All Alliances are Multilateral:

Rethinking Alliance Formation

Benjamin Fordham*

Paul Poast†

Word Count: 10,991

Abstract

Alliance formation is a multilateral process. The vast majority of alliance relations are created via multilateral alliances. Moreover, leaders assess the alliance as a whole, not just each prospective partner. Any alliance could have three or more members, so one must understand not just why third parties were included in multilateral alliances, but why they were excluded from bilateral alliances. Unfortunately, current research treats alliance formation as a bilateral process: it theorizes about bilateral alliances and tests hypotheses using dyadic research designs. Reconceptualizing all alliances as originating from a multilateral process reveals that a long-neglected theory, William Riker's size principle, illuminates the role of power in alliance formation. Using k-adic data to analyze multilateral processes, we find strong support for Riker's claim about minimum winning coalitions in world politics. Our argument and findings, by highlighting how a fundamental state behavior like alliance formation follows a multilateral process, suggests rethinking much of international relations research.

Acknowledgments: The authors would like to thank Brett Benson, Koji Kagotani, Ashley Leeds, Neil Narang, Emily Ritter, Gary Uzonyi, Alex Von Hagen-Jamar, and Scott Wolford, as well as participants in the World Politics Workshop at Binghamton University, and the 2011 annual meeting of the Peace Science Society for their comments on earlier versions of this paper. Any remaining errors and omissions are our responsibility.

---

*Professor, Department of Political Science, Binghamton University

†Assistant Professor, Department of Political Science, Rutgers University

1 Introduction

"Wherever in recorded history a system of multiple sovereignty has existed, some of the sovereign units, when involved in conflicts with others, have entered alliances" (Wolfers 1968, 269). As Wolfers' reference to multiple sovereignty suggests, alliances are about seeking security in a multi-state system. Indeed, leaders often felt that the key to security in a multi-state system was via the formation of *multilateral* alliances. Bismarck, in reference to the formation of the Three Emperors' League between Russia, Austria, and Germany, observed how "in a world governed by five powers, try to be a trois."[1]

In fact, the vast majority of alliance relations stem from multilateral alliances. Although just 16 percent of the alliances documented in the Alliance Treaty Obligations and Provisions (ATOP) dataset between 1815 and 2003 were multilateral, nearly 90 percent of allied dyads were united through only a multilateral pact. Only 5 percent of allied dyads had a bilateral tie alone. An additional 5 percent had both types of alliances (Leeds, et al. 2002).

This fact is underappreciated in existing alliance research, which is dominated by studies of bilateral processes.[2] This dyadic focus is particularly surprising since the multilateral conception of alliances figures prominently in many important works on international relations. Well known claims about the balance of power implicitly assume the possibility of multilateral alliances. Snyder (1997, 156-8) notes that states' ability to adjust the size of their alliance – something that is only possible in a multilateral setting – is critical to the balance of power. Multilateral alliances are central to Walt's "balance of threat" theory, while theoretical phenomena like "chain-ganging" and "buck-passing" make far less sense without a multilateral conception of alliances (Walt 1987; Christensen and Snyder 1990).

We argue that concepts and arguments formulated in terms of bilateral processes do not easily translate into a multilateral setting. The dyadic world is an analytical convenience, not a realistic depiction of the international system. If researchers lose sight of this fact, the

dyadic setup can fetter understanding of many important international processes.

Scholars should reconceptualize *all* alliances, bilateral and multilateral, as originating from a multilateral process. Even when states only form a bilateral alliance, the alliance, in principle, could have contained additional states. Like consumers selecting among "baskets" of goods in standard consumer choice theory (Besanko and Braeutigam 2011, 75), leaders in a multi-state system must assess an alliance as a whole, not just each individual prospective partner. A particular ally might make little sense except in conjunction with other partners. Alternatively, the viability of an alliance may hinge on the exclusion of a particular state.

With this multilateral setting in mind, we present an empirical test of a classic idea about alliance formation, Riker's "size principle." Riker claimed that the logic of the minimum winning coalition, widely applied to legislative politics, also pertained to military alliances. In the context of world politics, "size" refers to aggregate power rather than the number of coalition members, but the logic remains the same. The neglect of Riker's argument in recent research is puzzling because it focuses on power, something that is clearly important in alliance formation, but that rarely appears in empirical analyses of the phenomenon. We suspect that the difficulty of testing the size principle in a dyadic framework has contributed to its neglect. Riker's argument focuses on characteristics of the potential alliance as a whole, rather than the features of isolated pairs of prospective allies. In practice, these characteristics frequently involve more than two states.

The remainder of this paper is organized as follows. In the next section, we examine the role of bilateral processes in research on alliance formation and explain why moving to a multilateral setting is important. In the third section, we review Riker's size principle and objections to it. In the fourth section, we set out the multilateral research design needed to test our theoretical claims. While others, such as Croco and Teo (2005), highlight the insufficiencies of dyads as a unit of anlaysis, Poast (2010) offers an approach for empirically testing multilateral processes. The fifth section presents our results. We find that, even

after accounting for the average size of threats facing a group of states and for the number of states in that group, increasing capabilities will increase the probability of alliance formation, but only to a point. Specifically, the probability of alliance formation peaks when the total capabilities in the group of states are near 0.3 using the *Correlates of War* Composite Index of National Capabilities. Beyond this point, the probability of alliance formation declines. Alliances of this size have the capabilities to compete with any state in the system. The sixth section concludes.

2 The Need to Rethink Alliance Formation

Bilateral processes are the primary focus in much of the recent theoretical and empirical work on alliance formation. Two developments in the field have contributed to this analytical focus. First, theoretical examinations of alliance formation turned to game-theoretic models in the early 1990s, all of which focused on bilateral alliance formation. For example, examining the tradeoff between security and policy autonomy, Morrow (1991) finds that a relatively weak partner receives security while giving up autonomy to a more powerful partner. Although there are three players in the games employed by Smith (1995) and Fearon (1997), the potential alliance they consider is strictly bilateral, with the third state playing the role of motivating threat. Even Snyder (1997, 55-65), who explicitly stresses the importance of multilateralism in alliance formation elsewhere in his book, focuses on bilateral alliances when seeking to formalize his arguments.

Second, the dyadic turn in empirical international relations research began around the same time (e.g., Bremer 1992). Researchers focused on pairs of states when testing hypotheses about interstate conflict. Improvements in data management technology like EUGene made dyadic research designs even more attractive and accessible (Bennett and Stam 2000). When combined with the field's bilateral theoretical focus, it is not surprising that empirical

research on alliance formation in the last two decades has been almost exclusively dyadic.

Consider, for example, the following hypotheses from Lai and Reiter (2000):

- A pair of democracies is more likely to be allied than are other pairs of states.
- A pair of states with similar regime type is more likely to be allied than a pair of states with differing regime types.
- Two states that are members of a common cultural group are especially likely to be allied.
- The more threatened are the two states within a dyad, the more likely that they will be allied.
- Two states that face a common threat are more likely to be allied.
- Two states that have a conflictual relationship are less likely to be allied.
- A pair of states is more likely to be allied if one or both are major powers than if neither is a major power.

While the Lai and Reiter piece is over a decade old, nearly all recent empirical research on alliances employs a similar dyadic research design (e.g., Benson 2011; Crescenzi, et al. 2012; Fordham 2010; Gibler and Sarkees 2004; Gibler and Wolford 2006; Gibler 2008; Johnson and Leeds 2011; Leeds, et al. 2002; Long, Nordstrom, and Baek 2007; Siverson and Emmons 1991; Simon and Gartzke 1996).

Even recent approaches using network analysis focus on monadic and dyadic relationships (e.g., Maoz, et al. 2007; Warren 2010; Cranmer, Desmaris, and Menninga 2012), since individual states comprise the nodes (or constitutive elements) of the network. The core characteristic of network data, a node-to-node "edge" list, is inherently dyadic. This is not to say that key network analysis concepts, such as homophily or centrality, are useless outside a

dyadic framework. However, they do not obviate the need to evaluate multilateral processes as more than the sum of the individual parts. Network analysis, by reporting which states are most central or which states are most connected, can explain some state behavior (e.g., Cao 2012). A similar network position can even suggest whether countries constitute a group. But positional similarity is not equivalent to the states' decision to become a group. Instead, network analysis seeks to identify the emergence of groups, which may or may not result from deliberate policy decisions (Hafner-Burton and Montgomery 2012, 259).

