The Duration of Interstate Wars, 1816–1985

D. SCOTT BENNETT *The Pennsylvania State University*
ALLAN C. STAM III *Yale University*

We present a model of war duration which incorporates both realpolitik and domestic political variables. We hypothesize that strategy, terrain, capabilities, and government type, among other variables, will play key roles in determining the duration of war. We test these hypotheses using hazard analysis and find empirical support for our key arguments. We find that the realpolitik variables play a greater role than regime behavior and type in determining war duration. We also find that historically, on average, mobilization and strategic surprise have little effect on war duration and that wars are not duration dependent.

The decision by leaders to initiate or join a war is fundamentally a political decision. The impact of political decision making on international conflict does not end when a war begins, however. The decision to *continue* fighting a war is also a fundamentally political decision. Because the processes by which wars are fought are powerfully influenced by politically driven choices, we believe this to be an important area for research. In this paper we seek to understand why wars last as long as they do. We test several hypotheses about the causes of war duration using a statistical hazard model. We investigate four specific questions. (1) What factors cause some wars to last longer than others, and do domestic political factors significantly affect expected war durations when added to a realpolitik model? (2) Is war duration-dependent, that is, does the length of time a war has lasted influence the time it is likely to continue as conflict becomes institutionalized? (3) Can we accurately predict the duration of interstate wars, and in doing so use only ex ante data? (4) Does mobilization, often discussed as a key strategic factor in wartime, make a difference in determining war duration?

Answering these questions and understanding the duration of interstate war is important for several reasons. War duration is a key factor influencing the costs of war. A simple regression of war duration on battle deaths using the Small and Singer (1982) list of interstate wars since 1816 reveals that 24,000 soldiers die (on average) every month a war continues. In a more political vein, wars and their duration have important effects on leaders' popularity (Russett 1990) and on the stability of national regimes (Bueno de Mesquita, Siverson, and Woller 1992). Anticipating the outcome, duration, and costs of possible wars, leaders choose what wars to fight, leading to possible selection bias in the analysis of international conflict (Bueno de Mesquita and Siverson 1995). An improved model of war duration will help us better understand the nature of state behavior and interstate conflict.

This question is also important in helping analyze the role of military mobilization in war. An important body of literature suggests that military mobilization is a significant component in the process by which wars are fought in general (e.g., Trachtenberg 1991; Van Evera 1984). Others argue that states rarely jump through windows of opportunity such as those created by different mobilization rates (Lebow 1984; Reiter 1995; Sagan 1986). While we do not have data which allow us to test claims about mobilization effects (or their anticipated effects) at the very outset of a war, we do test whether mobilization across the course of a multiyear war affects war duration.

Little empirical work on war duration exists. That which does exist relies on questionable methodology or is limited in the independent variables examined. To explore how system polarization affects the duration of war, Bueno de Mesquita (1978) used an ordinary least squares (OLS) regression model. Given that the dependent variable of war duration is positive by construction, this technique was inappropriate for statistical reasons. More recently, Vuchinich and Teachman (1993) used an underspecified hazard model to compare the duration of wars to the duration of strikes, riots, and family arguments, predicting (and finding) that wars will become institutionalized over time. Without a general theoretical model to guide the inclusion and exclusion of independent variables, however, little weight can be given to the finding of duration dependence as an important factor in war duration because such variables could alter estimates of the hazard curve's shape. Other studies have included no independent variables (e.g., Horvath 1968; Morrison and Schmittlein 1980), simply fitting curves to war duration data. We believe that these data-driven models are seriously misspecified and offer little theoretical insight into war.

We draw from two bodies of work that point to theoretical models of war duration. First, some quantitative studies have tried to predict battle duration using a basically realpolitik approach. Epstein (1987) builds upon Lanchester differential equations approaches and develops a model of battles which includes military exchange rates, relative capabilities, strategy, and skill and technology levels. Second, studies of war termination (starting with Iklé 1971) have focused on the political nature of the decision to surrender in wartime.

D. Scott Bennett is Assistant Professor of Political Science, The Pennsylvania State University, University Park, PA 16802, and Allan C. Stam III is Assistant Professor of Political Science, Yale University, New Haven, CT 06520-8301.

An earlier version of this paper was presented at the 28th annual meeting of the Peace Science Society, University of Illinois, Champaign, November 4–6, 1994. The order of the authors' names is alphabetical and is not intended to connote principal authorship. We thank Scott Gartner, Kelly Kadera, Dan Reiter, Randolph Siverson, Gerald Sorokin, and Suzanne Werner for their helpful comments. The data set is available from the authors.

Although the theoretical treatment of war termination has gradually become more formalized (e.g., Mitchell and Nicholson 1983; Wittman 1979), the testing of both realpolitik and political hypotheses about war termination has remained informal, and war-termination studies have examined only a few cases. We believe that it is important not only to consider both realpolitik factors and regime behavior and type in analyzing war duration but also to test our predictions empirically using a population of wars.

THEORETICAL APPROACH

The General Model

We assume that states are headed by rational leaders. By rational we mean that leaders act to maximize the net gains resulting from a deliberate process of decision making. Once a war has begun, the two or more leaders involved periodically choose either to continue fighting or to seek to end their participation in that war by accepting some negotiated settlement or surrendering. A war ends when all the involved leaders have chosen to end their participation in the war rather than continue fighting. The decision to try to end a state's participation in a war will be made when the costs of continuing to fight become greater than the benefits. This quite general conception matches that of Wittman (1979), Morrow (1985), Stam (1996), and Vuchinich and Teachman (1993).

We assume that wars occur over some issue, with some fundamental benefit at stake, such as a piece of territory, economic influence over an area, or some government policy. We also assume these issues to be fixed and exogenous, despite the possible manipulation of the apparent stakes by political leaders. Unlike benefits, the costs of continuing a war rise over time. Different types of states may have different abilities to absorb these accumulating costs. Whether a war continues at some point is then determined by the benefits at stake, the costs of fighting, and states' ability to deal with costs. The causes of longer war durations are factors that increase the stakes, decrease the expense, and increase countries' ability to deal with the costs of war.

Time and the Costs of War

A crucial element in our approach is the relationship between time and the costs to key parts of the state during a war. Militarily, the passage of time may allow states to reduce the costs of fighting by adapting to their opponents' technology or strategy and as such may be viewed by some states' militaries as an asset. Although time can be the best friend of a state trying to adapt to its opponent's war fighting apparatus, it can be a politician's worst nightmare when attempting to maintain public and elite support for the war effort. Russett argues: "Governments lose popularity in proportion to the war's cost in blood and money" (1990, 46). He finds that wars produce a brief spurt of increased popular support followed by increasing levels of social dislocation, dispute, and violent political protest. This suggests that withdrawal may ultimately be forced on leaders if the war drags on past the stage of initial enthusiastic public support, a strong argument against negative duration dependence, which we will test below (negative duration dependence suggests that the probability of a war ending will decrease as the war continues and becomes entrenched).

The problems of military adaptation over time and leaders' inability to maintain indefinite popular support for war create incentives for leaders to limit the amount of time that their states will be involved in a war. Leaders accomplish this in large part by choosing particular military strategies. We believe that it is not necessarily organizational biases which lead militaries to prefer massive and complex offensive strategies (Posen 1984; Snyder 1984b; Van Evera 1984). Rather, we argue that such strategies lead to quicker outcomes and are a rational response to the possible effects of strategic adaptation and rising domestic unrest.

Realpolitik Approach

Strategy and Doctrine. Military strategy has frequently been ignored in political scientists' quantitative studies of international conflict. Previous analyses suggest that strategy plays a key role in determining war outcomes (Dupuy 1979; Epstein 1987; Stam 1993). Strategy is also a main way in which leaders can affect the duration of a war, which is a key concern of decision makers. Since strategy is determined by a choice which can be made in the relatively short term, strategy is a main way states have of directing their power to influence other states.[1] We define military *strategy* as the general way in which a state uses its military forces in a war, and classify three basic types of strategy: *maneuver, attrition*, and *punishment*. Strategy differs in our terminology from *doctrine*, which refers to states' foreign policy *goals* and plans for attaining them. Doctrine can be fundamentally *offensive* or *defensive* depending upon whether state objectives are to maintain or to alter the status quo. Regardless of doctrine, states can use any of the three strategy types.

Maneuver Strategy. Maneuver strategies are those whereby states focus on the use of speed and mobility to disarm the opponent through disrupting the opponent's ability to organize forces effectively. These are often referred to as "blitzkrieg" strategies. When used under an offensive doctrine, the attacker seeks to occupy as much of the defender's territory as possible through a series of sequential actions in which physically occupying enemy territory may be more important than destroying enemy forces. Germany defeated France in World War II with this strategy. When used under a defensive doctrine, the defending state will use mobile forces to seek out, destroy, or disarm attacking forces wherever they are, rather than set up a relatively static defensive line.

[1] Several other variables in the realpolitik approach, such as terrain or available forces (see below), represent *constraints* on leaders' choices of how to use power. While conceptually distinct, realists discuss both choices and constraints or resources as part of the same model. The results of the analysis do not depend upon both types of factors being included or excluded from the analysis.

Attrition Strategy. Attrition strategies seek to destroy or capture opposing forces, making them incapable of continuing to fight, without necessarily using mobility to achieve this. Typically, an attrition strategy seeks large confrontations with the enemy (Mearsheimer 1983, 34). On offense, Hitler adopted an attrition strategy in the Soviet Union, ordering his forces to take as many prisoners and to destroy as many forces as possible rather than bypass large battles, as they had in France. On defense, an attrition strategy typically employs a forward defense, as exemplified by the French Maginot line in World War II.

Punishment Strategy. Punishment strategies attempt to inflict such high costs on an opponent that the attack ceases or the opponent surrenders even though its military forces may not actually be defeated in battle. Unlike maneuver and attrition strategies, punishment is not necessarily a counter-force strategy. The punisher simply counts on being able to break the resolve of the enemy. In the West, punishment strategies developed from the theoretical work of Douhet and Mitchell (e.g., Mitchell 1921) and are exemplified by the logic behind the Mutual Assured Destruction (MAD) nuclear strategy. Another type of offensive punishment strategy is guerrilla warfare. Mao Tse-tung maintained that guerrilla forces should seek to avoid large military confrontations with the enemy and instead inflict damage only when they could avoid suffering serious costs themselves.

Military strategy has an important effect on the length of a war. Maneuver strategies offer the potential for, and anticipation of, the quickest outcomes. They concentrate on a narrow point of attack. Once a narrow salient has been opened, mobile forces move quickly toward rear areas, where they have the potential to cause enormous disruption to the opponent. Mobile strategies entail critical risks, however, because the advancing columns and their following logistical trains are themselves quite vulnerable (Luttwak 1987, 100–101). A maneuver strategy will tend to end relatively quickly, as its rapid success or failure becomes clear.

Attrition strategies rely upon wearing down the opponent and are not predicated on speed in the same way as maneuver strategies. There is less need to move quickly as long as damage is being inflicted on the opponent. Attrition strategies do not depend on the success or failure of any individual battle to as great an extent as a sequential movement strategy, and as a result they are less risky than maneuver strategies.

States using punishment strategies count on their opponents having difficulty adapting to the long-term imposition of costs during a war. The aim is for the target state ultimately to suffer political defeat through the erosion of political resolve among elites and/or mass publics. As the strategy counts on erosion of support rather than destruction or defeat of forces, wars in which one side or both use punishment strategies will last longest of all.

