

# **PUNCTUATED EQUILIBRIUM, REALISM, AND STALEMATE IN INTERNATIONAL RIVALRIES**

PAUL F. DIEHL

University of Illinois at Urbana-Champaign

and

GARY GOERTZ

University of Arizona

An earlier version was presented at the Annual Meeting of the American Political Science Association, San Francisco, 2001. This research is supported in part by a grant from the National Science Foundation (SES-0004279). The authors would like to thank Brad Jones for his advice.

# **PUNCTUATED EQUILIBRIUM, REALISM, AND STALEMATE IN INTERNATIONAL RIVALRIES**

## Abstract

The vast majority of all militarized disputes end without either side prevailing. With rare exceptions, the international conflict literature has ignored the numerous militarized disputes in which the outcomes fall short of war. We explore two potential explanations of stalemates in international conflict. One relies on a realist, power-based logic: if the two sides are relatively equal, the match might end in a tie. The punctuated equilibrium approach to international conflict provides a second theoretical perspective on the prevalence of stalemate. The results consistently suggest that realism's emphasis on prevailing through power preponderance is limited in its scope. Power preponderance along with positive expected utility had no impact on the vast majority (70-80%) of militarized disputes; even when states had a significant military advantage over their opponents, most disputes still ended in stalemates. Realist factors were better able to predict outcomes in the subset of the most severe disputes, including full-scale wars. Overall, the results were more supportive of the punctuated equilibrium model. Stalemates became more common as the rivalry matured, and capabilities and other variables only attained significance once their rivalry context was factored in.

## **PUNCTUATED EQUILIBRIUM, REALISM, AND STALEMATE IN INTERNATIONAL RIVALRIES**

When US President George Bush launched the Persian Gulf War and achieved a lightening victory, few expected that over a decade later the US would still be engaged in a protracted military conflict with Iraq, with no end in sight. The US-Iraq rivalry illustrates many of the key issues we consider in this paper. Although the new systematic data are not yet available, it is clear that the US-Iraq rivalry has involved a large number of militarized incidents since 1991. It is also just as evident that with the exception of the war itself, all or almost all of these confrontations have ended in stalemate.

As we note below, the US-Iraq rivalry is not unique in its large preponderance of stalemate outcomes (i.e., neither side is able to gain satisfaction and the dispute remains unresolved): the vast majority of all militarized disputes end without either side prevailing as well. This poses a puzzle for theories of international conflict: how are we to explain this predominance of stalemates? Are there theoretical approaches to international conflict that predict and explain stalemate outcomes? Unfortunately, with rare exceptions (e.g., Maoz 1983), the international conflict literature has ignored the numerous militarized disputes in which the outcomes fall short of war. Even then, the focus is most often with which side wins or loses. For most scholarly analyses, stalemate outcomes are of little interest, ignored or treated as unwelcome noise in models of victory/defeat.

In this paper we explore two potential explanations of stalemates in international conflict. One relies on a realist, power-based logic (actually we look at different variations of this approach). It regards stalemates in disputes as similar to draws in sporting events. If the two sides are relatively equal, the match might end in a tie. Analogously, if the two states are evenly matched in power (or more specifically military capabilities), then stalemates might be quite common. Nevertheless, the US-Iraq example indicates that repeated stalemates can occur between countries of significantly different sizes and capabilities.

The punctuated equilibrium approach to international conflict (Diehl and Goertz, 2000) provides a second theoretical perspective on why we would expect to see stalemates as frequent outcomes of crises and disputes. As is well-known, a small percentage of dyads (enduring rivalries) accounts for a large percentage of militarized disputes. Stalemates in enduring rivalries make a great deal of sense within the punctuated equilibrium framework. A stalemate in a dispute at time  $t$  means that neither side prevails, and therefore neither has achieved its goals. Hence, we would expect a new dispute might well arise at time  $t+1$  and the rivalry continues

because the issues were not resolved by the previous dispute. If most disputes end with the victory of one side or a negotiated compromise, then we would not expect to see many recurring disputes and enduring rivalries.

In short, we examine two plausible explanations for the existence of stalemate outcomes in militarized disputes and wars. We do so by examining the outcomes almost 3000 dyadic militarized disputes embedded in 1166 rivalries over the period 1861-1992. The realist explanation claims that stalemate arises from power symmetry. The punctuated equilibrium framework provides another take on the question. In that framework, because issues are not easily resolved, stalemates result. This produces more stalemates in the future, and hence the existence of a substantial number of enduring rivalries. We begin with a discussion of the realist and punctuated equilibrium approaches, with special attention to how they account for dispute outcomes, and stalemates in particular.

## **THEORETICAL APPROACHES**

### **REALISM**

Legro and Moravcsik (1999) presented a strong case for what we consider a minimal realist position. They argue that a realist theory of international politics consists of three assumptions, which are necessary and sufficient to define realist theory:

Assumption 1 -- the nature of the actors: rational, unitary political units in anarchy.

Assumption 2 -- the nature of state preferences: fixed and uniformly conflictual goals.

Assumption 3 -- international structure: the primacy of material capabilities.

Material capabilities or power is the only variable in the Legro and Moravcsik view of realism – literally, it is the only factor that “varies” over time and space, as the other two assumptions involve constant factors. We shall call realism that uses only material capabilities as the “minimal” realist hypothesis about stalemates. Thus, in this view, power determines conflict outcomes. To some extent, this is tautological (if a state “wins” a dispute, by definition it was more powerful in that context). Yet, more broadly realists expect that the winners and losers of militarized confrontations can be predicted by reference to each side’s material capabilities. In realist formulations (e.g., Claude, 1962; Kissinger, 1979), power is essentially formulated as equivalent to military capability. Quite simply, the stronger side should win militarized encounters. Stalemates should occur when neither side has a capability advantage. This is quite similar to many military models of war-fighting (e.g., Lanchester models) or economic models of conflict outcomes (e.g., Hirshleifer, 2000) in which capabilities are the dominant influence.

There are a number of caveats to this formulation. First, even the most dedicated realist would acknowledge that dispute or crisis outcomes are influenced by a number of factors, some unanticipated at the outset of the confrontation. Indeed, there is a significant literature on how “big states lose little wars” (Mack, 1975; Arreguin-Toft, 2001); still, even these articles treat such cases as exceptions and such research focuses only on full scale wars. Thus, at best, capabilities are unlikely to predict outcomes in *all* cases, although they should be the primary influence. Second, there are a number of modifications necessary to get a full picture of how a state transfers its capabilities into actual power; some of these include the “loss of strength gradient” (Boulding, 1962), society’s willingness to suffer (Rosen, 1992), and the state’s extractive capacity (Organski and Kugler, 1980). Third, there may be a selection effect operating. That is, certain capability distributions between states may make confrontations more or less likely to occur. Specifically, one might argue that when a disagreement arises between a very strong state and a considerably weaker counterpart, the conflict will be resolved through non-militarized means, as the strong state can use other forms of coercion to get its way and the weak state will be reluctant to challenge its stronger rival in a military confrontation it is highly likely to lose.<sup>1</sup> Still, for the purposes of this study, we consider only the simple formulation that capabilities determine outcomes, while recognizing the limitations of this approach.

With respect to dispute outcomes, a minimal realist model does not necessarily predict one particular outcome to be more common than another. Outcomes depend on the capabilities of the rivals, and therefore one might expect stalemates and victories in frequencies proportional to the capability distributions that produce them. *A priori*, the model doesn’t provide any clues to those distributions. Indeed, we know that international rivalries include pairs of states that are nearly equal in capabilities as well as those with wildly divergent capability profiles. Generally though, the minimal realist model would predict that equal capabilities are associated with stalemate outcomes, as neither side can be expected to defeat the other. It is conceivable that parity would also produce compromises, or outcomes satisfying both sides; yet realpolitik models are notably deficient in explaining cooperation and it is hard to derive a rationale for this particular outcome from realist thinking. If anything, realpolitik might predict few compromises. When a rival is preponderant over another, realpolitik suggests that the stronger side will prevail in the dispute.