Focusing on bilateral alliance ties in dyadic or network approaches might simplify analysis, but it comes at a price. Because alliance formation is actually a multilateral process, some real-world pairs of allies make little sense in bilateral terms. Poast (2010, 403) observed that their limited military capabilities and large geographic distance made Belgium and Turkey unlikely allies apart from their membership, alongside 26 other states, in the North Atlantic Treaty Organization (NATO). Their alliance tie arguably has more to do with their relations with the United States than with each other. Theoretical arguments about the characteristics of these states and their bilateral relations are unlikely to explain their alliance tie. Even considerations that genuinely influence bilateral alliance formation might find no support in dyadic empirical analyses that include unlikely allies linked though a multilateral pact. Nor will excluding multilateral alliances from empirical analyses solve the problem – this will treat as non-aligned many states that might have formed a bilateral pact if they were not members of a broader alliance.

Unlikely allies linked by multilateral pacts is just one problem confronting theories of alliance formation grounded in bilateral processes. Another is explaining why actual bilateral alliances do not include other states. Theoretical arguments about alliance formation that restrict attention to dyadic processes will always fall short because even bilateral alliances are formed in a potentially multilateral context. When it comes to bilateral alliances, a theory of alliance formation must explain not only why a particular ally was attractive but

also why other states were not also included. For example, the United States and Spain signed a series of bilateral agreements beginning in 1953, leading to a defense pact in 1963. In this case, the most interesting question is why Spain did not join NATO until 1982, an obvious substitute for its bilateral pacts with the United States. Indeed, why did the NATO alliance initially include only twelve members in 1949, and not Spain as well? If, as seems likely, other NATO members' distaste for Francisco Franco is part of the answer, this case illustrates how the positions of third-party states can influence the formation of a bilateral alliance. If these third parties had viewed Franco's regime differently, a multilateral alliance might have formed instead.

Conceptualizing all alliances as originating from a multilateral process enables scholars to think anew about the influences on alliance formation. In particular, we can reconsider an important but neglected consideration in alliance formation: the coalition's aggregate military power.

3 The Size Principle, Power, and Multilateral Alliance Formation

How much power should prospective members concentrate in an alliance? Should the alliance keep adding members to increase its aggregate military capabilities? An important but neglected argument on alliance formation, William Riker's "size principle", suggests an answer. As we shall see, this answer can only be tested in a multilateral framework.

In his 1962 book *The Theory of Political Coalitions*, Riker argues that minimum winning coalitions are prevalent in many important social situations. A certain size secures the coalition's goals, but larger coalitions reduce the benefits to each member. Riker formally demonstrates the logic behind this argument, which he calls the "size principle."

> In social situations similar to n-person, zero-sum games with side-payments, participants create coalitions just as large as they believe will ensure winning and no larger (Riker 1962, 47).

Although the size principle has been applied most often in research about legislative coalitions, Riker devotes roughly equal attention in his book to its role in shaping alliances in world politics. In Riker's account, the size principle explains rarity of overwhelming coalitions in world politics, even at historical moments where they have appeared possible. This is an important phenomenon. The inability of states to maintain encompassing alliances after winning global wars sustains the anarchic structure of the international system, undermines efforts to establish institutions for collective security, and can sow the seeds of future conflict. Riker suggests two distinct causal mechanisms: one arising from the aggregate capabilities of the coalition, the other from the diversity of interests within it. In the remainder of this section, we will consider each of these two causal mechanisms.

3.1 "Size" and Military Capabilities

In world politics, the optimal size of a coalition is not simply a matter of numbers, but of capabilities. The aggregate military capability of an alliance constitutes its "size." Contrasting world politics to legislative institutions, Riker (1962, 67) acknowledged that "[t]he players are the nations, but otherwise the rules are rather vague." Although he understood that the weight accorded each nation was difficult to specify precisely Riker (1962, 47-8; 77-80), Riker nevertheless suggested that there exists an optimal level of overall military capabilities for an alliance. States considering an alliance will strive to reach this level, but will not aggregate beyond it. Riker's position echoes realist thinking. As Morgenthau (1948, 102) remarked, "a policy of alliances is, then, a matter not of principle but of expediency. A nation will shun alliances if it believes that it is strong enough to hold its own unaided or that the burden

of the commitments resulting from the alliance is likely to outweigh the advantages to be expected."

Riker's argument about aggregate military capabilities follows directly from the more general logic of the size principle. Powerful coalitions should not invite additional allies when they add little or nothing to the probability of victory and diminish the benefits each member receives. This claim applies to any coalition where members must divide a finite stream of benefits. More partners in an offensive alliance implies that each gets a smaller share of the spoils of victory. Defensive alliances are a tougher case because the safety they provide might be a public good to their members. As we will explain in more detail shortly, though, the pure public good of defense is not the only consideration in forming a defense pact. The size principle applies to both offensive and defensive alliances.

Riker uses the decline of American hegemony after World War II to illustrate how concern about overall military capacity pushes alliances toward an optimal level of aggregate power and limits their growth beyond that point. He points out first that maintaining minor allies was simply not worth the necessary side payments for the United States: "[i]n concentrating on the rehabilitation of its European allies, the United States in effect decided that it would not pay very much to keep China, which perhaps seemed relatively worthless in military potential and would have been an extremely expensive ally to maintain" (Riker 1962, 225). At the same time, because the Soviet-led coalition was too small to insure victory, the value of additional allies was relatively greater to the Soviet Union, which was willing to outbid the United States in many cases: "So it seems that the Soviet Union will continue to gain allies, both neutrals and former allies of the West, until in the world's opinion the two great coalitions are roughly equal in size" (Riker 1962, 228).

The size principle thus implies the following hypothesis about military capabilities:

Hypothesis 1: *The probability of alliance formation has an inverted U-shaped relationship to the total military capabilities of the potential coalition, rising to a certain point, then declining.*

Power once played a central role in theoretical explanations for alliance formation. Indeed, most of these arguments treated alliances primarily as means of aggregating military capabilities against a common threat. Surprisingly, power plays a relatively minor role in recent research on bilateral alliances. The capabilities ratio, so common in models of interstate conflict, is absent from most alliance formation models (e.g., Lai and Reiter 2000; Gibler and Wolford 2006; Crescenzi et al 2012). Military power plays no clear role in dyadic analyses because either a weak or a powerful ally might make sense depending on that state's role in the alliance as a whole. Still, the omission of this consideration rests uneasily with most longstanding explanations for alliances.

Military capabilities reemerge as an important consideration in a multilateral setting. Providing security to members requires the alliance, *as a whole*, to compare reasonably well to potential adversaries. A weak alliance simply will not provide security and might even provoke attack. Because only the strength of the complete alliance matters, dyadic capabilities are irrelevant. It is total alliance capabilities that one must consider when assessing whether a potential multilateral alliance will form.

As we noted above, one important objection to this hypothesis concerns defensive alliances. In an early critique of Riker's work, Russett (1968, 291) objected that defense and deterrence are public goods for members of the alliance. Additional members do not diminish the current members' enjoyment of these goods. This line of argument suggests that more is always better in a defensive alliance.

In spite of this consideration, state leaders have at least two reasons to prefer a minimum winning coalition even in a defensive alliance. First, additional members' security concerns

could entrap other alliance members in wars where they have no stake. In principle, alliance contracts might avoid this problem by limiting when members could invoke them. Unfortunately, these contingencies are not always easy to forsee. Moreover, authoritative judgments, even about simple things like which state initiated a military conflict, are notoriously difficult in world politics. In anarchy, members may use ambiguous or abstract limits on the alliance to evade their obligations. Highly conditional contracts, which might be necessary to avoid entrapment, are thus not worth as much to their members.

