Defense Pacts and Deterrence: Caveat Emptor

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

In a previous study, we provided a novel empirical test that indicated that alliance formation is often a poor means of deterring militarized disputes or preventing war. In this issue, Leeds and Johnson respond, arguing that our results differ from their previous work because we analyze the time surrounding alliance formation, which they claim is inappropriate when testing deterrence theory. Here we outline the reasons why examining alliance formation is critical to the analysis of deterrence. We also demonstrate that the empirical concerns raised by Leeds and Johnson are largely inconsequential by reproducing our results using the same design implemented in existing studies that uncover evidence of deterrence success. We find additional support for our primary conclusion that alliance formation often fails to deter adversaries and instead increases likelihood of militarized conflict.

The main contribution of Kenwick, Vasquez, and Powers (2015) is to offer a new test and some evidence about the ability of defensive alliances to deter militarized interstate disputes (MIDs) and war. In addition, a valuable aspect of our design is to examine whether alliance formation increases or decreases the existing pattern of MIDs and wars. We are grateful to Leeds and Johnson (2016, in this issue) for their careful reading and thoughtful critique of our work. While we deal with Leeds's (2003) and Johnson and Leeds's (2011) work because of its importance, the tests and conclusion of our paper are broader than that and not just a critique of their research. Nor do we disagree with all aspects of their work; the Alliance Treaty Obligations and Provisions (ATOP) data and Leeds and Johnson's studies have made invaluable contributions to our understanding of not only deterrence but alliance behavior generally.

Nevertheless, we disagree with their claim that we cannot draw conclusions about deterrence theory by examining alliance formation. Leeds and Johnson (2016) object to our decision to examine only the five years after formation. We show here that this decision was not consequential by reproducing our results using Johnson and Leeds's (2011)

directed dyad year data. We find that whenever a potential target forms a new defensive alliance, it increases the likelihood of dispute initiation, regardless of how much time has passed since the alliance was signed. In fact, we find that the majority of defensive alliance observations in the Johnson and Leeds data are associated with an increased risk of conflict. This offers additional support for our original conclusion that alliance formation is often a poor means of inducing deterrence or preventing conflict.

WHY FORMING ALLIANCES MIGHT INCITE CONFLICT

Classic deterrence theory posits that states pay the costs of forming an alliance to signal their intent to intervene in a military conflict initiated by an outside adversary. If this signal is credible, potential opponents should be less likely to initiate armed conflict due to the heightened costs of fighting. While this logic is intuitively appealing, history is littered with instances of deterrence failure. To understand why deterrence might fail, we briefly outline the theoretical links to alliance formation and the onset of military conflict.

We begin by examining the effect of alliance formation on the distribution of bargaining power between the poten-

The Journal of Politics, volume 79, number 1. Published online November 10, 2016. http://dx.doi.org/10.1086/686700

tial challenger and the target. For the target, the presence of a recently formed alliance improves its bargaining position by reducing the leverage a potential opponent may have obtained from its offensive military capacity. While Leeds and Johnson claim that this will cause the challenger to abandon its demands, it is more likely that it will simply recalibrate demands in accordance with the new alliance. The alliance therefore only deters by the extent to which it reduces information asymmetries between the two parties. The alliance may also allow the target to adopt an intransigent bargaining position or to increase its demands (e.g., Fearon 1997, 84). While the third party might prefer that the target remain moderate, there is often difficulty observing the latter's bargaining tactics and determining which side is responsible for the onset of hostilities. These issues encompass a moral hazard problem associated with alliance formation. While states might try to anticipate or even harness moral hazard when constructing an alliance (Benson, Meirowitz, and Ramsay 2014), there is always the risk that forming the alliance may increase uncertainty and inflate the risk of deterrence failure. It is therefore not enough that alliances increase the costs of conflict—they must also reduce uncertainty between opponents.

Whether an alliance serves this function depends on how potential challengers respond to the act of alliance formation. Canonical models of extended deterrence often allow only a snapshot of the interaction between the target and the challenger, allowing the latter to either attack or accept the status quo once an alliance is formed. The Steps-to-War explanation examines responses by the challenger in explaining how defensive alliances become linked to onset of war (Vasquez 1993). Committed states will seek to counteract defensive alliances through counteralliances, arms buildups, or similar measures. As these interactions take place, each side continuously re-evaluates the reliability of the alliance in the changing security environment, which may increase uncertainty, itself a cause of war according to Fearon (1995). In short, the formation of alliance alters the balance of bargaining power and has the potential to incite escalation and increase uncertainty.