The minimal realist model implies a stable distribution of outcomes over time. The outcome of each dispute is context specific, and therefore independent of other disputes, past and future. Nevertheless, we also know that capability distributions between rival states are relatively constant over time; except for a few power transition cases, rivals maintain their relative positions

as well as the rough size of their capability gaps over the life of a rivalry. Thus, outcomes later in a rivalry should be similar to those at the outset.

Finally, we know that disputes vary in their severity, with many ending without a use of military force by either side whereas some escalate to full-scale war. A realist approach might predict that capabilities will have a stronger effect on outcomes as the dispute becomes more severe. For low-level disputes, capabilities will certainly influence resolve, bargaining strategies, and the like. Yet issue salience and other factors will also determine state behavior and the resulting outcomes. Once disputes escalate to the use of force, however, capabilities should be more important in outcomes. Most obviously, capabilities can be translated into military power and used to influence the resolution of the dispute directly. In addition, once military force is used, and particularly if war occurs, then bargaining strategies and other extraneous factors become less significant.

### **Other Realist Formulations**

We agree that the centrality of power as a major explanatory variable distinguishes realist theories. Unlike Legro and Moravcsik, however, we consider any theory that includes power as a central variable, or part of the theoretical core, as realist. For example, the power transition theory is a realist for us because the main variable involves power.<sup>2</sup> Most notable of the realist variants are theoretical approaches that also include a preference component. Legro and Moravcsik's criticisms to the contrary, many versions of realism include power combined with some type of preference variable. For example, Walt (1987) argues that states balance against threat not power; something else, what might be termed relative policy positions or preferences, must distinguish states that are threats from those that are not. It is no longer merely powerful countries that are threats. Similarly, power transition theory (Kugler and Lemke, 2000) has two variables, (1) shift in power and (2) dissatisfaction with the status quo. The latter is the power transition version of the preference variable. According to power transition theory, for example, the US and the EU will not fight a war because both are content with the status quo.

A test of "enhanced" realist models in the context of stalemate outcomes would involve both power and preferences factors. An obvious choice for us is Bueno de Mesquita's expected utility approach (1981). The two core variables in *The War Trap* model and its derivations are power ratios and preferences (alliance portfolio similarities).<sup>3</sup> In the original formulation, Bueno de Mesquita (1981) argued that states initiating war would have a positive expected utility toward their opponents; positive expected utility was a necessary, but not sufficient, condition for war. The same basic logic should apply to disputes short of war: governments should not initiate them unless they expect that it "pays" to do so (Morrow, 1986).

If states correctly project the payoffs from military action (assuming complete information), then a logical corollary to the expected utility model is that most disputes should result in clear victories for the initiator. In reality, we know that predictability of outcomes is less than perfect. Still, one might say that a prediction of the expected utility model is that the predominant outcome is victory by the initiating side, just as the punctuated equilibrium model expects a high number of stalemates. Stalemates in the expected utility framework would seem to be a result of miscalculation; a state bears the risks and costs of initiating military action and yet its demands are not met. At best, a marginally positive expected utility might result in a stalemate, but this may only occur when a state is risk acceptant. A compromise is more puzzling in the expected utility model. If two states could find a middle ground to resolve their disagreement, with perfect information they should have avoided the militarized confrontation and its attendant costs and made the agreement beforehand. Thus, an expected utility model would predict many clear victories, few stalemates, and perhaps no compromises. Even the introduction of significant incomplete information would not necessarily change this distribution radically.

Similar to the minimal realist model, the expected utility or enhanced realist model would postulate no specific patterns over time within a rivalry. Because each dispute and accompanying decisions are unique, there is no memory or history in the process; the outcome of one dispute does not influence subsequent interactions. This is not to say that patterns in outcomes will not appear. They might, but essentially they would be regarded as epiphenomenal. Any patterns would be the result of repeated conditions over time within a given rivalry. The expected utility model provides no rationale for why one rivalry would exhibit similar patterns over time to other rivalries.

## **PUNCTUATED EQUILIBRIUM**

The punctuated equilibrium model of rivalries (Diehl and Goertz, 2000; see also Cioffi-Revilla, 1998) starts with a longitudinal view of international relations, which does not atomize disputes and wars and rip them from their historical context. Fundamentally, the unit of analysis is not the dispute, crisis, or war but the rivalry relationship (which may be long or short). A key aspect of our approach to rivalry is to see linkages between conflicts flowing from an organizational, policy model of decision-making. The punctuated equilibrium model expects repeated conflicts as government lock into conflict policies of various sorts. Hence, within a policy model of international conflict, initial policy decisions are crucial because they shape future decisions in powerful ways. When faced with a problem, organizations and governments

(which we consider as organizations) generally try to devise a policy to deal with it. Internally, the policy model has future and past dimensions. A policy is a commitment -- implicit or explicit -- to act in a certain way in the future. The concept of “precedent” is key in looking backwards. Hence, historical actions provide reasons to act the same way in the present. Core to the character of governments is the need to act consistently over time and space. The policy may prove to be dysfunctional or have negative unexpected consequences. Nevertheless, the government is constrained to follow preestablished policy. As Wilson indicated, “the standard operating procedure (SOP) is not the enemy of organization; it is the essence of organization.” (Wilson 1989: 221). Precedent also works in the forward sense. Because governments use reasoning based on precedent and because they are concerned with consistency, they hesitate to take what are good choices in the present if they set bad precedents for future situations. The inertia character of governments thus supports the basic punctuated equilibrium framework, with its emphasis on stasis most of the time.

Based on the punctuated equilibrium model, there are a number of expectations about the outcomes of rivalry disputes. First and most simply, stalemates should be the predominant outcome (as opposed to compromises, clear victories, and the like in which at least one party is satisfied). The punctuated equilibrium model postulates that rivalry processes are characterized by stasis, which are largely repeated behaviors over time. Thus, we do not anticipate that rivalries will be characterized by oscillations of compromises, victories, or stalemates in some random pattern. Rivalries should roughly exhibit similar patterns of outcomes over time; indeed, Leng (2000) reports that states tend to repeat the same influence strategies from one crisis to the next, so perhaps it is not surprising that outcomes remain the same as well. Diehl and Goertz (2000) demonstrated that over three quarters of enduring rivalries exhibit consistent patterns of behavior over time (with respect to dispute severity); that is, few show random patterns or those of increasing or decreasing hostility over time. Our expectation with respect to dispute outcomes is consistent with that overall pattern.

Thus, we expect consistent outcomes to disputes, but why do we anticipate that those outcomes will be stalemates? If disputes consistently resulted in compromises, one might expect there to be no reason for militarized conflict to occur or continue. Furthermore, the punctuated equilibrium sees rivalries and resulting cooperation as coming abruptly (empirically confirmed in Diehl and Goertz, 2000). Stalemates provide a natural linkage mechanism between disputes. If a dispute ends in stalemate one would not be surprised that another dispute arises. One of the characteristics that make rivalries “enduring” is that there is repeated failure of conflict resolution between the parties; neither side is able to be victorious and an agreed settlement is not

forthcoming. Thus, stalemates in a rivalry beget more stalemates. Under those circumstances, when confrontation persists, we might expect that stalemates are the most likely outcome.

To some extent, the punctuated equilibrium model begs the question of how the first stalemate occurs that sets in motion the repeated process of stalemates and policy inertia. This is usually treated as exogenous to the model. Nevertheless, recent studies give some clues to the origins of this process. Tir and Diehl (2002) find that enduring rivalries are disproportionately composed of disputes over territorial issues. We know from past conflict research that territorial issues are the most conflict prone and often are difficult to resolve. Indeed, Vasquez (2003) argues that when states resort to military force over territory, a highly salient concern, they are less willing to compromise and repeated stalemates are the likely outcomes. Similarly, Stinnett and Diehl (2001) find territorial issues associated with the onset of enduring rivalries; they also find a number of other conditions associated with enduring rivalries and by implication stalemate outcomes; these include major power status, non-democratic dyads, and recently independent states. Thus, the conditions that prompt an initial stalemate, and the policies are underlying it, are related more to the structural aspects of the rivalry rather than with the behavioral elements associated with the dispute itself.

