treaty: 647 (627) observations, (c) entente: 3,531 (3,528) observations, and (d) no formal agreement: 187,424 (188,074) observations. In the multivariate analyses, Bremer made no distinction between the different formal agreements. Hence, the dummy alliance variable was assigned a value of 1 if any security agreement was present and 0 otherwise. The replication dataset uses v3.0 of the COW alliance dataset.

Freedom

To control for regime type, Bremer used Chan's (1984) dichotomous Comparative Assessment of Freedom indicator. In this dataset, each state is defined as comparatively free or comparatively unfree, based on a comparison with the systemic distribution of regime types. The descriptive statistics indicate a high degree of similarity between the original and replication datasets: 99,580 (96,574) units have no free states, 78,349 (78,785) contain one free state, while there are 21,644 (21,754) free dyad-years. The bivariate regression suggested that dyads consisting of two unfree states were significantly more prone to war. Thus, the freedom dummy was coded 1 if there were no free states in the dyad.

Democracy

In the bivariate analyses, Bremer also tested a second regime variable, the Polity II democracy variable (Gurr, Jaggers, & Moore, 1989). The democracy index was dummy-coded where states with a value of at least 5 on the 0 to 10 scale were coded as democratic (1), otherwise as undemocratic (0). The replication democracy variable is based on the Polity IV dataset. However, one amendment was carried out; in order to preserve as many cases as possible, some of the missing democracy values (those with –66 on the democracy index, reflecting regime interruption) were recoded to 0 on the dichotomous democracy variable, since such regimes by definition fall outside the democracy classification. Table 1 again suggests a fairly high degree of similarity between Bremer's data and the replication dataset: 80,801 (85,993) dyads contain no democracies, 80,668 (83,178) include one democracy, whereas there are 22,859 (23,297) democratic dyads.

Bremer reported the Polity II and Chan variables to have a Yule's Q correlation coefficient of 0.93. In the replication dataset, the freedom and democracy variables appear to be just as comparable, having a Yule's Q of 0.92.[17] Despite this apparent similarity, Bremer's bivariate analyses produced somewhat different results where the Chan democracy indicator seemed more supportive of a monadic democratic peace. Because of this and because the Polity measure led to the exclusion of more wars due to missing data, only the Chan variable was included in the multivariate analysis. In this paper, both the freedom and democracy dummies are tested in order to assess whether the choice of regime type proxy might influence the results.

Development

As with the relative power variable, Bremer used the COW National Material Capabilities dataset for the development measure. A simple dichotomous variable was computed, measuring whether each state's share of system-wide economic capability (iron and steel production and energy consumption) was larger than that of demographic capability (urban and total population). The dyads were then assigned to one of three groups: (a) both more developed: 7,160 (7,501) dyad-years, (b) one more developed: 61,823 (63,618) dyad-years, and (c) both less developed: 128,939 (131,762) dyad-years. The dummy development indicator was assigned a value of 1 if both states in the dyad were developed, denoting the expected most war-prone situation. The replication dataset uses the updated v3.0 of the COW capability dataset.

Militarization

Following the logic of the development indicator, a state is considered more militarized if its share of global military capability (number of military personnel and military expenditure) is larger than its share of demographic capability. This variable, too, is based on the COW National Material Capabilities dataset. In all, some 29,366 (29,395) dyad-years contain two more militarized states, 87,720 (88,655) observations include one more militarized state, while 76,467 (80,071) observations include only less militarized states. Based on the findings of the bivariate analysis, Bremer expected a dyad of two militarized states to have the highest war proneness, hence the dichotomous militarization variable is coded 1 if both states are more militarized.

[17] Using Pearson's R as an alternative measure of similarity, I find a correlation of 0.72 between the Chan and Polity variables.