Do Alliances *Really* Deter?

Michael R. Kenwick, Pennsylvania State University
John A. Vasquez, University of Illinois at Urbana-Champaign
Matthew A. Powers, University of Illinois at Urbana-Champaign

The scholarly literature is still divided on the relationship between defensive alliances and interstate conflict. While some scholars argue that defensive alliances can deter conflict, others posit that alliances accelerate its approach. This article seeks to make headway in the debate by using a research design that examines whether the recent formation of defensive alliances leads to increases or reductions in militarized disputes and war. We find that this relationship differs in the pre and post-nuclear era. In the prenuclear era, alliance formation is positively associated with both the initiation of militarized disputes and war onset. In the nuclear era, however, forming certain types of alliances reduces the likelihood of militarized dispute initiation, but has no effect on whether war occurs. This suggests assertions that defensive alliances will consistently deter conflict should be tempered and that alliance formation can sometimes undermine efforts to preserve peace.

The question of deterrence has been a central part of the international relations literature since the end of the Second World War (Kahn 1960; Schelling 1960) and remains so today amid questions of how the introduction of nuclear weapons to the international system might affect alliance politics (Claire 2013; Fuhrmann and Sechser 2014). Despite the importance of this issue, scholars are divided on whether the formation of alliances with defensive or deterrent provisions is more commonly linked with peace or war. Much of the recent research on extended deterrence posits that alliances can promote peace by deterring potential opponents that may otherwise challenge alliance partners in international disputes (Benson 2012; Johnson and Leeds 2011; Leeds 2003; Smith 1995). Others argue that the formation of military alliances is part of a larger set of policy practices, or "steps-to-war," that heighten insecurity among potential disputants and perpetuate hostile spirals that increase the risk that conflict will arise (Senese and Vasquez 2008; Vasquez 1993). Given these competing expectations, no consensus has been reached on whether alliance formation is an effective policy for inducing deterrence and promoting peace.

In this article we question the extent to which alliance formation can do either. We advance previous research by conducting a novel analysis of deterrence in relation to whether the formation of defensive alliances is associated with an increase or reduction in both the initiation of militarized disputes and the occurrence of war. Focusing on the period around alliance formation allows for an appropriate test of the competing explanations linking alliances to conflict, since it accommodates both the deterrence and steps-to-war expectations and permits an examination of whether the probability of militarized disputes and wars changes after alliances are formed. We also examine whether the effect of alliance formation has changed over time to address arguments that the introduction of nuclear weapons in the international system has fundamentally altered the potential costs of war and the logic of deterrence (Geller 2012; Herz 1950; Kahn 1960).

The study is structured as follows. In the first section, we outline different theoretical arguments linking alliances to both war and peace. We then discuss how research design choices in the recent literature on military alliances limit the

Michael R. Kenwick (mrk245@psu.edu) is a doctoral student at the Department of Political Science, The Pennsylvania State University, Pond Lab, University Park, PA 16802. John A. Vasquez (vasqueja@illinois.edu) is the Mackie Scholar of International Relations at the Department of Political Science, University of Illinois at Urbana-Champaign, 420 David Kinley Hall, MC-713, Urbana, IL 61801. Matthew A. Powers (mpower5@illinois.edu) is a doctoral student at the Department of Political Science, University of Illinois at Urbana-Champaign, 420 David Kinley Hall, MC-713, Urbana, IL 61801.

Data and supporting materials necessary to reproduce the numerical results in the paper are available in the *JOP* Dataverse (https://dataverse.harvard.edu/dataverse/jop). An online appendix with supplementary material is available at http://dx.doi.org/10.1086/681958.

The Journal of Politics, volume 77, number 4. Published online July 7, 2015. http://dx.doi.org/10.1086/681958

ability to adjudicate between these competing explanations. The third section outlines our own research design, highlighting the ways it addresses the issues discussed in the previous section. The fourth section presents the empirical results. We conclude with a brief summary of findings and discussion of future research.

THEORETICAL EXPECTATIONS

We begin by focusing on whether defensive alliances function as effective deterrents against the onset of interstate conflict. Classic deterrence uses a theoretical framework that often begins with the assumption that states form alliances as a means of signaling the intention to follow through with the terms of their provisions. While the terms of defensive alliances vary, this usually entails two or more states agreeing to come to the aid of another should an external challenger attack one of the signatories. These alliances are often argued to be credible to the extent that alliance partners pay sunk costs in forming the alliance or face potential costs for reneging on their commitments (Fearon 1997; Leeds 2003, 428). When states are successful in signaling their resolve through alliance formation, potential adversaries will expect that an attack on one alliance partner will lead others to honor their commitments and come to the target's defense. An attack on a state with a defensive alliance is therefore expected to be more costly than an attack on a state without such an alliance. Thus, deterrence occurs when the terms of alliances are credible and the potential costs associated with attacking an allied state are sufficiently high (Leeds 2003; Smith 1995).

Initially, this logic did not comport with empirical evidence that suggested alliances actually increased the likelihood of interstate war (e.g., Senese and Vasquez 2008; Siverson and Starr 1991; Vasquez 1993). These studies, however, do not differentiate between alliances that were intended to coordinate offensive action, secure mutual defense incase of an attack, guarantee neutrality, or serve some other purpose. Leeds, Long, and McLaughlin Mitchell (2000) remedy this by disaggregating alliances in terms of the functions they are meant to achieve. The authors argue that if one examines the actual text of alliance treaties, defensive alliances, which obligate a state to intervene in case of an attack, are typically upheld and that allies are generally reliable.

More recent studies have built upon this work and found that states with defensive alliances are less likely to have disputes initiated against them than states without defensive alliances, which is inferred to be evidence of deterrence (Leeds 2003; Wright and Rider 2014). Moreover, evidence is uncovered that disputes initiated against states with defensive alliances are actually less likely to escalate a dispute to war (Johnson and Leeds 2011).[1] From this, the authors conclude that defensive alliance formation is an effective policy for states seeking to induce deterrence and secure peace internationally. While not examining the act of alliance formation specifically, the underlying logic of these studies and deterrence theory more generally suggests that signing an alliance will reduce the chances that a signatory party will be the target of a militarized dispute or become involved in an interstate war. This expectation of deterrence success is stated as:

> **The Deterrence Hypothesis:** Forming defensive alliances, which clearly obligate allies to intervene to defend an ally, will have a successful deterrent effect on the onset of interstate disputes and war.

