Explaining and predicting interstate war deaths

Andreas Beger
Florida State University

June 2012

dissertation.v1.2

> Bombardment, barrage, curtain-fire, mines, gas, tanks, machine-guns, hand-grenades–words, words, words, but they hold the horror of the world.
>
> – Erich Maria Remarque, *All Quiet on the Western Front*

> It is forbidden to kill; therefore all murderers are punished unless they kill in large numbers and to the sound of trumpets.
>
> – Voltaire, *War*

1 Introduction

Between 1816 and 1997, there were 79 interstate wars involving 281 states as combatants that led to more than 31 million military deaths (Sarkees, Wayman and Singer, 2003). The causes and conduct of these wars have been extensively studied, addressing questions like why and when states fight, how they fight, and how one can explain the outcomes of a war. What is sometimes obscured is the sheer human toll these conflicts cause. Even when it comes to "legitimate" victims of these wars—uniformed soldiers fighting on behalf of organized states—the human cost is large. More than 31 million soldiers lost their lives in battle in this time period. There is however large variation in the wars that contribute to this figure. The Falklands War in 1982 led to 964 battle deaths, while near the other extreme, World War 1 led to more than 8.5 million military fatalities. What explains this variation in the human cost of war? And, given that wars occur, why are some so much deadlier than others?

The answers to these questions are interesting for several reasons. First, presumably an important part of the reason that we study war in the first place is their destructive nature, yet there is large variation in the cost of individual wars that has received comparatively little attention in the quantitative study of wars. If the study of war is interesting because they destroy things and people and cost money, then the question of fatalities should also be intrinsically interesting. The focus in this project is on battle-deaths due to limits in the data available, and these are admittedly only part of the true cost of war. These same limitations, along with the more fundamental problem of quantifying the cost of human death, make it difficult to ascertain how much battle deaths contribute to the overall cost of a war. It seems safe to presume that battle-deaths are a significant part of the true cost of war, and maybe even to presume that they are correlated with the overall cost of a war. And regardless, it provides at least a practically useful starting point for distinguishing wars by their severity.

Second, examining the determinants of war fatalities can inform existing research on war by (1) empirically evaluating the implications such theories have for fatalities, and (2) due to the particular way in which war onset, termination, and duration are operationalized. Bargaining theories of war have implications for the level of fatalities in war in the sense that these constitute part of the cost of war that states bear while fighting. Although the primary motivation of this paper is *not* to evaluate bargaining theories of war (rather, the focus is on systematic prediction of fatalities), the results can be informative for the original theories.

More abstractly, research on war fatalities is related to research on other aspects of war like onset as well due to the way in which the operational definitions of these concepts depend on fatality levels. In any given conflict, war onset is conventionally only coded if fatalities surpass a certain threshold of fatalities, usually 1,000, war termination presupposes that fatalities levels are zero or very low, etc. Research that examines these aspects of war thus also indirectly makes claims about fatality levels, and hence research like this can be useful for it. For example, one way to interpret questions of conflict escalation is that they attempt to determine, given that there is some form of conflict between two states, whether total fatalities will surpass 1,000 or not.[1]

And third, the answers to the questions above also hold potential significance for policy-making. Presumably most state leaders confronted with situation that could lead to war will care about how costly and deadly such a war would be if it were to occur. The problem is that predicting how costly wars will be is to my knowledge not straightforward. Before the Persian Gulf War, estimates for the number of fatalities were far higher, up to an order of magnitude, than the approximately 30,000 that did die. Lacina and Gleditsch (2005) give a best estimate of 29,171 for the conflict, whereas Cioffi-Revilla (1991) predicted between 100,000 and 1,000,000 fatalities. Conversely, one may wonder whether the prospect of a Second Gulf War would have encountered more resistance in 2003 had it been obvious that the aftermath would drag out for more than 6 years and cost the lives of more than 4,200 U.S. soldiers and a large but unknown number of Iraqi soldiers, police, militia, and civilians.

The focus in this project is to answer the questions posed in the first paragraph, and doing so ultimately is interesting to the extent that one can actually predict fatality levels with some measure of accuracy. If the goal is to explain variation in the levels of fatalities across wars, raw coefficients and significance levels are of secondary interest to substantive effects and how well statistical models can actually predict fatalities. What conditions lead to particularly deadly wars? Thus one of the goals is to identify theoretically reasonable statistical models that perform well in terms of prediction.

2 How long will states bear the costs of war?

On a fundamental level, the level of casualties and other costs in a war is a function of the extent to which states will tolerate further loss, since hypothetically they could end a war at any moment by surrendering or making large concessions. In that sense, the question of how many fatalities a war will produce is theoretically similar to the question of how long a war will last(Bennett and Stam III, 1996; Slantchev, 2004). Bargaining theories of war and previous work on war duration suggest that states will bear the costs of war either until the underlying cause of a war, e.g. information asymmetries about relative capabilities, or commitment problems, have been sufficiently addressed, or until a military resolution to a conflict is reached (one side defeats the other). Since fatalities are presumably a significant component of war costs, one should also expect that these factors influence war fatalities as well (Wagner, 2000).

[1] The empirical analysis below lacks a temporal dimension, mainly due to the questions it is meant to address, but also due to lack of data. Thus it is not directly comparable to empirical work that examines war onset in cross-sectional time-series data. But it could be used to study the issue of conflict escalation, for example, and more generally the point that there is a link between the study of fatality levels in general and other aspects of war remains.

2.1 Incomplete Information

The incomplete information explanation of war focuses on the role that uncertainty among state leaders about relative capabilities and resolve plays in their expectations. As long as war is costly, there should always be a bargaining outcome that both states in a dispute should prefer(Fearon, 1995). Of course in practice obstacles such as difficulty in observing material and less tangible capabilities and incentives for state leaders to misrepresent information can lead to unrealistic expectations among states about the proper share due to them, which in turn can lead to costly war (Slantchev, 2010). In this sense, war serves as a mechanism to reveal information about the combatants and ends when it loses its informational value (Filson and Werner, 2002; Powell, 2004; Slantchev, 2003*b*; Smith, 1998).

Wars can convey information through two mechanisms, the battlefield outcomes they entail, and the strategic behavior of the combatants during the war. Because battlefield outcomes are in practice determined independent of negotiations (i.e. they are not easily subject to strategic manipulation), state behavior and negotiation offers are usually more informative in regard to the establishment of a set of reasonable expectations (Slantchev, 2003*b*). They are so however exactly because war is costly, which helps to separate weak from strong states. The fatalities states sustain in a conflict, as a major component of costs, are a key part in allowing war to convey useful and credible information to uncertain combatants.