Although the punctuated equilibrium model predicts that stalemates will be the modal outcome, this should not suggest that all disputes will result in stalemates or that no variation exists over the course of a rivalry. The punctuated equilibrium model distinguishes between the initial phases in which policy is established and later phases in which the policy is carried out in the fashion typical of organizations. Rivalry creation occurs as states make major policy changes and lock into conflictual patterns of behavior. At that time, conflict generating policies as well as conflict management ones can be implanted. Depending on the character of the issues and how they are handled, conflict management at the early stage can become conflict resolution (that is, the rivalry ends in the nascent stage). Here it is the question of whether isolated conflicts or proto-rivalries occur or become enduring rivalries or not. The mid-life of a rivalry corresponds to the standard operating procedures phase of the policy life-cycle. Given that rivalries, once established, are resistant to change, conflict management may be neither frequent nor easy to achieve (Bercovitch and Diehl, 1997) during this phase. Rivalry termination is again a period of large policy shifts, followed by another period of stability and less- or non-militarized relations. Thus, one might predict that compromise and other outcomes will be more common in early stages of rivalries. An overwhelming victory by one side or a compromise outcome may lead to the end of a rivalry; that is, the initial outcomes help determine whether a rivalry becomes enduring or not (Stinnett and Diehl, 2001). It is also in the nascent stages of rivalry that policies

are being formulated and the “lock-in” phase of rivalry behaviors and policies are not yet complete; thus, we might expect somewhat more variation in the early disputes in a rivalry and certainly fewer stalemates.

Consistent with this, the punctuated equilibrium model would predict that stalemates would become more frequent over time in a rivalry; that is, the probability of a stalemate should increase with the maturity of a rivalry. Once established, states will find it increasingly difficult to change policy. Recall that the punctuated equilibrium model sees change in rivalry occurring rarely and only in accordance with a significant “political shock” that alters the decision-making environment. If anything, states adopt more coercive bargaining strategies later in protracted conflicts and rivalries (Leng, 1983; Brecher, 1984, Hensel, 1996). Under those conditions, rivals are less likely to pursue cooperative outcomes.

Thus, overall, the punctuated equilibrium model predicts that stalemate outcomes will be quite numerous in rivalries, more frequent than other outcome types. The punctuated equilibrium model, however, also expects that the prevalence of stalemates will vary over the course of a rivalry, less frequent at the outset of the rivalry, but increasingly common as the rivalry matures.

## **EMPIRICAL EVIDENCE**

Few studies have focused on explaining dispute outcomes, and none to our knowledge account for stalemates in particular. Not surprisingly, scholars have been most concerned with war outcomes and the role that capabilities have in predicting victory or defeat. A number of studies (e.g., Stam, 1996) have determined that military capabilities and other realpolitik factors are important elements in determining war outcomes; Bennett and Stam (1998) report similar findings, although they note that capabilities and related measures of power vary over time in their impact on war outcomes. The expected utility model has also received some verification with respect to war (Bueno de Mesquita, 1981; Bueno de Mesquita and Lalman, 1992), but even then, initiators at best win only 2/3 of wars (see Wang and Ray, 1994). It is not clear whether any of the findings on war would be replicated in lower level conflicts. Indeed, Ayres (2000) reports that stalemates increase as the severity level of conflicts decrease.

Studies of conflict outcomes below the war threshold are relatively rare. To the extent that they have dealt with conflict outcomes, their central theoretical focus has been on democracy, and not power-based explanations. A few studies concentrate on explaining compromise outcomes in disputes, generally ignoring other, more frequent outcome types. Mousseau (1998) does not include a measure of relative capabilities in his model of dispute compromises.<sup>4</sup> In a follow-up study, Bremer (2000) finds that relative capability does not have much impact on the

probability of negotiated settlements or compromise outcomes in militarized disputes. Dixon and Senese (2002) reach similar conclusions.

More broadly, Maoz's (1983) study is perhaps the only one that considers a range of outcomes to militarized disputes. Consistent with the findings above, however, he reports that military capabilities had little or no effect on dispute outcomes; rather bargaining "resolve" was better predictor of who won or lost such confrontations. Nevertheless, capabilities nonetheless have a greater impact in wars than in lesser confrontations. Yet Maoz's conclusions were drawn from a small sample (N=164) of disputes from an early version of the COW dispute data set, and he uses only military personnel as an indicator of capability. The former calls into the question that robustness of the findings, while the latter raises questions about the validity of the findings beyond the 19<sup>th</sup> century. Indeed, Diehl (1985) has found that capabilities correlate with outcomes (e.g., stronger side wins more often than chance) in the presence of an arms race.

In summary, previous studies have not directly tested realist propositions with respect to stalemate outcomes. What work has been done suggests the absence of a relationship between power and conflict outcomes, something not in line with general realist expectations. Yet, consistent with realist notions, capabilities may have a stronger impact on the most severe confrontations, evidenced by consistent findings on power and war outcomes.

There essentially have been no tests of the propositions derived from the punctuated equilibrium model. Yet, several studies offer some hints about the validity of that model. Mousseau (1998) notes that previous hostility between states (e.g., earlier disputes or what we would call rivalry) was negatively related to compromise outcomes. Similarly, Brecher (1993) argues that in more protracted conflicts (crises that tend to repeat, analogous to rivalries), there are fewer agreements between protagonists and more ambiguous and stalemated outcomes in individual crises.

## **RESEARCH DESIGN**

### **RIVALRIES AND DISPUTES**

In order to test the predictions made by the different models, we first need a population of rivalries on which to conduct the analysis. We use a set of international rivalries provided by Diehl and Goertz (2000), comprised of 1166 militarized rivalries over the period 1816-1992. The data are constructed from the Correlates of War Project Militarized Dispute (MID) data (Jones, Bremer, and Singer 1996). Diehl and Goertz rely on time-density to identify the connected disputes that comprise each particular rivalry. Analyzing all possible dyads from 1815 to 1992, dyadic disputes are considered to be part of the same rivalry if they are no more than 15 years

apart. A dispute is considered to be part of a rivalry if it occurs within 11 years of the first dispute in that rivalry, 12 years after the second, and so on through 15 years after the fifth (Diehl and Goertz 2000). This is similar to coding efforts made for events within the same crisis, albeit there in a narrower time frame (e.g., see Brecher and Wilkenfeld 1997).

Within each rivalry, the unit of analysis is the dyadic militarized dispute (N=2949), whose outcome we will analyze. This generally corresponds to Maoz's dyadic dispute dataset in which multilateral disputes are decomposed into a series of dyadic ones according to a set of rules (see Diehl and Goertz 2000 for a complete explanation).

We will also focus on the location of a given dispute in a rivalry sequence; that is, we identify whether a dispute occurs in the infancy (first three disputes), adolescence (fourth through sixth disputes), mid-life (seventh through twelfth disputes), or the final life phase (the thirteenth dispute and beyond) of a rivalry as per some of our theoretical expectations.

## **DISPUTE OUTCOMES AND STALEMATES**

Our next step is to identify the occurrence of stalemate outcomes that occur for the disputes within our rivalries. We turn to the COW militarized dispute data set (Jones, Bremer, and Singer, 1996). The COW data set provides information on what is referred to there as the "outcome" of the dispute. This variable has nine categories of dispute outcomes: (1) Victory for Side A, (2) Victory for side B, (3) Yield by side A, (4) Yield by side B, (5) Stalemate, (6) Compromise, (7) Released, (8) Unclear, and (9) Missing.<sup>5</sup>

One way to think about this scheme is in reference to the number of satisfied parties in each of these categories. Compromise usually indicates that both sides achieved some goals and are relatively happy. Victories, as well as the two categories of yield, indicate that one party is reasonably content with the outcome while the loser/yielder is not. Stalemate signifies that neither side is especially satisfied with the results. To the extent that there is a clear initiator and target, the target maybe more content, but a stalemate indicates that it has not permanently pushed back the challenge. Similarly, with the relatively special category of released (exclusively for fishing boat and related seizures), this rarely indicates that the underlying reasons for the seizure have been successfully dealt with. An examination of the unclear and missing cases also revealed that they too involve unresolved disputes.