Conversely, there is much in the literature and some empirical evidence to suggest that alliances, including defensive alliances, are associated with the onset of conflict, and war in particular. From this perspective alliances become an additional variable in a multivariate processes and/or they contribute to a step function that increases threat perception and subsequent hostile interactions. This is accompanied by increases in the political influence of domestic hardliners, which itself enhances the probability of repeated crises and the kinds of crises that are more likely to escalate to war. Senese and Vasquez (2008, chap. 1) refer to these responses as a series of steps to war.

This theoretical framework employs the logic of the security dilemma, which suggests that defensive alliances may not increase security but instead have the opposite effect. According to Herz (1950) and to Jervis (1976), actions taken by a state to increase its national security can prove counterproductive by producing insecurity among states and spurring reciprocal actions by potential opponents; that is, if a state expands its military forces, a potential opponent is likely to respond with similar action, as it fears a decline in its ability to defend itself against the former's newly expanded military. In either event, the interactions between these two states may feed a hostile spiral that fuels competition and strains interstate political relations (Snyder 1984). As a result, defensive alliance formation might result in a counter-alliance, increased threat perceptions, and even deterrence failure because it results in rivalry and commits decision mak-

1. The authors are cognizant that there may be alternate pathways between defensive alliances and war. Leeds (2005) notes that reliable alliances do tend to produce multiparty disputes, which themselves may be more likely to escalate to war.

ers to certain hand-line stances and conditions not supportive of compromise.

Another theme in the literature is that a defensive alliance can create a moral hazard in that the presence of a defensive alliance emboldens a weak ally to take "aggressive" action because it fears of reprisal are lessened or because it is emboldened to take action because it can rely on the support of a powerful ally (Snyder 1984).[2] Alternatively, one could argue that even if the weaker state is not emboldened to take aggressive action it will simply be less restrained and compliant when bargaining with potential challengers. The security dilemma, the steps-to-war, and moral hazard offer an explanation as to why defensive alliances might increase the probability of conflict rather than deterrence. This expectation of deterrence failure is stated as:

> **The Steps-to-War Hypothesis:** Defensive alliance formation will more often be associated with deterrence failure than deterrence success in regard to the onset of interstate disputes and war.

Each of these hypotheses makes contradictory claims. Supporters of a broad conception of deterrence would expect the deterrence hypothesis pass testing and the steps-to-war hypothesis to be falsified. The steps-to-war hypothesis embodies the logic of the security dilemma and more specifically, Senese and Vasquez's (2008) claim that defensive alliances, like all alliances, increase threat perception and are more apt to be followed by war than cases where no alliances are present. It would therefore be expected not only that the deterrence hypothesis would be rejected but that the steps-to-war hypothesis will pass testing. In the event both hypotheses uniformly fail and a null finding is uncovered, this would suggest that alliances are ineffective deterrents, but do not share a positive association with interstate conflict, as the steps-to-war explanation would posit.

TESTING FOR A DETERRENCE EFFECT

Determining whether deterrence actually occurs is fraught with many problems of inference (see Levy 1989). While recent studies have made considerable headway in addressing some of these problems, we outline four issues common to this recent body of work that preclude the ability to adjudicate between the competing expectations outlined in the previous section. The first of these issues revolves around a tendency to conflate the effect of having an alliance with the effect of forming an alliance. Much of the recent work on alliances examines deterrence by assessing whether states with particular alliance types are more or less likely to have disputes initiated against them when compared to states without such alliances (e.g., Benson 2011; Johnson and Leeds 2011; Kang 2012; Leeds 2003). Though intuitively appealing, this approach is problematic when testing the steps-to-war hypothesis because it fails to take into account the specific mechanism linking alliance formation with increases in the probability of conflict. The logic of this hypothesis suggests that it is not necessarily the presence of a defensive alliance that will prove threatening, but rather the formation of these alliances that alter the status quo, may be perceived as a hostile action and, perpetuate a security dilemma between potential combatants. These explanations therefore focus on alliance *formation* and do not necessarily make predictions about the long-term effects of *having* alliances. The deterrence hypothesis, on the other hand, predicts that defensive alliances should always deter, regardless of the time since formation. The relevant period of observation for adjudicating between these competing expectations is therefore the time around formation, and not the time encompassed by the whole lifespan of an alliance.

The second problem endemic to existing studies of extended deterrence is the commonplace assumption that the relationship between alliances and deterrence has remained unchanged over the past two centuries. Indeed, there are strong theoretical grounds to expect that the introduction of nuclear weapons at the conclusion of World War II fundamentally altered the logic of deterrence. Prior to the nuclear era, alliances represented a means by which a state could suddenly and dramatically increase its ability to wage *and win* a conventional war. Even so, the potential attacker in these cases—seeing that its intended target is actively seeking to increase its power—may successfully offset the alliance by seeking out its own alliance partners to assist it in attacking the target (Snyder 1984). In this conventional framework, deterrence occurs when a state increases its capabilities by such an extent as to make the prospects of victory too low for potential opponents.

By comparison, the logic of nuclear deterrence is clearer on why war should be prevented. In contrast to the conventional framework, where deterrence is induced by shifts in conventional military capabilities, nuclear deterrence rests on the logic of mutually assured destruction and the notion that nuclear armaments make wars unwinnable, effectively removing the Clausewitzian rationale for pursuing

2. Benson, Meirowitz, and Ramsay (2014) argue that, under certain conditions, the moral hazard dynamic can actually induce deterrence among potential challengers. This notwithstanding, recent empirical evidence seems to support the traditional formulation of the moral hazard problem formulated in Snyder (1984) and elsewhere (see Kang 2012).

war as a means of attaining political objectives. No goal is worth the complete destruction of a society, and if neither side can attain its goals, then the irrationality of such a war is even more evident (Kahn 1960). Total nuclear war therefore brings about a revolution in the way international politics is conducted and makes such wars unthinkable. In short, failing to control for nuclear weapons applies the more convincing logic of nuclear deterrence to a domain where it may be less applicable.