SELECTION BIAS AND DETERRENCE

Leeds and Johnson argue that selection issues cause us to underestimate the deterrent effect of alliances. Even if one assumes that alliances do reduce the probability of conflict, any positive association between their formation and dispute onset suggests that this effect is not sufficient to offset the risk factors that drove states to sign the alliance in the first place. In Kenwick et al. (2015), we attempt to address this issue by matching observations on a set of covariates

associated with alliance formation and the latent risk of conflict. While not a complete solution, matching is a step in the right direction.

Further, we reject the claim that nothing can be learned about deterrence by examining the time surrounding alliance formation for several reasons. The first centers on when alliances are most credible and therefore most likely to deter. While states may be more likely to seek alliance partners when they face a security threat, we have little reason to believe that alliances will fail during this time. On the contrary, many scholars argue that alliances will be most credible when they are first signed but that they will become increasingly unreliable as the preferences of the signatories and international circumstances change (Fearon 1997, 82; Smith 1995, 418). Defense pacts should therefore have their greatest deterrent capacity shortly after formation.

Second, because forming a defense pact is often costly (Fearon 1997; Morrow 1994; Smith 1995), we should not expect that states would sign agreements they expect to fail. If the latent risk of conflict is high at the time of formation, states have already accounted for this and have decided that the prospects for successful deterrence are sufficiently high.

Third, disregarding analyses of alliance formation undercuts our ability to evaluate whether signing an alliance is an effective policy in difficult security environments. While we may not be able to make causal claims about alliance formation and the occurrence of conflict, any association between the two provides valuable insights from a policy perspective.

Fourth, Leeds and Johnson argue that, due largely to selection issues, ours is an analysis of immediate, not general, deterrence. This seems unlikely since we did not select observations based on whether there is an existing challenge or crisis. More importantly, their theory does not articulate why alliances should have disparate effects in these environments, so an analysis in either domain is a valid test. If we do not observe a negative association, we must conclude that alliances either have no deterrent effect or that this effect is too small to counteract unobserved factors driving states to conflict.

THE IMPORTANCE OF DOMAINS

In our original analysis, we found that the effect of alliance formation varies by historical period—1816–1945 vs. 1946–2000—and by dependent variable—MIDs or wars. In their reply, Leeds and Johnson do not recognize these domains. They say they do not agree with the need to pull out the post-1945 period, and they continue to use solely the full period 1816–2000 in their reply. We show, however, that the three time periods have different results. They say, "We are

not convinced that nuclear weapons have fundamentally changed the effects of other variables on dispute propensities, and therefore we do not find the argument that the samples should be analyzed separately convincing" (Leeds and Johnson 2016, **XXX**, in this issue). Yet there is a marked difference in the results.

We should point out that the logic of nuclear deterrence is more convincing than the logic of conventional deterrence. In the former, war is deterred because nobody can win under conditions of mutual assured destruction (MAD)—in a sense, both are deterred. In the latter, war is deterred only if one side agrees that the other has so much power that it cannot win and therefore does not initiate. The logics are different in the two eras, and our results are consistent with the claim that it is more difficult for deterrence to work prior to 1946.

Neither do Leeds and Johnson (2016) address our findings on war. They concentrate on MIDs, which is legitimate given their previous work. Our article, however, is not mainly a critique of their work but an analysis of deterrence. From a policy perspective, what is the point of deterring MIDs if the same logic cannot prevent war? As Leeds and Johnson (2016) and others recognize, deterrence failure is not uncommon. Leeds (2005) recognizes that if a MID occurs despite a defensive alliance, then allies are reliable and join that dispute. As she notes, multiparty MIDs have a higher probability of escalating to war than bilateral MIDs.

Our findings show that preventing war is more difficult than deterring MIDs and that more hostile MIDs are more difficult to deter than lower-level ones. This is an important contribution of our analysis that should not be forgotten. We turn now to additional data analysis to address some of Leeds and Johnson's (2016) empirical critiques.