H_{Maneuver}: *Wars in which one or both sides use a maneuver strategy will be the shortest.*

$H_{\text{Punishment}}$: *Wars in which one or both sides use a punishment strategy will be the longest.*

Terrain. Dense terrain (heavy woods, jungles, swamps, or mountains) hinders the movement of both invading and defending forces. In dense terrain, military forces can hide quite easily. While this makes ambushes possible, if fighting begins to inflict high costs on a military force, it can move out of the battle and hide, lowering the costs of fighting. We argue that since costs can be reduced in dense terrain, wars fought in such terrain will last longer. In contrast, flat terrain will lead to shorter wars, as forces can move quickly and fight decisively with little time lost in finding the opponent. Terrain (as well as the capability factors discussed next) can be viewed as a constraint on leaders, influencing such choices as what strategy to employ.

H_{Terrain}: *Wars fought on flat, open terrain will be shorter than those fought on inhospitable terrain.*

Strategy and Terrain Interaction. Terrain and strategy choice also interact. Decision makers and military leaders typically anticipate that terrain and strategy will affect their ability to inflict costs on the opponent, and they will pick a strategy to match the terrain in which they are fighting. For example, states with rough terrain, such as jungles or mountains, will be more likely to fight with punishment rather than maneuver strategies, since the rugged terrain hinders rapid movement but allows forces to inflict costs upon their opponent with relative impunity.

It is possible, however, for leaders to choose an "inappropriate" strategy for given terrain. If a state tries to use a punishment strategy in open terrain, for example, then typically it will find it difficult to sustain the secrecy necessary to fight with a Maoist-style strategy. Without the cover and concealment of rough terrain, more time will have to elapse between operations as military units struggle to maintain the necessary level of secrecy in the open. Given the mismatch of strategy and terrain, paradoxically the war will be longer than when punishment can be applied almost continuously. Similarly, if a state uses a maneuver strategy in terrain that cannot support the quick transportation of soldiers and supplies, then it will have difficulty implementing its strategy, leading to a longer duration. In all cases, wars fought with punishment strategies should last longer than wars fought using movement strategies, but within the context of a single strategy choice, a mismatch of strategy with terrain will slow the war.

$H_{\text{Interaction}}$: *Wars fought with a strategy appropriate to the terrain will be shorter.*

Balance of Capabilities. When one side is stronger than its adversary, it will be able to inflict more damage on the adversary in a given period than the adversary is able to inflict on it, and the opponent may be quickly overwhelmed. This will lead to the weaker side surrendering sooner than it otherwise would, as it recognizes the probability of winning the ongoing conflict is small. When two armies are relatively equal in size, however, neither side is likely to collapse quickly, leading to a longer war.

$H_{\text{Balance of Capabilities}}$: *The greater the imbalance of the two sides, the quicker is the war.*

Total Military Capabilities. The absolute size of military forces also affects war duration. When countries have larger militaries, they can absorb more damage before they are forced to give up or otherwise seek to end the war. In addition, it simply takes more time for any military to fight through a large opposing army than through a small opposing army.

$H_{\text{Total Capabilities}}$: *The greater the total military forces involved in a war, the longer is the war.*

Population. When countries are involved in a war, they draw upon both material and human resources to fight the war. Larger countries have more human resources on which they can draw than do smaller countries. They can thus absorb more casualties and mobilize more people than can smaller countries before reaching their limit of fighting.

$H_{\text{Total Population}}$: *The greater the total population involved in a war, the longer is the war.*

As with military forces, it may be the case that an imbalance of population will lead to a quicker end to a war, as one side can draw upon more human resources than its adversary.

$H_{\text{Population Ratio}}$: *The greater the disparity in the size of the opposing countries' populations, the quicker is the war.*

Military Quality. When one state's military has better-trained troops and equipment of higher technology than that of the other, it can inflict relatively more damage on the adversary in a given amount of time than the adversary can inflict on it. A better-equipped military can also absorb more damage than a lower-quality military before collapsing. As a result, when the disparity between the adversaries' military quality is high, the war will end more quickly.

H_{Quality}: *The greater the difference in the two sides' military quality, the shorter is the war.*

Surprise. Military strategists often discuss the importance of surprise (Axelrod 1979; Clausewitz 1976; Dupuy 1979; Sun Tzu 1991). Strategic surprise occurs when a state is forced to change how it has planned to fight the overall war based on a sudden military move by its opponent. For example, the enemy could suddenly increase its military commitment or change strategy. Note that both sides can achieve strategic surprise at different times in a single war. In the Korean War, the North Koreans surprised the United States and South Korea with their initial attack, the United States achieved surprise with the Inchon landing, and the Chinese surprised the United States with their sudden entry into the war.

H_{Surprise}: *If a state achieves strategic surprise, then the war will be shorter.*

Mobilization. States' ability to mobilize their resources for war is perceived to be a key factor in their ability to prevail in a conflict (Trachtenberg 1991). Differences in mobilization rates have also been blamed as a cause of war, particularly in World War I (Snyder 1984a; Van Evera 1984). Yet, the argument that mobilization has a substantial influence on war outcomes or duration has, to our knowledge, not been broadly tested. While we cannot test propositions about preemptive motives resulting from anticipated short-term advantages at the outset of a war or preventive motives prior to a power transition, we can test the effect of longer term mobilization over the course of a war.

We argue that the mobilization of resources which varies during a war should not affect war duration for two reasons. First, most interstate wars are quite short. The mean war duration in our data set is approximately 15 months, the median less than that. Since most wars are won or lost in a short time, the year-to-year mobilization of capabilities during a war should be expected to have relatively little aggregate effect. Second, when a war does continue for more than a short period, states can adapt to their situation and to any mobilization problems that they have.

$H_{\text{Mobilization}}$: *Military mobilization during a war will have little effect on war duration.*

Issue Salience. Different wars are fought over different issues reflecting different potential benefits. The greater the benefits at stake in a conflict, the more costs rational leaders will be willing to accept in that conflict (Zimmerman 1973), and the longer state leaders will be able to convince domestic audiences to continue fighting. Realists in particular distinguish salient security and territorial issues from issues of trade or extended possessions.

H_{Salience}: *The more salient the issues at stake, the longer is the war.*

Domestic Political Approach: Regime Behavior and Type

Somewhat paradoxically, we believe that both highly repressive and highly democratic regimes will be associated with shorter wars. States at neither extreme will be involved in longer wars.

Repression. Leaders in highly repressive regimes may choose to fight in wars with objectively low chances for victory but potentially large gains to the repressive elite if the state pulls off an improbable victory. These leaders also know that they can remain in power by repressing the opposition, even if they lose a war. Leaders who cannot suppress the opposition as efficiently will be acutely aware of the postwar domestic costs of losing (Bueno de Mesquita and Siverson 1995). As a result, they will be less likely to become involved in highly risky wars. In most cases risky wars are mismatches. Consequently, they are likely to end more rapidly than wars in which the opponents are closely matched in capabilities, and so repressive states will be more likely to get involved in mismatched wars which end quickly.

$H_{\text{Repression}}$: *Highly repressive states will select risky wars which are likely to end quickly.*

Democracy. Two plausible arguments support the hypothesis that democratic states are more likely to be involved in shorter wars than are nondemocratic states. First, because citizens are freer to voice their opposition to government policy in democracies than in nondemocracies, opposition to a war will surface more quickly in democracies. Democratic leaders may be less able to repress this dissent than are leaders in autocratic regimes. Therefore, democratic countries should be involved in shorter wars than are nondemocracies as leaders respond to actual or anticipated opposition by ending their war involvement sooner than they might otherwise. Second, wars involving democracies will be shorter because of choices made prior to the outbreak of war. Because they face more certain sanctions following a flawed choice, leaders in highly democratic states are likely to avoid, whenever possible, wars in which they anticipate long duration and the associated high political costs. This aversion to risky wars may account for the observed difference in duration.

$H_{\text{Democracy}}$: *The more democratic the states involved, the shorter is the war.*

Empirically, then, we expect to see the same type of effect for both the political behavior of repression and the institutional constraint posed by democracy. The democracy variable measures state type, while the repression variable measures a form of state behavior. By choosing this particular internal political behavior, we believe that state leaders indirectly manipulate the observed duration of wars. Democratic states may be constrained in their ability to repress opposition to leaders' choices, but in theory there is no reason that both authoritarian and participatory regimes could not choose or threaten to violently repress political dissent during war.

Other Approaches

Several other explanations germane to war duration are discussed in various areas of the international conflict literature, but they do not clearly fit under the labels "realpolitik" or "regime behavior and type." Without capturing the entire scope of other approaches, we present three important arguments here.

Duration Dependence. Duration dependence refers to whether the future endurance of an event is related directly to its past length. Nearly all the previous work on war duration (e.g., Vuchinich and Teachman 1993) has found that the relative likelihood of a war ending at any particular time decreases the longer the war lasts, and it is argued that such negative duration dependence is a basic characteristic of wars. Three theoretical arguments support this position. First, large organizations (such as militaries) stockpile resources in anticipation of conflict, which reduces the costs of war to states. Second, large organizations have the ability (through propaganda) to affect the *perceived* benefits of victory so that mass publics continue to support the organization's actions. Finally, states adapt over time to one another's strategy. The longer a war lasts, the more time states have to adapt their strategies to allow them to continue fighting. All these factors help stretch out wars by reducing the acceleration in the costs of fighting.

In contrast to the existing literature, we do not believe there is a theoretical reason for wars to be duration dependent. Rather, such apparent dependence is a result of underspecification. We believe the independent variables in our model capture the arguments underlying war's apparent duration dependence. Our strategy variables directly capture the impact of strategic choice and adaptation. The effect of organizational stockpiling and propaganda is captured by our inclusion of variables measuring the size of the participants' population and military forces. With these controls, we expect no duration dependence.

$H_{\text{Duration Dependence}}$: *The length of war is not duration dependent.*

Previous Disputes. Literature on enduring rivalries has recently begun to emphasize that interstate conflicts are likely to be connected over time (e.g., Goertz and Diehl 1993). For example, it has been shown that past disputes between states have significant effects on conflict initiation (Huth, Bennett, and Gelpi 1992). Previous disputes also may have important effects on conflict duration. Repeated disputes between states are likely to raise the salience or stakes for the two sides. In addition, as they recur domestic constituencies are likely to form in opposition to the opponent (as was the case over time during the Cold War in the United States). The longer the history of conflict between states, the easier it will be for leaders to justify continued fighting to domestic audiences. Previous disputes thus may institutionalize conflict.

H_{Disputes}: *The more numerous the previous disputes between states, the longer is the war.*

Number of Actors. One emphasis of institutional approaches to international politics is that international coalitions (most commonly alliances, but also groups of states fighting together) must deal with collective action problems, which have implications for large war-fighting coalitions. As the number of participants in a war increases and the benefits of potential victory become distributed among more actors, the potential gains from victory for any individual state shrink. When many states are involved in fighting, some of them may receive no tangible benefits at all from victory. Because of smaller and possibly more abstract benefits, the leader of an individual state in a coalition thus will be less able to convince domestic audiences to continue fighting than if the state were fighting alone. We expect that the more states who fight in a war, the more quickly a coalition will fall apart due to collective action problems, leading to a shorter war. Note that our argument is the opposite of that made by Blainey (1988, 197) and Vasquez (1994, 258–60). Blainey argues that balanced capabilities, multiple fronts, and no danger of outside intervention make multistate wars last longer. Because we control for these variables, and because Blainey fails to consider the collective action problem of coordinating continued military action, we believe the opposite.