A third problem centers on the common use of militarized disputes as the exclusive means of testing deterrence arguments (e.g., Benson 2011; Kang 2012; Leeds 2003). The issue here is that the logic of deterrence is most often linked with whether a potential challenger will attack a target holding an alliance (Smith 1995). Indeed, defensive alliances are most often operationally defined as those that promise a state will come to the aid of another "in the event of attack on the partner's sovereignty or territorial integrity" (Leeds et al. 2002, 241). The reason these alliances should also deter low-level military actions such as threats to use force, displays of force, or actions such as border fortifications is not clearly specified.

A fourth, related problem has to do with the meaning of "deter." When war, or minimally a full-scale military attack, is the dependent variable the meaning is clearer—only the absence of these events can be considered potential cases of deterrence success. But with militarized interstate disputes (MIDs), do all militarized challenges have to be prevented for deterrence to occur or just a percentage? In classic deterrence scholars were trying to figure out whether nuclear weapons would *prevent* a state from challenging the status quo with the use of force. The point of deterrence was to prevent a challenge by threatening retaliation, especially nuclear retaliation, which again raises the question of the difference between nuclear and conventional deterrence. Rather than look at complete prevention, dyad-year studies often look at a *reduction* in the likelihood of MID initiation and label this a deterrent effect. This is better defined as a "muting effect" to distinguish it from prevention. Theoretically, the implications of what a muting effect means are unclear. Does it mean that deterrence works, works only sometimes, or has some impact of which the precise nature is unknown? Given the aggregate nature of these analyses, it is not easily ascertainable how many instances there are of actual prevention. This poses a problem because it permits the nonprevention of MIDs to count as evidence consistent with successful deterrence, which again broadens the base of evidence consistent with the theory.

The above suggests that the concept of extended deterrence and its logic have been stretched in terms of the kinds of situations to which it is applied (nuclear vs. nonnuclear), the dependent variable that is employed (MIDs vs. wars), and the evidence used to test the theory (reduction or prevention of conflict). In particular, we are concerned that the body of evidence consistent with finding a deterrent effect has been expanded, which means that the body of evidence that would falsify the theory has been narrowed. These tendencies are hardly unique to the recent literature on alliances. These are only the most recent; similar problems go back to Russett (1963). Nevertheless, these studies follow typical trends in deterrence research of collapsing the difference between alliance formation and the aggregate presence of alliances and the difference between nuclear and conventional deterrence.

RESEARCH DESIGN—AN ALTERNATE TEST

We attempt to address each of the issues by adopting a novel research design that is centered on examining changes in the amount of conflict between two states in the time surrounding defensive alliance formation. More specifically, our primary sample of observations encompasses every defensive alliance initiation year from 1816 to 2000 where an alliance is introduced to a directed dyad where none had previously existed; in other words, each observation records a case where a directed dyad goes from a period when it has no relevant defensive alliance to one where it does. For example, the directed dyad with France as the potential challenger and Germany as the target appears twice in our data set, once in 1915 and again in 1925. This indicates that Germany formed a defensive alliance relevant to an attack from France in each of these years following a period where no alliance was present. We then include the five years before 1924 and 1937 and the four years after in the sample, creating a total of 10 years surrounding each alliance formation.

We limit our sample of formation years to those where a defensive alliance was present at least four of the five years following initiation and absent at least four of the five years preceding it. The most straightforward and common type of observation in our data is one where a potential challenger makes a smooth transition from a five-year period with no defensive alliances to a five-year period where one is present. We also include observations with a single anomalous year in the time before or after alliance formation. These would be years in the pre-alliance formation period where a defensive alliance is present or years in the post-alliance formation period without a defensive alliance.

Taking this approach, we are able to examine how conflict behavior varies in the period before alliance formation to the period after. The choice of a five-year window to ex-

amine conflict occurrence is not uncommon (Senese and Vasquez 2008, 63–64). It is also questionable whether forming a defensive alliance can be considered to have deterred, or failed to deter, an action undertaken more than five years after its formation, as the alliance was already in place when the action was to be contemplated in the first place. In this scheme the formation of an alliance can be seen as a *nonrandom* "treatment" with us examining conflict behavior before and after to see if the "treatment" made a difference in the frequency of conflict.

In testing the hypotheses we are concerned with whether *any* form of defensive alliance is capable of exerting a deterrent effect; we are agnostic as to which specific type of alliance is most likely to do so. We therefore obtain data on alliance formation from two recent studies that have found evidence consistent with deterrence success. The first are the sample of defense pacts that Johnson and Leeds (2011) find reduce the probability of dispute initiation. These are alliances where one or more states agree to come to the aid of another if attacked by an adversary (Leeds et al. 2002). The second sample of cases is drawn from Benson (2011), who finds that conditional deterrent alliances are most effective in reducing the likelihood of dispute onset. Like the previous alliances, these alliances commit a third party to defend an alliance partner should it be attacked by an adversary, but here the third party is exempt from these obligations should the target of the attack attempt to revise the status quo.[3] These alliances are only observed for directed dyads for which their terms are relevant.

After having identified each alliance initiation year, we then code the two main sets of conflict variables that we use in the analysis. The first set, drawing on Kenwick's (2011) analysis, is the number of MIDs that were initiated in the five years before and in the five years after the defensive alliance was formed. We choose only to record MIDs where the challenger in a directed dyad acts as the primary initiator and the target is the primary target within a particular dispute. We obtain these data from the Maoz (2000) dyadic data set.

The second set of variables (also drawn from the Maoz dyadic data) is the number of wars that occur five years before and after the defensive alliance was signed. Here we focus on war occurrence and not war initiation. We do this both because we cannot infer the side that initiated a MID is responsible for escalating that dispute to the level of war and because war is an inherently dyadic phenomenon. Unlike MID initiation, war only occurs at the consent of each party involved. Even in cases where one side instigated and initiated the war, the other side is making the decision to fight instead of acceding to the initiators' demands. Applying directionality to the study of war onset is therefore inappropriate in this case. Note that, for both our measure of dispute initiation and war occurrence, we omit conflicts that took place in anomalous years, by which we mean years in the pre-formation period with a defensive alliance, or post-formation years without a defensive alliance.