We collapse these outcomes into three basic categories, according to the number of satisfied parties:

*Stalemate* -- no satisfied parties -- stalemates, released, unclear, and missing

*Victory* -- one satisfied party -- victories or yields

*Compromise -- two satisfied parties, -- compromises*

## **OTHER VARIABLES**

In order to test expectations from the expected utility and simple realpolitik models, we need additional variables. For analyses of power, we examine both pure military capability as well as an economic measure of capability, relying on data from the Correlates of War Project (Singer, Bremer, and Stuckey, 1972).<sup>6</sup> Military power is of direct relevance in militarized disputes and war; hence, we use military personnel and military expenditure indicators to signify military capabilities. Economic capabilities are represented by the COW economic indicators: iron/steel production and energy consumption; these are strongly correlated with various demographic, military, and other economic indicators, providing a basis for assessing the strength of the state. Each of these four individual indicators has theoretical and practical advantages and disadvantages. For example, iron and steel production is a good measure of industrial strength for large countries from about 1870-1950, but its value is more limited in the last few decades and for many Third World countries with no production of these products. Military personnel works well as a measure of military power until the latter part of the 19th century, but becomes less valuable thereafter. Military expenditures are more valid across the entire 1816-1992 time period, but they involve the very tricky problems of exchange rates for different currencies.

Since we are dealing with dyads, we take a simple ratio of the indicator concerned where the largest value of the dyad is put into the numerator. Because we are concerned with outcomes, we take an average of the two indicators for each dimension (economic and military) respectively for the last year of the dispute. This means that all relative power indicators have a minimum of 1.00. We adopt the common dichotomization of power relations into “preponderance” or “parity,” with the former defined as values greater than 3.00. This corresponds to the conventional 3:1 rule (Epstein, 1989; Mearsheimer, 1989) used by scholars and policymakers to signify a meaningful military advantage by one side in a conflict. We also considered different thresholds (i.e., 1.5:1, 2:1; 5:1), but found few, if any, significant changes in the results reported. We initially distinguished between bilateral and multilateral disputes. Because the rivalry data breaks up multilateral disputes into a series of dyadic ones, the relative power ratios may not be reflective of actual dispute dynamics. As a result, we conduct analyses on (1) all dyads, including multilateral disputes, and (2) bilateral-only disputes and wars, in order to assess whether the multilateral disputes confound the analyses. Yet, except where noted otherwise, we report only the results of using the former given that there was little difference in the results across the two

sets. For the multivariate logit analyses, we included a variable to account for multilateral disputes, employing the number of parties to the dispute.

We use expected utility as defined by Bueno de Mesquita (1981). These data (as generated by EUGene – Bennett and Stam, 2000) express the expected utility of conflict for one state toward the other and vice versa. Where relevant, we look at the initiator's expected utility, with the initiator defined as the state that was on "Side A" and an initial participant in the dispute according to the MID 2.1 data set (Jones, Bremer, and Singer, 1996).<sup>7</sup> We tested several variations of the basic expected utility score including those used in Bueno de Mesquita and Lalman (1992) and Signorino and Ritter (1999). The results were not substantially changed, and thus we report the results below only with the simple, original measure of expected utility. For our analyses, we sought to distinguish between strongly negative/positive expected utility score and ones that were close to zero. Accordingly, a positive score was defined as  $> 0.5$ , a negative score as  $< -0.5$ , and all other scores ( $-0.5 < X < 0.5$ ) defined as zero expected utility. For the concluding logit analyses, we used the continuous measure.<sup>8</sup>

For analyses of severity, we adopt the Diehl and Goertz (2000) approach, which scales severity on a 0-200 scale based on the "level of hostility" exhibited by both states in the dispute and the number of fatalities in the confrontation. For some analyses, we divided the cases into four categories along this scale, based on inductively derived breakpoints. We also use a modified version of the standard Small and Singer (1982) definition of war as a conflict that resulted in 1,000 or more battle-related fatalities.<sup>9</sup>

For the logit analyses, we use democracy as a control, and include a variable indicated by the maximum democratic score of the two states in the dispute. Note that we do not use the common "weakest-link" or minimum democracy score for the dyad. The literature finds that democracies are likely to win the wars that they fight. Hence, if a dyad includes a democracy – indicated by the maximum of the two democracy scores – then we would be more likely to see one side, the democracy, prevail. Democracy is measured using the Polity 4 (<ftp://isere.colorado.edu/pub/datasets/p4/p4vksg.asc>) scale of  $-10$  (authoritarian) to  $+10$  (democracy) and then rescaled to 0-20 for the statistical analyses. We also inserted a variable on whether the dispute in question involved territorial issues or not. The territory variable is based on the MID data set coding of "revision type." Each case is coded as "1" if at least one of the states in the pair was seeking a territorial revision and "0" otherwise.

## EMPIRICAL RESULTS

### GENERAL PATTERNS IN DISPUTE OUTCOMES

Our first concern is with the general pattern of dispute outcomes. Each of the models has different predictions on the frequency of different kinds of outcomes, specifically stalemates (none satisfied), victories (one satisfied), and compromises (both satisfied). Table 1 offers those results.

(Table 1 about here)

Immediately evident is the predominance of stalemate or indecisive outcomes; they constitute over two-thirds of the cases (68%).<sup>10</sup> This is consistent with expectations of the punctuated equilibrium model. Only two full-scale wars end in a stalemate, but the vast majority of lower level conflicts end without either side achieving a significant portion of its goals. Victories by one side occur only in approximately 26% of the cases, far fewer than might be suggested by either the minimal or enhanced realist models. Consistent with all explanations, however, the number of compromise outcomes is remarkably small, barely more than 6% of the cases; thus, when states face each other in a militarized confrontation, it is rare that both sides leave the conflict satisfied.

The results in Table 1 replicate what Jones, Bremer, and Singer (1996) reported in their presentation of the MID dataset. Yet no one has gone beyond these basic findings to examine the puzzle posed by the predominance of stalemate outcomes.<sup>11</sup> The results in Table 1 strongly support the punctuated equilibrium model, but the two realist models may provide alternative explanations for those patterns. Two expectations arise from realist models. The first is that we should expect to see few disputes between countries of asymmetric power. Indeed, "(m)ost recent research supports the view that preponderant power is more likely to discourage military conflict than is a balance of national capabilities." (Russett and Oneal, 2001: 88) Following from this is the second expectation that the high stalemate rate is a result of power symmetry of the rivals: neither state can impose its will on the other and the confrontation ends in a tie.

To test these assertions, we first look at the minimal realist model and attempt to predict dispute outcomes based on the capabilities of the two sides. Table 2 provides those results, for both military and economic capabilities.

(Table 2 about here)

Most notable at the outset is that a clear majority of disputes (approximately two-thirds of the cases, according to either indicator) involve disputants with disparate capabilities.<sup>12</sup> This suggests immediately that the minimal realist model will not be able to account for the predominance of stalemates, and indeed that is the case. The stalemate rate is very consistent

across different capability distributions. It makes little difference whether the initiating side or the target rival is stronger; most disputes still end in stalemates. In the cases in which a victory does occur, the stronger side is more likely to win, sometimes more than twice as likely. This does support realist conceptions, but one must balance this admittedly modest finding against the prevalence of stalemates when the stronger side should have won according to the model. Put another way, the minimal realist model only correctly predicts victories by the stronger side at best 20% of the time.

A modification of the minimal realist approach, the expected utility or enhanced realist model, is no more successful in accounting for the prevalence of stalemates. This model predicts that rational states will initiate disputes when it is in their interests and when they expect to prevail. Yet the first clue that the expected utility may not be successful in predicting outcomes is the finding that only 28% of initiators have a positive expected utility score; almost 43% actually had a negative score. Bueno de Mesquita (1981) found that the vast majority of wars were initiated by states with positive expected utility. Such is clearly not the case, however, for militarized disputes at lower levels. More importantly, even when initiators have a positive expected utility, they are victorious only 17% of the time. Paradoxically, victory is most likely in the presence of a *negative* expected utility score. Thus, neither of the central elements of the expected utility model's predictions on dispute outcomes are borne out: empirically, states don't always (or even commonly) initiate dispute when they can expect victory, and even in the limited instance in which this is the case, victory is actually relatively rare.