DATA RESPONSE

Our original design sampled observations of alliance formation and matched these to control cases to determine how alliance formation affects MID initiation in the following five years. We chose to restrict our sample in this way because it delimits the empirical domain to an area where deterrence theory and the Steps-to-War explanation make opposite predictions. The Steps-to-War argument predicts that forming an alliance has the potential to incite conflict, while deterrence theory always predicts a reduction in conflict initiation, making no distinction between recently formed and well-established alliances.

Leeds and Johnson objected to some elements of this design, correctly noting that our sample does not contain the full population of directed dyad years or defensive alliance observations, that we only examine the first five years after

alliance formation, and that we exclude cases where targets form alliances in addition to those they already hold. We believe that our decisions were justified in order to test each theory on common ground. Nevertheless, we address each of these issues by reproducing our findings using the design employed by Johnson and Leeds (2011), and we find additional support for our original conclusions.

To implement this new design, we begin with the Johnson and Leeds (2011) data, which contain the population of directed dyad years 1816–2000. We categorize all defensive alliance observations in the data into two groups: those that have been in place since a directed dyad first entered the data and alliances that were formed afterward. Note that states may hold both types of alliances simultaneously. It is also possible for an alliance to be coded as existing from birth in one dyad while being formed after birth in another. Consider, for example, the Organization of American States (OAS). When the United States signed this agreement in 1948, it was coded as having a defensive alliance formed after birth for all directed dyads where the United States is the target state. When new states entered the system after this date, such as Jamaica in 1962, this alliance is coded as existing "from birth."

We group alliances in this way as an alternate strategy for adjudicating between competing expectations. The Steps-to-War hypothesis predicts that alliances incite escalation when they alter the distribution of power. It would therefore predict that alliances formed after birth would increase MID initiation. Alliances that have been in place since the inception of a dyad, however, do not alter bargaining leverage and should therefore have little potential for escalation. Deterrence theory, on the other hand, does not make this distinction and would predict that each set of alliances will deter.

Of the directed dyad years that have a defensive alliance, 40% contain defense pacts that have been held since birth, while 67% contain defense pacts formed after birth.[1] Figure 1 displays alliances based on how frequently they appear in the directed dyad year data.[2] The horizontal axis displays the proportion of total "from birth" defensive alliance observations in which a specific alliance is present. The vertical axis reports the same for alliances formed after birth.

Large organizational alliances comprise a relatively large proportion of each category. The OAS, for example, is present in about a quarter of all observations where the target holds a defensive alliance that has been present since the

---

1. These two do not sum to 100 since some observations have both types of defensive alliances.

2. We are grateful to an anonymous reviewer who recommended we list influential cases in each category.
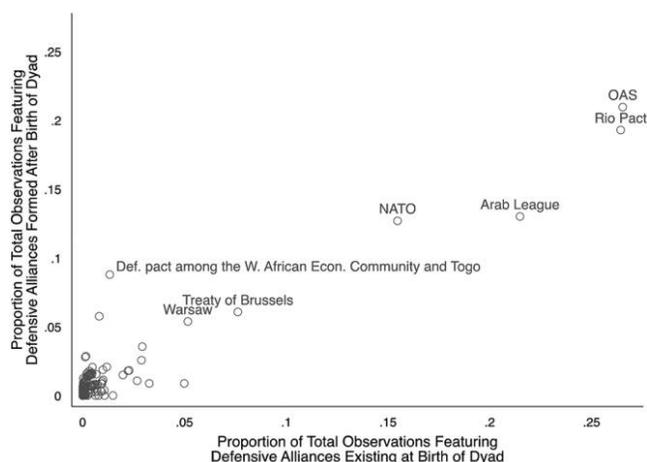
Figure 1. Alliances by frequency of appearance in directed dyad year data. The horizontal axis reports the proportion of times a given alliance is present in directed dyad years featuring an alliance that existed at the birth of a dyad. The vertical axis reports the same for alliances formed after birth.

birth of a dyad and in a fifth of all observations where an alliance was formed after birth. When examining such cases, it is important to note that alliances serve purposes other than deterrence, acting instead as a conflict resolution mechanism or a means by which states coordinate and develop shared institutions (Bennett and Stam 2007, 74; Gibler 1996). In this way, alliances like the OAS and the Arab League are forms of regional governance similar to the Concert of Europe (see Barnett 2003, 232–35) as well as defense pacts.