Unlike our sample of MID initiations, we also choose to include joiners with respect to war. Omitting war-joiners from the analysis would reduce widespread cases of international conflict to a single observation, potentially causing deterrence failure to appear less frequent than it actually is. Furthermore, wartime alliances would often be, by definition, identified as cases of deterrence success (i.e., defensive alliance observations that aren't followed by conflict), as joiners to the ongoing war would never be observed if one were to only examine MID originators. To illustrate each of these points, consider the directed dyad with France as the challenger and Germany as the target in the time surrounding World War I. Here, Germany was a member of the triple alliance with Italy and Austria-Hungry in the lead-up to 1914 and was subsequently part of a defense pact with Austria-Hungary and Bulgaria beginning in 1915. If one were to exclude joiners from the analysis, this case would offer evidence consistent with deterrence success, since France never *initiated* a conflict against Germany, only joined the ongoing conflict that began between Austria-Hungary and Serbia. This issue with joiners is particularly acute when studying war, since a state may join a MID prior to its escalation to war. In other words, a state can be a joiner to a MID while still being an originator in the war that resulted. We believe these cases should be included, as war-joining behavior is both a common and important way in which deterrence is observed to have failed.[4]

Descriptive analysis

We perform two complementary procedures in order to test the deterrence and steps-to-war hypotheses. In the first, we simply examine changes in conflict frequency in the period before and after each instance of alliance formation within the data. If the deterrence hypothesis is correct, we expect

3. While Benson does not specifically label them as such, we refer to these alliances as being "defensive" in the sense that they seek to induce deterrence among potential adversaries.

4. In the appendix we also conduct a series of robustness tests. These modifications generally do not alter the results, but we note when they do.

cases featuring fewer MID initiations or wars after alliance formation to outnumber those with more. The key prediction is that there should be a statistically significant reduction in MIDs or wars in order to infer successful deterrence. If the steps-to-war hypothesis is correct, we expect the opposite. We also compare the distribution of cases in each of these categories to a set of control observations, or 10-year directed dyadic observations where no defensive alliances were present. If either hypothesis is correct, we would expect that these distributions to differ, as forming an alliance is expected to either increase or reduce the frequency of conflict initiation.

While descriptive in nature, this approach has several advantages. First, within this 10-year timeframe, several factors that may also affect the occurrence of conflict can be expected to remain relatively constant. Some of these are easily controlled for, such as contiguity and GDP, whereas others are not and often go unobserved in dyad-year analysis, such as the shared dyadic history. Second, it makes for a clear and simple procedure for making inferences—change in behavior before and after treatment. Third, and finally, it remains close to the historical record, so each observation can be given a name and examined individually.

As stated above, the relationship between alliance formation and deterrence may change as a function of whether nuclear weapons are present in the international system. To account for this potential source of temporal variation we conduct this analysis on three historical eras: 1816–2000; 1816–1945; and 1946–2000. This will provide some intuition as to whether the prospects for extended deterrence change as a function of historical era.

Matching analysis

While the descriptive analysis will allow us to conduct a thorough examination of the data and identify historical cases of deterrence success and failure, it is not sufficient to conduct a comprehensive test of the presented hypotheses. Doing so requires one first confront the selection problems known to plague studies of extended deterrence (see Levy 1989, 117–20). The issue here is that many of the same factors that may be responsible for causing states to select into defensive alliances may also be positively associated with deterrence failure. That is, states are often thought to form defensive alliances when they are attempting to deter a conflict that they fear may otherwise occur. As a result, states with defensive alliances may be more conflict-prone than those without because of the latent risk of conflict that caused them to join the alliance in the first place. Thus, if one does not control for the factors associated with the nonrandom distribution of defensive alliances among observations fully, then the estimated relationship between defensive alliances and conflict is likely to be biased.

We address this issue by utilizing matching techniques. Matching is commonly used to adjust for nonrandom treatment assignment in observational data (Ho et al. 2007; Iacus, King, and Porro 2012). Of course, matching simply allows us to ensure balance exists among observed features of treatment and control observations, and does not adjust for unobserved factors that may be associated with selection into the treatment. Nevertheless, it provides one means to adjust for the nonrandom distribution of defensive alliances, and has been used to address selection issues in other recent work on alliances (Benson 2012).

In order to implement a matching design, we compare MID and war onset in the post-treatment period (i.e., years 6–10) of our 10-year treatment observations to the latter half of 10-year control observations that never feature the presence of a defensive alliance. Like the treatment observations, which feature alliance formation, these control cases are generated from the directed dyad year Johnson and Leeds (2011) and Benson (2011) replication data.[5] We match on a series of variables we believe are associated with selection into defensive alliances and the latent probability of conflict. These include the presence of other alliances, joint democracy, major power status, strategic rivalry, and history of MIDs.[6] With respect to the alliance indicators, we match on whether the potential challenger or target holds any other alliances outlined in the Johnson and Leeds (2011) or Benson (2011) typologies.[7] For the alliance, democracy, and joint democracy indicators, directed dyads are matched on the values at the time of treatment allocation (i.e., defensive alliance formation, or year 6 among control observations). For strategic rivalry and history of MIDs, an ordinal variable is created counting the number of MIDs or rivalry years in the pretreatment window (years 1–5 for control observations). For example, a treatment observation featuring two major power democracies with no offensive alliances or neutrality pacts with no history of MIDs or strategic rivalry in the five

5. We eliminate instances of overlapping observations. For example, if a directed dyad never featured a defensive alliance from 1960 to 1980, we simply extract two observations, 1960–1970 and 1970–1980, and not all possible 10-year periods.

6. Joint democracy is operationalized as each state having a Polity2 score above 5. These and ATOP variables are from the Johnson and Leeds (2010) replication data. Rivalry data is from Thompson (2001).

7. When using the Johnson and Leeds defense pacts, we match on whether the potential challenger or target has any offensive alliances or neutrality pacts. When using the Benson operationalization, we match on whether either actor has an unconditional or conditional compellent alliance and whether the target has an unconditional deterrent or probabilistic deterrent alliance.

years prior would be matched with a control observation with the same covariate values.[8]

We match on the exact values of the aforementioned variables. We also choose to match multiple control observations to a single treatment observation when this is possible, assigning the weights to adjust for the latter observations.[9] We then estimate a series of logit models, assigning the weights generated in the previous step. We regress two dichotomous dependent variables on our treatment indicator (defined as whether a given observation experienced defensive alliance initiation in year 6). The first is whether any MID initiation occurred in years 6–10 of the longitudinal window, while the second is whether war onset occurred during this time. The treatment indicator is the only covariate in each of these models because we have matched all confounding factors on their exact values. If the deterrence hypothesis is correct, we expect to uncover a negative association with the treatment indicator and both dependent variables; if the steps-to-war hypothesis is correct, we expect to uncover a positive relationship.