$H_{\text{Number of Actors}}$: *The more states involved in a war, the shorter is its duration.*

METHODOLOGY

To test our general theory and hypotheses of war duration, we apply a hazard model to the population of interstate wars which began between 1816 and 1985.

Hazard Analysis

The appropriate way to analyze the duration of events (such as wars) is to use a hazard model.[2] The use of OLS regression on data which take the form of some finite length of time is inappropriate. Since duration data are strictly positive, the assumption of normally distributed errors made by OLS regression is violated. Applying a linear model to duration data would introduce specification error, biased coefficients, and the possible prediction of negative values for expected war duration (Hanushek and Jackson 1977; King 1989).

Hazard models focus on the *hazard rate*, which is roughly the instantaneous rate at which events terminate at duration t, given that they have survived until time t. Technically, given observed duration times t, which are drawn from a random variable T, the hazard rate is the probability that an event will end in an interval of time $t + \Delta t$, as the interval goes to 0, given that the event has survived up to time t. The hazard rate is defined as

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

When applied to our data, a hazard model estimates the instantaneous rate at which transitions occur from a state of war to a state of peace as a function of a set of variables, including time. It is often intuitive to think of the hazard rate as the conditional probability of event termination at some time t. There are two important differences, however, between a hazard rate and a probability of event termination. First, since the hazard rate is instantaneous, the probability of event termination at any particular instant is infinitesimal (although it is possible to integrate the hazard rate over an interval, and so to estimate the probability that an event will end before a particular duration is reached). Second, the hazard rate is not bounded between 0 and 1, technically having no upper limit (Allison 1984, 23).

Hazard rates need not be constant over time, and one important use of hazard models is in identifying the duration dependence of a process. If the hazard rate in a distribution is increasing over time, and so an event is expected to end more quickly the longer it continues, the distribution is said to show positive duration dependence. If the hazard rate is decreasing over time, the distribution shows negative duration dependence. Negative duration dependence would occur if an event or situation tended to become institutionalized over time. Finally, if the hazard rate is independent of the time a situation has continued, there is no duration dependence.[3]

Several different parametric functional forms can be specified as the basis for a hazard model, with the exponential, Weibull, normal, log-normal, and gamma distributions representing only a few of many possibilities. There are few firm guidelines for selection of model form; as Kiefer (1988, 661) states, the family of duration distributions may be chosen "on the basis of a particular economic theory, convenience, and perhaps some preliminary plotting of data." For various reasons, we use a Weibull specification in our hazard analysis. The Weibull form allows for positive, negative, or no duration dependence; an exponential model assumes no duration dependence and would not allow us to test for it. The Weibull model permits the inclusion of independent variables and easy statistical estimation even when time-varying covariates (discussed below) are included, which the normal or log-normal model would not. It is also possible to add a specification for heterogeneity between cases; we test for this and find a Weibull model with gamma heterogeneity statistically indistinguishable in overall fit from the simpler Weibull. The distribution of survival times in our final data set is displayed in Figure 1.

Without any covariates, the basic functional form of the hazard rate $h(t)$, using a Weibull specification,[4] is

$$h(t) = \lambda p(\lambda t)^{p-1}.$$

The function $h(t)$ is the hazard rate at some time t. The parameter λ, on the right-hand side of the model, is a constant. The parameter p determines the shape of the hazard function and so accounts for duration dependence. When $p = 1$, there is no duration dependence, and the hazard rate $h(t)$ equals the constant rate λ. When $0 < p < 1$, the hazard rate decreases monotonically over time. When $p > 1$, the hazard rate increases monotonically.

Covariates (independent variables) can be added into the model as influences on the hazard rate by specifying

$$\lambda = e^{-\beta X i}.$$

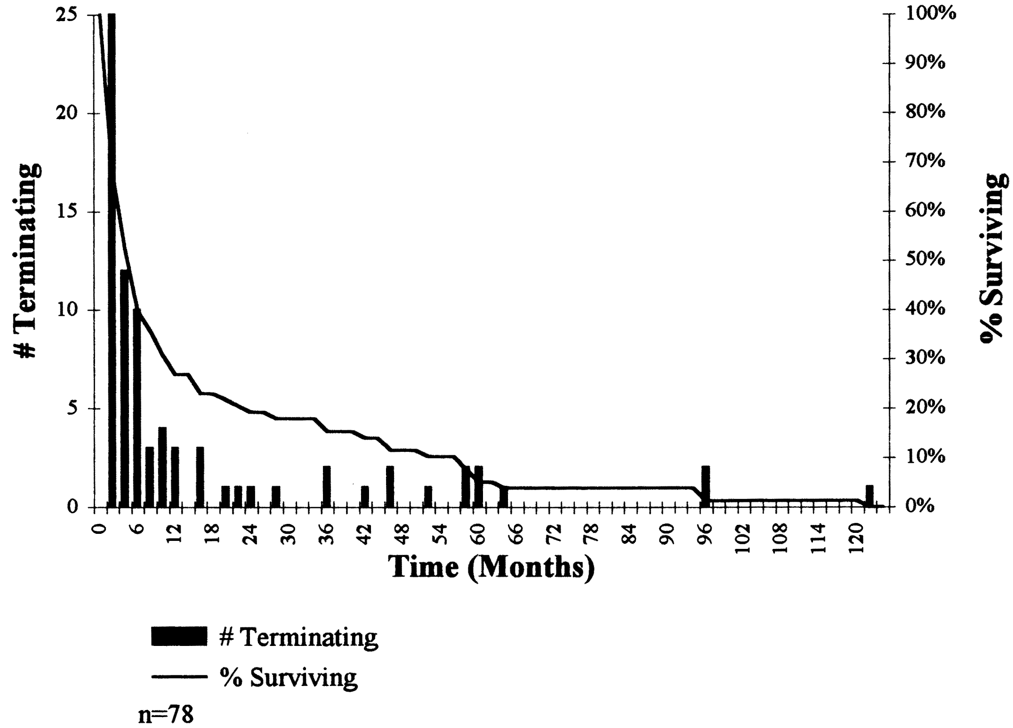
Direct interpretation of the final estimated β coefficients is difficult due to the presence of e and nonlinearity. The effect on duration of unit changes in the

[2] Hazard models are also referred to as survival models, duration models, or event history models. A basic introduction to such analysis is provided by Allison (1984). Comprehensive sources for duration analysis include Greene (1993), Kalbfleisch and Prentice (1980), Kiefer (1988), and Lancaster (1990). Recent applications of hazard analysis to political science include Bueno de Mesquita and Siverson (1995), Vuchinich and Teachman (1993) and Warwick (1992).

[3] Another useful feature of duration models is that they can be applied to data with censoring, that is, data for which the actual ends of some events are not observed. We do not take advantage of this, however, as all the wars in our population have ended; we would have done so if any wars included in our data set were still ongoing.

[4] It is important to note that hazard model notation varies significantly among sources, resulting in some confusion, as the same model may appear with different signs and parameter normalizations (Greene 1992, 725). Our presentation here primarily uses the notation of Greene (1993).

FIGURE 1. Number of Wars Ending and Percentage Surviving over Time, 1816–1985.



independent variables is not constant over the range of the independent variables, as they would be in an OLS model; the effect changes depending both upon the value of the independent variable in question and the value of the other independent variables. The expected survival time for any given case (set of values x_i), however, given a Weibull distribution, is $E[t|x_i] = \exp(\beta' x_i)*\Gamma(1/p+1)$.[5] The best way to interpret the magnitude of the coefficients is to use this formula to calculate and compare the values of hypothetical cases. For example, if we calculated the expected duration of a war to be 16 months given an attrition strategy, but only 12 months given a maneuver strategy, while holding other variables constant at their means, we would estimate the effect of the strategy change on duration as −4 months, holding the other variables at their mean values. At the simplest level, positive coefficients in β predict longer duration, and negative coefficients predict shorter durations.

Analyses to be Performed

Hazard models are typically estimated on data consisting of one observation per event, using variables measured at the beginning of each event. Variable values are then constant within each event. In our case, this approach would use a data set with one observation per war, with variables measured at the beginning of the war. We use this type of model as a means of making ex ante predictions about war duration. The results of this analysis will help us investigate how accurately decision makers could have estimated war duration, since we use only information that could have been observed by leaders when the war began. Because the values of the independent variables (covariates) do not change over time under this method, it is often referred to as a "time-invariant covariate" model.

While reasonable in many situations, a model using time-invariant covariates may ignore important information if variables change over time within cases. For example, if the balance of forces in a war changes dramatically at some point, affecting expected war duration from that point forward, that information is lost. This problem is sometimes handled by creating variables which measure average values across an event (e.g., the mean of the balance of forces across several years of war). Using average values might make better predictions (albeit ex post) than would the use of initial values, since the effects of changes are indirectly included. A better approach is available, however. Relatively recent work has extended the estimation of hazard models to allow inclusion of *time-varying covariates*, that is, variables which vary independently over the course of an event (Greene 1993; Lancaster 1990; Petersen 1986). Each event can be broken into discrete intervals that have different values on the independent variables; each interval contributes to the log-likelihood function computed for an observation. Since it uses all available information without losing any to averaging, this type of

[5] Greene (1993, 722) states that $E[t|x_i] = \exp(p\beta' x_i)$. This has been verified by Greene to be incorrect (via personal communication). The correct equation involves the Gamma function. The correct formulas for expected duration in Greene (1993) should be (p. 720) $E[t] = (1/\lambda)*\Gamma(1/p+1)$, and by substituting for λ (p. 722), $E[t|x_i] = \exp(\beta' x_i)*\Gamma(1/p+1)$.

model should provide a better fit to the data than a one-observation-per-event model using either ex ante or average variable values.

Current data sets in international relations contain variables with annually measured values. As a result, we can fit a statistical model incorporating time-varying covariates to our data. We subdivide wars which cross a year boundary into multiple observations and refer to the resulting data as the time-varying covariate (TVC) data set. This includes the same wars as the time-invariant (hereafter non-TVC) data but has additional information. For example, for World War I, we have one observation in the non-TVC data set, but in the TVC data set we have data for 1914, 1915, 1916, 1917, and 1918. The model with time-varying covariates is substantively important, because significant variables may well change during a war as alliances are formed and broken, additional capabilities are mobilized, opposition to the war is voiced or repressed, and individual countries opt to continue fighting or stop their participation in the war. We compare the results of the TVC model to those from the non-TVC model to evaluate our ability to perform ex ante versus ongoing prediction of war duration. We use the statistical routines found in LIMDEP (Greene 1992) to perform our estimations.

OPERATIONALIZATION

We created two data sets. For the TVC analysis, the data were coded for each year of each war. For the non-TVC analysis, only the values for the first year of each war were used. Below, we discuss the operational details of the variables in the data sets.

Population of Cases

Our population of cases for analysis is all interstate wars begun between 1816 and 1985. To identify these, we use the Correlates of War (COW) project's database of interstate wars (Small and Singer 1982, and updates) and the Dupuy and Dupuy (1986) *Encyclopedia of Military History*. We made two changes to the COW data set for theoretical reasons. We disaggregated World War II and the Vietnam War into multiple wars because we believe that each consisted of more than one distinct military conflict. We believe that decision makers rarely anticipate or think in terms of large systems of wars, but instead think in terms of sequences of opponents. We split single wars apart when Dupuy and Dupuy (1986) identified separate and distinct negotiated settlements between various participants or identified fundamentally independent war-fighting efforts between or among the participants.[6]

Consider the two modified cases. The COW data set treats World War II as one multilateral six-year war which included the United States, Germany, France, Poland, and others. Following Dupuy, we break the war down into a series of sequential smaller wars, treating Germany versus Poland in 1939 as one war, Germany versus France in 1940 as a separate war, and so forth. These wars were quite distinct militarily and in time, with each war being violent and brief as far as the military organizations were concerned. In 1940 France formally surrendered to Germany in World War II, thereby ending the war between France and Germany. To include France as a continuing participant in World War II until 1945 is a misstatement. Similarly, in Vietnam the United States and North Vietnam reached a negotiated settlement in 1973. Following that, the North Vietnamese routed the South in what we consider to be a separate war two years later.