We regress MID initiation using an identical model specification to Johnson and Leeds (2011), but we include two defensive alliance indicators using the categories outlined above. We again estimate separate models for the 1816–2000, 1816–1945, and 1946–2000 periods to control for the introduction of nuclear weapons and other systemic shifts occurring in 1945. In the appendix (available online), we provide more details pertaining to these results and an additional robustness test that disaggregates alliances formed after independence by age.

The quantities of interest from these models are reported in figure 2. While the effect sizes differ, we find consistent results across all three domains. Consistent with deterrence theory and with Leeds and Johnson (2016), we find that defensive alliances that are in place from the first year a target or challenger enters the system significantly reduce MID initiation. These alliances are relatively rare in the pre-nuclear era, accounting for only 20 of the total observations with defensive alliances. In the nuclear era, they are more common, accounting for 41% of observations. As previously stated, regional alliances comprise a large proportion of these cases, and it is unclear whether their pacific effect is attributable to

deterrence or this regional governing function. The results pertaining to these cases are consistent with the predictions of deterrence theory.

Consistent with the Steps-to-War hypothesis, alliances introduced in a directed dyad after birth significantly increase the probability of MID initiation.[3] Note that these results are stronger than our original findings pertaining to alliance formation and that they operate across all three samples, though the magnitude of the effect sizes and overall probability of conflict varies considerably across the periods. One explanation for this strengthened positive effect is that the directed dyad year design is poorly equipped to adjust for the latent risk of conflict when compared to our original design. Even if alliances are not causing conflict per se, these results suggest that, at minimum, their deterrent effect is insufficient to offset unobserved factors driving states to conflict.

This modified design addresses the major empirical criticisms of Leeds and Johnson (2016). These data contain the population of directed dyad years, expand beyond the first five years after formation, and record all instances of alliance formation after independence. We find that alliances sometimes reduce the likelihood of MID initiation but that, in the majority of cases (i.e., alliances formed after independence), they have the opposite effect. While making these changes in our research design produced different results, they only strengthen our primary finding—forming a new defensive alliance often proves to be an ineffective policy for leaders seeking to induce deterrence.

In summary, our findings show that within certain domains forming defensive alliances is not followed by a reduction of MIDs or wars. Leeds and Johnson argue that our emphasis on alliance formation and our research design eliminates too many cases. In this response, we employ a dyad-year design more attune to their analyses. We get similar results when we split the population by time period, which they do not do in their analysis. We also uncover a new finding—that only alliances present from birth deter MIDs.

CAVEAT EMPTOR: NUCLEAR DETERRENCE AND NUCLEAR PROLIFERATION

How much evidence of deterrence failure can we tolerate to still maintain the policy? Our evidence suggests that it is more difficult to deter wars than MIDs and that alliance formation

---

3. Although their confidence intervals overlap in fig. 1, cases with alliances formed after independence are significantly more likely to experience MID initiation than those without alliances. Examining overlap in confidence intervals is not sufficient to determine if two estimates are different (Schenker and Gentleman 2001).