This leaves open the question of what influences uncertainty and information asymmetries in the first place. Formal models have picked up on a long line of reasoning about balance of power and preponderance of power systems to examine the potential role of different power distributions on information asymmetries. Such a relationship is by no means straightforward, but there is some evidence to suggest that parity is linked to uncertainty (Reed, 2003). Although observable capabilities may be taken into account in pre-war bargaining by fully informed actors, *given* that two states already are at war, observable capabilities should be related to uncertainty because they influence how important unobservable capabilities are (Gartzke, 1999). Reed gives the example that, regardless of how large it is, Denmark's resolve will probably not influence perceptions of the likely course of a war with the USA, whereas resolve might change things quite a bit if two states are more or less equally matched in observable capabilities. Furthermore, battlefield outcomes themselves are less informative when observed capabilities are near parity and as a result it will take more battles and more deaths for credible information to be conveyed (Slantchev, 2004, 816). Thus when observed capabilities are at or near parity, states will face a relatively high level of uncertainty and will tolerate more costs:

H 1. *Parity in observable capabilities increases war deaths.*

Abstract discussion of war tend to focus on conflict between two states, but a significant number of wars involve more than two states. This adds additional sources of uncertainty to power calculations. There will be strategic interaction among states fighting on the same side which might influence how observable capabilities correlate with the unobservable true capability of one side. At the very least they will face common collective action problems. Thus wars fought between multiple states should increase uncertainty due to the potential strategic behavior involved, which in turn should make the main combatants willing to sustain higher fatalities than they otherwise might.

H 2. *As the number of states involved in a war increases, so will war deaths.*

2.2 Commitment Problems

The second major rationalist explanation of war focuses on commitment problems which arise when states cannot reach a peaceful settlement because it is known or believed that one of the states involved will have incentives to deviate from the settlement in the future. War in this case arises because although a peaceful settlement is obvious to both sides, at least one of them for some reason is known or thought to have an incentive to renege on the settlement in the future because it will think that it can gain a more favorable bargain by doing so. As a result, neither side can credibly commit to upholding the current settlement under negotiation.

The archetypical reason underlying states' belief that they face a commitment problem is shifts in bargaining power over time, e.g. changes in observed capabilities or resolve (Powell, 1996, 2004). Other situations that may lead to commitment problems include significant surprise or first-strike advantages, or bargaining over issues which themselves can change the balance of power between adversaries (e.g. the Golan Heights between Israel and Syria) (Powell, 2006). The key in all three of these situations is that they hold the potential for rapid shifts of power that underlie commitment problems. When the stakes are high enough and large enough changes in power are looming, war arises because states will just not be able to resist the temptation of fighting to get all they want (Leventoğlu and Slantchev, 2007). Knowing this, states choose to engage in war instead.

Credible information provided through battlefield outcomes and state's behavior in response to them plays less of a role in resolving wars caused primarily through commitment problems. Rather, war incidentally resolves the problem by destroying enough of what states are fighting over to eventually make peace possible.(Leventoğlu and Slantchev, 2007, 767). Instead of dividing the pie, states fight war long enough to destroy so much of it that both can credibly commit to a peaceful settlement because there just is not enough left to make war an attractive gamble. This suggest that in situations where there is a credible commitment problem, states will be willing to endure fatalities and other costs of war in proportion to what is at stake. The more that is at stake the longer it will take to destroy enough of it to "sour" the prospect of continuing war. This is is consistent with the argument that the more important the issues underlying the dispute, the more states will be willing to sacrifice in order to get their share (e.g. Hensel et al., 2008):

H 3. *High stakes in war increase deaths.*

2.3 Military decision

War can resolve obstacles to peaceful bargaining, but ultimately it is also a military contest that exposes states to the risk of collapse or military defeat (Reiter, 2003, 30). Regardless of what caused World War 2 in Europe, the occupation of almost all of Germany before the surrender ended it. Although most wars terminate in negotiated settlements, a significant number are decided militarily (Pillar, 1983). This suggests that military resolutions to a war are important for models of war termination or the costs of war—some wars simply stop because one side looses the ability to fight. The fatalities in a war should thus be constrained by factors that influence how long a state can keep fighting.

The most straightforward way to eliminate another state's ability to continue a war is to destroy or sufficiently weaken its armed forces. Rough terrain generally should make this harder to achieve since it makes large, conventional battles more difficult (Bennett and Stam III, 1996). This can allow combatants more time to capitalize on their reserves or to fight prolonged wars using non-conventional means (Fearon and Laitin, 2003). Wars fought in some types of rough terrain such as urban areas are also associated with higher levels of fatalities.

H 4. *As the proportion of rough terrain increases, so will deaths.*

Undermining the willingness of an enemy to continue fighting is an alternative to this brute force approach, and a very common argument relating to the ability of some states to bear the costs of war concerns the aversion of democracies to casualties (Bueno de Mesquita and Siverson, 1995; Filson and Werner, 2004; Gartner, 2008). If casualties become too high, continuing a war becomes politically infeasible under certain regime types. There is empirical evidence to support this argument in the case of the United States and it seems to also be the case in other democracies (Carson et al., 2001; Gartner, 2008). As a result of this loss aversion, democracies will not only be unable to fight costly wars, but they might also self-select into fighting only short, easily winnable wars:

H 5. *Deaths will be lower if one of the combatants in a war is a democracy.*

Another potential indication of a state's ability to bear or inflict costs lies in its selection of strategy to fight the war. Strategies aimed at producing quick military victories often indicate the unwillingness or inability of a state to wage prolonged warfare (e.g. Germany in World War 2), while strategies based on prolonged insurgency warfare can only be used by states with a high capacity for absorbing costs (e.g. North Vietnam during the Vietnam War):

H 6a. *Wars in which at least one of the combatants uses a blitzkrieg-style strategy will produce less fatalities.*

H 6b. *Wars in which at least one of the combatants uses a guerilla-style strategy will produce more fatalities.*

3 Data and methods

The data consist of 90 interstate wars from 1815 to 1991 with 910 or more battle deaths. The list of wars is based on the Correlates of War list of interstate wars (79 wars), with two changes (Sarkees, Wayman and Singer, 2003). It includes two additional wars that are consistent with COW state system membership (Slantchev, 2004), the Pastry War and Uruguayan Dispute. Furthermore, World War 2, the Vietnam War, and the Persian Gulf War are disaggregated into multiple smaller wars totaling 15 in number. The reasoning underlying this change is twofold. First, it makes this work empirically more consistent with previous work on war duration that uses similar or the same set of cases. Second, Bennett and Stam III (1996) argue that the cases which were split really consisted of distinct wars in the sense that actors in each distinct episode probably did not take into account the events that would later occur in the larger war. For example, when Germany invaded Poland in 1939, Hitler almost certainly did not expect the UK and France to respond forcefully in an effective manner. From the resulting 93 wars, 3 drop out of the final sample because fatalities were well below the traditional threshold of 1,000.

In addition to listing one major combatant for each of the two sides in a war, the data also include other combatants on each side from the list of states in the COW interstate war participants data and other historical references (Sarkees and Schafer, 2000; Dupuy and Dupuy, 1986; Holsti, 1991). The list of combatants is largely consistent with the COW interstate war participants list, unless Slantchev (2004) or Bennett and Stam III (1996) listed a state that was not included in the COW participants list as one of the major combatants. Historical sources were used to identify combatants for the added or disaggregated wars (e.g. World War 2) and to determine on which side of a conflict a state fought.