Despite its advantages, even this approach does not fully resolve the nonrandom selection issues surrounding alliance formation.[10] Even after matching on observables, the assumption of as-if random treatment assignment is tenuous at best. We therefore do not infer causal effects with respect to alliance formation. This does not mean that such observational studies are unimportant in providing evidence of association or discriminating between contradictory hypotheses as the ones offered here about the whether defensive alliances deter. Indeed, even associations between alliance formation and deterrence success or failure are of interest to scholars of international relations and policy makers alike.

By conducting both the descriptive analysis of the data and implementing the matching design across multiple time frames, we hope to construct as rigorous a test as possible to examine these competing hypotheses. The examination of whether there is an increase or reduction in conflict in the period around alliance formation provides a sense of alliance formation that relates to conflict in terms of within-case (here directed dyad) variation. Nevertheless, taking this approach precludes the ability to draw inference from the set of observations that never experience conflict. This might be viewed as problematic due to the rarity of conflict in the international system. The matching analysis addresses this, as it incorporating observations without defensive alliance formation and exploiting cross-sectional variation while improving balance between treatment and control observations.

RESULTS

Dispute initiation

Our first results examine the frequency of MID initiation using the ATOP and Benson classifications of defensive alliances, respectively. Each observation in the datasets is sorted into categories of "Fewer MIDs after Formation," "No MID Occurrence Before or After," "Same Quantity of MIDs Before and After," and "More MIDs after Formation." We also examine the frequency of observations in each category for both "treatment" and "control observations," with the latter included for reference. Recall that treatment observations feature defensive alliance formation in year 6 of a particular decade, whereas control observations never feature any defensive alliances. Thus, for treatment observations the more and fewer categories capture change in the frequency of MID initiation following alliance formation, whereas for control observations this simply pertains to changes after year 5 of the decade-long spell.

We conduct two tests using these data. First, we assess whether the distribution of treatment observations across these categories differs from that of the control observations using a simple Pearson's χ^2 test and report the p-value of the test statistic with each table. Here, statistical significance indicates that the distribution of treatment and control cases are different from one another as predicted by both the deterrence and steps-to-war hypotheses. The former expects these distributions to differ because of an influx of cases with fewer MIDs after formation among the treatment observations, while the latter does so because of an expected influx of cases with more MIDs after formation. If, however, the distribution of treatment and control observations were not statistically different from one another, this would be consistent with the null hypothesis, suggesting alliance formation has no effect on MID initiation.

We then determine whether there is a significant difference in the distribution of cases with more and fewer MIDs after alliance formation. This is done using a binomial probability test. If the null hypothesis is rejected and there is a preponderance of "fewer MIDs after" cases, this would offer evidence in support of the deterrence hypothesis.[11] If there

8. We do not match on variables in the post-treatment window because variables that could be a result of treatment allocation should not be controlled for (see Ho et al. 2007, 202).

9. The *cem* package in Stata is used to match and assign weights (Blackwell et al. 2009).

10. See Gerber and Green (2012) on the importance of random assignment.

11. The test determines the probability of observing a distribution between the "more" and "fewer" categories if the probability of observing

are significantly more MIDs after, this would offer evidence consistent with steps-to-war hypothesis, indicating that defensive alliance formation may actually encourage conflict or are at least associated with it. Finally, if the null is accepted this indicates that the alliance formation had no effect.

Table 1.1 (reported in the appendix, available online) details the results obtained from the ATOP defensive alliances for 1816–2000. We find that distribution of treatment observations is significantly different from that of the control observations for this sample. The reason for this is most likely the preponderance of treatment observations featuring more MIDs after alliance formation relative to the time before. Among these observations, 78 cases (1.12%) feature more instances of MID initiation after formation, while 54 cases (0.79%) feature less. Conducting a binomial test confirms that there are more observations experiencing an increase in MID initiation than one would expect if alliance formation had no effect. This evidence suggests that alliance formation is typically associated with an increase in dispute initiation, as predicted by the steps-to-war hypothesis.

We next split the temporal range of the data following the conclusion of World War II and the first use of nuclear weapons. Henceforth, we refer to these two periods as the prenuclear and nuclear era, which pertains to 1816–1945 and 1946–2000, respectively. As expected, the relationship between defensive alliances and deterrence success or failure differs across these two time periods. In many ways the results from the prenuclear era parallel those of the full temporal span, but with an even stronger relationship between alliance formation and increases in conflict frequency. Here, the distributions of treatment and control cases again differ from one another, with the former having an even larger preponderance of cases in the "more MID initiations after" category. There are now 47 cases of treatment observations experiencing an increase in conflict, relative to only 27 observations that experience a decrease; the difference between these categories is again statistically significant (see tables 1.2 and 1.3 in the appendix). The nuclear era, on the other hand, displays evidence consistent with the null hypothesis. During this time, the difference in the distribution of treatment and control cases is no longer statistically significant, and only a slight difference in the number of treatment observations experiencing more or fewer MID initiations (24 vs. 19) after alliance formation. These results offer support for the notion that alliance formation is often followed by deterrence failure in the prenuclear era, but neither deters nor instigates conflict thereafter.

We next conduct the same tests when using Benson's classification of conditional deterrent alliances as the treatment.[12] These results generally comport with those obtained using the ATOP classification (see tables 2.1–2.3 in the appendix). Once again, in the full time sample and prenuclear era there is a significant difference between the distribution of treatment and control cases across each category and that alliance formation leads to a significant increase in MID initiation among treatment observations. In the nuclear era, we again find no significant relationship between alliance formation and MID initiation. The fact that these results obtain using two separate classifications of defensive alliances offers additional support for the steps-to-war hypothesis prior to the conclusion of World War II and the null hypothesis thereafter.