(Table 3 about here)

Overall, the general patterns of dispute outcomes support the punctuated equilibrium model. In addition, we find that a large percentage of disputes are initiated by governments with negative expected utility. This combined with a majority of disputes between power asymmetric dyads means that the various realist models have not empirically lived up to expectations. Thus, not only has the punctuated equilibrium model performed well in predicting a large percentage of stalemates, but the realist models have performed poorly in their counter-arguments. We turn next to other hypothesized patterns in dispute outcomes.

## **PATTERNS IN DISPUTE OUTCOMES OVER TIME**

In this section, we are concerned with whether stalemates are more common at different stages of a rivalry. Realist models, along with Gartzke and Simon (1999), do not see victory or stalemate depending on the position of the dispute within the rivalry life-cycle. In contrast, the punctuated equilibrium model sees a dramatic difference between the initial disputes of a rivalry

and later ones (evolutionary models of rivalry – Hensel, 1999 -- would make similar claims). The vast majority of rivalries die out after the first or second dispute. According to the punctuated equilibrium model, we would expect to see a larger percentage of victories in these initial disputes. Once the rivalry is established, however, we should see that stalemates become more frequent. As the rivalry progresses governments learn how to avoid defeat.

(Table 4 about here)

Table 4 indicates that the likelihood of stalemates increases the later in the rivalry that a dispute occurs.<sup>13</sup> Even in the first few disputes, stalemates are the most likely outcome, occurring in more than two-thirds of the cases. Yet, the likelihood of a stalemate increases gradually until over 85% of disputes end in stalemate during the advanced phase of enduring rivalries. These findings are again consistent with the predictions of the punctuated equilibrium model. That model argues that stasis forces strengthen over the course of a rivalry, and it is increasingly difficult for states to break out of their patterns of interaction.

The predictive capacity of the capability distribution is never strong with respect to dispute outcomes. Still, it is strongest at the outset of a rivalry. The success rate of the stronger side declines by over 87% with respect to military capability and almost 54% with respect to economic capability when a dispute occurs in the latest stage of a rivalry as compared to the earliest phase. In the advanced stage of rivalry, the stronger side is virtually never successful in militarized confrontations. This is also consistent with the prediction of the punctuated equilibrium model that the lock-in phase of rivalries will produce a broader variety of outcomes and some clear victories will precipitate the end of the rivalry. The findings on expected utility mirror those with respect to capability. States with positive expected utility have a higher success rate – more than three times as great – at the outset of a rivalry as compared to its tail end. Nevertheless, in the vast majority of cases at any stage of a rivalry, the initiator with a positive expected utility score does not unilaterally succeed in the dispute.

## **PATTERNS IN DISPUTE OUTCOMES ACCORDING TO SEVERITY**

In this section, we look at outcomes of disputes of varying severities, and in particular the most severe disputes – wars.

(Table 5 about here)

Most militarized disputes are low severity ones, with almost 60% involving no battle-related fatalities. Table 5 shows the distribution of dispute outcomes according to severity. There is a modest relationship between the severity of a dispute and its outcome. Yet this relationship is largely driven by the impact of high severity disputes. The percentage of

stalemates and victories is relatively stable (never varying more than 6% in any case) across the first three levels of severity; typically, stalemates constitute 77-83% of the cases at these levels. Yet, in the high severity cases, the rate of stalemates drops to only 27%, with victories now the preeminent outcome at over 70%. These results are mirrored for an analysis of wars, with victories (77%) again strongly associated with the most severe disputes, those that escalated to war. Thus, the more severe the dispute/war the more likely it is to end with one side prevailing. Conversely, this means that low-level disputes are particularly likely to end in stalemates.

Although the minimal realist and expected utility models have fared poorly in our previous analyses, they may do better in predicting outcomes to the most severe disputes. The minimal realist model predicts that the stronger side should prevail in a conflict, especially one involving the use of military force. Regardless of capability indicator, the stronger side rarely prevails (generally less than 20% of the time) in low-level disputes. Its success rate is considerably higher, sometimes four times higher, in the most severe disputes; still, the overall success rate hovers around 50% in that highest severity category (actually 41-53% depending on capability indicator). Similar results are present for wars; the stronger side is five times more likely to prevail in full-scale wars than in conflicts short of that threshold. Clearly our basic capability measures contain some error, and this may account for the inability to predict more outcomes accurately. Nevertheless, the results are consistent with the expectations that power matters more in the most severe disputes.

Again, the findings on expected utility are similar to those for the simpler version of the realist model. States with a positive expected utility score win 36% of the most severe disputes and 35% of the wars, far higher success rates (approximately 2-4 times as great) than lower-level conflicts. In many ways, this reinforces the findings of Bueno de Mesquita (1981) and others that expected utility calculations have better predictive power for war. At the same time, however, they are, at best, weakly related to predicting outcomes in lower-level militarized disputes. Overall then, the realist models have better explanatory power for the most severe confrontations.

## **COMBINED MODELS OF STALEMATE OUTCOMES**

The previous analyses provided insights into the patterns of stalemate outcomes generally, across time, and with respect to different severities. Although the results were consistent with expectations of the punctuated equilibrium model, there were no direct attempts to predict stalemate or victory outcomes in particular disputes. Therefore, as a final analysis, we conduct a logit analysis that incorporates variables from the punctuated equilibrium and realist

models to assess their explanatory power in general and relative to one another in predicting stalemate outcomes.

(Table 6 about here)

The "basic model" in Table 6 looks at dispute outcomes using just power, expected utility, and control variables.<sup>14</sup> This model is a pure cross-sectional analysis of dispute outcomes, not taking into account the stage of the rivalry dyad. Implicitly, the hypothesis is that the various realist factors exercise a constant effect, independent of rivalry history. The results confirm what we found above in the descriptive tables. Neither power variable -- economic or military capabilities -- is significantly associated with prevailing in militarized disputes. The signs of the coefficients for the capabilities variables, as well as for the expected utility variable, are positive but nowhere near statistically significant.

Our descriptive analyses showed a very dramatic impact of dispute severity on the likelihood of one side prevailing. This is confirmed by the results of the "basic model." The severity variable is statistically very significant and has a very large odds ratio. The odds ratio value of 4.58 means that if one moves from a low severity dispute to a high severity dispute, the chances of one side winning increase 458 percent. We included several control variables in the model. The literature on the democracy and war has indicated that democracies are more likely to win wars than non-democracies (e.g., Reiter and Stam, 1998). In our analyses we find no such effect for militarized disputes in general. Multilateral disputes are more likely to end with one side prevailing. Similarly, and perhaps counter-intuitively, territorial disputes are also associated with victory. Yet, in both cases, this may largely be a function of the severity effects. Multilateral disputes and those involving territory are, on average, more severe than other disputes, although the relationship is not perfectly uniform.

(Table 7 about here)

The "basic model" in Table 6 did not include any consideration of rivalry. Thus, we reformulated the analyses from that table, taking into account rivalry factors; the results are reported in Table 7. First, our descriptive analyses showed that as dyads become locked into rivalries the probability of prevailing declines. Accordingly, we included a variable "rivalry order" into the basic model as one means to take rivalry into account.<sup>15</sup> In statistical terms, we no longer assume -- as in the "basic model" -- that the intercept is constant across all disputes and wars. In the rivalry model, we now have the intercept varying depending on the stage of the rivalry. The results indicate that this has a strong impact on the likelihood of prevailing in a dispute. The odds ratio of .52 means that as one moves to the next stage of rivalry development, the probability of prevailing decreases by 52 percent.