Out of the total 93 wars, 35% (33 wars) were fought by 3 or more combatants, with the Korean War having the largest number of combatants (16–2 on the side of North Korea, and 14 on the UN side). The disaggregated World War 2 conflicts account for a large number of the remaining multilateral wars.

There are two rationales underlying this coding. First, as historical sources and the COW participants list indicate, a substantial number of wars genuinely were multilateral in the sense that more than 2 states contributed significant manpower and matériel to the conflict—e.g. consider the First World War or any of the Arab-Israeli wars. Thus only counting the material capabilities, population, etc. of a single state could be a misrepresentation of the true military situation prior to the outbreak of the war. Second, the addition of more combatants to a conflict likely influences the uncertainty of information for decision makers on either side. As a result, the explanatory variables dealing with material capabilities are calculated not only for the two major combatants, but for all states involved in a conflict.

The dependent variable consists of fatality data from three sources. Lacina and Gleditsch (2005) provide the most recent fatality data. These particular data only reach back to 1900, which covers 49 of the 93 wars. To obtain fatality estimates for all wars in the sample I used those from Slantchev and the COW project for the remaining 44 wars (Sarkees and Schafer, 2000; Slantchev, 2004). The correlation between fatality estimates from the two sources is 0.98 for the 49 wars in which the data overlap, suggesting that this is reasonable to do. The resulting fatality data range from a low of 12 (the German invasion of Denmark in WW2) to slightly above 10 million (the Great Patriotic War between Germany and the USSR in World War 2). Usually a conflict is coded as a war when battle deaths surpass 1,000. The Falklands War is coded as having 910 fatalities by COW, and using this as the threshold for annual battle deaths, the number of wars in the sample is reduced from 93 to 90.

3.1 Estimation strategy

Fatality figures are unlikely to be the result of a normal data generating process, and instead seem to follow a power law or similar type of distribution, where more serious wars will be less frequent and *vice versa* (Cederman, 2003; Levy and Morgan, 1984; Clauset, Shalizi and Newman, 2007). Existing studies of (civil) war fatalities use OLS regression of the log of fatalities to take this into account (Cioffi-Revilla, 1991; Lacina, 2006). This allows the explanatory variables to have what effectively is an exponential effect on the number of fatalities in a war: if $\ln y = x_i\beta$, then $y = e^{x_i\beta}$.

However, the fatality data used here are also truncated in the sense that only conflicts in which fatalities exceed 910 battle deaths are included. Regular statistical models like OLS that do not take this into account can produce biased and misleading results. A truncated regression model corrects this issue by rescaling the normal distribution to account for the truncation (where τ is the truncation point, ϕ is the standard normal pdf, and Φ is the standard normal cdf) (DeMaris, 2004; Greene, 2008, 756-761):

$$f(y|y>\tau) = \frac{\frac{1}{\sigma}\phi(\frac{y-x_i\beta}{\sigma})}{1-\Phi(\frac{\tau-x_i\beta}{\sigma})} \quad (1)$$

This in turn leads to the log-likelihood function for truncated normal regression:

$$lnL = \sum_{i=1}^{N}\left(\ln\left[\frac{1}{\sigma}\phi\left(\frac{y-x_i\beta}{\sigma}\right)\right] - \ln\left[1-\Phi\left(\frac{\tau-x_i\beta}{\sigma}\right)\right]\right) \quad (2)$$

Because the sample size of 90 is small, I use bootstrapped estimates to evaluate statistical significance. Bootstrapping does not rely on the asymptotic assumption of normality for errors and instead uses repeated re-sampling with replacement from the observed data to estimate errors. Significance levels for coefficients are determined by using 95% and 90% bias-corrected accelerated (BC_a) confidence intervals, which correct for bias and skewness in the bootstrap samples (DiCiccio and Efron, 1996).

Finally, war duration is a confounding factor that is related to both fatalities (longer wars are deadlier on average) and some of the independent variables. Since war duration is observed only after a war has concluded, it cannot figure in states' decision to continue fighting (Slantchev, 2004). Instead, it is possible to model and estimate war duration itself using observed conditions at the outset of a war. These include parity and total military personnel, parity and total population, number of combatants, rough terrain, and whether a combatant was democratic (table 2). Predicted war duration from this model in turn is used as a further independent variable in the fatality models (Slantchev, 2004).

The overall estimation strategy thus consists of three distinct steps: (1) estimate a duration model and generate estimates of war duration, (2) estimate a truncated normal regression model of war deaths and generate estimates of war deaths, and (3) repeat the first two steps 1,000 times to obtain bootstrap estimates for all parameters and model predictions.

3.2 Explanatory variables

Explanatory variables are measured during the beginning year of a war, except the number of states involved in a war and the resulting capabilities measures.

Parity in observed capabilities. Parity in observed capabilities is measured with two variables: active military personnel (in thousands) and COW's composite index of national capability (CINC). Both come from the COW National Material Capabilities data (Singer, 1987; Singer, Bremer and Stuckey, 1972). Military personnel are one of the easiest observable indicators of a state's power prior to a war, although they do not take quality into account. The composite index of national capability is more sensitive to this issue since it is based on military personnel, military expenditures, industrial capacity, and several other relevant variables. Each variable is used to construct a separate measure of parity. In the case of wars that have multiple combatants on a side, I add the active military personnel or CINC figures for all states fighting on the same side in a war. Parity is calculated as the difference between observed capabilities of the stronger state and the weaker state relative to the sum of their capabilities. A 0 indicates complete superiority by one side, and a 1 indicates perfect parity.[2] The average war was initiated by a state or coalition with 430,000 military personnel (median) against a target state or coalition with 153,000 military personnel at the starting year of the war. The mean values for parity in military personnel and parity in CINC scores are 0.44 and 0.43 respectively.

Number of states. The total number of states involved in a war, as coded by the COW interstate war participants list (Sarkees and Schafer, 2000). A few wars have additional combatants if it appears that those states made a significant contribution in manpower or logistics (Dupuy and Dupuy, 1986). The variable ranges from 2, i.e. wars fought solely between two states, to 16 combatants for the Korean War.

Issue salience. Issue salience is identified for each side in a war with the following coding scheme: high salience if the issue is regime or state survival, national liberation, or autonomy; medium salience if the issue involves territory, integrity of state, or honor/ideology; and low salience if the issue is maintaining an empire, commercial disputes, or policy (Slantchev, 2004). Each of the two variables measures the salience of the issue underlying the conflict for one of the two major combatants.

Rough terrain. This variable identifies the extent to which a war was fought over rough terrain, e.g. heavy woods, jungles, swamps, and mountains (Slantchev, 2003*a*; Stam III, 1999). Higher values for the variable indicate that a war featured more combat in such types of terrain.

Democratic combatants. I measure democracy using data from the Polity IV project (Marshall and Jaggers, 2002). First I created a polity score that equals the democracy score minus the autocracy score for each of the two major combatants in a war in the year that the war started. The resulting index ranges from -10 to 10. In cases where there was a regime transition or other turmoil, I use the democracy or autocracy score from the last year available. The final measure is dichotomous and indicates a democracy if a state had a polity score of 6 or higher, otherwise an autocracy. Thirty-two wars involved a democratic state as one of the two main combatants.