Second, powerful alliances create an important commitment problem because they could potentially be turned against the interests of their members. The Delian League of ancient Greek city-states, formed to oppose the Persians, is one example. As Kagan (2003, 8) notes, it started out as a "voluntary alliance" but "[i]t gradually became an empire under Athenian command, functioning chiefly to the advantage of Athens." Measures that increase the alliance's credibility and military effectiveness can increase this danger to its own members. Provisions for stationing allied troops on members' territory and integrating military command structures (Leeds 2003, 808) increase members' vulnerability to their alliance partners. Even without immediate conflicts of interest with their allies, states cannot be sure that none will emerge. As Robert Jervis (1978, 168) observed, "minds can be changed, new leaders can come to power, values can shift, new dangers and opportunities can arise."

Is Russett's public-goods objection enough to eliminate the usefulness of the size principle for explaining defensive alliances? Ultimately, this is an empirical question. Defensive alliances are clearly a difficult test for Riker's argument.

3.2 Alliance Heterogeneity and Alternatives to the Size Principle

In his explanation of the causal mechanism behind the size principle, Riker referred not only to considerations arising from the military capabilities of large alliances, but also to the heterogeneity of interests in overwhelmingly large coalitions. The diversity of members'

interests in large coalitions raises the risk of entering conflicts where many members have little at stake. At the limit, the members of a very diverse coalition might end up fighting to protect states with whom they have hostile relations. For instance, Riker used this argument to explain American difficulty in recruiting allies among (then) recently independent countries, such as Egypt. "To ally with the United States has often meant that the ex-colonies ally with the ex-colonial powers and this has often been an unpalatable contract" (Riker 1962, 226).

While this discussion of alliance heterogeneity also suggests limits on alliance size, it does not follow from Riker's core argument about power and the minimum winning coalition. Though both considerations arise in large alliances, diversity of interests is unrelated to the aggregate military capability of the alliance. The problems it suggests could arise in alliances too weak to achieve military victory. Because the diversity of interests does not follow from the size principle but suggests a similar pattern in alliance politics, it is actually a competing explanation. Tests of the size principle must thus control for the diversity of potential allies. Otherwise, the aggregate military power of the alliance might proxy the diversity of interests within it, providing spurious support for the size principle. All of these considerations are intrinsically interesting and suggest processes beyond the size principle that influence multilateral alliance formation. However, we can only discuss them briefly here. Our primary interest is in avoiding spurious support for the size principle.

3.2.1 Foreign Policy Similarity

Perhaps the most obvious way to control for heterogeneity of interests is through a measure of foreign policy similarity. States with similar foreign policies should have fewer conflicts of interest, and ought to have an easier time forming an alliance. As the potential alliance gets larger, it will be increasingly difficult to find additional prospective members with so much in common. This consideration thus captures the type of heterogeneity that might replace

the size principle as an explanation for the absence of overwhelming coalitions. Some other studies of alliance formation have used such a variable to identify likely allies (e.g., Lai and Reiter 2000; Gibler and Wolford 2006; Gibler 2008; Fordham 2010).

Hypothesis 2: *Groups of states with similar foreign policies should be more likely to form an alliance.*

One problem with controlling for the heterogeneity of the prospective alliance partners is that the most widely used measure of foreign policy similarity, the S-score, is based on alliance portfolio similarity (Signorino and Ritter 1999). It is thus arguably related to alliance formation in a nearly tautological way. This would be a serious problem if we were using this variable to test our primary hypothesis. However, because it embodies an alternative explanation, the bias only serves to make our test of the size principle more demanding.

3.2.2 Regime Type

Regime type is another potential source of heterogeneous interests within an alliance. As alliances get larger, they will eventually have to include states whose different regime types might make them more difficult allies.

Previous research makes several claims about the influence of regime type on alliance formation, each suggesting a different alternative to the size principle. One possibility is that similar regimes should be more likely to form alliances. Russett (1968, 286) argues that "frequently ties of common ideology, similar political or economic institutions...are mentioned" to show that "countries ally with others that share certain common interests." Because politically similar states might share more common interests, they may be better allies. Reaching a similar conclusion through a different line of argument, Leeds (1999) suggests that democratic states' greater credibility and autocratic states' greater policy flexibility makes similar regimes more attractive partners.

Hypothesis 3: *Groups of states with similar regime types should be more likely to form an alliance.*

A second proposed linkage between regime type and alliance formation focuses exclusively on democracies (Siverson and Emmons 1991). In theory, democracies should make better allies because they can generate higher audience costs and thus make more credible commitments. Democratic leaders will incur public disapproval, perhaps losing office, if they renege on public commitments. They are thus more attractive alliance partners. Autocratic leaders face no such constraint, so their commitments are less reliable. This argument differs from claims about regime similarity because it does not imply that autocracies should be more likely to align together. The claim is simply that democracies should be more likely to form an alliance.

Hypothesis 4: *As a group of states becomes more democratic, the probability of alliance formation increases.*

Others present results contradicting this claim about democratic alliance formation. Gibler and Wolford (2006) argue that democracy follows alliance formation: alliances reduce territorial conflict, making it possible for states to adopt more democratic institutions. They find a negative relationship between joint democracy and dyadic alliance formation (Gibler and Wolford 2006, 130). Democracies might also be less likely to form an alliance simply because they have a more rigorous and politically costly ratification procedure (Morrow 2000, 80; Gibler and Wolford 2006, 132). This leads to our final hypothesis:

Hypothesis 5: *As a group of states becomes more democratic, the probability of alliance formation decreases.*

3.2.3 Geographic Distance and the Size Principle

Like capabilities, geographic distance has long been recognized as a key determinant of alliance formation. Recent research finds that distance is negatively associated with the probability of dyadic alliance formation (e.g., Lai and Reiter 2000; Gibler and Wolford 2006; Fordham 2010). Pooling military resources is less costly across shorter distances. Geography is another source of heterogenous interests within an alliance because states in different locations will have different preferences about the placement of the alliance's military resources.

Like foreign policy similarity and regime type, geographic distance offers an alternative explanation for the absence of overwhelming coalitions in world politics. As a proposed alliance becomes larger, its members will necessarily cover more geographic space and longer distances. If distance between the allies, rather than the logic of the minimum winning coalition, renders large alliances unattractive, then omitting distance from the model could produce spurious support for the size principle.

In a dyadic setting, distance is a relatively simple matter. There is only one relevant conceptual variable to measure when considering just two states. In a multilateral setting, however, the number of potentially relevant distances rises rapidly with the number of states. As noted earlier, selecting the best measure of distance within a multilateral setting is a theoretical issue, not a technical one. The key theoretical issue here is the difficulty of moving military forces to protect distant members of the alliance, since this might make a large pact unattractive. States must consider whether a potential member could actually come to the defense of its allies. The longest distance between prospective members of the alliance is a relevant quantity because it represents the most difficult feat of power projection that the alliance could entail.[3] This suggests a relatively straightforward hypothesis:

Hypothesis 6: *The longest distance between prospective alliance partners should be negatively related to the probability of alliance formation.*

A second consideration is the geographic configuration of the prospective allies. Some states are unlikely alliance partners because they do not share a border. States with limited power projection capability might only be able to assist a contiguous ally. Compactness is thus potentially important. Including a prospective alliance member that shares a border with most or all other prospective members poses fewer military difficulties than adding a geographically distant member. Major powers may be able to project power over long distances, but even these states should prefer relatively compact sets of allies. Negotiating a single agreement with a geographically compact set of states is simpler than extending the same commitment to a less compact set. Promises made in one area of the world might lack credibility elsewhere. This might explain why the United States established separate alliances in different parts of the world, rather than extending the North Atlantic Treaty to Japan, Australia, or Brazil.

Hypothesis 7: *The proportion of a group of states that share a border will be positively related to the probability of alliance formation.*

That geographic distance modifies power in a multilateral setting is sometimes overlooked. For example, Hemmer and Katzenstein (2002, 581) argue that material capabilities are insufficient for explaining why the United States formed a multilateral alliance in Europe after World War II, but a series of bilateral alliances in Asia. They attribute this difference to racism and xenophobia: "Shaped by racial, historical, political, and cultural factors, U.S. policymakers saw their potential European allies as relatively equal members of a shared community. America's potential Asian allies, in contrast, were seen as part of an alien and, in important ways, inferior community" (Hemmer and Katzenstein 2002, 575).