The wars included in our analysis and their participants, dates, and durations are listed in Appendix A. For inclusion in the population, 87 wars initially met our criteria. Because of missing military expenditure data, however, our analysis (and Appendix A) includes 78 wars when we use annual data, and 77 when we measure variables only in the first year of war. The non-TVC analysis involves 77 data points for the 77 wars, while the TVC analysis involves 78 wars, which include 169 separate (annual) data points.

War Duration

The duration data come from the COW interstate war data set, updated with information found in Clodfelter (1992), Dupuy and Dupuy (1986), and Langer (1980).

Independent Variables

Strategy. We coded strategy as a set of dummy variables following Stam's (1996) procedures. We first coded which side sought to alter or preserve the status quo (determining offensive or defensive doctrine), based on Dupuy and Dupuy (1986) and Holsti (1991). We then coded each state's military strategy using Dupuy and Dupuy (1986), Dupuy (1983), and Clodfelter (1992). If the historical consensus was that a state used a blitzkrieg strategy or the actor encircled and divided its opponent's forces, then we coded those instances as maneuver.[7] An attrition strategy was coded when states fought meeting

[6] While we believe that this coding change is theoretically appropriate, we did check the sensitivity of our results to this change. Our results were quite similar even if we dropped the changed wars (dropped because we believe they are currently miscoded). Two differences emerged. First, since the Vietnam War is the only war in the population with the strategy OPDA (offensive punishment, defensive attrition), we are forced to drop that variable. Second, because this procedure involves dropping 14 wars (12 in World War II), *all* the standard errors increase slightly. The coefficient estimates were stable, however.

[7] Note that the coding of strategy as described in the historical record leads to a possible problem for our ex ante argument in that the coding is based on an ex post evaluation of the strategies used. In a few cases states initially tried to execute maneuver strategies but were forced to switch to attrition. An example is Germany's World War II strategy on the Eastern front, namely, a switch from maneuver to attrition as the blitzkrieg ran out of steam in the Soviet Union. Unfortunately, purely ex ante coding is impossible without access to military strategic plans. In the empirical record, however, there are relatively few instances of strategy shift among our categories, and so the problem affects only a few cases. When states did change strategy, they tended to do so soon after the beginning of a war, and so we believe that the magnitude of this problem is not severe. It does mean that our future predictions will have to be contingent on revealed strategy.

engagements against their opponents. Instances in which a state followed a Maoist guerrilla strategy or in which civilians were the principal military target were coded as punishment. If states used multiple strategies, then we coded the strategy which absorbed the majority of the state's military assets. If more than one country fought on a side in a war, then we coded the strategy of the largest state (in terms of capabilities).[8] In making the codings, our interpretation was as broad as possible. For example, a counterforce attrition strategy of searching out enemy submarines and a countervalue attrition strategy of searching out enemy shipping are both coded as attrition strategies.[9]

Attrition is the most common strategy, accounting for about two-thirds of all cases. Punishment strategies are the least common. With an offensive and defensive side each having three strategy choices, nine combinations are possible. The variables marking the possible strategies are:

0. OMDM: offensive maneuver, defensive maneuver;
1. OMDA: offensive maneuver, defensive attrition;
2. OADM: offensive attrition, defensive maneuver;
3. OADA: offensive attrition, defensive attrition;
4. OMDP: offensive maneuver, defensive punishment;
5. OPDM: offensive punishment, defensive maneuver;
6. OADP: offensive attrition, defensive punishment;
7. OPDA: offensive punishment, defensive attrition;
8. OPDP: offensive punishment, defensive punishment.

The 0 to 8 labels reflect the strategies' ordinal value in order of increasing expected duration, although we include a set of dummy variables in the analysis. Compared to attrition or maneuver, punishment strategies are always expected to make war longer. For the other orderings, wars in which the offensive state uses a faster strategy are expected to be shorter than wars in which the defending state uses a faster strategy because the distances the two sides must cover are not symmetrical. The offensive force typically must venture farther than the defender since the war is being fought into the defender's forces. When the offense is using a faster strategy, we expect the war to be shorter than if the attacker were using a slower strategy, which would take additional time to cover the necessary distances.

Historically, we observe the strategies 1, 2, 3, 6, and 7. Since these are dummy variables, we include four of them in the final analysis: OADM, OADA, OADP, and OPDA. OMDA was omitted, as we expect it to lead to the quickest wars.

Terrain. Terrain codes, which come from *The New York Times World Atlas* (1983), correspond to the location of the majority of the battles fought during the war. Terrain types were scaled to predicted movement times, using data from Dupuy (1979), which estimated movement speeds on various types of terrain.[10] The variable was coded between 0 (open or rolling terrain) and 1 (impassable terrain).[11]

Interaction of Terrain and Strategy. To measure the interaction of terrain and strategy, normally we would simply multiply our included strategy and terrain variables. With dummy variables marking strategy, however, the inclusion of each dummy interacted with terrain creates multicollinearity problems. As a result, we multiply a single ordinal-scaled strategy variable by terrain. This strategy variable is coded 0 through 8 in rank-order of increasing expected duration for the strategies OMDM through OPDP, as listed above. We expect a negative coefficient on the interaction.[12]

Balance of Capabilities. We use the COW composite capabilities index (Singer, Bremer, and Stuckey 1972) as an indicator of states' capabilities.[13] We define the balance of capabilities as the ratio of the larger side's total capabilities to the total capabilities of all participants. We discount states' capability scores according to the distance from a country to the point at which fighting with the adversary took place, using Bueno de Mesquita's procedures (1981, 105).[14]

Total Military Capabilities. We measure military capabilities as the sides' total military personnel, using the COW project's national capabilities database, coded as millions of soldiers.[15]

[8] In three cases there was a lack of consensus among military historians about the strategy choice (Germany in World War I, Germany versus the United States and Britain in World War II, and Israel in the Yom Kippur War). In these cases we coded the modal strategy of attrition.

[9] Gartner's *Strategic Assessment and War* (1992) contains an extended discussion of tactical and strategic change during wartime. The types of strategy change he addresses do not affect the broad codings we employ here.

[10] Because Dupuy coded the ease of tactical movement in different types of terrain based on a different data set of *battles*, it is not circular to use the variables as a predictor of the duration of *wars*. For detailed discussions about data codings on a battle by battle basis see Dupuy (1983).

[11] The highest observed value for the terrain variable was 0.75, marking very rough terrain, such as jungles or extremely mountainous regions. In cases with more than two actors on one side, and where the terrain on which different states fought might have been quite different, the coded value was an average of terrain scores weighted by the size of the forces fighting in particular terrain.

[12] Including the scale or individually multiplied dummy variables makes little difference in the overall results, as the model fit remains nearly identical.

[13] This index ranges potentially from 0 to 1 and indicates a country's share of total international military personnel, military expenditures, energy consumption, iron/steel production, total population, and urban population. When wars continued past 1985, we updated the capability data (the published data end in 1985) and supplemented it with interpolation of data and additional research using International Institute for Strategic Studies (IISS) and Stockholm International Peace Research Institute (SIPRI) data.

[14] When multiple countries fought together, we summed their discounted capabilities to obtain the total capabilities on each side of a war. When countries were involved for only part of a year, we adjusted their contribution to capabilities based on the portion of the year in which they were involved. For example, since Russia ended its involvement in World War I in early December 1917, we included 11/12 of Russia's capabilities in the 1917 Entente capabilities. When a country was involved in a two-front war, which we coded as separate wars, we discounted its total capabilities by the amount of forces allocated to each conflict. We coded a 50-50 allocation of resources between fronts with one exception: For U.S. participation in World War II, there is clear evidence that the United States fought primarily a holding action in the Pacific until the war in Europe was over (Dupuy and Dupuy 1986). We allocated U.S. capabilities in the proportions 75 and 25 to the European and Pacific wars, respectively.

[15] When countries were involved for only part of a year, we adjusted

Population. We measure the total population of the combatants using data from the COW National Capabilities data set, coded as billions of people.[16] We also code the ratio of the total population on the larger side to that of the smaller side to create the variable POPULATION RATIO.

Military Quality. We estimate the quality of military forces for a country as military expenditures (in constant U.S. dollars) divided by number of military personnel. Following Huth, Bennett, and Gelpi (1992) and Stam (1996), we argue that additional spending per soldier reflects additional troop training, equipment, and relative technology endowment.[17] We then create a ratio of the superior side's quality to the inferior side's quality.

Surprise. The surprise codings sought to capture strategic surprise at any time during the war. The coding is based on secondary historical sources as presented in Stam (1996). Since both sides can achieve surprise, we create a variable which attempts to capture not only the difference in surprise over the entire course of the war but also the seriousness of that surprise, given capabilities and how much of each side's forces were surprised. The variable ranges from 0 (no or symmetrical surprise) to 1 (large and asymmetrical surprise). Given states 1 through i on sides A and B, we code the variable as

$$\text{Net Surprise} = Abs\left(\frac{\sum_i Capability(A_i) * Surprise(A_i)}{\sum_i Capability(A_i)} - \frac{\sum_i Capability(B_i) * Surprise(B_i)}{\sum_i Capability(B_i)}\right).$$

Mobilization. We do not need to include a separate variable to test the argument regarding mobilization. Support for or against the hypothesis will be given by a comparison of the results of the analyses of the two different data sets. The non-TVC data use the values for capabilities measured in the first year of each war. Therefore, results using this data set will reflect no information about mobilization over the course of the war, only the forces available for the initial battles. The TVC model measures capability values annually throughout the war, and results using this data set will reflect the results of mobilization. On the one hand, if longer term mobilization is the key to determining war duration, we should observe a substantively and statistically significant coefficient estimate for the balance of forces variable in the second (TVC) analysis but not in the first. On the other hand, if initial forces are important, with later mobilization being a less important factor, similar coefficient estimates should be found for both models.

Issue Salience. We code the issues at stake in each war using Holsti's (1991) categorization. We follow arguments in Stam (1996) and use Zimmerman's (1973) conception of the pole of power and pole of indifference to code these issues as salient or nonsalient. Given an inability to make a more sophisticated differentiation, we set up a dummy variable for each side marking conflicts involving survival, territory, unification, reputation, and autonomy as salient. Conflicts involving policy, empire, and trade are coded as nonsalient.[18] We then create a final interval variable indicating that the issues were salient to neither side (0), salient to one side (1), or salient to both sides (2).

Repression. We use the Gurr, Jaggers, and Moore (1989) "competitiveness of participation" variable as our measure of the repressiveness of a state government (repression). For each state, Gurr's variable ranges from 1 (no significant opposition activity permitted) to 5 (significant and regular political competition in opposition to ruling leaders).[19] Given only five categories, this is a somewhat crude measure of state behavior. In its favor, however, are its availability and comparability both historically and cross-nationally, unlike other possible measures of repression, such as those based on deaths in political violence. We multiply the variable by -1, so that higher values represent more repressive regimes, and summed the repressiveness of each side to obtain the final variable.