Strategy. Bennett and Stam measure strategy using a fairly complex variable that captures which side pursued and offensive/defensive doctrine and whether the particular strategy consisted of maneuver warfare, attrition, or punishment (e.g. insurgency) (Bennett and Stam III, 1996). I used their measure as well as historical references to construct two dummy variables indicating whether any side used either a maneuver strategy (19 wars) or a guerrilla strategy (5 wars) (Dupuy and Dupuy, 1986). The majority of wars were fought between opponents who both used attrition strategies (69 out of 93 wars), and there were no instances were at least one of the combatants did not employ an attrition strategy.

Control Variables. All regression models include three variables to control for exposure effects ((e.g. deaths cannot exceed the total population of all combatants) and relative population size: ***total observed capabilities***, i.e. the sum of observed capabilities, either military personnel or CINC score, for all combatants involved in a war; ***parity in total population***, i.e. how close to parity in population each side in a war is; and the natural log of***total population***, the sum of all combatant states' prewar population.

Duration model. The dependent variable is the duration of a war in days (Slantchev, 2004). The

[2] Specifically, $1 - \frac{p_{stronger} - p_{weaker}}{p_{stronger} + p_{weaker}}$, where p is either military personnel or CINC.

covariates used are similar to those in Slantchev's full model. They include parity in military personnel, parity in total population, total military personnel, total population, the number of actors, rough terrain, contiguity, and whether a democratic state was involved as principal combatant. I use the natural log of total military personnel and total population as those produced better fitting models. The purpose of this duration model is to generate predicted duration values for each war to use in the main regression models.

4 Empirical results

The duration model which was used to generate predictions that are used as an independent variable in the war deaths models is shown in table 2. Based on AIC scores, a log-normal form provided the best-fitting model. As with the war deaths models, the estimates are bootstrapped due to the small sample size. The results, in terms of specific associations and significance, are roughly consistent with previous models of war duration (Bennett and Stam III, 1996; Slantchev, 2004), suggesting this is a theoretically reasonable model. Parity in military personnel, the number of states involved, and terrain are associated with significant increases in war duration.

Moving on to war deaths, table 3 shows estimates for the two main truncated regression models, using either military personnel or CINC to measure parity. Table 4 shows coefficient signs and significance for several further models to check for robustness of results.

Raw coefficients in a truncated regression model provide the marginal effect of a variable in the underlying, unobserved population of cases, which here would mean all conflicts including those short of war (Greene, 2008, 867–869). Calculating marginal effects for just the truncated sample (i.e. wars) requires adjustment by a parameter that is bound by 0 and 1, depending on the extent of truncation. Thus the magnitude, but not the direction of effects from these results would be somewhat closer to zero in regard to traditionally-defined wars. Exponentiated coefficients for those that are significant show the factor change in observed fatalities associated with a one unit change in the independent variable.

4.1 Regarding the hypotheses

Two variables stand out for having highly significant and consistent effects across both models (and other specifications; table 4)—democracy and issue salience for the initiating state. Wars that involve at least one democracy appear to be consistently and significantly less deadly, with an average reduction in fatalities by a factor of about 0.2. Whether it is because of self-selection or internal politics, democracies tolerate far less deaths in wars than autocratic regimes.

Higher issue salience for the initiator increases deaths on average by a factor of between 5 and 7 times for a one unit increase in salience (it ranges from 0 to 2). However, the coefficient for issue salience for the target state is always insignificant. So it appears that issue salience for the initiating state matters, while issue salience for the target state does not. Maybe states that have a lot at stakes also tend to be the ones that initiate wars, to exploit the element of surprise or some other first-strike advantage. The data suggest otherwise. In most wars (82%), the target has more or roughly as much at stake as the initiator. War deaths thus seem to be driven by the extent to which the initiator values the issue over which it is fought, regardless of the target states stakes.

The two hypotheses relating uncertainty and fatalities overall receive only mixed support. The coefficients for both parity in observable capabilities and the number of states in a conflict are positive and significant when using military personnel, but not with the CINC score. One reason for this may be related to the fact that the variables are all measured in the first year of the conflict, so that if their values and hence state's expectation change over the course of a conflict, the models here are unable to account for any effect these changes have. Another alternative has to do with the measures themselves: while military personnel is a fairly straightforward measure of military capabilities, the CINC score is more broad and less observable in real time.

What about military resolutions to war? Support for a relationship between factors that influence the military conduct of a war and consequent fatality levels are also mixed. There is a very clear indication that wars involving a democracy as major combatant, all else being equal, have lower fatalities than those fought purely between non-democratic regimes. On the other hand, terrain has a significantly positive impact in the first model, but not in the second and strategy does not seem to have a significant impact when controlling for other factors.

4.2 Robustness checks

The main results—significant effects for democracy and issue salience, a possible effect for parity in military personnel and the number of states—remain consistent across several different model specifications.

The first set of robustness checks both change the estimation strategy by either using a negative binomial regression instead of truncated regression in the second step of modeling war deaths, or does away with the first step, generating war duration predictions, away altogether by using actual duration. Both approaches simplify model estimation. Using a negative binomial regression (models 3 and 4 in table 4) to model counts of war dead leads to a significant effect for rough terrain, but otherwise leaves the results largely unchanged. The in-sample predictions become notably worse however. Using actual duration renders military personnel parity insignificant, and leads to a positive and significant effect for maneuver strategies and observed war duration.

The variable used to measure issue salience in the base models is ordinal with 3 possible values. Using it as is assumes that the effect is linear, i.e. going from a low value to moderate value (0 to 1) is equivalent to going from a moderate to high value (1 to 2). Using dummy variables for moderate and high value issues removes this assumption. Both of the dummy variables for initiator issue salience have a significant, positive effect on war deaths, while neither of the two dummy variables for the target state has a significant effect. However, parity using either military personnel or CINC, the number of combatants, and terrain now have significant effects and are associated with deadlier wars.

The last four models look at how sensitive the results are to the inclusion of the World War 2 conflicts and World War 1, since these represent the major changes to the data and some of the deadliest wars in the data as well. The main change is that the number of combatants in the new subsample of wars has a consistent positive effect on casualties. Out of the wars excluded from this subsample are 4 conflicts with more than 3 combatants, of which 3 have a million or more casualties. Both parity variables, using military personnel or CINC, also show positive and significant effects.

4.3 Regarding overall model fit

Theoretical contributions aside, the actual fit of statistical models is of interest here for pragmatic reasons. Being able to accurately forecast war fatalities is practically useful to know. To assess how well the models do in predicting actual data, I estimated (1) in-sample predictions (which are *ex post*), as well as (2) out of sample forecasts for the Eritrean-Ethiopian War from 1998–2000 and the Kargil War in 1999. Fatality estimates are calculated during the bootstrapping process, which deals with the problem of calculating accurate fitted values in log-normal models using point estimates (Duan, 1983).