Geographic distance offers a simpler explanation. When the Southeast Asia Treaty Organization (SEATO) was formed in 1954, its non-Atlantic members were Australia, New Zealand, the Philippines, Thailand, and Pakistan. None of these states are contiguous. The

shortest distance between any of them is 577 miles (Pakistan to Thailand). The longest distance is between Pakistan and New Zealand (12,338 miles), while even Australia and New Zealand are separated by more than 1,000 miles of ocean. If, in fairness to Hemmer and Katzenstein, one considers only the "non-Western" states (i.e. excluding Australia and New Zealand), the longest distance is still 2,909 miles (Pakistan to the Philippines). By contrast, of the 10 European states that formed NATO in 1949, 9 were contiguous with at least one other member. (The number rises to 10 if one considers Iceland to be contiguous with Britain.) The longest distance between any two of those states was 2,409 miles (Iceland to Italy). The European allies of the United States are geographically more suited to a multilateral alliance than are the more widely dispersed allies of the United States in Asia.

3.2.4 Threat

International threats should also influence alliance formation. As Liska (1962, 12-13) famously observed, "alliances are against, and only derivatively for, someone or something." Threats are closely related to capabilities because the more powerful the threats, the greater the capabilities needed to ensure victory. As Waltz (1979, 181) makes clear "weaker powers could improve their positions through alliance, by adding the strength of foreign armies to their own" (See also Morgenthau 1948, and Organski 1968.) While Walt (1987) argues that material capabilities are not sufficient for identifying a threat and Morrow (1991) points out that some members may acquire non-security benefits from an alliance, the need for protection is still a primary reason for alliance formation. Otherwise, the states would form an alternative type of agreement. This leads to the following hypothesis:

Hypothesis 8: *The probability of alliance formation is positively related to the level of threat facing the potential coalition.*

4 Research Design

To test these hypotheses, our unit of analysis must capture groups of states. We also require variables capturing our key explanatory factors and that control for potentially confounding factors. This section explains how we approached each of these challenges.

4.1 Unit of Observation and Dependent Variable

To study alliance formation, we must compare groups of states that formed alliances with groups of states that did not. Comparing groups rather than pairs of states requires rethinking the standard dyadic research design. As detailed in Poast (2010), using dyadic data to analyze multilateral events will lead to flawed inferences. Instead, scholars should use what Poast calls 'k-adic' data. A k-ad is an observation with k members, where $k \geq 2$. Hence, a dyad is a special case of k-adic data, where $k = 2$ for all observations.

Creating a k-adic dataset requires two steps. First, we construct a data set containing only k-ads for which the event of interest occurred. Here, the event is when all of the states in a group, regardless of the number of participants, form an alliance in year t. These are the "event" k-ads. Hence, one event k-ad is the 1832 bilateral alliance between England and France, while another event k-ad observation is the 1854 trilateral alliance between England, France, and Prussia. It is possible (even likely) that the formation of the earlier bilateral alliance between England and France may affect the probability of forming the later trilateral alliance. As a robustness check, we will control for the proportion of states in the k-ad that were previously allied. However, since the inclusion of this variable does not alter our findings, we report these results in the Online Appendix.

Second, one needs to acquire the "non-event" k-ads. Here, a non-event k-ad is a set of states that did not form an alliance. For example, the triad England, Russia, and the United States in the year 1912 is a non-event because these states did not form alliance in that

year. How many non-event k-ads do we require? We could apply the analog to dyadic data, i.e., including all possible combinations of k-ads in the data set. In practice, however, this would lead to large data sets that are computationally infeasible. For example, if one had 100 countries in a dataset, all combinations of 100, 99, 98, 97, ..., down to 2 countries would result in a dataset of 1.26765×10^{30} observations!

We therefore sample from the population of non-event k-ads. When conducting our sampling, we stratify the sample by k. A purely random sample of non-event k-ads is biased because the event k-ads tend to have small values of k, whereas a random sample of non-event k-ads would include a disproportionately large number of k-ads with large values of k.[4] Therefore, we apply choice-based sampling to create a stratified sample (King and Zeng 2001). Choice-based sampling works as follows. First, we consider the distribution of k in the event k-ads. Suppose 80 percent of the event k-ads have $k = 2$, 10 percent of the event k-ads have $k = 3$, five percent of the event k-ads have $k = 4$, three percent of the event k-ads have $k = 5$, and two percent of the event k-ads have $k = 6$. Given this distribution, the analyst then creates a random sample of non-event k-ads that follow the same distribution: 80 percent of the non-event k-ads should have $k = 2$, 10 percent of the non-event k-ads should have $k = 3$, five percent of the non-event k-ads should have $k = 4$, three percent of the non-event k-ads should have $k = 5$, and two percent of the non-event k-ads should have $k = 6$. In other words, choice-base sampling entails creating a random sample of non-event observations that are stratified according to the distribution of observations in which the event occurred.

How many non-event observations should the analyst collect? According to King and Zeng (2001, 702), one can collect anywhere from two to five times more 0's than 1's, though one should attempt to collect as many 0 values as is computationally feasible. Consider a simple example. Suppose the event k-ads dataset contains 100 dyads that formed an alliance (i.e. there were 100 bilateral alliances formed) and 50 triads that formed an alliance (i.e.

there were 50 trilateral alliances formed). If this is the case and the analyst wishes to use 10 times as many non-event observations as event observations, the non-event sample could contain 1000 dyads that did not witness the formation of an alliance and 500 triads that did not witness the formation of an alliance. Stated differently, if one is working with a binary dependent variable (where the dependent variable, Y, equals 1 when the event occurred, zero otherwise), then the final dataset will have 100 dyads where $Y = 1$, 50 triads where $Y = 1$, 1000 dyads where $Y = 0$, and 500 triads where $Y = 0$.[5]

To construct a k-adic dataset, we begin by using the ATOP dataset to identify alliance formations (Leeds et al 2002). Alliance formation is captured in the binary dependent variable *Alliance Formation*, which equals 1 in year t for k-ad i when k-ad i formed a defense pact in year t. The possibility that collective defense is a public good makes defensive alliances an especially hard test of the size principle. There are 155 dyads that witnessed the formation of unique bilateral alliances, 29 triads that witnessed the formation of unique trilateral alliances, 12 four-ads that witnessed the formation of unique quadrilateral alliances, 10 five-ads that witnessed the formation of pentalateral alliances, while 60 other k-ads formed alliances with 6 or more members (the largest was an alliance of 48 members). This gives us a total of 266 k-ads that formed alliances.

Next, we create 2*266 = 532 k-ads that did not witness the formation of an alliance (in the Online Appendix, we consider alternative sized samples of non-event k-ads and alternative random draws of the non-event k-ad members). The distribution of these 532 k-ads is approximately the same as the k-ads that witnessed alliance formations. To make this point clear, Table 1 reports the relevant number of event and non-event k-ads of various sizes.

[Table 1 about here]

When creating this k-adic dataset, we wish to retain a panel structure. Therefore, the unit of observation is now the k-ad-year. It should be noted that our panel dataset is

unbalanced, as each k-ad only has yearly values for the years in which each state in a k-ad was a member of the international system. We have data for each year in which the k-ad existed from 1816-2001. This gives us a total of 26,600 observations, though missing data on some of our independent variables somewhat reduces the sample size.

To further clarify the structure of our k-adic data set, Table 2 provides a sample of the actual data. It shows the first five years of k-ad number 5 (1975 through 1979) and the first five years of k-ad number 743 (also from 1975 through 1979). K-ad number 5 has two members (the United States and Panama), while k-ad number 743 has three members Luxembourg, Finland, and Romania). The final column indicates that the United States and Panama formed an alliance in 1977.

[Table 2 about here]

We note that k-adic data do not preclude using other methods to capture extra-dyadic concepts, such as network analysis. For example, it is possible that homophily (i.e. nodes sharing one or more attributes in common are more likely to form links than nodes having different attributes) between states in a k-ad can influence alliance formation. It might also prove insightful to consider the network connections between k-ads (where a k-ad is a node). Though it is not the focus on this paper, future work should explore this approach.