Democracy. We use the Gurr, Jaggers, and Moore (1989) "institutionalized democracy" variable as our measure of democracy in a state government. For a single country, the variable is a scale ranging from 0 (few constraints on the chief executive, closed executive recruitment procedures, and noncompetitive political participation) to 10 (many constraints, open recruitment, active opposition).[20] We then summed the democracy value of each side to obtain the final variable. Democracy is negatively correlated (Pearson's $R = -0.64$) with repression, but not perfectly so. This level of correlation raises some concern that the results observed may be

their contribution to total personnel accordingly and then summed the personnel of the two sides to obtain total military personnel of all combatants.

[16] When countries were involved for only part of a year, we adjusted their contribution to total population accordingly.

[17] Data on military expenditures and personnel were obtained from the COW National Capabilities data set. Exchange rates and an inflation deflator were obtained from Gurr, Jaggers, and Moore (1989). When there were multiple countries on a side, we created a weighted average of quality on each side whereby a state's contribution to its side's overall quality was proportional to its capabilities.

[18] When there was more than one country on a side, we coded a side's salience as the salience of the war to the largest state (in terms of capabilities).

[19] When there were multiple countries on a side, we again weighted each state's contribution to the overall value by its capabilities. There were significant problems of missing data; when data were lacking for a country, we substituted the mean of the variable across the rest of the data set, or 2.36.

[20] We weighted each state's contribution to the overall value by its capabilities when there were multiple countries on a side. There was significant missing data on the variable; to allow estimation, when data were missing for a country we substituted the mean of the variable across the rest of the data set, or 2.61.

TABLE 1. Hazard Model Coefficient Estimates, Effects on War Duration

Variable	Model 1 VT model, TVC	Model 2 Realpolitik, TVC	Model 3 Regime, TVC	Model 4 Complete, TVC	Model 5 Complete, non-TVC
Constant	2.480 (0.678)	2.470 (1.470)	1.750 (0.598)	2.260 (1.540)	1.260 (1.500)
Realpolitik					
Strategy: OADM	—	2.480 (0.978)***	—	2.750 (0.940)***	2.870 (0.947)***
Strategy: OADA	—	3.250 (0.938)***	—	3.200 (0.875)***	3.330 (0.801)***
Strategy: OADP	—	7.010 (2.000)***	—	6.270 (1.870)***	6.280 (2.050)***
Strategy: OPDA	—	11.600 (3.850)***	—	10.970 (3.580)***	7.990 (1.706)
Terrain	—	6.600 (4.120)*	—	5.020 (3.750)*	2.990 (3.650)
Terrain × Strategy	—	−2.020 (1.050)**	—	−1.690 (0.976)**	−1.240 (0.995)*
Balance of Forces	—	−5.030 (1.360)***	—	−4.760 (1.300)***	−3.980 (1.280)***
Total Military Personnel	—	0.061 (0.041)*	—	0.124 (0.047)***	0.273 (0.088)***
Total Population	—	0.753 (0.567)*	—	0.707 (0.586)	0.162 (0.716)
Population Ratio	−0.024 (0.021)	0.001 (0.019)	—	0.007 (0.017)	0.007 (0.022)
Quality Ratio	—	0.013 (0.013)	—	0.010 (0.013)	0.007 (0.012)
Surprise	—	−0.012 (0.746)	—	−0.204 (0.807)	−0.219 (0.636)
Salience	—	0.336 (0.327)	—	0.420 (0.380)	0.427 (0.399)
Regime					
Repression	—	—	−0.281 (0.160)**	−0.200 (0.144)*	−0.246 (0.168)*
Democracy	—	—	−0.130 (0.070)**	−0.100 (0.063)*	−0.118 (0.066)**
Other Approaches					
Previous Disputes	—	—	—	−0.008 (0.070)	0.016 (0.069)
Number of States	0.064 (0.111)	—	—	−0.189 (0.095)**	−0.135 (0.098)*
Year	−0.001 (0.005)	—	—	—	—
p (duration param.)	0.629 (0.060)	0.907 (0.088)	0.629 (0.061)	0.965 (0.108)	0.942 (0.122)
Log-likelihood	−273.6	−244.7	−273.1	−242.1	−126.1
Mean error (months)	17.0	15.8	17.5	13.0	12.2
SD of mean error	18.2	23.7	18.0	20.5	16.2
Median error	12.8	6.2	11.1	5.1	5.0
Mean error (as % of war length)	22.8%	5.6%	17.4%	4.9%	4.4%
Number of wars	78.0	78.0	78.0	78.0	77.0
No. of data points (war-years)	169.0	169.0	169.0	169.0	77.0

Notes: Standard errors are in parentheses. Significance tests are one-tailed. The constant-only model on the one-observation-per-year data set estimates $p = 0.629$, log-likelihood −274.8, mean error 17.3 months (25.9%), and median error 13 months. The constant-only model on the one-observation-per-war data set estimates $p = 0.618$, log-likelihood −176.7, mean error 18.3 months (27.7%), and median error 15.1 months. Year is included in model 1 because of its inclusion in Vuchinich and Teachman (1993).
$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

due purely to multicolinearity or correlation of measurement error. Tests below show that this is not the case.

Previous Disputes. We count the average number of reciprocated disputes in the ten years before each war between all pairs of states on the opposing sides of each war as given by the COW project's data set of militarized interstate disputes. We count only disputes that lasted a minimum of 30 days to avoid including minor conflicts (such as "fishing boat incidents") which would inflate the count. We then normalize by dividing this count by the number of states in each war (so that large wars would not automatically have more disputes).

Number of Actors. We count how many states were involved in the war in each year as given by the COW project's data set of interstate wars.

Duration Dependence. The presence of duration dependence is given by the p parameter, estimated as part of the hazard analysis.

RESULTS

Table 1 presents the results of five hazard analyses. Models 1 through 4 are estimated on the time-varying covariates data set, while model 5 is estimated on the non-TVC data set. Model 1 includes only the Vuchinich and Teachman (1993) variables estimated on our data set. It is important to estimate this model because our population of wars is somewhat different from that of Vuchinich and Teachman.[21] Without a comparison of the data sets, any differences in our results could be attributed solely to differences in the population. In fact, the results obtain with our population are very similar to those of Vuchinich and Teachman. In particular, the p

[21] Vuchinich and Teachman (1993) analyze 118 interstate and extrasystemic wars. We limit our analysis to interstate wars because we believe that including both types may lead to two problems. First, data for extrasystemic wars are much more uncertain than for interstate wars. Second, the process by which each type is fought may be quite different.

parameter is significantly less than one, indicating negative duration dependence when other variables are omitted. We believe that this suggests the comparability of the results we obtain in other models.

Models 2 and 3 use the realpolitik and regime variable subsets, respectively.[22] These allow us to compare the relative explanatory power of the realpolitik approach to the regime behavior and type approach as well as build up to our complete theory. Models 4 and 5 estimate our full theory on the TVC and non-TVC data sets, respectively.

Overall Fit and Model Predictions

An examination of log-likelihood values shows that the more completely specified model presented here is better than either the Vuchinich and Teachman (1993) model (model 1)[23] or a "naive" constant-only model ($p = 0.01$). The results of a series of specification checks are contained in Appendix B. In particular, the realpolitik model adds substantial explanatory power when predicting war duration, even after taking into account the number of variables it contains ($p = 0.01$). Although the individual variables in the regime behavior and type model, as well as number of states from the "other approaches" subset, are statistically significant at the 0.05 level (or, as in the case of democracy, just miss the standard level of statistical significance, $p = 0.06$), the improvement in overall model fit in moving from model 2 to model 4 does not meet standard conventions of statistical significance ($p = 0.20$). The same holds true for adding just the two regime variables to the realpolitik model (log-likelihood for the realpolitik plus regime model is -243.7). Thus, using likelihood ratio tests, we cannot distinguish between the simpler realpolitik model and the more complex and completely specified models 4 and 5.[24] While the realpolitik model is a substantial improvement over simpler approaches, adding variables for regime behavior, regime type, and the number of actors does not significantly improve the overall fit. We believe this suggests that, while the regime and other factors may have some effect on war duration, the influence is likely to be small and remains somewhat uncertain, relative to realpolitik factors. This finding supports the position of realists who argue that, once war begins, regime type and willingness to repress mass publics are less important in determining state behavior than are military considerations.

While log-likelihood values provide a relative measure of fit between nested models, maximum likelihood models generally do not have an overall measure of fit analogous to the R^2 statistic in an OLS model. As a result, we give three auxiliary indicators of prediction error in an attempt to provide a feel about model fit in a "real-world" sense.[25] We take the absolute value of the difference between the actual duration of each war and the models' prediction of duration. We then calculate the mean absolute error, median absolute error, and mean of the absolute error as a percentage of each war's length (the reason for this last indicator is discussed below). We present multiple measures because any single indicator of overall fit is likely to be imperfect (if there were a good indicator, then it would be in widespread use), especially given that the distribution of war durations is skewed. Since mean war duration is longer than median war duration, a model which predicts durations close to the median will perform well on measures of the median error but poorly on measures of mean error, while the opposite is true of models whose predictions are close to the mean.

The complete TVC model (model 4) predicts with a mean absolute error of 13 months. Our errors are quite skewed, however, as are the duration data. The median error produced by the TVC model is only 5.1 months. For comparison, a naive "constant-only" model has a four-month greater mean prediction error than model 4, and an eight-month greater median prediction error. It is important to place neither too little nor too much emphasis on this prediction comparison. A naive prediction could perform quite well given a skewed data set. In this case, since most wars are short, a prediction of the median length (five months) would predict within five months of the true duration for all wars shorter than ten months. Despite this "accuracy," such a prediction is theoretically uninteresting, as it provides no way to make differentiated predictions. Such a naive model will *always* miss predicting long wars, which are the most interesting since they are typically the least expected.

Our model *can* (and does) predict long wars. This is reflected by the third error measure, the mean of the absolute error as a percentage of each war's length. A model which accurately predicts only short wars while making large errors in predicting long ones will have a higher value on this measure than a model which can more accurately predict both types. Our theoretical models do much better on this measure (4.9% and 4.4%, respectively) than do the less complete model 1 and a naive constant-only model (22.8% and 25.9%, respectively). Finally, while the log-likelihood values of model 4 versus model 2 did not suggest a statistically significant

[22] We do not present estimates for a separate "other approaches" model. It offers no overall statistical improvement over the constant-only model, and neither previous disputes nor number of states are statistically significant in the separate analysis.

[23] The log-likelihood value for model 4 is not directly comparable to model 1 as we present it because the models are not nested. We omit the variable "year" in our models because we know of no theoretical reason for its inclusion. When a truly nested model is estimated, model 4 is still significantly better than model 1. The same is true if the Vuchinich and Teachman model is estimated on the non-TVC data set and compared to model 5.

[24] Beck and Katz (1992) argue that too much reliance on likelihood ratio tests may be misleading. They maintain that, in addition to the standard tests, prediction errors also should be used for model assessment and choice. Because of this, we do not reject the hypothesis that domestic variables are an important factor in war duration, even though the likelihood ratio test supports the inclusion of the variables only at the 0.20 level. Below, we discuss the changes in prediction accuracy as Beck and Katz suggest and find that democracy and political repression do improve our estimates of expected duration.