The in-sample predictions consist of point estimates for fatalities as well as the 95% interval for those point estimates.[3] Point predictions for these data are not the only information that is interesting since the chances that a prediction will absolutely match actual fatalities is near zero. Intervals, if accurate, on the other hand can provide a better sense of the likely range fatalities span through their size. For example, two wars might have identical point fatality estimates, but if one has a much smaller interval than the other we can be much more confident in the general range that war is going to produce. The 95% interval for estimated fatalities from model 1, in relation to observed fatalities, are shown in figure 1. The y-axis shows fatality figures (on a logarithmic scale) for each of the 90 wars in the estimation sample. The red dot corresponds to the observed fatalities for that war, while the blue lines indicate the 95% interval for the predicted number of fatalities. The list of wars is sorted by the observed number of deaths, with less deadly wars on the left and deadlier wars on the right.

Statistical models of war, e.g. war onset or termination, tend to be limited in their ability to accurately predict observed values. Here, the 95% intervals of predicted values tends to, on average, do fairly well in capturing observed fatalities within it. The predicted deaths intervals bracket observed fatalities 69.89 percent and 65.59 percent of the time for models 1 and 2 respectively. The point estimates for fatalities, using median value, are correlated with observed fatalities at 0.63 and 0.60 respectively.

On the other hand, the intervals are fairly wide. In three-quarters of the wars they span between 1 and 2 orders of magnitude (log_{10}), which can translate to a difference between 100 and 10,000 deaths at the low end of the scale. One of the better predictions is for the Yom Kippur War 1973. Here the median prediction is 5,800 with an interval from 900 to 22,700—observed fatalities were 11,084. But at the other extreme the interval for predictions for the Korean War spans 5.4 orders of magnitude with a nonsensical upper estimate of deaths. So there certainly is room for much improvement in the quality of these estimates.

Compared to earlier efforts to predict the magnitude of war, the estimates seem reasonable in accuracy and variance. Similar efforts to predict war deaths in the Persian Gulf War in 1991 for example produced a best estimate of 100,000 and a few million (Cioffi-Revilla, 1991). Actual, observed fatalities were far lower, around 29,171 (plausible range of 28,945 to 44,271). Table 5 shows the fatality estimates produced by the two base models. Both models bracket observed fatalities, although the first model has a very high upper limit close to 1.2 million. Neither model,

[3] Estimates are linear predictions, i.e. $\ln \hat{y} = X\beta$.

looking at the upper limit, would have entirely ruled out some of the very large fatality figures that were heard in the lead up to the war. The median, or guest guess, estimates of 46 and 13 thousand dead are close to the observed 29 thousand fatalities however.

Part of the reason for the over-prediction from the two base models lies with the large number of states that participated in this war as members of the US-led coalition. Fatality estimates for the Korean War, which also had a very large number of combatants, are similarly large relative to observed fatalities, and in fact they are by far the highest estimates generated by the models in absolute terms. The Korean War and Gulf War have respectively the largest and third largest number of combatants out the wars in the sample with 16 and 14 respectively. But compared to other wars with many combatants, both wars stand out as being fought primarily on one side by the U.S. at the forefront of coalitions sanctioned by the United Nations, and so the apparent large number of combatants distorts actual contribution of effort.

4.4 Out of sample predictions

How well would these models actually do in forecasting fatalities in a future war? Clearly, we would have a little bit of a problem in evaluating forecast accuracy for wars that have not occurred yet. But since the data used so far end in 1997, we can still do out of sample forecasts but with the benefit of having observed actual events since then. In other words, we can pretend it is roughly 1998/99 instead and create out of sample forecasts for hypothetical wars between Eritrea and Ethiopia (Eritrean-Ethiopian War, 1998–2000) and India and Pakistan (Kargil War, 1999). Using the statistical models above and data collected in the beginning years of these two wars, how accurately would we have predicted war deaths?

The first step is to code some missing data for our independent variable. The Eritrean-Ethiopian War was fought over disputed territories (issue salience = 1, moderate, for both sides) on the countries' arid, mountainous border (rough terrain = 1) and both countries had conventional militaries (strategy = attrition). Ethiopia had both a larger military, with 200,000 against 100,000 troops (parity = 0.67), larger capabilities, and larger population with 50 million compared to Eritrea's 3 million. Neither side had a democratic government. With these inputs, the two base models provide median estimates of 21 and 15 thousand deaths, with a range from the low thousands to 80 and 96 thousand deaths, as shown in Figure 2 (see table 5 for the exact numbers). With the benefit of hindsight, actual deaths in the war were 50 thousand, and both models correctly bracket that number.

The Kargil War between India and Pakistan in 1999 was fought over a Pakistani incursion into mountainous terrain near Kashmir (issue salience = 1, rough terrain = 1). Both states used an attrition strategy. India had a democratic regime and was superior to Pakistan on all measures of capabilities and size (parity = 0.62). The resulting median estimates are 15 and 7 thousand respectively. The observed number of fatalities is between 884 and 4,527+ deaths, which provides a mean estimate of 2705 fatalities.[4] Both models bracket the number of 2705 fatalities, but generally over-predict fatalities in the war.

Notably however, the Kargil War is the only war to date that has been directly fought between nuclear powers. Fighting in the war was restricted to a relatively small geographic area and it did

[4]There are no clear fatality figures available for the Kargil War, mainly because Pakistani casualties are unclear. India claims 527 fatalities of its own, but Pakistan claims anywhere from 357 to 4,000 plus.

not escalate into a general conflict between India and Pakistan. One can speculate that fatalities were low because both sides were reluctant to escalate and risk fighting a nuclear war. Since this is the only war so far fought directly between two nuclear powers, there is no way to take something like mutually assured destruction into account. In an interesting bit of speculation, maybe the predicted fatalities listed in the table are close to the level of fatalities we would have seen in a 1999 Kargil War between Pakistan and India had they not been nuclear-armed.

5 Discussion and conclusion

The statistical models employed here were derived from theoretical bargaining models about conflict. As I argued in the introduction, although they do not directly make arguments about war dead, they do include the cost of conflict. If one can accept that deaths are a significant part of war costs, then by extension these models should have implications for the level of war dead as well. Furthermore, those that deal with concepts like war onset and duration, some of which have been empirically tested in these contexts, also implicitly assume arguments about war dead due to the fact that onset, termination, and duration are typically operationalized using fatality counts.

The positive, significant, and substantively interesting relationship between democracy and deaths is consistent with colloquial and academic arguments about the inability of democracies to tolerate war deaths. The causal mechanism is open to question but two possible explanations are that democracies either end wars that prove to be costly in human life (at least that of their own citizens/soldiers), or that they self-select into only fighting certain wars that are less likely to produce large fatalities. Indeed, democracies appear to fight shorter wars, although they are not more likely to win.[5] Another factor that may explain this result is that democracy, aside from influencing the ability of a state to sustain high fatalities and continue fighting a war, may also help resolve information asymmetries surrounding war through the mechanism of audience costs and the transparency inherent in democratic procedures.[6]

Issue salience for the initiating side is also strongly associated with fatalities. Although it is probably difficult to distinguish which side in a conflict will initiate a war beforehand, the finding does suggest that particular attention should be paid to evolving conflicts between states that involve highly salient issues on either side. Further, the role that issue salience plays in prolonging wars fought over commitment problems suggests that third-party mechanisms to address these problems, like monitoring or enforcement, may be particularly apt for avoiding the escalation of such conflicts into full-scale war.