4.2 Key Independent Variable: K-adic Capabilities

To test Hypothesis 1, our main hypothesis, we use the total Composite Index of National Capabilities (CINC) score of the k-ad and its square.[6] Finding that the total CINC score of the k-ad is positively related to the probability of alliance formation and that its square is negatively related will support our hypothesis. One potential problem with this test is that the number of states in the alliance is likely to be correlated with its total military capabilities. The transaction costs of setting up an agreement involving a large number of

states might produce the same pattern in the data as the size principle. To guard against such spurious support for the size principle, we control for the number of states in the k-ad and the square of this number. Overall, the total CINC score of the k-ad has a minimum value of 0, a maximum value of 0.89, a mean value of 0.07 and a standard deviation of 0.12.

4.3 Control Variables

We also include a variety of control variables.[7]

Foreign Policy Similarity: To operationalize foreign policy similarity within the k-ad, we use the average S-score of all the dyads within the k-ad (Signorino and Ritter 1999).

Regime Type: To operationalize democracy k-adically, We employ two measures. First, we apply the "weakest link" principle of Oneal and Russett (1997) by using the lowest of the 21 point *Polity* score of states in the k-ad (ranging from -10, complete autocracy, to 10, complete liberal democracy) (Marshall, Jaggers, and Gurr 2013). This variable enables us to test hypotheses 4 and 5. Second, we use the largest difference in polity scores between members of a k-ad. This allows us to test hypothesis 3. To ensure that our findings are not simply a product of our decision to apply the "weakest link" principle, we conduct a series of tests using different indicators of regime type. These include the mean polity score of the k-ad, the proportion of states with a polity score above the traditional democracy cut point of polity=6, and the maximum polity score in the k-ad. These tests, which do not affect our principal findings, are reported in the Online Appendix.

Geographic Distance: We consider two measures of distance that make use of capital-to-capital distance, treating contiguous states as having a distance of zero.[8] The first measure is the maximum distance between any of two states in the k-ad. This enables us to test hypothesis 6. The second measure is "Compactness", which is the proportion of states in the k-ad that are contiguous to one another. This enables us to test hypothesis 7.

Threats: To capture the magnitude of possible threats, we use the measure of inter-

national threats proposed by Leeds and Savun (2007). The Leeds and Savun measure of threats facing state i is created in three steps. First, one begins with the set of all major powers and the set of states contiguous to state i. Second, members of this "politically relevant international environment" that had no alliance with state i and had an S-score with state i that was lower than the median for all politically relevant dyadic relations in the international system during the 1816-2000 period (0.775) are considered potential threats. Third, one sums the CINC scores of all potential threats facing state i. Having created this measure for all states inthe system, we convert it into a k-adic measure of threat by using the average score of all the states in a k-ad. The average becomes higher as more members of the potential coalition face greater threats and fewer potential members face relatively little threat. This captures the role of threat better than other aggregations such as total level of threat, which always rises with the number of states in the k-ad. This allows us to test hypothesis 8.

5 Empirical Analysis

Since we have a binary dependent variable and event data, we estimate a rare-event logit model with cubic splines (Beck, Katz, and Tucker 1998). The rare event logit applies a post-estimation correction to the constant term to account for how sampling on the dependent variable can artificially inflate the prominence of observations where the dependent variable equals 1.[9] We also include the time since the emergence of the k-ad or the collapse of its last alliance, and the number of previous alliances in the k-ad.

5.1 Power and the Size Principle

Table 3 presents the results of two models that test our primary claim about the size principle, hypothesis 1. The first model includes only total capabilities and its square, as well as controls

for time dependence. The second adds the number of states in the k-ad and its square. These additional variables produce only very small changes in the coefficient estimates for the total CINC score and its square, suggesting that the transaction cost of setting up a very large alliance is not driving our findings concerning military capabilities.

[Table 3 about here]

The coefficient on the total military capabilities of the k-ad is positive and statistically significant in both models. The coefficient on the square of total capabilities is negative and statistically significant. These results provide strong support for the size principle. To illustrate, consider Figure 1, which uses the first model in Table 3 to depict the relationship between total capabilities and the probability of alliance formation. The probability peaks when the total capabilities of the states in the k-ad are near 0.3.[10] Beyond this point, the probability of alliance formation declines. This result makes substantive sense. An alliance of this size would have had the capabilities to compete with any state in the system. Only the United States in the 20th Century, Britain in the 19th, and Germany in 1941 had CINC scores exceeding 0.2. External threat would rarely have required an alliance with total capabilities in excess of 0.3. By contrast, alliances with capabilities less than 0.2 would find it difficult to deal with an opposing coalition containing one or more major powers. The data suggest that neither sort of alliance is especially likely. Instead, as the size principle suggests, states tend to form alliances that are militarily strong but not overwhelming.

[Figure 1 about here]

Table 4 presents three more models of multilateral alliance formation. These models attempt to capture various forms of heterogeneity among potential alliance partners. As noted above, heterogeneity within the prospective alliance is a competing explanation for the size principle's account of how power influences alliance formation. Such heterogeneity could

independently account for the absence of "overwhelming majorities" in world politics (Riker 1962, 66). Larger alliances should usually face greater geographic distances and contain an increasingly diverse array of political regimes and foreign policies. If these considerations accounted for the absence of very large alliances, then including these variables should diminish or eliminate the effect of military capabilities.

[Table 4 about here]

Table 4 indicates that these controls modify the effect of capabilities but do not eliminate it. Figure 2 employs the coefficients in the third model to illustrate our findings. It depicts the relationship between total capabilities and the probability of alliance formation when the other variables included in this model are held at their mean values (the vertical axis in this figure is identical to that in Figure 1 in order to facilitate comparison). Including controls does not change the total CINC score associated with the highest predicted probability of alliance formation; it remains at roughly 0.3. However, the the presence of these additional variables attenuates the marginal effect of a move from a total CINC score of 0.1 to 0.3. This move increased the probability of alliance formation by 78 percent in the model without controls, but by only 26 percent when they are included.

[Figure 2 about here]

In principle, it is possible that better measures of common interests might still eliminate the findings in support of the size principle. However, unless these measures are unrelated to considerations already in the model – geographic proximity, foreign policy similarity, and regime type – they probably will not eliminate the relationship between total capabilities and alliance formation. It is more likely that power really does have the effect suggested by Riker, Morgenthau, and others. Realists stress the need to behave cautiously in the anarchic international system and argue that states should be reluctant to contribute to

an overwhelming coalition. Such a coalition could be dangerous because even states that appear friendly today could become enemies in the future. Our results are consistent with the claim that state leaders' concerns about the power of their prospective allies are enough in themselves to inhibit alliance formation even after controlling for similarity with these potential partners. Put differently, our results suggest that the size principle is truly about power, not the diversity of states' immediate interests or regime types. Power is not the only important consideration, but it does make a substantial difference.

To gain further confidence in our findings regarding the size principle, we conducted a series of additional robustness checks. While all of the results are available in the Online Appendix, a few are worth highlighting here. First, to show that our claims do not apply to only defensive pacts, we estimated the model using all alliances. The substantive results were the same. Second, there are proportionately fewer non-event k-ads with major powers than event k-ads with major powers. To address this concern, we created a dataset that oversampled major power k-ads. We adjusted the non-event generating algorithm so that it generated a random number between 0 and 1 when it identifies the first member of a k-ad. When this number was equal to or greater than 0.5, the first member of the k-ad was drawn from among the major powers. When the number was less than 0.5, the first member of the k-ad was drawn from the list of all countries, including the major powers. In our regular dataset, 41 percent of the k-ads contained at least one major power. In the oversampled dataset, 62 percent of all the k-ads had at least one major power and 59 percent of the "non-event" k-ads had at least one major power. The results remained largely unchanged, but were sensitive to the selection of a threshold for selecting a major power k-ad. For example, lowering the threshold to 0.3 results in a few more observations with major powers (69 percent) and causes the coefficients on total CINC and total CINC squared to be statistically insignificant at the 0.90 confidence level.