[25] One problem of maximum likelihood models generally is the absence of a measure of fit analogous to R^2. If one were to estimate a technically inappropriate OLS using the model 5 variables and data set with duration as the dependent variable, then the R^2 would be 0.56, and the mean error would be 15.7 months.

improvement in model fit, it is also useful to note that the prediction errors of model 4 are smaller (4.9%) than those from the realpolitik model 2 (5.6%).

Perhaps even more important than the overall predictions are the inferences possible by comparing the more completely specified models to the less theoretical models, especially regarding duration dependence. Once realpolitik variables are included in models 2, 4, and 5, the apparent duration dependence of war disappears. The simplest Vuchinich and Teachman model (as well as the constant-only model) suggests with high confidence that negative duration dependence exists, as did prior work on war duration. The controls introduced in the regime model (model 3) do not change this conclusion. Model 2, however, and especially models 4 and 5 suggest that the p parameter is not clearly distinguishable from one. Better model specification leads to a conclusion regarding duration dependence that is quite different from previous work, namely, that wars are not duration dependent.

The prediction errors of models 4 and 5 can be compared to assess our ability to make ex ante predictions. The ex ante model 5 predicts just as well as model 4. This suggests that accurate prediction of war duration is quite possible even with primarily ex ante data. One caveat is in order here. Our strategy codings are the result of historical observation of how states actually fought and are not based on information known before the wars occurred. As previously noted, this is because states tend to be quite secretive about their strategies and tactics before war. The effect of strategy is thus inferred from ex post information gathered after the wars had occurred. Forecasts based on our results must be contingent in that they must rely upon assumptions about what strategy a side will use during some hypothesized war.

Prediction error is important to testing our hypothesis about mobilization. If the results of the non-TVC model 5 and the TVC model 4 were quite different in their ability to make predictions, and if the balance of forces variable was much weaker in model 5, then it would suggest that annual changes in mobilization during a war play a key role in determining war duration. That variable is statistically significant and of a similar substantive magnitude in the two analyses, however, and the models' predictions are similar. In addition, although total military personnel has a large and significant effect, total population (representing mobilizable resources) does not. We believe this suggests that while mobilization at the outset of a war may affect its duration, mobilization and the potential for it after the initial battles do not play the overwhelming role that some might attribute to it. This provides support for Lebow's (1984) argument that windows of opportunity are not as tempting to states as many fear and for Reiter's (1995) argument that preemptive attacks to gain short-run advantages before the opponent can mobilize reservists or the general population are unnecessary.

We performed one additional test to evaluate our predictive ability. We predicted the duration of the Persian Gulf War, which is not in our population of wars. By our criteria, there were two separate wars in the Persian Gulf, the first when Iraq occupied Kuwait, the second when the United States attacked Iraq in January 1991. The duration of the first was one day (0.03 months); the strategy was OADA. The duration of the second can be coded in two ways. The eventual ground war, in which the U.S. mobile strategy was of the OMDA type, lasted four days (0.13 months). If the war is also coded to include the bombing of Iraqi military targets starting on January 16, an OADA strategy, the duration was 43 days (1.39 months). Using the expected duration formula given earlier, and using the model 4 coefficients, our model predicts a duration of 6.9 months for the Iraq/Kuwait war, 0.70 months for the U.S./Iraq war with a U.S. OMDA strategy, and 17 months for the U.S./Iraq war with an OADA strategy. We believe this compares favorably with regard to what actually happened and to other models, many of which predicted a longer duration for the U.S.-Iraqi conflict (Record 1993, 58–59).[26]

Individual Effects

Important individual elements of both models had substantive and statistically significant effects as anticipated. A simple observation of the coefficients, however, is not adequate to compare the influence of different variables. Because the estimated models are nonlinear, the coefficients have no direct linear interpretation, and the marginal effects of variables differ depending on the values of other variables. The only way to gauge the relative effect of different variables is to estimate and compare expected durations given different variable values. This is done by estimating the expected duration of a hypothetical case (a war with a given set of parameters) as $E[t|x_i] = \exp(\beta' x_i)^*\Gamma(1/p+1)$, as discussed earlier. We have done this in Table 2, which presents the effects of changes in the variables with other variables held constant at typical observed values. Table 2 uses the coefficient estimates from the ex ante model 5.

Balance of forces appears to be a key variable in determining the duration of wars. The more out of balance the opponents' capabilities, the faster the war progresses. Note in Table 2 that the most dramatic changes in war duration occur when the force ratio moves just away from parity. This finding supports those, such as Organski and Kugler (1980), who argue that small shifts in the balance of power when the states are closely matched can have a major impact on international politics. Once capabilities have moved from parity, additional increments bring declining marginal returns in terms of shortening a war's duration.

Strategy appears to have a major and consistent influence on the duration of war, and all our expectations regarding the effect of different strategies are met.

[26] For those interested in replicating this exercise, the values of the variables are (represented as Iraq-Kuwait value/Iraq-U.S. air-plus-ground value/Iraq-U.S. ground-only value): strategy = OADA/OADA/OMDA; terrain = 0/0/0; terrain × strategy = 0/0/0; balance of forces = .945/.823/.823; total military personnel = 1.02/3.12/3.12; total population = .021/.268/.268; population ratio = 9.37/13.04/13.04; quality ratio = 5.70/10.33/10.33; surprise = 1/1/1; salience = 2/2/2; repression = 3/−7/−7; democracy = 1/11/11; previous disputes = 2/2/2, number of states = 2/2/2.

TABLE 2. Marginal Effect of Variables from One-Observation-per-War Model on War Duration

Variable	Change in Independent Variable: From	To	Change in War Duration (months)
(Base Case)			(11.2 months)
Balance of Forces	1:1	2:1	−18.4
	2:1	3:1	−5.5
	3:1	4:1	−2.5
Strategy	OMDA	OADM	+12.4
	OADM	OADA	+0.4
	OADA	OADP	+155.9
	OADP	OPDA	+439.1*
Terrain, given strategy OMDA	0.1	0.5	+1.5
given strategy OADM	0.1	0.5	+9.9
given strategy OADA	0.1	0.5	+2.9
given strategy OADP	0.1	0.5	−54.7
given strategy OPDA	0.1	0.5	−548.7*
Total Military Personnel	1.0 million	2.4 million (mean)	+10.4
	2.4 million	5.0 million	+34.0
Democracy	0.0 (least democratic)	6.1 (mean)	−28.9
	6.1	17.0 (max. observed)	−19.9
Repression	−9.0 (least repressive obsvd.)	−5.0 (mean)	−45.4
	−5.0	−1.0 (max. observed)	−17.0
Number of States	2.0	3.0	−3.8
	3.0	4.0	−3.3
	4.0	8.0	−9.6

Note: The changes in expected war duration are calculated from the estimated coefficients in model 5 by setting all other variables to values of a base case consisting of typical (mean or median) values on other variables. In particular, the base case has strategy = OADA, balance of forces = 0.75, terrain = 0.34, number of states = 3, total population = 165 million, total military personnel = 2.4 million, population ratio = 2:1, quality ratio = 2:1, salience = 0, democracy = 6.3, repression = 5.1, and surprise = 0.5, with expected duration 13.5 months. Expected duration is then estimated as $E[t|x_i] = \exp(\beta' x_i)^*\Gamma(1/p+1)$ (corrected version of Greene 1993).
*Problematic estimate. See the text for discussion.

The increasing magnitudes of the strategy coefficients are reflected in the increasing expected length of war shown in Table 2. The most dramatic change occurs when attacking states move from maneuver to attrition strategies, or when either state shifts to a punishment strategy. These findings cast some doubt on the assertion of the "Cult of the Offensive" scholars, who argue that militaries choose mobile strategies for organizational reasons. Our results imply that a mobile strategy choice is rational and understandable for either military organizations or civilian leaders as a way of shortening a war. Despite the advantages of such a choice, however, states may not choose a mobile strategy because of constrained internal politics, and so we observe the suboptimal attrition and punishment strategies (many other nonrealpolitik factors affect strategic choice as well; see, for example, Kier [1995], Rosen [1995], and Stam [1995, 1996]).

An important caveat is in order regarding our results on punishment strategies. We observe only one war in our data set (the Vietnam War) in which the strategy combination is OPDA, and only a few cases of OADP. This results in difficulty with obtaining precise point estimates for these strategy variables. In model 5's non-TVC analysis, OPDA has so little variation and is therefore so colinear with other variables that a huge standard error is obtained. With the additional observations (and variation within the Vietnam War) included in model 4, a more reliable point estimate is obtained, but we do not feel confident in making a strong claim about the degree to which the OPDA coefficient is greater than the OADP coefficient. Given the uncertain estimates, prediction of the marginal effects of OADP and OPDA strategies is also difficult. The predicted jumps from OADA to OADP and OPDA may well be too high.[27] Unfortunately, given the historically observed distribution of strategies, we do not know of a better way to obtain more precise estimates of punishment strategies' effect on war duration.

The terrain on which a war is fought and its interaction with strategy also affect duration. Fighting on rougher terrain results in longer wars when the observed strategy combination is OMDA, OADM, or OADA, and there is a "turnaround" with slow punishment strategies, as expected. As with strategy, colinearity appears to be leading to a statistically "insignificant" coefficient estimate for terrain in model 5. Yet, if terrain is dropped from the analysis, the overall fit as measured by log-likelihood decreases quite a bit, and so we believe this variable indeed has an effect on war duration.[28]

[27] The overall predicted length of the Vietnam War is not off by very much because other variables work in the direction of shortening the anticipated length of that war.

[28] If other apparently nonsignificant variables are dropped from the analysis, then terrain as well as some of the variables significant at the 0.1 level become significant at the 0.05 level. In the absence of other research, however, dropping other variables raises the possibility of introducing misspecification by choosing our model to fit the data. As a result, we have chosen to report all our included variables, even though this weakens the apparent significance of some of them.

The effects of democracy and repression are quite interesting, as anticipated. We find that both are associated with shorter wars.[29] Holding all other variables at their respective means, we expect war between highly repressive regimes to be 15 months shorter on average than a war involving moderately repressive regimes, while we expect a 21-month difference between wars involving moderately democratic states and wars between states without democratic institutions. Yet, two possible theoretical explanations remain as to why democracy and repression are both associated with shorter wars, as advanced in our selection and public-constraint arguments above.[30] Some additional evidence that helps sort out this argument can be obtained by examining the average duration of the wars initiated by highly democratic or repressive states (selected) versus the wars they did not initiate (the hazard model does not differentiate between who starts the war).

For wars in which highly democratic states participated, the average length of those they initiated was 9.2 months, while the average of those they did not initiate was 16.7 months. The comparable figures for highly repressive (risk acceptant) states were 11.8 months and 18 months, respectively. Grouping the two extreme types of states together, the average duration of wars started by either type was 11.2 months. Wars initiated by states neither highly democratic nor repressive lasted an average of 22.4 months, which suggests that the selection argument has strong merit (p = 0.02 level).[31] We can examine the auxiliary question of whether risk aversion and risk acceptance are driving these initiation decisions by looking at the wars' outcomes. If the risk argument holds, then we would expect risk-averse, highly democratic states to select easy, short wars, which frequently they should win. Highly risk-acceptant, repressive states would be expected to initiate short wars, but they would be less likely to win them. These expectations are strongly supported by the data. Ten of the 12 wars (83%) initiated by highly democratic states were won by them. Highly repressive states won only 20 of the 40 wars (50%) they initiated (the chi-squared test of independence statistic is 6.9, $p < 0.001$). While these results are not conclusive, they point strongly in the direction of the selection bias arguments of Bueno de Mesquita and Siverson (1995).