Variables related to information asymmetry, i.e. parity in observable capabilities, were less clearly related to war deaths. There is some indication that parity in military personnel increases deaths, but using CINC scores there is no relationship. Empirically, this might be because CINC scores are a more complex measure of military strength. Theoretically however the absence of a clear statistical relationship could be because the effect of parity is conditional on other factors like relative war costs.[7] and the extent of commitment problems.[8]

[5] Bennett and Stam III (1996); Slantchev (2004)

[6] Fearon (1994); Schultz (1999); Slantchev (2010)

[7] Wittman (2009)

[8] Wolford, Reiter and Carrubba (2011)

Overall, the coefficients and statistical significance proceeded by these bargaining-derived statistical models overall thus show limited evidence for relationships between bargaining conflict theories and fatalities. But looking at how these models match reality through in sample fit and out of sample forecasts does show that in the aggregate bargaining theories provide a useful starting point for specifying statistical models of war dead. Underlying this point is the fact that statistical significance can be as much a result of sample size as of a true association, and that by themselves, significant coefficients do not mean that a model matches the real world well.[9] The fact that the model predictions here do seem to be accurate and useful for forecasts, compared to the standards of conflict research on war onset for example, speaks to their usefullness to both theory and applied research.

From both an academic and future policy perspective, the most important shortfalls of this effort are that war dead are aggregated and do not distinguish among combatants, and that the model predictions are broad, typically spanning 1 or 2 orders of magnitude. Both of these issues are driven by lack of finer-grained data. Including lower-scale conflicts and disputes would do much to reduce uncertainty, assuming the same mechanisms were at work in producing deaths as in large-scale wars. Practical issues with missing and inaccurate data on fatalities makes this a difficult, but possible, future effort.

Despite these shortcomings, from an applied perspective, being able to forecast war dead even if only roughly is useful. Faced with a potential conflict, where would one even begin in predicting the likely human toll? This effort provides a reasonably accurate starting point that can be used to anchor more educated guesses specific to any given case.

[9] Schrodt (2010)

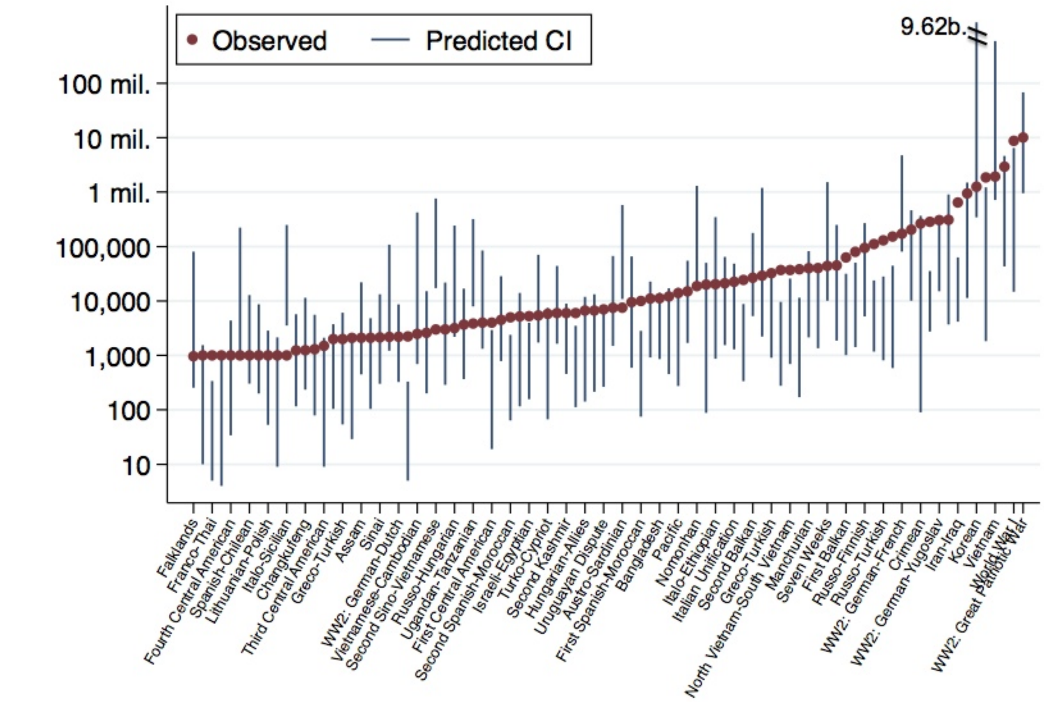

Figure 1: Predicted and observed war deaths for the best-fitting model. The x-axis lists the wars in the sample, with the observed number of war deaths indicated by the red dot. The list is ordered by the total number of deaths, ranging from the Falklands war to the eastern front of World War 2. The number of war deaths is on a logarithmic scale. The blue bars show the interval containing 95% of the predicted war deaths in each bootstrap iteration. The intervals correctly bracket observed deaths for 69% of all wars.

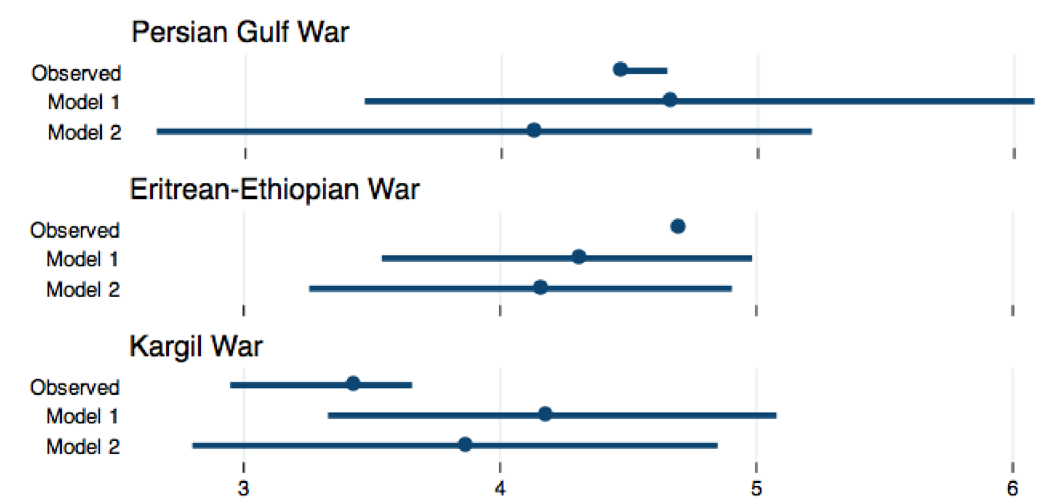

Figure 2: Predicted and observed war deaths for selected wars. The predictions for the Eritrean-Ethiopian and Kargil wars are out of sample forecasts. Bars show low and high estimates for observed fatalities and 95% range for model predictions. Points show best or median estimate for observed and predicted fatalities respectively. The scale is logarithmic with base 10, so for example $10^3 = 1,000$ deaths.