5.2 Testing the Alternative Hypotheses

Although geographic distance does not explain the pattern indicated by the size principle, it does have an important effect on alliance formation. Using the estimates from the first model in Table 4, an increase in the longest dyadic distance in the k-ad from 0 to the mean value in our sample (3,268 miles) decreases the probability of alliance formation by nearly 70 percent, from 0.013 to 0.004. An increase in the proportion of contiguous dyads in the k-ad from 0 to 0.19, the mean value in our sample, increases the probability of alliance formation by roughly 10 percent, from 0.0040 to 0.0044. These effects are nearly identical when we include other control variables.

As expected, we find that the average level of threat has a positive and statistically significant relationship with alliance formation. In a supplemental analysis, we sought to determine whether the average threat level modified the relationship between capabilities and alliance formation. The optimal aggregate capabilities of the alliance might be greater when the average threat is higher. We tested this hypothesis by (1) splitting our sample into subsets with observations facing threats above and below the mean on the threat variable, and then (2) estimating model 1 from Table 1 on each subset. As the results in the Online Appendix indicate, the level of aggregate capabilities associated with the highest probability of alliance formation is nearly the identical in both subsets.

Our results concerning regime type are less robust than those concerning threat or geographic distance. Model 2 indicates that joint democracy, measured as the lowest polity score in the k-ad, makes alliance formation less likely, while political similarity, measured as the largest difference between polity scores of states in the dyad, makes alliance formation more likely. In other words, sets of highly autocratic states are most likely to form an alliance, while more heterogeneous groups, especially those including some democracies, are less likely to do so. Using our model, a completely homogenous set of consolidated autocracies has a 0.18 probability of alliance formation, while a homogenous set of consolidated democracies

has a 0.005 probability of forming an alliance. A set of states with both consolidated democracies and a consolidated autocracies have a 0.006 probability of alliance formation. These values represent the maximum possible effects associated with regime type and they are not robust to the addition of the control variables included in model 3 in Table 4.

To ensure that our relatively weak results concerning regime type are not simply a product of our decision to apply the "weakest link" principle, we conduct a series of tests using different indicators of regime type. These include the mean polity score of the k-ad, the proportion of states with a Polity score above 6 (the most commonly used threshold for democracy), and the maximum polity score in the k-ad. Each of these tests produced essentially the same results as the models in Table 3. Regime type has been a central concern of dyadic research on alliance formation. Our analysis suggests that researchers might fruitfully shift their attention to other sources of explanation, such as the size principle.

6 Conclusion

Alliance formation is a multilateral process. In principle, any alliance could have three or more members. Leaders in a multi-state system must assess the value of the alliance as a whole. Whether a particular partner makes sense could depend on the inclusion or exclusion of other states from the pact. For this reason, *all* alliances are (potentially) multilateral. Taking this point seriously requires rethinking much previous research.

Because it is impossible to test in a dyadic framework, Riker's argument about the role of power has received little attention in recent research. Our results show that this neglect is unjustified. Conceptualizing alliance formation in a multilateral way makes clear that the size principle offers a strong explanation for the absence of overwhelming coalitions in world politics. Empirically, we found an inverted-U shaped relationship between the total power of a group of states and the probability that the group will form an alliance. The probability of

alliance formation is greatest when the combined capabilities of the alliance are larger than any one major power, but not large enough to dominate the system – much weaker or much stronger alliances are much less likely to form. Such a finding is not possible using a dyadic research design.

The size principle suggests that power itself is the principal explanation for this pattern. As realists such as Morgenthau have suggested, the aggregate capabilities of a prospective alliance affect the value of the pact to each of its members, and thus the probability that each member will agree to it. While the logic of power behind the size principle might be less intuitive than the need to manage interest heterogeneity within a large alliance, the evidence offers greater support for the size principle than for other considerations. Military capabilities do not function as a proxy for the transaction costs, geographic distances, political diversity, or divergent foreign policies associated with large alliances. The number of states in the prospective alliance, their geographic configuration, and the similarity of their foreign policies all influenced alliance formation but none was substantively more important than power. Regime type has arguably received more attention than any other influence on alliance formation in recent research. Ironically, we found that it had far less effect than did military power, a consideration that recent research on alliance formation has largely overlooked.

We stress the role of power because recent research largely neglects this consideration, not because it is the only influence on multilateral alliance formation. As our control variables suggest, the diversity of interests matters, too. Indeed, this diversity, rather than considerations of military capabilities, probably explains Spain's late entry into NATO, one of our opening examples. Many concepts explored in previous research on dyadic alliance ties are also relevant in a multilateral setting, including moral hazard (Yuen 2009; Benson, Meirowitz, and Ramsay 2014), the costs of peacetime coordination (Morrow 1994), and the costs of public goods provision (Olson and Zeckhauser 1966; Sandler and Hartley 2001). These issues deserve further attention in a multilateral framework. However, when our em-

pirical results differ from those of dyadic analyses, the k-adic findings should be given priority because the phenomenon of alliance formation is multilateral. While our results suggest that previous dyadic research might be on the right track, there is no reason to assume that this will always be the case.

We emphasize that switching from a bilateral to a multilateral setting is not merely a technical change in modeling technique. Instead, it requires revisiting the logic leading to hypothesized relationships in the first place. This theoretical exercise cannot be avoided. Dyadic statements about multilateral processes like alliance formation are necessarily artificial. Until they can be understood in a multilateral setting, their relationship to the phenomena they purport to explain will remain tenuous.

Notes

[1]Quoted in Remak 1967, 5.

[2]See section 2.

[3]One could also make a case for average distance. These two numbers are highly correlated and produce essentially the same results.

[4]This is because, a k increases, the number of possible k-ads in a given set of countries increases exponentially.

[5]To alleviate concerns of receiving an odd sample of non-event k-ads, the Online Appendix presents results from 1,000 alternative draws of non-events.

[6]Singer, Bremer, and Stuckey 1972; Singer 1987; Correlates of War Project 2005.

[7]Summary statistics are in the Online Appendix. We assembled portions of our dataset using EUGene 3.203 (Bennett and Stam 2000).

[8]Singer and Small 1982. We also conduct our analysis using the minimum distance metric developed by Weidmann, Kuse, and Gleditsch (2010). Minimum distance is measured by the smallest distance between the borders of the two states (represented by polygons).

[9]See the Online Appendix for an explanation for why we do not use inverse weighting.

[10]We note that with a mean value of 0.07 and a standard deviation of 0.12, this value is within two standard deviations of the mean.

References

Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42(4): 1260-88.

Bennett, D. Scott, and Allan Stam. 2000. "EUGene: A Conceptual Manual." *International Interactions* 26: 179-204.

Benson, Brett. 2011. "Unpacking Alliances: Deterrent and Compellent Alliances and Their Relationship with Conflict, 1816-2000." *Journal of Politics* 73(4): 1111-27.

Benson, Brett, Adam Meirowitz, and Kristopher Ramsay. 2014. "Inducing Deterrence Through Moral Hazard in Alliance Contracts." 58(2): 307-335.

Besanko, David and Ronald Braeutigam. *Microeconomics*. Wiley. 2011

Bremer, Stuart. 1992. "Dangerous Dyads: Conditions Affecting the Likelihood of Interstate War, 1816-1965." *Journal of Conflict Resolution* 36(2): 309-41.

Carter, David and Curtis Signorino. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis*. Summer 2010. pp. 271-292.

Christensen, Thomas J., and Jack Snyder. 1990. "Chain Gangs and Passed Bucks: Predicting Alliance Patterns in Multipolarity." *International Organization* 44 (2): 137-68.

Correlates of War Project. 2005. National Material Capabilities Data, version 3.02. Online dataset, http://www.correlatesofwar.org/. Accessed 12 January 2012.

Cranmer, Skyler, Bruce A. Desmaris, and Elizabeth J. Menninga. 2012. "Complex Dependencies in the Alliance Network." *Conflict Management and Peace Science* 29(3): 279-313.

Crescenzi, Mark J.C., Jacob D. Kathman, Katja B. Kleinberg, and Reed M. Wood. 2012. "Reliability, Reputation, and Alliance Formation." *International Studies Quarterly* 56(2): 259-74.

Croco, Sarah and Teo, Tze Kwang. 2005. "Assessing the Dyadic Approach to Interstate Conflict Processes: A.k.a. "Dangerous" Dyad-Years." *Conflict Management and Peace Science*. 22: 5 - 18.

Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands and Sinking Costs." *Journal of Conflict Resolution* 41(1): 68-90.

Fordham, Benjamin O. 2010. "Trade and Asymmetric Alliances." *Journal of Peace Research* 47(6): 685-96.

Gibler, Douglas M., and Meredith Reid Sarkees. 2004. "Measuring Alliances: The Correlates of War Formal Interstate Alliance Dataset, 1816-2000." *Journal of Peace Research* 41 (2): 211-22.

Gibler, Douglas M., and Scott Wolford. 2006. "Alliances, Then Democracy: An Examination of the Relationship between Regime Type and Alliance Formation." *Journal of Conflict Resolution* 50 (1): 129-53.

Gibler, Douglas M. 2008. "The Costs of Reneging: Reputation and Alliance Formation."*Journal of Conflict Resolution*. 53(3): 519-542.

Hafner-Burton, Emilie M. and Alexander H. Montgomery. 2012. "War, Trade, and Distrust: Why Trade Agreements Don't Always Keep the Peace." *Conflict Management and Peace Science*. July: 257-278.

Hemmer, Christopher and Peter J. Katzenstein. "Why is There No NATO in Asia? Collective Indentity, Regionalism, and the Origins of Multilateralism." *International Organization*. Summer 2002. pp.575-607.

Hoff, Peter and Ward, Michael. 2004. "Modeling Dependencies in International Relations Networks." *Political Analysis*. 12:2: 160 - 175.

Jervis, Robert. "Cooperation Under the Security Dilemma." *World Politics*. January 1978. pp. 167-214.

Johnson, Jesse C., and Leeds, Brett Ashley. (2011), Defense Pacts: A Prescription for Peace?" *Foreign Policy Analysis* 7(1): 45-65.

Kagan, Donald M. 2003. *The Peloponnesian War*. New York: Viking.

Lai, Brian, and Dan Reiter. 2000. "Democracy, Political Similarity, and International Alliances, 1816-1992." *Journal of Conflict Resolution* 44 (2): 203-27.

Leeds, Brett Ashley. 1999. "Domestic Political Institutions, Credible Commitments, and International Cooperation." *American Journal of Political Science* 43(4): 979-1002.

Leeds, Brett Ashley. 2003. "Alliance Reliability in Times of War: Explaining State Decisions to Violate Treaties." *International Organization* 57(4): 801-27.

Leeds, Brett Ashley, Jeffrey M. Ritter, Sara McLaughlin Mitchell, and Andrew G. Long. 2002. "Alliance Treaty Obligations and Provisions." *International Interactions* 28 (3): 237-60.

Leeds, Brett Ashley, and Burcu Savun. 2007. "Terminating Alliances: Why Do States Abrogate Agreements." *Journal of Politics* 69(4): 1118-1132.

Liska, G. 1962. *Nations in alliance: The limits of interdependence.* Johns Hopkins Press.

Long, Andrew, Timothy Nordstrom and Kyeonghi Baek. 2007. "Allying for Peace: Treaty Obligations and Conflict between Allies." *Journal of Politics* 69: 1103-1117.

Maoz, Zeev, Lesley Terris, Ranan Kuperman, and Ilan Talmud. 2007. "What Is the Enemy of My Enemy? Causes and Consequences of Imbalanced International Relations, 1816-2001." *Journal of Politics* 69(1): 100-15.

Marshall, Monty G., Keith Jaggers, and Tedd Robert Gurr. *Polity IV Project.* http://www.systemicpeace.o Accessed 27 February 2013.

Morgenthau, Hans. 1948. *Politics Among Nations.* New York: McGraw Hill.

Morrow, James D. 1991. "Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances." *American Journal of Political Science* 35 (4): 904-933.

Morrow, James D. 1994. "Alliances, Credibility, and Peacetime Costs." *Journal of Conflict Resolution.* 38: 270-297.

Olson, Mancur and Richard Zeckhauser. 1966. "An Economic Theory Alliances." *The Review of Economic Statistics.* August. pp. 266-279.

Oneal, John R., and Bruce M. Russett. 1997. " The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950-1985." *International Studies Quarterly* 41(2): 267-93.

Organski, AFK. 1968. *World Politics.* New York: Alfred A. Knopf.

Poast, Paul. 2010."(Mis)Using Dyadic Data to Analyze Multilateral Events." *Political Analysis.* Autumn. 403-425.

Riker, William H. 1962. *The Theory of Political Coalitions.* New Haven: Yale University Press.

Remak, Joachim. 1967. *The Origins of World War I, 1871-1914.* New York: Holt, Rinehart, and Winston.

Russett, Bruce M. 1968. "Components of and Operational Theory of Alliance Formation." *Journal of Conflict Resolution* 12(3): 285-301.

Sandler, Todd and Keith Hartley. 2001. "Economics of Alliances: The Lessons for Collective Action." *Journal of Economic Literature.* Sepetember. pp. 869-896.

Signornio, Curtis and Jeffrey Ritter. "Tau-b or Not Tau-b: Measuring the Similarity of Foreign Policy Positions." *International Studies Quarterly.* 43 1999. pp. 115-144.

Singer, J. David. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985." *International Interactions.* 14: 115-32.

Singer, J. David, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." in Bruce Russett, ed., *Peace, War, and Numbers.*Beverly Hills, CA: Sage. 19-48.

Singer, J. David, and Melvin Small. 1982. *Resort to Arms.* Beverly Hills, CA: Sage Publications.

Siverson, Randolph M., and Juliann Emmons. 1991. "Birds of a Feather: Democratic Political Systems and Alliance Choices in the Twentieth Century." *Journal of Conflict Resolution* 35 (2): 285-306.

Simon, Michael W., and Erik Gartzke. 1996. "Political System Similarity and the Choice of Allies: Do Democracies Flock Together, or Do Opposites Attract?" *Journal of Conflict Resolution* 40 (4): 617-35.

Smith, Alastair. 1995. "Alliance Formation and War." *International Studies Quarterly.* 40(1): 133-53.

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

Waltz, Kenneth. 1979. *A Theory of International Politics.* New York: McGraw-Hill.

Weidmann, Nils B., Doreen Kuse and Kristian Skrede Gleditsch. 2010. "The Geography of the International System: The CShapes Dataset." *International Interactions.* 36(1):86-106.

Wolfers, Arnold."Alliances." in David L. Sills, (ed). *International Encyclopedia of the Social Sciences*. Vol. I. New York: Macmillian and The Free Press. 1968.

Yuen, Amy. 2009. "Target Concessions in the Shadow of Intervention." *Journal of Conflict Resolution*. 53(5):727-744.

Table 1: Number of Event and Non-Event k-ads, by k-ad Size

k-ad Size	Number of k-ads in Dataset
Event k-ads	
k=2	155
k=3	29
k=4	12
k=5	10
k=6	6
k=7	4
k=8 or more	50*
Non-Event k-ads	
k=2	310
k=3	58
k=4	24
k=5	20
k=6	12
k=7	8
k=8 or more	100*

Note: This table does not include the yearly observations for each k-ad.
* Complete distribution for all categories found in the Online Appendix

Table 2: Sample of K-adic Data

Kad-id	Year	Number of Members	Member 1	Member 2	Member 3	Member 4	Alliance Formation?
5	1975	2	USA	Panama	.		No
5	1976	2	USA	Panama	.		No
5	1977	2	USA	Panama	.		Yes
5	1978	2	USA	Panama	.		No
5	1979	2	USA	Panama	.		No
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
743	1975	3	Luxembourg	Finland	Romania	.	No
743	1976	3	Luxembourg	Finland	Romania	.	No
743	1977	3	Luxembourg	Finland	Romania	.	No
743	1978	3	Luxembourg	Finland	Romania	.	No
743	1979	3	Luxembourg	Finland	Romania	.	No

Table 3: Capabilities, Size, and Alliance Formation, Rare Events Logit

	Model 1	Model 2
Total CINC of k-ad	8.62***	8.26***
	(1.25)	(1.19)
Total CINC of k-ad	-14.30***	-14.89***
	(3.00)	(2.78)
Number of Members		0.10***
		(0.03)
Number of Members Squared		-0.00
		(0.00)
Base Controls		
Time Since Previous Defense Pact	-0.13***	-0.10**
	(0.04)	(0.04)
Number of Previous Alliance Failures	-1.33***	-1.38***
	(0.22)	(0.23)
Constant	-4.07***	-4.52***
	(0.17)	(0.21)
Number of Observations	24,611	24,611