The empirical effects of several other variables also matched our expectations. Total military personnel appears to have a major positive effect on war duration. The salience variable does not meet standard levels of statistical significance in both models 4 and 5 (p = 0.13, but it was consistently positive through a variety of specification tests). Once controls for other factors are introduced, the number of states involved appears to have a negative effect on war duration; note, however, that the substantive effect is small relative to other factors (Table 2). Finally, as previously stated, we find no evidence of duration dependence in the wars we analyzed once we include the variables in the realpolitik model. In combination with our strong findings on other variables, we believe that the apparent duration dependence of war can be accounted for by measuring underlying factors.

There are a few surprising results in the analysis. The estimated effects of total population, the ratio of states' populations, and the ratio of force quality were near zero. Previous disputes also show a coefficient near zero. When these variables are dropped from the analysis, the log-likelihood function changes very little, suggesting that they are not adding much information to the model. We hypothesize (post hoc) that the dispute finding is due to our controlling for issue salience (salience is one key element of the dispute argument; the two variables are positively correlated at the $p < 0.05$ level). Similarly, we hypothesize post hoc that the effect of total population is captured by the total military personnel variable, and the effects of population ratio and quality ratio are captured by the balance of forces ratio, which incorporates these components. Finally, in retrospect, we also can make a strong theoretical argument based upon our mobilization finding that total population is not a key variable. The population must be mobilized somehow (whether by induction into the military or by relocation to areas where they produce war materials) in order to have an effect on war duration. Given that mobilization does not appear to have a dramatic effect on war duration, we should expect that only the actual military variables, and not the raw component of population size, will affect a war's length.

Finally, in contrast to our expectations derived from consideration of military strategy, we find that surprise plays little role in determining the duration of war. While the sign on the estimate was negative (as predicted), the standard deviation was relatively large, and the true effect of the variable is quite uncertain.

[29] We cannot be quite as confident of these results as we are of several others. The probability that democracy does not lead to shorter wars is 0.055 in the TVC model, and the probability that repression does not lead to shorter wars is 0.08 and 0.07 in the TVC and non-TVC models, respectively. These estimates miss the 0.05 cutoff that many analysts use as a benchmark of significance, and the regime behavior and type model does not substantially improve the overall log-likelihood (p = 0.20). The 0.05 cutoff represents an elementary rule of thumb. While recognizing that we have exceeded it, we believe that repression and democracy are likely to have some effect on duration, both because the magnitude of the variables remained relatively unchanged through a variety of other specification tests, and because the mean expected effect of the variables on war duration appears substantial (Table 2). It is important to note, however, that because of the wider confidence interval around democracy (and number of states, discussed below), the effect on duration could be somewhat farther off for these variables—either larger or smaller—than for others.

[30] Additional analysis showed that our results were not simply the result of colinearity; if either repression or democracy is dropped, then the magnitude of the remaining regime variable remains similar. Because the institution of democracy and behavior of repression are negatively correlated, an alternative explanation could be that there is a curvilinear relationship between a more general "regime type" variable and war duration, with extremely liberal states (strong institutional constraints on elite behavior) and extremely authoritarian states both fighting short wars, while "moderate" states fight longer wars. We checked for this possibility by estimating a model without repression (behavior) and substituting a democracy-squared term. While the democracy term remains approximately the same as in the other models, the squared term has no substantive or statistical effect.

[31] We define as "highly democratic" those states coded by Gurr as scoring 8 or more on the democracy scale. We define as "highly repressive" those states scoring 1 on Gurr's PARCOMP index of repression.

CONCLUSIONS

From our detailed analysis of the duration of interstate war, we conclude that a number of variables not usually considered in empirical work by political scientists, especially various strategies and terrain, have important effects. Since the anticipated duration of war affects leaders' decisions to initiate wars, we believe that models of war initiation and termination should consider incorporating these factors. While the realpolitik model clearly has the most explanatory power, it appears that a number of other variables—particularly democracy, repression, and the number of actors involved—may affect war duration.

We also found that some of the variables often assumed to influence the length of war played a minor role. Neither long-term mobilization nor surprise had a major effect. War also does not appear to be duration dependent. This finding differs from previous studies and is attributable to our inclusion of several independent variables key to the duration dependence argument. As a result, we can suggest that ongoing wars are not subject to continuation forever, as each additional increment of time spent at war does *not* further perpetuate conflict. Rather, identifiable factors suggest which wars are likely to continue.

Finally, we conclude that an accurate model of the duration of interstate war is possible, even using ex ante data and a small number of contingent variables specified at the outset. We predicted war duration with a mean error of thirteen months and a median error of five months. Our ex ante model made predictions that were within one month of our alternative model, which incorporates ex post information. Clearly, additions and refinements need to be made, in particular, those relating to the precise effect of punishment strategies and terrain interactions, but we have taken a major first step toward understanding how long and why wars will last.

APPENDIX A

TABLE A-1. Interstate Wars, 1816–1985

War Name	Actors	Start Year	End Year	Length (Months)
Franco-Spanish	France vs. Spain	1823	1823	4
Mexican-American	United States vs. Mexico	1846	1847	22
Austro-Sardinian	Austria-Hungary vs. Italy/Sardinia, Modena, Tuscany	1848	1849	16
1st Schleswig-Holstein	Germany/Prussia vs. Denmark	1848	1848	6
Roman Republic	France, Austria-Hungary vs. Papal States, Two Sicilies	1849	1849	2
La Plata	Brazil vs. Argentina	1851	1852	12
Crimean	United Kingdom, France, Italy/Sardinia, Russia vs. Turkey/Ottoman Empire	1854	1856	28
Anglo-Persian	United Kingdom vs. Iran/Persia	1856	1857	6
Italian Unification	France, Italy/Sardinia vs. Austria Hungary	1859	1859	5
Italo-Roman	Italy vs. Papal States	1860	1860	10
Italo-Sicilian	Italy vs. Two Sicilies	1860	1860	2
Franco-Mexican	Mexico vs. France	1862	1867	58
2nd Schleswig-Holstein	Germany/Prussia, Austria-Hungary vs. Denmark	1864	1864	6
Lopez	Brazil, Paraguay vs. Argentina	1864	1869	63
Spanish-Chilean	Peru, Chile vs. Spain	1866	1866	6
Seven Weeks	Germany/Prussia, Mecklenburg, Italy vs. Austria-Hungary, Hanover, Bavaria, Baden, Saxony, Wurtemburg, Hesse-Electoral, Hesse-Ducal	1866	1866	1
Franco-Prussian	France vs. Bavaria, Germany/Prussia, Baden, Wurtemburg	1870	1871	10
Russo-Turkish	Russia vs. Turkey/Ottoman Empire	1877	1877	9
Pacific	Peru, Bolivia vs. Chile	1879	1883	58
Central American	Guatemala vs. El Salvador	1885	1885	4
Serbo Bulgarian	Serbia vs. Bulgaria	1885	1885	3
Sino-Japanese	China vs. Japan	1894	1895	9
Greco-Turkish	Greece vs. Turkey/Ottoman Empire	1897	1897	5
Spanish-American	United States vs. Spain	1898	1898	4
Boxer Rebellion	United States, United Kingdom, France, Russia, Japan vs. China	1900	1901	15
Russo-Japanese	Russia vs. Japan	1904	1905	16
Central American	Guatemala vs. Honduras, El Salvador	1906	1906	3
Central American	Honduras vs. El Salvador, Nicaragua	1907	1907	11
Italo-Turkish	Italy vs. Turkey/Ottoman Empire	1911	1912	12
First Balkan	Serbia, Bulgaria, Greece vs. Turkey/Ottoman Empire	1912	1913	7
Second Balkan	Serbia, Greece, Turkey/Ottoman Empire, Rumania vs. Bulgaria	1913	1913	2
World War I	Serbia, Russia, France, Belgium, United Kingdom, Japan, Italy, Portugal, Rumania, United States, Greece vs. Austria-Hungary, Germany, Turkey/Ottoman Empire, Bulgaria	1914	1918	52

TABLE A-1. Continued

War Name	Actors	Start Year	End Year	Length (Months)
Russo-Polish	Poland vs. USSR	1920	1920	6
Hungarian-Allies	Hungary vs. Czechoslovakia, Rumania	1919	1919	5
Greco-Turkish	Greece vs. Turkey	1919	1922	41
Sino-Soviet	USSR vs. China	1929	1929	4
Manchurian	China vs. Japan	1931	1933	19
Chaco	Bolivia vs. Paraguay	1932	1935	36
Sino-Japanese	China vs. Japan	1937	1945	96
Changkufeng	USSR vs. Japan	1938	1938	1
German-Czech	Germany vs. Czechoslovakia	1938	1938	0.033
German-Austrian	Germany vs. Austria	1938	1938	0.1
Nomohan	USSR, Mongolia vs. Japan	1939	1939	4
Russo-Finnish	USSR vs. Finland	1939	1939	4
World War II				
German-Polish	Germany vs. Poland	1939	1939	1
German-Belgian	Germany vs. Belgium	1940	1940	0.11
German-Netherlands	Germany vs. Netherlands	1940	1940	0.1
German-Danish	Germany vs. Denmark	1940	1940	0.033
German-Norwegian	Germany vs. Norway	1940	1940	2
German-French	Germany vs. France	1940	1940	1.5
Italo-Greek	Italy vs. Greece	1940	1940	2
Pacific	United States vs. Japan	1941	1945	45
Western	United States, United Kingdom vs. Germany, Italy	1942	1945	60
Eastern	Germany vs. USSR	1941	1945	46
German-Yugoslav	Germany vs. Yugoslavia	1941	1941	0.33
German-Greek	Germany vs. Greece	1941	1941	0.67
Franco-Thai	France vs. Thailand	1940	1940	3
1st Kashmir	India vs. Pakistan	1947	1948	24
Palestine	Iraq, Egypt, Syria, Lebanon, Jordan vs. Israel	1948	1948	8
Korean	United States, S. Korea vs. China, N. Korea	1950	1953	36
Russo-Hungarian	Hungary vs. USSR	1956	1956	1
Sinai	United Kingdom, France, Israel vs. Egypt	1956	1956	1
Sino-Indian	China vs. India	1962	1962	1
Vietnamese I	United States, S. Vietnam vs. N. Vietnam	1964	1973	121
Second Kashmir	India vs. Pakistan	1965	1965	5
Six Day	Egypt, Syria, Jordan vs. Israel	1967	1967	0.2
Israeli-Egyptian	Egypt vs. Israel	1970	1970	0.25
Football	Honduras vs. El Salvador	1969	1969	0.15
Bangladesh	India vs. Pakistan	1971	1971	2
Yom Kippur	Egypt, Iraq, Syria, Jordan, Libya vs. Israel	1973	1973	3
Turko-Cypriot	Turkey vs. Cyprus	1974	1974	1
Vietnamese II	North Vietnam vs. South Vietnam	1975	1975	3
Ethiopian-Somalian	Cuba, Ethiopia vs. Somalia	1977	1978	8
Ugandan-Tanzanian	Uganda, Libya vs. Tanzania	1978	1978	6
Iran-Iraq	Iran vs. Iraq	1980	1988	96
Falklands	Argentina vs. United Kingdom	1982	1982	3
Israeli-Syria (Lebanon)	Syria vs. Israel	1982	1982	2
Sino-Vietnamese	China vs. Vietnam	1985	1990	60

APPENDIX B: SPECIFICATION CHECKS

We did auxiliary analyses to check the specification of the model in several ways, none of which change the major substantive conclusions. An analysis using an exponential specification produces nearly identical coefficient estimates, as expected, given our Weibull model's estimate of p close to 1.0. To check for the effect of the few longest (outlying) wars, we estimated a Cox proportional hazard model on the non-TVC data set and obtained very similar estimates on all variables except for terrain × strategy, for which the statistical significance level dropped from 0.09 to 0.19. We also performed a Weibull analysis on that data set after simply dropping the three extremely long wars. In that analysis, the magnitude of the coefficients remained approximately the same, but some standard errors increased, and statistical significance dropped from 0.07 to 0.12 for repression, from 0.05 to 0.07 for democracy, and from 0.10 to 0.35 for number of actors.