Table 1: Variable descriptions and sources

Variable	Indicator	Source
War deaths	Number of military battle deaths.	(Lacina and Gleditsch, 2005)
Parity military personnel	Ranges from 0 (one side dominant) to 1 (equally strong). Calculated using active duty military strength for all combatants at beginning of the war.	(Singer, 1987; Singer, Bremer and Stuckey, 1972)
Parity CINC	Ranges from 0 (one side dominant) to 1 (equally strong). Calculated with composite index of national capabilities (CINC), which itself uses military personnel, military expenditures, and industrial capacity as inputs.	(Singer, 1987; Singer, Bremer and Stuckey, 1972)
Number of states	Number of states that fought in the war.	(Sarkees and Schafer, 2000; Dupuy and Dupuy, 1986)
Rough terrain	Ranges from 0 (open country) to 1 (rough).	(Slantchev, 2003*a*; Stam III, 1999)
Democratic combatant	0 or 1. For one of the major combatants, was Polity Score (democracy-autocracy) >6 at start of war?	(Marshall and Jaggers, 2002)
Issue salience	2 if state/regime survival, 1 if territory or ideology, 0 if maintaining empire, commercial, policy disputes	(Slantchev, 2004)
Strategy	1 if at least one side used maneuver or punishment strategy, 0 otherwise; reference is pure attrition war	(Bennett and Stam III, 1996; Dupuy and Dupuy, 1986)
Total military personnel	Sum of military personnel for all combatants.	(Singer, 1987; Singer, Bremer and Stuckey, 1972)
Total CINC	Sum of all CINC scores.	(Singer, 1987; Singer, Bremer and Stuckey, 1972)
Parity population	Parity in total population.	(Singer, 1987; Singer, Bremer and Stuckey, 1972)
ln Total population	natural log of sum of each combatant's population.	(Singer, 1987; Singer, Bremer and Stuckey, 1972)
Contiguity	0 or 1	(Singer, 1987; Singer, Bremer and Stuckey, 1972)
Predicted duration	War duration in days from duration model (table 2).	

Table 2: Log-normal regression of war duration

	β	95% CI
Parity military personnel	1.570**	(0.483, 2.741)
Total military personnel^{a}	-1.138	(-0.437, 0.179)
Parity total population	-0.947	(-2.085, 0.182)
Total population^{a}	0.134	(-0.218, 0.492)
Number of States	0.118*	(-0.008, 0.266)
Rough terrain	3.186**	(2.009, 4.433)
Democratic combatant	-0.201	(-0.813, 0.412)
Constant	2.305*	(-0.423, 4.791)
σ	1.247**	(1.168, 1.406)
No. observations	90	
Wald χ^2	72.30**	

Notes: ^{a} natural log used. Log-normal duration model of war duration in days as dependent variable. Bootstrapped bias-corrected accelerated (BC_A) confidence intervals. Significance levels (two-tailed): * $p \leq 0.10$, ** $p \leq 0.05$.

Table 3: Truncated normal regression of war fatalities

	Model 1			Model 2		
	β	exp(β)	95% CI	β	exp(β)	95% CI
Parity military personnel	2.494**	12.11	(0.690 , 6.142)			
Parity CINC				2.165		(-0.560, 5.455)
Number of States	0.290**	1.34	(0.025 , 0.663)	0.190		(-0.116, 0.389)
Rough terrain	2.762*	15.83	(-0.007 , 8.404)	1.934		(-1.465, 5.454)
Democratic combatant	-1.570**	0.21	(-3.081 , -0.135)	-1.727**	0.18	(-3.112, -0.556)
Issue salience, initiator	1.639**	5.15	(0.723 , 2.787)	1.975**	7.21	(0.963, 3.048)
Issue salience, target	-0.039		(-1.207 , 0.974)	0.102		(-0.894, 1.132)
Maneuver strategy	0.787		(-0.787 , 3.010)	1.114		(-0.328, 3.056)
Punishment strategy	1.589		(-1.427 , 5.468)	1.614		(-1.659, 4.775)
Total military personnel	0.0002*	1.0002	(-0.0001, 0.0005)			
Total CINC				6.493**	660.50	(2.506, 12.733)
Parity total population	-1.644*	0.19	(-4.823 , 0.249)	-2.404*	0.09	(-5.398, 0.185)
ln Total population	0.464*	1.59	(-0.028 , 1.048)	0.345		(-0.120, 0.851)
Contiguity	-0.336		(-2.532 , 0.988)	0.402		(-1.368, 1.901)
Predicted duration	-0.001		(-0.007 , 0.001)	-0.0004		(-0.005, 0.002)
Constant	-0.773		(-9.631 , 4.928)	0.265		(-6.602, 6.045)
σ	1.826**		(1.642 , 2.503)	1.785**		(1.594, 2.324)
No. observations		90			90	
Wald χ^2		63.79**			75.22**	

Notes: Dependent variable is ln of fatalities. Bootstrapped BC_a confidence intervals. Significance levels (two-tailed): * $p \leq 0.10$, ** $p \leq 0.05$.

Table 4: Variable significance

Coefficient	1	2	3	4	5	6	7	8	9	10	11	12
	Base		Negative Binomial		Actual Duration		Dummies for Issue		Excluding WW2		Excluding WW1/WW2	
Milper parity	++		++				++		+		+	
Total milper	+						++					
CINC parity				+				+		+		
Total CINC		++		++				++				
Pop parity	−	−		−−				−−	−−	−−	−	−
Total pop	+		++	−	++	++			++		++	++
Number of states	++		++				++	++	++	+	++	++
Terrain	+		++	++			++	+				
Dem	−−	−−	−−	−−	−−	−−	−−	−−	−−	−−	−−	−−
Salient 1	++	++	++	++	++	++	++	++	++	++	++	++
Salient 2												
Maneuver					+	++						
Punishment									+		+	
Contiguous												
Duration				−	++	++						++
Observations	90	90	90	90	90	90	90	90	80	80	79	79
Fit	69.89	65.59	54.84	58.06	66.67	63.44	72.04	63.44	64.52	67.74	62.37	68.82

Notes: Signs denote whether a coefficient had a significant positive or negative effect, with one sign for 90% significance and two for 95% significance, e.g. ++ denotes a positive effect significant at 95% level. Significance levels are derived from bootstrapped BC_a confidence intervals.