The above descriptive analyses report the specific number of cases consistent with the two hypotheses we are testing. A more rigorous test is to employ matching and these are reported in figure 1. Panel *A* reports the initial results for the matching tests and shows the effect that alliances have on the percent change in the odds of MID initiation in the five years following defensive alliance formation.[13] Recall that directed dyads are matched based on other types of alliances held by the potential challenger and target, democracy, major power status, and the number of disputes and rivalry years in the first five years of the decade long observation. Like the descriptive analysis, when using the full sample (1816–2000), ATOP defensive alliance formation is positively associated with MID initiation. Conditional deterrent alliances (as measured by Benson), however, do not have a statistically significant effect, as evidenced by the confidence interval overlapping zero. Also consistent with the descriptive analysis, both types of defensive alliance formation produce significant increases in the odds of MID initiation in the prenuclear era. In the nuclear era, on the other hand, the relationship between ATOP defensive alliance formation and MID initiation is negative, but only attains statistical significance when employing a $p < .10$ cutoff. Conditional deterrent alliances, on the other hand, have no effect on the likelihood of MID initiation during this time. Taken together, these results suggest that the effect of alliance forma-

each is equal (i.e., $pr(\text{Fewer MIDs after}) = pr(\text{More MIDs after}) = 0.5$), as one would expect if alliance formation was unrelated to conflict.

12. Note that there are fewer observations using these data than were obtained using ATOP. This is because these data are obtained from the Benson (2011) replication data, which samples only politically relevant dyads.

13. Tables containing the number of treatment and control observations, coefficient estimates, and standard errors pertaining to figure 1 are reported in the appendix.

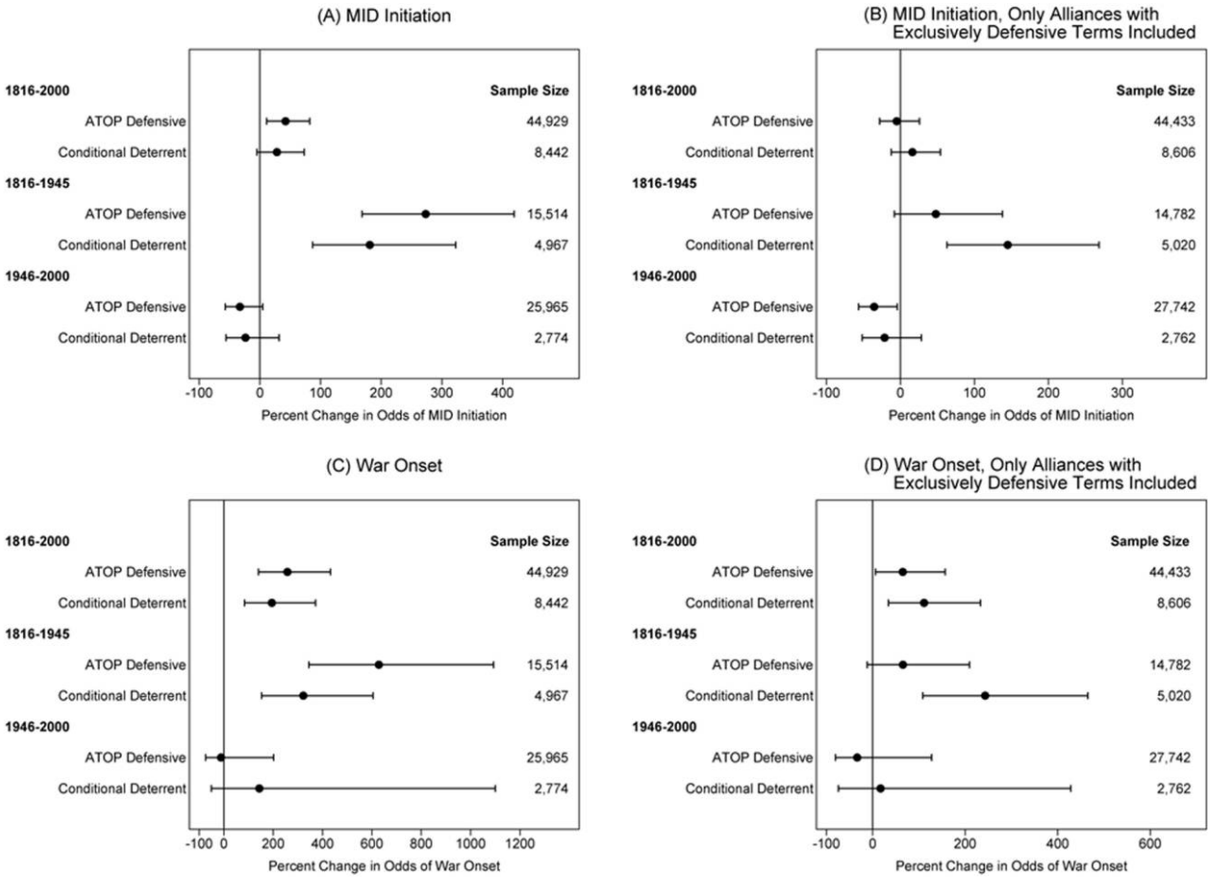

Figure 1. Effect of defensive alliance formation on conflict onset. Results obtained from matching analysis. The figure reports percent change in odds of conflict onset with 95% confidence intervals.

tion does change after the introduction of nuclear weapons, and that the bulk of evidence supports the steps-to-war hypothesis in the prenuclear era, and that weak evidence is found for the deterrence hypothesis in the nuclear era.

As previously stated, one of the factors we matched on was whether a particular alliance had any provisions that may serve to increase the probability of dispute initiation. To control for this issue further, we also generated an alternate version of our data that included only observations for "pure" defensive alliances, or those that do not contain provisions thought to heighten the risk of conflict.[14] The results we obtain after making this adjustment are reported in panel *B* of figure 1. We now find that neither type of alliance has a statistically significant effect in the 1816–2000 sample. In the prenuclear era, the effect of ATOP defensive alliance formation no longer has a significant effect, but that conditional deterrent alliances still increase the odds of MID initiation. In the nuclear era, the negative effect of ATOP defensive alliances formation is now statistically significant, though conditional deterrent alliances still have no effect. Restricting our treatment observations to "pure" alliance formations therefore weakened evidence in support of the steps-to-war hypothesis in the prenuclear era and strengthened results consistent with the deterrence hypothesis in the nuclear era.