Second, we “contextualized” all the variables except the severity variable.<sup>16</sup> We created interaction terms between these and the four point categorical variable indicating where in the rivalry sequence that the dispute falls. Thus, unlike the realist model which says power has a constant and uniform impact on disputes, the contextualized variable means that the impact of power superiority varies as the rivalry develops over time. For example, in some preliminary analyses, we found that for economic capabilities there was a clear, but modest, trend whereby preponderance in economic capabilities mattered more in early disputes of a rivalry than in later ones. Unlike in the basic (realist) model, once we take into account rivalry stage, economic capabilities are significantly associated with prevailing. Also in agreement with our descriptive analyses, however, the impact is modest. The odds ratio is only 1.20, signifying that in general once rivalry is taken into account preponderance only increases the likelihood of victory by 20 percent. While including the rivalry context does increase the significance level of the military capability variable, its sign becomes negative. The contextualization of the expected utility variable has no impact, and that variable remains very insignificant. This is perhaps not surprising since we found that positive expected utility really had no impact on the outcome of disputes. In fact, if any it appeared that negative expected utility states were more likely to initiate than positive expected utility ones.

The control variable of multilateral conflict remains significant when we take into account rivalry stage as does the territorial one. Their odds ratios hardly change from those of the “basic model”. The democracy variable becomes statistically significant (at .10) once rivalry is included, but the sign is “wrong.” Its odds ratio is .99, however, signifying that substantively it has little impact. Not fully evident in our logit analyses is the importance of capabilities in severe conflicts. We tried including interaction terms with capability and severity, but the severity variable overwhelms everything and one cannot see the impact of capability clearly. Our next analysis makes this relationship clearer

(Table 8 about here)

Another way to see the impact of rivalry on outcome is to examine the predicted probabilities of prevailing for each stage of rivalry with some key configurations of the independent variables. Table 8 shows these for four types of disputes and disputants. In cases of preponderance, we chose values for the independent variables that should predict victory according to realist thought. For the parity cases, we set the values at the other extreme, predicting stalemates. Because we have found that conflict severity has such an important influence, we have considered both high and low severity dispute situations. Table 8 gives these

results. Even in the optimal conditions when victory is predicted (preponderance and positive expected utility in a high severity dispute), the likelihood of prevailing declines significantly as the rivalry matures (from 78% to 53% from first to last phase). In *every* scenario, rivalry effects are clear and consistent. As the rivalry develops, the likelihood of victory decreases and the prospects for stalemate increase. This is consistent with the punctuated equilibrium expectation that rivalry processes lock-in policies and make victory more difficult the longer the rivalry lasts. Preponderance makes victory more likely than does parity and the effect is uniform across rivalry phases, but the differences are consistently quite modest. The severity variable still exercises the greatest impact. Fundamentally, except for high severity disputes and wars, most militarized conflicts end in stalemate.

## CONCLUSION

When states clash in militarized confrontations, the outcome more often than not (approximately 2/3 of the time) is indecisive. The purpose of this paper was to assess the extent to which two theoretical models – the realist and punctuated equilibrium models respectively – account for the prevalence of stalemates in militarized disputes. We looked at almost 3,000 disputes embedded in 1166 rivalries over the period 1816-1992.

In testing its explanatory capability, we explored both minimal and enhanced versions of the realist model. The results consistently suggest that realism's emphasis on prevailing through power preponderance is limited in its scope. Power preponderance along with positive expected utility had no impact on the vast majority (70-80%) of militarized disputes; even when states had a significant military advantage over their opponents, most disputes still ended in stalemates. Realist factors were better able to predict outcomes in the subset of the most severe disputes, including full scale wars; even then, however, successful predictions of victories were far from perfect. Material capabilities, and by implication the realist model as a whole, are best suited to explaining militarized confrontations in which those resources are actually used. Lower order confrontations in which military force is only threatened, displayed, or employed in a very limited fashion do not engender many advantages to the more powerful state.

It is also evident that the utility of military preponderance declines over the life of a rivalry relationship; as states lock into long-term conflict they are less able to use the tools of realpolitik effectively; the chances of prevailing decline significantly as a rivalry matures. Paradoxically, Leng (2000) has found that states are more likely to use coercive or realist bargaining approaches in recurrent crises, even if they are less effective over time. In terms of power politics and deterrence models, one needs to see that the existence of rivalry itself is often

a signal that deterrence has failed. That most disputes take place between states with power asymmetries means that power preponderance has failed to deter the weaker side.

Our results indicate that as long as the weaker side avoids war and high fatality conflicts, it can keep the issues on the agenda of the stronger side. For the weaker side, stalemate might be considered as limited success; although it has not achieved its goals, it is incurring costs to the more powerful side and perhaps keeping the more powerful side from achieving its goals. One short-term conclusion that a realist might draw from our analyses is that one needs to push the dispute to that level at which power preponderance matters. Our analyses do suggest that the more powerful is more likely to get a decisive victory in severe conflicts. Nevertheless, we began this article with the case of the Persian Gulf War and this should be a powerful reminder that a decisive outcome in a war does not often end a rivalry. This is also evident in the Israeli experience in the aftermath of its victory in the 1967 war.

Overall, the results were more supportive of the punctuated equilibrium model. Punctuated equilibrium stresses stasis in rivalries. One aspect of stasis is that when confrontations occur, little is resolved. This produces the next dispute, but does not produce victory or any fundamental change in the rivalry. Accordingly, the prevalence of stalemates is consistent with the expectations of this model; enduring rivalries are the product of repeated failures of states to impose their wills or come to some kind of accommodation with one another. Also consistent with this model is that stalemates become more common as the rivalry matures; punctuated equilibrium indicates that rivalries undergo an initial “lock-in” period after which policies and strategies are solidified and it is increasingly difficult for rivals to escape their patterns of behavior. It was also worthy of note that capabilities and other variables only attained significance once their rivalry context had been factored in. This is confirmation that such context matters, and further undermines claims (Gatzke and Simon, 1999) that disputes are merely unconnected or random events.

As much as the results seemed to support the punctuated equilibrium, there are a number of gaps in that model with respect to dispute outcomes. First, as noted above, the model is better able to account for the general pattern of stalemate outcomes than it is to explain outcomes of individual disputes. Some greater attention to situational factors and the characteristics of the rivals, largely outside of the punctuated equilibrium model, may be necessary to understand how the cycle of stalemates begins and then is sustained in enduring rivalries. Second, there must be some dynamic that keeps rivalries alive (and eventually leads some to blossom into enduring rivalries). Further study might consider how stalemate outcomes play into that dynamic. The punctuated equilibrium posits that stalemate outcomes are critical, but this relationship still needs

better specification and empirical confirmation. Third, the severity level of the dispute was uniformly the strongest predictor of dispute outcomes. This largely begs the question of what leads disputes to be more severe. Thus, one needs to develop a model (such as Vasquez, 1996) of when rivalry conflict escalates to severe levels including war. Yet, as we know from many rivalry studies, such severe conflicts are interruptions in a steady pattern of lower level disputes with indeterminate outcomes.

As we write (May 2002), we can see the patterns we have described in actions in Israel and Palestine. Israeli Prime Minister Sharon has employed a similar strategy in 2002 that he developed as Defense Minister in the early 1980s. Sharon was the key figure behind the Israeli invasion of Lebanon. All observers and the Israelis themselves viewed this action as a failure (though again in accordance with the punctuated equilibrium model it took the Israelis several years to pull out). Twenty years later Sharon is trying the similar strategy and although the dust has not settled, we suspect that it will not be more successful this time than the last and the cycle of stalemates will continue.