Table 4: Capabilities, Size, Distance, Regime Type and Alliance Formation, Rare Events Logit

	Model 1	Model 2	Model 3
Total Capabilities	3.87***	8.39***	4.26***
	(1.33)	(1.19)	(1.39)
Total Capabilities Square	-7.70**	-14.95***	-7.81**
	(3.04)	(2.68)	(3.15)
Number of Members	0.21***	0.12***	0.37***
	(0.04)	(0.04)	(0.06)
Number of Members Square	-0.00	-0.00*	-0.01***
	(0.00)	(0.00)	(0.00)
Distance Controls			
Maximum Distance b/w any two states in k-ad	-0.00***		-0.00***
	(0.00)		(0.00)
"Compactness": Proportion of Contiguous States	0.40**		0.04
	(0.18)		(0.19)
Other Controls			
Time since last defpact	-0.13***	-0.11***	-0.14***
	(0.04)	(0.04)	(0.04)
Number of Previous Alliance Failures	-1.86***	-1.35***	-1.87***
	(0.23)	(0.23)	(0.23)
Maximum Polity Difference in k-ad		-0.06***	-0.02*
		(0.01)	(0.01)
Minimum Polity Score		-0.07***	-0.03*
		(0.01)	(0.02)
Total Threat in k-ad			0.32
			(0.39)
Average S-score			1.47***
			(0.35)
Constant	-3.71***	-4.47***	-5.07***
	(0.26)	(0.23)	(0.42)
Number of Observations	22,713	23,507	21,855

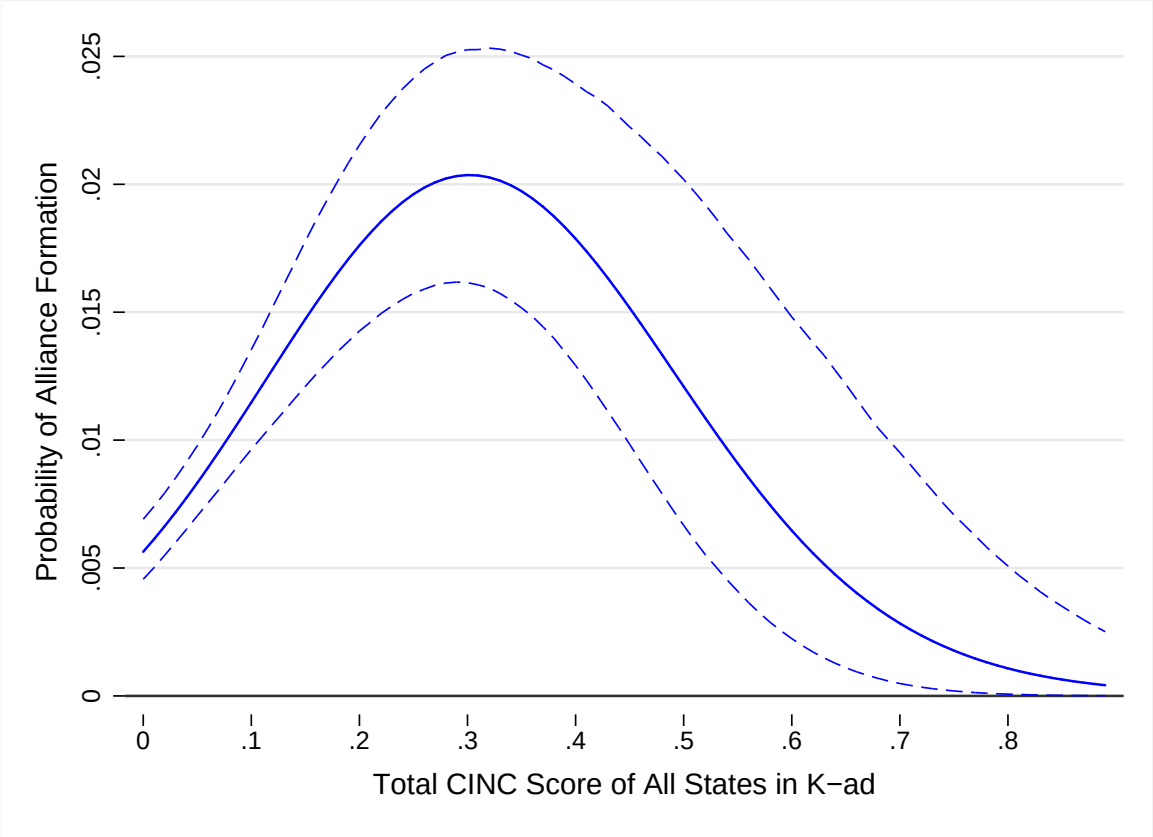

Figure 1: Relationship Between Total Capabilities and Alliance Formation.

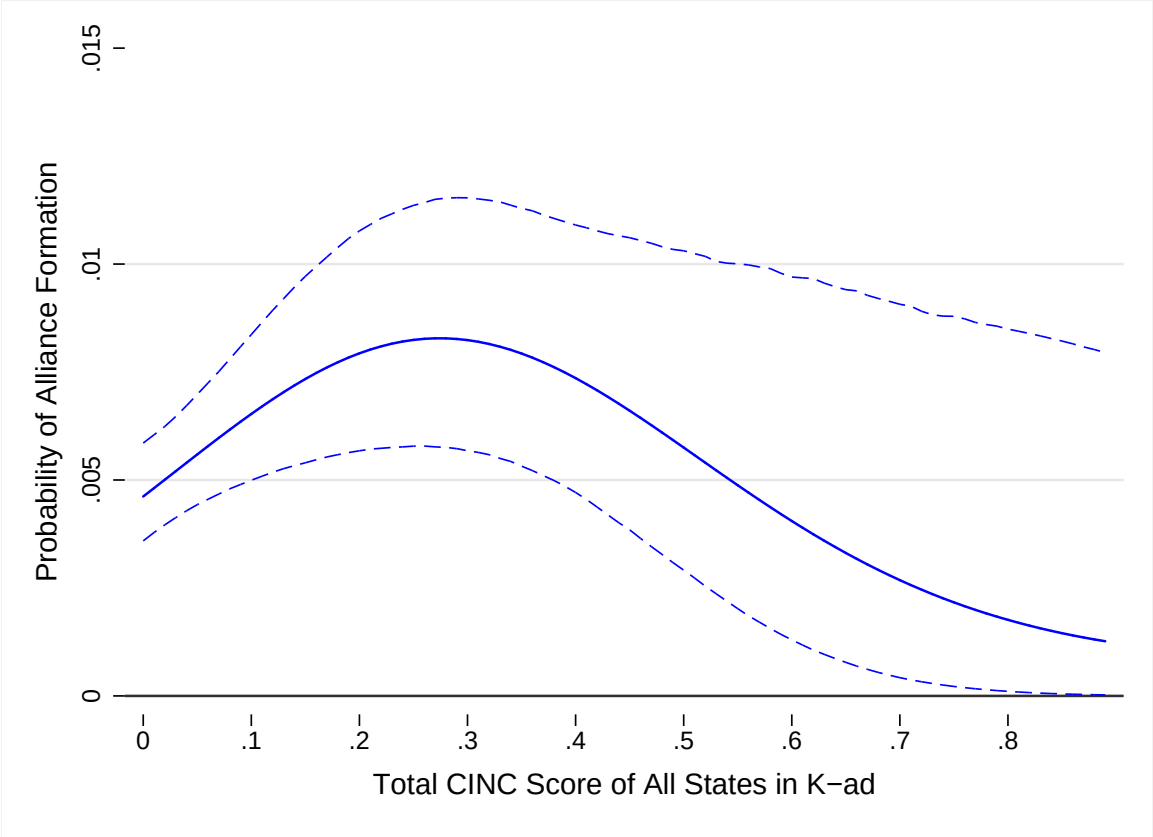

Figure 2: Relationship Between Total Capabilities and Alliance Formation, with Controls.