1816-2000

No defensive alliance

Defensive alliance from birth of dyad

Defensive alliance formed after birth of dyad

0   .001   .002   .003   .004

1816-1945

No defensive alliance

Defensive alliance from birth of dyad

Defensive alliance formed after birth of dyad

0   .001   .002   .003   .004

1946-2000

No defensive alliance

Defensive alliance from birth of dyad

Defensive alliance formed after birth of dyad

0   .001   .002   .003   .004
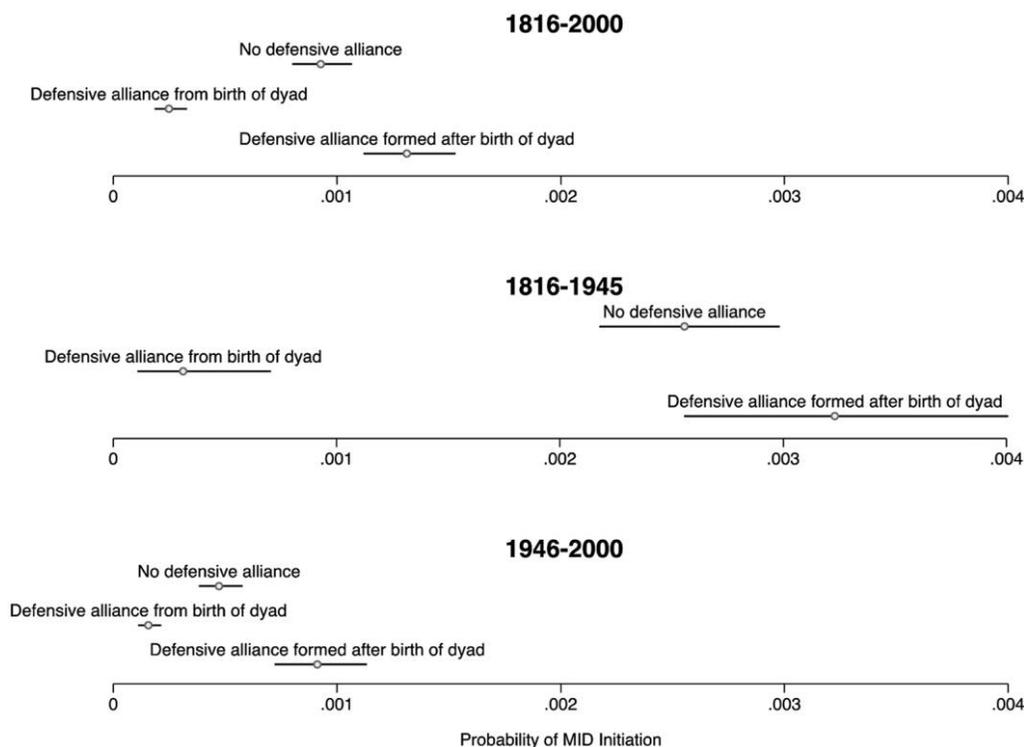
Probability of MID Initiation

Figure 2. The effect of defensive alliances on MID initiation. Point estimates are reported with 95% confidence intervals.

is followed by an increased risk of militarized conflict. Even if defensive alliances have a deterrent impact in some circumstances, these findings suggest that this effect is insufficient to systematically reduce the risk of conflict. Leaders seeking to promote peace by forming new alliances should therefore beware of the unintended consequences this might have and the significant potential for deterrence failure.

Generally, most adherents of nuclear deterrence believe that it is easier to deter nuclear than conventional wars, especially given the Cold War record. One failure, however, could be disastrous. With the spread of nuclear weapons, whether deterrence will work at the most dangerous levels is a crucial question. We think we have provided evidence that should give both scholarly analysts and policy makers pause about deterrence's effectiveness in the future. Our findings and those of Leeds and Johnson have been jointly beneficial because they have uncovered the complexity of empirical patterns and difficulties of how to investigate them. While there are still disagreements, the differences have narrowed at least somewhat. What is important is that we all agree that careful and scientific empirical analysis is how such disagreements should be handled.

ACKNOWLEDGMENTS

REFERENCES

Barnett, Michael. 2003. "Alliances, Balances of Threat, and Neorealism." In John A. Vasquez and Colin Elman, eds., *Realism and the Balancing of Power*. Upper Saddle River, NJ: Prentice-Hall, 222–46.

Bennett, D. Scott, and Allan C. Stam. 2007. *The Behavioral Origins of War*. Ann Arbor: University of Michigan Press.

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.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49 (3): 379–414.

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

Gibler, Douglas M. 1996. "Alliances That Never Balance: The Territorial Settlement Treaty." *Conflict Management and Peace Science* 15 (1): 75–97.

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

Kenwick, Michael R., John A. Vasquez, and Matthew Powers. 2015. "Do Alliances Really Deter?" *Journal of Politics* 77 (4): 943–54.

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 (3): 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, and Jesse C. Johnson. 2016. "Theory, Data, and Deterrence: A Response to Kenwick, Vasquez, and Powers." *Journal of Politics* 79 (1): **XX–XX**.

Morrow, James D. 1994. "Alliances, Credibility, and Peacetime Costs." *Journal of Conflict Resolution* 38 (2): 270–97.

Schenker, Nathaniel and Jane F. Gentleman. 2001. "On Judging the Significance of Differences by Examining the Overlap between Confidence Intervals." *American Statistician* 55 (3): 182–86.

Smith, Alastair. 1995. "Alliance Formation and War." *International Studies Quarterly* 39 (4): 405–25.

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