## **References**

- Arreguin-Toft, Ivan. 2001. How the weak win wars: A theory of asymmetric conflict. *International Security* 26 (1):93-128.
- Ayres, R. William. 2000. A world flying apart?: Violent nationalist conflicts and the end of the cold war. *Journal of Peace Research* 37(1):107-17
- Bennett, D. Scott, and Alan Stam. 1998. A combined model of war outcomes and duration: Modelling the declining advantages of democracy. *Journal of Conflict Resolution* 42:344-366.
- Bennett, D. Scott, and Alan Stam. 2000. EUGene: A conceptual manual. *International Interactions* 26:179-204
- Bercovitch, Jacob, and Paul F. Diehl. 1997. conflict management of enduring rivalries: Frequency, timing, and short term impact of mediation. *International Interactions* 22:299-320
- Boulding, Kenneth. 1962. *Conflict and defense*. New York: Harper and Row.
- Brecher, Michael. 1984. International crises and protracted conflicts. *International Interactions* 11:237-298.
- Brecher, Michael. 1993. *Crises in world politics: Theory and reality*. Oxford: Pergamon Press.
- Brecher, Michael, and Jonathan Wilkenfeld. 1997. *A study of crisis*. Ann Arbor: University of Michigan Press.
- Bremer, Stuart. 2000. Resolving conflicts: Conditions favoring negotiated compromise in militarized interstate disputes. Draft paper, State College, PA.
- Bueno de Mesquita, Bruce. 1982. *The war Trap*. New Haven: Yale University Press.
- Bueno de Mesquita, Bruce, and David Lalman. 1992. *War and reason*. New Haven: Yale University Press.
- Claude, Inis. 1962. *Power and international relations*. New York: Random House.
- Cioffi-Revilla, Claudio. 1998. The political uncertainty of interstate rivalries: A punctuated equilibrium model. In *The dynamics of rivalries* ed. by Paul Diehl. Urbana: University of Illinois Press, pp.64-97.
- Diehl, Paul F. 1985. Armaments without war: An analysis of some underlying effects” *Journal of Peace Research* 22:249-259.
- Diehl, Paul F., and Gary Goertz. 2000. *War and peace in international rivalry*. Ann Arbor: University of Michigan Press.
- Dixon, William, and Paul Senese. 2002. Democracy, disputes, and negotiated settlements. *Journal of Conflict Resolution* (forthcoming).

- Epstein, Joshua. 1989. The 3:1 rule, the adaptive dynamic model, and the future of security studies. *International Security* 13 (4):90-127.
- Frazier, Derrick, and Gary Goertz. 2001. Militarized interstate disputes: outcome and settlement in non war disputes and their implications for conflict management. Paper presented at the Workshop on Conceptualizing and Measuring Conflict Management, Halifax, Nova Scotia.
- Gartzke, Erik, and Michael Simon. 1999. Hot hand: A critical analysis of enduring rivalries. *Journal of Politics* 63:777-798.
- Hensel, Paul. 1996. *The evolution of interstate rivalry*. Ph.D dissertation, University of Illinois.
- Hensel, Paul. 1999. An evolutionary approach to the study of interstate rivalry. *Conflict Management and Peace Science* 17 (2):175-206.
- Hirshleifer, Jack. 2000. The macrotechnology of conflict." *Journal of Conflict Resolution*, 44:773-792.
- Jones, Daniel, Stuart Bremer, and J. David Singer. 1996 Militarized interstate disputes, 1816-1992: Rationale, coding rules, and empirical patterns *Conflict Management and Peace Science* 15:163-213.
- Kissinger, Henry. 1979. *White House years*. Boston: Little, Brown, and Co.
- Kugler, Jacek, and Douglas Lemke. 2000. The power transition research program: Assessing theoretical and empirical advances. In *Handbook of war studies II*. Ed by Manus Midlarsky. Ann Arbor: University of Michigan Press, pp.129-163.
- Legro, Jeffrey, and Andrew Moravcsik. 1999. Is anybody still a realist?" *International Security* 24 (2):5-55.
- Leng, Russell. 2000. *Bargaining and learning in recurrent crises*. Ann Arbor: University of Michigan Press.
- Leng, Russell. 1983. When will they ever learn?: Coercive bargaining in recurrent crises *Journal of Conflict Resolution* 21:379-419.
- Mack, Andrew. 1975. Why big nations lose small wars: The politics of asymmetric conflict." *World Politics* 27 (2):175-200.
- Maoz, Zeev. 1983. Resolve, capabilities, and the outcome of interstate disputes, 1816-1976. *Journal of Conflict Resolution* 27 (2):195-229.
- Maoz, Zeev. 1989. Power, capabilities, and paradoxical conflict outcomes. *World Politics* 41, (2):239-266.
- Mearsheimer, John. 1989. Assessing the conventional balance: The 3:1 rule and its critics. *International Security* 13 (4):54-89.

- Morrow, James. 1986. A spatial model of international conflict. *American Political Science Review* 80 (4):1131-1150.
- Mousseau, Michael. 1998. Democracy and Compromise in militarized interstate conflicts, 1816-1992. *Journal of Conflict Resolution* 42:210-230.
- Organski, A.F.K., and Jacek Kugler. 1980. *The war ledger*. Chicago: University of Chicago Press.
- Reiter, Dan, and Allan Stam. 1998. Democracy, war initiation, and victory. *American Political Science Review* 92:377-390.
- Rosen, Stephen. 1992. War power and the willingness to suffer. In *The scientific study of war and peace*. Ed. by John Vasquez and Marie Henehan. New York: Lexington Books, pp.255-273.
- Russett, Bruce, and John Oneal. 2001. *Triangulating Peace*. New York: W.W. Norton.
- Signorino, Curtis, and Jeffrey Ritter. 1999. Tau-b or not Tau-b: measuring the similarity of foreign policy positions. *International Studies Quarterly* 43 (1):115-144.
- Singer, J. David, Stuart Bremer, and John Stuckey. 1972. Capability distribution, uncertainty, and major power war, 1820-1965. In *Peace, war, and numbers*, ed. by Bruce Russett. Beverly Hills: Sage.
- Small, Melvin, and J. David Singer. 1982. *Resort to arms*. Beverly Hills: Sage.
- Stam, Alan. 1996. *Win, lose, or draw*. Ann Arbor: University of Michigan Press.
- Stinnett, Douglas, and Paul F. Diehl. 2001. The path(s) to rivalry: behavioral and structural explanations of rivalry development *Journal of Politics* 63 (3):717-740
- Tir, Jaroslav, and Paul F. Diehl. 2002. Geographic dimensions of enduring rivalries. *Political Geography* 21:263-286.
- Vasquez, John. 1996. Distinguishing rivals that go to war from those that do not: A quantitative comparative case study of the two paths to war. *International Studies Quarterly* 40:531--58.
- Vasquez, John. 2003. The probability of war, 1816-1992. *International Studies Quarterly* (forthcoming).
- Walt, Stephen. 1987. *The origins of alliances*. Ithaca: Cornell University Press.
- Wang, Kevin, and James Lee Ray. 1994. Beginners and winners: The fate of initiators of interstate wars involving great powers since 1495." *International Studies Quarterly* 38:139-154.
- Wilson, James Q. 1989. *Bureaucracy: What government agencies do and why they do it*. New York: Basic Books.

**TABLE 1**  
**DISPUTE OUTCOMES IN INTERNATIONAL RIVALRIES**

<u>Outcome</u>	%	<u>N</u>
Stalemate	68.0	2029
Victory	25.7	768
Compromise	6.3	187

**TABLE 2**  
**DISPUTE OUTCOMES AND CAPABILITY**  
**DISTRIBUTIONS**

**Military Capabilities**

	<b>Stalemate</b>	<b>A Wins</b>	<b>B Wins</b>	<b>N</b>
Parity	77%	13%	10%	786
A Stronger	73%	18%	9%	980
B Stronger	76%	11%	13%	449

**Economic Capabilities**

	<b>Stalemate</b>	<b>A Wins</b>	<b>B Wins</b>	<b>N</b>
Parity	79%	10%	11%	630
A Stronger	76%	17%	7%	1155
B Stronger	79%	7%	14%	371

**TABLE 3**  
**DISPUTE OUTCOMES AND EXPECTED UTILITY**  
**SCORES**

Initiator's Expected Utility Score	<u>Stalemate</u>	<u>Victory- Initiator</u>	<u>Victory- Other</u>	<u>N</u>
Negative	66%	19%	15%	1,174
Zero	78%	12%	10%	808
Positive	75%	17%	8%	780