Finally, we performed a check of the Weibull specification assumption using procedures described in Greene (1993, 723) and Jaggia (1991b), which attempt to reject the hypothesis that another (unspecified) model is correct. Duration models do not produce residuals directly comparable to regression residuals. It is necessary to generate a "generalized residual" = $\varepsilon_i = [\exp(-\hat{\beta}x_i)t_i]^{\hat{p}}$ to test multiple moment restrictions. The test estimators are created by generating an $n \times J$ matrix M, consisting of J moment restrictions for the n cases, and an $n \times k$ matrix D, which is made up of all $\partial L/\partial\theta_i$, where L is the model likelihood function and $\theta_i(i = l..k)$ are the k parameters of the model. The second moment restriction (for example) is $\varepsilon^2 = 2$. Two test statistics can be generated, both of which are known to be distributed $\chi^2(J)$. The first statistic equals nR^2, where R^2 is obtained by regressing unity on the matrix (DM). The second statistic equals $i'M(M'M-M'D(D'D)^{-1}D'M)^{-1}M'i$. On the second moment test, our test statistics were 5.37 and 4.74, which slightly exceed the chi-square critical value of 3.81. The results fall into a gray area: While they do not indicate rejection of the Weibull, they do suggest that a more general specification could be appropriate.

We therefore estimated a gamma heterogeneity model, which allows heterogeneity among cases. Along with neglected duration dependence, the major source of misspecification in duration models is heterogeneity (Pagan and Vella 1989, S55). The final source of misspecification, exclusion of other relevant variables, can only be addressed by further theory. If heterogeneity is important, the new analysis would result in large coefficient changes, as it does in Jaggia's (1991a, 179) example. In our case, however, the gamma heterogeneity model did not yield an improved log-likelihood, did not estimate a heterogeneity parameter significantly different from 1, and most important did not affect our coefficient estimates and hence our inferences. In conjunction with the specification tests, this suggests that the Weibull is superior to the likely alternative distributions.

REFERENCES

Allison, Paul D. 1984. *Event History Analysis*. London: Sage.

Axelrod, Robert. 1979. "The Rational Timing of Surprise." *World Politics* 31(January):228–46.

Beck, Nathaniel, and Jonathan Katz. 1992. "Model Assessment and Choice via Cross-Validation." Presented at the 1992 Annual Meeting of the American Political Science Association, Chicago.

Blainey, Geoffrey. 1988. *The Causes of War*, 3d ed. New York: Free Press.

Bueno de Mesquita, Bruce. 1978. "Systemic Polarization and the Occurrence and Duration of War." *Journal of Conflict Resolution* 22(June):241–67.

Bueno de Mesquita, Bruce. 1981. *The War Trap*. New Haven: Yale University Press.

Bueno de Mesquita, Bruce, and Randolph M. Siverson. 1995. "War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Account Ability." *American Political Science Review* 89(December):841–55.

Bueno de Mesquita, Bruce, Randolph M. Siverson, and Gary Woller. 1992. "War and the Fate of Regimes: A Comparative Analysis." *American Political Science Review* 86(September):638–46.

Clausewitz, Carl Von. 1976 (1832). *On War*. Trans. Michael Howard and Peter Paret. Princeton: Princeton University Press.

Clodfelter, Michael. 1992. *Warfare and Armed Conflicts*. 2 vols. Jefferson, NC: McFarland and Co.

Dupuy, R. Ernest, and Trevor N. Dupuy. 1986. *The Encyclopedia of Military History from 3500 BC to the Present*. 2d rev. ed. New York: Harper and Row.

Dupuy, Trevor Nevitt. 1979. *Numbers, Predictions and War: Using History to Evaluate Combat Factors and Predict the Outcome of Battles*. Indianapolis: Bobbs-Merrill.

Dupuy, Trevor Nevitt. 1983. *Analysis of Factors That Have Influenced Outcomes of Battles and Wars: A Data Base of Battles and Engagements: Final Report*. Vols. 1–6. Dunn Loring, VA: Historical Evaluation and Research Organization, Division of T.N. Dupuy Associates.

Epstein, Joshua M. 1987. *Strategy and Force Planning: The Case of the Persian Gulf*. Washington, DC: Brookings Institution.

Gartner, Scott Sigmund. 1992. "*Strategic Assessment in War: A Bounded Rationality Model of How Organizations Evaluate Policy Effectiveness*." Ph.D. diss, University of Michigan.

Goertz, Gary, and Paul F. Diehl. 1993. "Enduring Rivalries: Theoretical Constructs and Empirical Patterns." *International Studies Quarterly* 37(June):147–72.

Greene, William H. 1992. LIMDEP, v. 6.0. New York: Econometric Software, Inc.

Greene, William H. 1993. *Econometric Analysis*. 2d ed. New York: MacMillan Publishing Company.

Gurr, Ted Robert, Keith Jaggers, and Will H. Moore. 1989. "Polity II data set, ICPSR 9263." Boulder, CO. Photocopy.

Hanushek, Eric A., and John E. Jackson. 1977. *Statistical Methods for Social Scientists*. New York: Academic Press, Inc.

Holsti, Kalevi J. 1991. *Peace and War: Armed Conflicts and International Order 1648–1989*. Cambridge: Cambridge University Press.

Horvath, William J. 1968. "A Statistical Model for the Duration of Wars and Strikes." *Behavioral Science* 13(January/February):18–28.

Huth, Paul, D. Scott Bennett, and Christopher Gelpi. 1992. "System Uncertainty, Risk Propensity, and International Conflict Among the Great Powers." *Journal of Conflict Resolution* 36(September):478–517.

Iklé, Fred Charles. 1971. *Every War Must End*. New York: Columbia University Press.

International Institute for Strategic Studies. Various years. *The Military Balance*. London: IISS.

Jaggia, Sanjiv. 1991a. "Specification Tests Based on the Heterogeneous Generalized Gamma Model of Duration: With an Application to Kennan's Strike Data." *Journal of Applied Econometrics* 6(April–June):169–80.

Jaggia, Sanjiv. 1991b. "Tests of Moment Restrictions in Parametric Duration Models." *Economics Letters* 37(September):35–38.

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

Kiefer, Nicholas M. 1988. "Economic Duration Data and Hazard Functions." *Journal of Economic Literature* 26(June):646–79.

Kier, Elizabeth. 1995. "Culture and Military Doctrine: France between the Wars." *International Security* 19(Spring):65–93.

King, Gary. 1989. *Unifying Political Methodology*. Cambridge: Cambridge University Press.

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

Langer, William L. 1980. *An Encyclopedia of World History*. Boston: Houghton Mifflin.

Lebow, Richard Ned. 1984. "Windows of Opportunity: Do States Jump Through Them?" *International Security* 9(Summer):147–76.

Luttwak, Edward N. 1987. *Strategy*. Cambridge: The Belknap Press of Harvard University Press.

Mearsheimer, John. 1983. *Conventional Deterrence*. Ithaca: Cornell University Press.

Mitchell, C. R., and Michael Nicholson. 1983. "Rational Models and the Ending of Wars." *Journal of Conflict Resolution* 27(September): 495–520.

Mitchell, William. 1921. *Our Air Force: The Keystone of National Defense*. New York: E.P. Dutton.

Morrison, Donald G., and David C. Schmittlein. 1980. "Jobs, Strikes, and Wars: Probability Models for Duration." *Organizational Behavior and Human Performance* 25(April):224–51.

Morrow, James D. 1985. "A Continuous-Outcome Expected Utility Theory of War." *Journal of Conflict Resolution* 29(September):473–502.

New York Times. 1983. *The New York Times World Atlas*. New York: Times Publishing.

Organski, A.F.K., and Jacek Kugler. 1980. *The War Ledger*. Chicago: University of Chicago Press.

Pagan, Adrian, and Frank Vella. 1989. "Diagnostic Tests for Models Based on Individual Data: A Survey." *Journal of Applied Econometrics* 4(supplement):S29–S59.

Petersen, Trond. 1986. "Fitting Parametric Survival Models with Time-Dependent Covariates." *Applied Statistics* 35(No. 3):281–88.

Posen, Barry R. 1984. *The Sources of Military Doctrine*. Ithaca: Cornell University Press.

Record, Jeffrey. 1993. *Hollow Victory: A Contrary View of the Gulf War*. McLean, VA: Brassey's (US).

Reiter, Dan. 1995. "Exploding the Powder Keg: Preemptive Wars Almost Never Happen." *International Security* 20(Fall):5–34.

Rosen, Stephen Peter. 1995. "Military Effectiveness: Why Society Matters." *International Security* 19(Spring):5–31.

Russett, Bruce M. 1990. *Controlling the Sword*. Cambridge: Harvard University Press.

Sagan, Scott D. 1986. "1914 Revisited: Allies, Offense, and Instability." *International Security* 11(Fall):151–75.

Singer, J. David, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820–1965." In *Peace, War and Numbers*, ed. Bruce Russett. Beverly Hills: Sage Publications.

Small, Melvin, and J. David Singer. 1982. *Resort to Arms: International and Civil Wars, 1816–1980*. Beverly Hills: Sage Publications.

Snyder, Jack. 1984a. "Civil-Military Relations and the Cult of the Offensive, 1914 and 1984," *International Security* 9(Summer):108–46.

Snyder, Jack L. 1984b. *The Ideology of the Offensive: Military Decision Making and the Disasters of 1914*. Ithaca: Cornell University Press.

Stam, Allan C., III. 1995. "Strategy, Class and Ethnic Bias: Their Mutual Effects on War Outcomes and Durations." Presented at the ISA annual meeting, Chicago.
Stam, Allan C., III. 1996. *Win, Lose, or Draw*. Ann Arbor: University of Michigan Press.
Stockholm International Peace Research Institute. Various years. *SIPRI Yearbook*.
Sun Tzu. 1991. *The Art of War*. London: Shambhala.
Trachtenberg, Marc. 1991. *History and Strategy*. Princeton: Princeton University Press.
Van Evera, Stephen W. 1984. "The Cult of the Offensive and the Origins of the First World War." *International Security* 9(Summer): 58–107.
Vasquez, John A. 1994. *The War Puzzle*. Cambridge: Cambridge University Press.
Vuchinich, Samuel, and Jay Teachman. 1993. "The Duration of Wars, Strikes, Riots, and Family Arguments." *Journal of Conflict Resolution* 37(September):544–68.
Warwick, Paul V. 1992. "Rising Hazards: An Underlying Dynamic of Parliamentary Government." *American Journal of Political Science* 36(December):857–76.
Wittman, Donald. 1979. "How a War Ends: A Rational Model Approach." *Journal of Conflict Resolution* 23(December):743–63.
Zimmerman, William. 1973. "Issue Area and Foreign-Policy Process: A Research Note in Search of a General Theory." *American Political Science Review*. 67:1204–12.