In general, these indicate that the effect of defensive alliance formation is, in fact, modified by both historical era and the specific terms within each alliance. In the prenuclear era, the bulk of evidence uncovered was in favor of the steps-to-war hypothesis, with both ATOP defensive and Conditional deterrent alliances increasing the possibility of MID initiation, though the former relationship becomes insignificant after subsetting cases to alliances with exclusively defen-

14. When employing the ATOP typology, we excluded defensive alliances that contained offensive terms or terms promising neutrality. In the Benson typology, we excluded alliances with unconditional compellent or deterrent terms.

sive terms. In the nuclear era, on the other hand, we find mixed support in favor of both the null and deterrence hypotheses. Specifically, while ATOP defensive alliances with exclusively defensive terms exert a deterrent effect on MID initiation, other types of defensive alliance formation have no effect.[15]

War occurrence

Our analysis of deterrence with respect to war occurrence parallels that of dispute initiation. We begin with the descriptive results using the ATOP typology to examine if the number of wars following the formation of a defensive alliance reduces (or increases) the number of wars compared to the prior period without a defensive alliance. When examining the whole time period, our results are generally consistent with the steps-to-war hypothesis in that observations with more war after alliance formation significantly outnumber those with fewer wars. In fact, at 48 to 23, there is nearly double the number of cases experiencing more wars after formation than fewer (see table 3.1 in the appendix). This distribution of treatment cases across each category is statistically different from that of the control cases.

The results for the prenuclear era are similar (see table 2.2 in the appendix). At 31 to 14 cases, there are significantly more wars after alliance formation than before and the distribution of treatment and control observations across each category is significantly different. Interestingly, we also find that treatment observations are more likely than control observations to be both preceded and followed by war. This suggests that alliance formation is tightly linked with war occurrence in the prenuclear era.

In the nuclear era, defensive alliance formation has a somewhat weak proclivity to lead to increases in the occurrence of war relative to reductions. As before, the distribution across treatment and control cases is statistically different from one another (see table 3.3 in the appendix). Unlike the prenuclear era, however, there are a similar proportion of treatment and control cases in the "fewer wars after" category and the difference in distributions appears to be an influx of cases in the "more wars after" category. Though we again find more than double the number of treatment observations featuring more wars after formation than fewer (17 vs.7), this difference is only statistically significant if one were to use a one-tailed significance test. We interpret this as weak evidence in favor of the steps-to-war hypothesis with respect to war following 1945.

These tests are repeated using conditional deterrent alliances (see tables 4.1–4.3 in the appendix). In the full sample, treatment observations featuring more war after formation outnumber those with less, but the difference between these two categories is statistically insignificant. Even so, while conditional deterrent alliance formation does not appear to lead to significant within-case increase in war occurrence, the treatment observations again tend to feature a larger proportion of cases in the "more wars after" and, to a lesser extent, "fewer wars after" category than do control observations. As a result of this, the distribution of cases among treatment observations is significantly different than that for control observations. This pattern is even more evident when restricting the sample to the 1816–1945. Among these observations, we again find there is no significant difference between the number of cases in the more and fewer categories among treatment observations. Nevertheless, there are roughly four times as many cases in each category as there are among control observations and the overall distribution of treatment and control cases is statistically different from one another.

The results for the nuclear era offer additional support for the null hypothesis, though the small number of treatment observations experiencing war likely drives this result. Only two treatment observations experience war onset and each time this occurs after the formation of a conditional deterrent alliance. Not surprisingly, this difference between the more and fewer categories is statistically insignificant, as is the difference between the distribution of observations across categories for the treatment and control groups.

Overall, the findings with respect to conditional deterrent alliances suggest that alliance formation is a poor deterrent against war onset. Nevertheless, while alliance formation does not result in within-case increases in the frequency of war onset as the steps-to-war hypothesis would expect, treatment observations tend to experience war at a higher rate than control observations. Put differently, this result provides some partial support for the steps-to-war hypothesis, but this evidence comes from variation between treatment and control cases and not variation within treatment observations.

The results from the matching analysis when employing war onset in the five years after formation as the dependent variable are reported in panels *C* and *D* of figure 1. Panel *C* reports the results obtained when examining any form of defensive alliance formation, while panel *D* reports the results for cases where alliances with exclusively defensive terms are introduced into a directed dyad. With respect

15. We also conducted this analysis using only "use of force" MIDs. We find that in the full time period conditional deterrent alliances now increase the odds of MID initiation, though ATOP defensive alliances still have no effect. The prenuclear results remain the same. The results for 1946–2000 become statistically insignificant. This implies that defensive alliances are less likely to deter the use of force.

to the former set of cases, formation of ATOP defensive and conditional deterrent alliances each share a positive association with war onset in both the full and prenuclear samples. The same results obtain when using the more conservative criteria for defensive alliance formation, save for ATOP defensive alliances in the prenuclear era, which now has an insignificant relationship with war onset in the prenuclear era. In general, these results suggest that alliance formation is a poor deterrent against war onset and that the formation of some defensive alliances increase the probability of war, as the steps-to-war hypothesis suggests.

During the nuclear era, on the other hand, we find that defensive alliances have no effect on war onset, as indicated by the wide confidence intervals that overlap zero. The fact that the confidence intervals are so large for this period, particularly with respect to conditional deterrent alliances, is a reflection of the dearth of observations featuring war after alliance formation. Recall that only two observations of alliance formation feature war after a conditional deterrent alliance has been formed. Regardless, this evidence suggests both the deterrence and steps-to-war hypotheses are to be rejected when restricting the sample to the nuclear era.

Taken together, neither the descriptive nor the matching analysis produced any evidence suggesting that forming a defensive alliance will deter adversaries from initiating or joining wars against potential targets. Instead we find that in the prenuclear era forming some types of defensive alliances—particularly those with nondefensive terms—actually heightens the risk two states will go to war with one another. In more recent history alliance formation appears to have little bearing on whether war will occur. Nevertheless, the fact that we do not observe a negative relationship suggests that forming an alliance may not be an effective policy for leaders seeking to induce peace with challengers aboard.[16]

CONCLUSION

Several conclusions follow from this study. First, previous findings suggesting that defensive alliance formation is an effective means of inducing deterrence must be qualified. Whatever their long-term effects, none of our tests in the prenuclear era suggests that alliance formation results in a systematic reaction in the likelihood an ally will experience conflict in the near future. Instead, alliance formation during this time very often leads to an increase in the likelihood a state will have a militarized dispute initiated against it, lending empirical support to the argument that alliance formation heightens hostility among potential adversaries. Conversely, in the nuclear era our analysis finds weak but inconsistent evidence that forming exclusively defensive alliances reduces the likelihood of conflict initiation, suggesting evidence of extended deterrence.