**TABLE 4**  
**DISPUTE OUTCOMES OVER TIME**

<u>Timing of Dispute in the Rivalry</u>	<u>Stalemate</u>	<u>Victory</u>	<u>N</u>
1 <sup>st</sup> -3 <sup>rd</sup>	69%	31%	1802
4 <sup>th</sup> -6 <sup>th</sup>	76%	24%	393
7 <sup>th</sup> -12 <sup>th</sup>	79%	21%	291
13 <sup>th</sup> +	85%	15%	311

**TABLE 5**  
**DISPUTE OUTCOMES BY DISPUTE SEVERITY**

<u>Dispute Severity</u>	<u>Stalemates</u>	<u>Victory</u>	<u>N</u>
0-40	83%	17%	1236
40-100	82%	18%	467
100-145	77%	23%	638
>145	27%	73%	456

**TABLE 6**

**MULTIVARIATE ANALYSIS OF DISPUTE OUTCOMES:  
BASIC MODEL, NO RIVALRY VARIABLES**

<u>Variable</u>	<u>Parameter Estimate</u>	<u>SE</u>	<u>Odds Ratio</u>
Intercept	-3.11*	0.26	
Expected Utility	0.01	0.03	1.01
Economic Capabilities	0.26	0.18	1.29
Military Capabilities	0.06	0.16	1.01
Democracy	-0.007	0.01	0.99
Multilateral	0.31*	0.04	1.37
Severity	1.52*	0.24	4.58
Territory	0.53**	0.17	1.70

**Chi-square = 239.25 (significant at .001)**  
**gamma = .48**

**\* significant at .001**

**\*\* significant at .01**

**TABLE 7**

**MULTIVARIATE ANALYSIS OF DISPUTE OUTCOMES:  
“RIVALRY” MODEL**

<u>Variable</u>	<u>Parameter Estimate</u>	<u>SE</u>	<u>Odds Ratio</u>
Intercept	-1.67*	0.15	-
Rivalry*Expected Utility	0.004	0.02	1.004
Rivalry*Economic Capabilities	0.17**	0.08	1.18
Rivalry*Military Capabilities	-0.06	0.08	0.94
Rivalry Order	-0.66*	0.14	0.52
Rivalry*Democracy	-0.008***	0.005	0.99
Rivalry*Multilateral	0.17*	0.02	1.18
Severity	1.66*	0.24	5.28
Rivalry*Territory	0.20**	0.08	1.23

**Chi-square = 259.78 (significant at .001)  
gamma = .51**

- \* significant at .001**
- \*\* significant at .05**
- \*\*\* significant at .10**

**TABLE 8**

**PREDICTED PROBABILITIES OF VICTORY IN DISPUTES**

<u>Capability</u>	<u>Severity</u> <i>Level</i>	<u>Rivalry</u> <i>Phase #1</i>	<u>Rivalry</u> <i>Phase #2</i>	<u>Rivalry</u> <i>Phase #3</i>	<u>Rivalry</u> <i>Phase #4</i>
Preponderance	Low	.41	.32	.24	.18
Preponderance	High	.78	.71	.63	.53
Parity	Low	.38	.27	.19	.12
Parity	High	.76	.66	.55	.42

## NOTES

<sup>1</sup> Much of the literature on power parity versus power preponderance has found that disputes are more likely the more evenly matched the two countries are (e.g., Kugler and Lemke, 2000). This is suggestive of just such a selection effect. Yet our results below run against what one would normally expect in a selection argument involving dispute and war outcomes. That there is little (economic preponderance) or no (military preponderance) relationship between power or expected utility and dispute outcome: this means that governments are not selecting disputes that they can win (of course they may *think* ex ante that they will prevail). Furthermore, we make causal inferences given that states have entered into disputes and rivalries: This is no different from those who study, for example, factors that influence educational test scores. They ignore that there are selection effects in who gets born and who survives until school age. We do not deny the potential importance of selection effects, but for us that means developing a theory of the origins of rivalry, something well beyond the limits of this paper.

<sup>2</sup> Kugler and Lemke 2000 note that the power transition theory is nonrealist in its emphasis on the hierarchical nature of the international system.

<sup>3</sup> More developed versions of the basic model (Bueno de Mesquita and Lalman, 1992) might incorporate risk and domestic political factors in the calculation.

<sup>4</sup> Mousseau (1998) does include a measure of joint economic development, which is arguably correlated with military capability. Still, this is an indicator, at best, of whether both disputants are powerful or developed, rather than measuring how powerful one side is vis-à-vis another; thus, it does not really provide a test of realist propositions.

<sup>5</sup> The MID dataset also contains a settlement variable that taps a dimension of unresolved conflict that particularly interests us. The settlement variable has four categories: (1) Negotiated, (2) Imposed, (3) None, and (4) Unclear. Clearly a negotiated settlement seems related to the compromise coding of the outcome variable. Similarly imposed settlements would appear to be related to victories and yields, while no and unclear settlements apparently match well with stalemate outcomes. Indeed, if we cross-tabulate the two dimensions (COW outcomes and settlement categories), we see that these suppositions hold up quite well empirically ( $\gamma = .85$ ). The major difference occurs with respect to the intersection of negotiated and compromise categories. Negotiated settlements appear more frequently and are therefore not confined to compromise outcomes. Apparently, many outcomes where one side prevails still involve negotiated solutions. For our analyses below, we use only the satisfaction variable based on the COW outcome coding. Nevertheless, we reran all the analyses with the settlement variable and essentially found the same results. See Frazier and Goertz, 2001 for more on the relationship between these two variables.

<sup>6</sup> Specific measurement details on the indicators are given in Singer, Bremer, and Stuckey, 1972 and are available from the Correlates of War Project.

<sup>7</sup> There are about 60 cases in which both sides are listed as the initiator; in those cases, we use the average of the two expected utility scores.

<sup>8</sup> Because the expected utility model deals with initiation, we took the expected utility score for the year before the dispute.

<sup>9</sup> We adopt the standard procedure of defining a dispute as a war if the level-of-hostility of one side is war – code=5 -- (there are a fair number of cases in the dispute data where the level-of-hostility of one side is but the other side has a nonwar level-of-hostility)

<sup>10</sup> We note in passing that this finding is not a result of collapsing cases of “unclear” and “missing” outcomes with the original “stalemate” coding in the MID data set. The former two categories constitute only about 10% of the outcomes, leaving almost 60% of the cases as stalemates even under the most restrictive definition of the term.

<sup>11</sup> For the purposes of the rest of this paper, we shall drop the compromise cases. They merit sustained attention (see Mousseau, 1998 and Frazier and Goertz 2001) but our main goal in this paper is comparing realist theories with the punctuated equilibrium model. The realist model connects much more with victory (through power superiority) than to compromise. Because compromises comprise only 6% of the cases, we can safely ignore them (in all of the analyses, we have checked our results and none are influenced by the inclusion of these few disputes).

---

<sup>12</sup> The results were not substantially different with the other thresholds for distinguishing preponderance from parity were used. Some cases were dropped from the analysis because of missing military (N=734) and economic (N=793) capability data respectively. Stalemates are slightly more common in the remaining cases than in the full set of cases.

<sup>13</sup> Another way of considering temporal patterns is to assess whether stalemates are less common in shorter rivalries (e.g., isolated rivalries) than in more developed rivalries (proto and enduring rivalries), which last longer and have a larger number of disputes. We tested for this possibility and the results mirror those in Table 4. That is, more advanced rivalries have a greater percentage of stalemate, consistent with the expectations of the punctuated equilibrium model.

<sup>14</sup> Introducing capability variables into the model results in a large reduction of cases due to missing data. Still, we reran the analyses without the capability variables, and therefore with a greater number of cases, and the results were nearly identical to those analyses with the capability variables included.

<sup>15</sup> We report the results using the four categories of rivalry sequence from Table 5. Nevertheless, the results were not different when the simple number (1-53) of the dispute in the rivalry sequence was used.

<sup>16</sup> The previous results suggest that conflict severity exercises an impact independent of rivalry position.