Table 5: Model predictions for selected wars

	Lower bound[a]	Median estimate	Upper bound[a]
Persian Gulf War, 1991; *observed: 29,171 (28,945 to 44,271)*			
1	2,219	45,685	1,198,118
2	450	13,435	162,274
Eritrean-Ethiopian War, 1998–2000; *observed: 50,000 (no interval)*			
1	3,450	20,617	96,032
2	1,797	14,543	80,202
Kargil War, 1999; *observed: 2,705 (884 to 4,527+)*			
1	2,131	15,146	119,597
2	630	7,382	70,669

Notes: [a] 95% confidence interval using percentiles.

References

Bennett, D. Scott and Allan C. Stam III. 1996. "The duration of interstate wars, 1816-1985." *American Political Science Review* 90(2):239–257.

Bueno de Mesquita, Bruce and Randolph M. Siverson. 1995. "War and the survival of political leaders: a comparative study of regime types and political accountability." *American Political Science Review* pp. 841–855.

Carson, J.L., J.A. Jenkins, D.W. Rohde and Mark A. Souva. 2001. "The impact of national tides and district-level effects on electoral outcomes: the US Congressional elections of 1862-63." *American Journal of Political Science* 45(4):887–898.

Cederman, Lars-Erik. 2003. "Modeling the size of wars: from billiard balls to sandpiles." *American Political Science Review* 97(1):135–150.

Cioffi-Revilla, C. 1991. "On the likely magnitude, extent, and duration of an Iraq-UN war." *Journal of Conflict Resolution* 35(3):387.

Clauset, A., C.R. Shalizi and M.E.J. Newman. 2007. "Power-law distributions in empirical data." *Arxiv preprint arXiv:0706.1062* .

DeMaris, A. 2004. *Regression with social data: modeling continuous and limited response variables.* Vol. 417 Wiley-Blackwell.

DiCiccio, Thomas J. and Bradley Efron. 1996. "Bootstrap confidence intervals." *Statistical Science* 11(3):189–212.

Duan, N. 1983. "Smearing estimate: a nonparametric retransformation method." *Journal of the American Statistical Association* 78(383):605–610.

Dupuy, R.E. and T.N. Dupuy. 1986. *The encyclopedia of military history: from 3500 BC to the present.* 2nd ed. New York: HarperCollins.

Fearon, James D. 1994. "Domestic political audiences and the escalation of international disputes." *American Political Science Review* 88(3):577–592.

Fearon, James D. 1995. "Rationalist explanations for war." *International Organization* 49(3):379–379.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, insurgency, and civil war." *American Political Science Review* 97(1):75–90.

Filson, Darren and Suzanne Werner. 2002. "A bargaining model of war and peace: Anticipating the onset, duration, and outcome of war." *American Journal of Political Science* 46(4):819–837.

Filson, Darren and Suzanne Werner. 2004. "Bargaining and fighting: the impact of regime type on war onset, duration, and outcomes." *American Journal of Political Science* 48(2):296–313.

Gartner, S.S. 2008. "The multiple effects of casualties on public support for war: an experimental approach." *American Political Science Review* 102(01):95–106.

Gartzke, E. 1999. "War is in the error term." *International Organization* 53(3):567–587.

Greene, HW. 2008. *Econometric Analysis.* 6th ed. Upper Saddle River, NJ: Prentice Hall.

Hensel, P.R., Sara McLaughlin Mitchell, Thomas E. Sowers and Clayton L. Thyne. 2008. "Bones of contention: comparing territorial, maritime, and river issues." *Journal of Conflict Resolution* 52(1):117–143.

Holsti, K.J. 1991. *Peace and war: armed conflicts and international order, 1648-1989.* Vol. 14 New York: Cambridge University Press.

Lacina, Bethany. 2006. "Explaining the severity of civil wars." *Journal of Conflict Resolution* 50(2):276.

Lacina, Bethany and Nils Petter Gleditsch. 2005. "Monitoring trends in global combat: a new dataset of battle deaths." *European Journal of Population/Revue Européenne de Démographie* 21(2):145–166.

Leventoğlu, B. and Branislav L. Slantchev. 2007. "The armed peace: a punctuated equilibrium theory of war." *American Journal of Political Science* 51(4):755–771.

Levy, Jack S. and T. Clifton Morgan. 1984. "The frequency and seriousness of war: an inverse relationship?" *Journal of Conflict Resolution* 28(4):731–749.

Marshall, Monty G. and K. Jaggers. 2002. "Polity IV Data Set." Available online.
URL: *http://www.systemicpeace.org/polity/polity4.htm*

Pillar, P.R. 1983. *Negotiating peace: war termination as a bargaining process.* Princeton, NJ: Princeton University Press.

Powell, Robert. 1996. "Stability and the distribution of power." *World Politics* 48:239–267.

Powell, Robert. 2004. "Bargaining and learning while fighting." *American Journal of Political Science* 48(2):344–361.

Powell, Robert. 2006. "War as a commitment problem." *International Organization* 60(1):169–203.

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

Reiter, Dan. 2003. "Exploring the bargaining model of war." *Perspective on Politics* 1(1):27–43.

Sarkees, M.R., F.W. Wayman and J. David Singer. 2003. "Inter-state, intra-state, and extra-state wars: a comprehensive look at their distribution over time, 1816–1997." *International Studies Quarterly* 47(1):49–70.

Sarkees, M.R. and P. Schafer. 2000. "The Correlates of War data on war: an update to 1997." *Conflict Management and Peace Science* 18(1):123–144.

Schrodt, Phil A. 2010. "Seven deadly sins of contemporary quantitative political analysis." Paper presented at the 106th Annual American Political Science Association (APSA) Meeting and Exhibition, Washington, DC.

Schultz, K.A. 1999. "Do democratic institutions constrain or inform? Contrasting two institutional perspectives on democracy and war." *International Organization* 53(2):233–266.

Singer, David J. 1987. “Reconstructing the Correlates of War Dataset on material capabilities of states, 1816–1985.” *International Interactions* 14(2):115–132.

Singer, J. David, Stuart A. 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 pp. 19–48.

Slantchev, Branislav L. 2003*a*. “The power to hurt: costly conflict with completely informed states.” *American Political Science Review* 97(1):123–133.

Slantchev, Branislav L. 2003*b*. “The principle of convergence in wartime negotiations.” *American Political Science Review* 97(4):621–632.

Slantchev, Branislav L. 2004. “How initiators end their wars: the duration of warfare and the terms of peace.” *American Journal of Political Science* 48(4):813–829.

Slantchev, Branislav L. 2010. “Feigning weakness.” *International Organization* 64(3).

Smith, A. 1998. “Fighting battles, winning wars.” *Journal of Conflict Resolution* 42(3):301.

Stam III, Allan C. 1999. *Win, lose, or draw: domestic politics and the crucible of war*. Ann Arbor, MI: University of Michigan Press.

Wagner, R. Harrison. 2000. “Bargaining and war.” *American Journal of Political Science* 44(3):469–484.

Wittman, Donald. 2009. “Bargaining in the shadow of war: when is a peaceful resolution most likely?” *American journal of political science* 53(3):588–602.

Wolford, Scott., Dan Reiter and Clifford J. Carrubba. 2011. “Information, commitment, and war.” *Journal of Conflict Resolution* 55(4):556–579.