Second, alliance formation appears to be a poor deterrent against the onset of war. We found no evidence that forming a defensive alliance reduces the probability a challenger will initiate or join a war against a potential target during any time period. As with the onset of militarized disputes, we instead find that alliance formation is often followed by a heightened risk of war in the prenuclear era. Unlike the dispute analysis, we did not find any evidence of deterrence success, even after the advent of nuclear weapons. This lends support to Senese and Vasquez's (2008) claim that outside alliances tend to increase the probability of war; that is, defensive alliances are not exempt from this general pattern in the 1816–1945 period. This also indicates that war is a primary means by which deterrence is often observed to fail.

Furthermore, our analysis implies that the relationship between alliance formation and international conflict may hinge critically on the international context in which the alliance is formed. One such factor appears to be historical era, as evidenced by the fact that the relationship between alliance formation and conflict appears to have reversed following the conclusion of World War II and introduction of nuclear weaponry to the international system.[17] This suggests that standard practices of collapsing samples from the nineteenth through twenty first century is indeed likely to mask meaningful variation concerning the effect of alliance formation.

Lastly, this study utilizes a longitudinal design to focus analysis around the time surrounding alliance formation. This allows for a more precise test of arguments centered on hostile spirals, which hinge primarily on forming, not holding, international alliances. In other words, this approach also keeps us from conflating the instantaneous effect of defensive alliance formation with the potentially different effect these alliances have after they have become entrenched.

REFERENCES

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

Benson, Brett V. 2012. *Constructing International Security: Alliances, Deterrence, and Moral Hazard*. Princeton, NJ: Princeton University Press.

16. In an axillary analysis, we also modified the war indicator to omit joiners. All relationships were rendered statistically insignificant, largely due to a dearth of observations experiencing war in both the treatment and control observations.

17. For an analysis that argues that the very nature of the Cold War alliance system may account for the reversal see Vasquez and Kang (2013).

Benson, Brett V., Adam Meirowitz, and Kristopher W. Ramsay. 2014. "Inducing Deterrence through Moral Hazard in Alliance Contracts." *Journal of Conflict Resolution* 58 (2): 307–35.

Blackwell, Matthew, Stefano Iacus, Gary King, and Giuseppe Porro. 2009. "Cem: Coarsened Exact Matching in Stata." *Stata Journal* 9 (4): 524–46.

Clare, Joe. 2013. "The Deterrent Value of Democratic Allies." *International Studies Quarterly* 57 (3): 545–55.

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

Fuhrmann, Matthew, and Todd S. Sechser. 2014. "Signaling Alliance Commitments: Hand-Tying and Sunk Costs in Extended Nuclear Deterrence." *American Journal of Political Science* 58 (4): 919–35.

Geller, Daniel. 2012. "Nuclear Weapons and War." In J. Vasquez, ed., *What Do We Know about War*? 2nd ed. Lanham, MD: Rowman & Littlefield, 139–63.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: Norton.

Herz, John. 1950. "Idealist Internationalism and the Security Dilemma." *World Politics* 2 (January): 157–80.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15:199–236.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis* 20:1–24.

Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.

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

Kahn, Herman. 1960. *On Thermonuclear War*. Princeton, NJ: Princeton University Press.

Kang, Choong-Nam. 2012. "Alliances: Path to Peace or Path to War?" In J. Vasquez, ed., *What Do We Know about War?* 2nd ed. Lanham, MD: Rowman & Littlefield, 27–44.

Kenwick, Michael. 2011. "Analyzing Deterrence: Defensive Alliance Formation and Dispute Onset." BA honors thesis, University of Illinois at Urbana-Champaign.

Leeds, Brett Ashley. 2003. "Do Alliances Deter Aggression? The Influence of Military Alliances on the Initiation of Militarized Interstate Disputes." *American Journal of Political Science* 47 (July): 427–39.

Leeds, Brett Ashley. 2005. "Alliances and the Expansion and Escalation of Militarized Interstate Disputes." In Alex Mintz and Bruce Russett, eds., *New Directions in International Relations*. Lanham, MD: Lexington, 117–34.

Leeds, Brett Ashley, Andrew G. Long, and Sara McLaughlin Mitchell. 2000. "Reevaluating Alliance Reliability: Specific Threats, Specific Promises." *Journal of Conflict Resolution* 44 (October): 686–99.

Leeds, Brett Ashley, Jeffrey M. Ritter, Sara McLaughlin Mitchell, and Andrew G. Long. 2002. "Alliance Treaty Obligations and Provisions, 1815–1944." *International Interactions* 28:237–60.

Levy, Jack S. 1989. "Quantitative Studies of Deterrence Success and Failure." In Paul C. Stern, Robert Axelrod, Robert Jervis, and Roy Radner, eds., *Perspectives on Deterrence*. New York: Oxford University Press, 98–133.

Maoz, Zeev. 2000. *Dyadic Militarized Interstate Disputes Dataset*. Version 2.0.

Russett, Bruce. 1963. "The Calculus of Deterrence." *Journal of Conflict Resolution* 73 (June): 97–109.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. New York: Oxford University Press.

Senese, Paul D., and John A. Vasquez. 2008. *The Steps to War: An Empirical Study*. Princeton, NJ: Princeton University Press.

Siverson, Randolph M., and Harvey Starr. 1991. *The Diffusion of War: A Study of Opportunity and Willingness*. Ann Arbor: University of Michigan Press.

Smith, Alastair. 1995. "Alliance Formation and War." *International Studies Quarterly* 39 (December): 405–25.

Thompson, William R. 2001. "Identifying Rivals and Rivalries in World Politics." *International Studies Quarterly* 45 (4): 557–86.

Snyder, Glenn H. 1984. "The Security Dilemma in Alliance Politics." *World Politics* 36 (4): 461–95.

Vasquez, John A. 1993. *The War Puzzle*. Cambridge: Cambridge University Press.

Vasquez, John A., and Choong-Nam Kang. 2013. "How and Why the Cold War Became a Long Peace: Some Statistical Insights." *Cooperation and Conflict* 48 (1): 28–50.

Wright, Thorin M., and Toby J. Rider. 2014. "Disputed Territory, Defensive Alliances, and Conflict Initiation." *Conflict Management and Peace Science* 31 (2): 119–44.