Alliance Participation, Treaty Depth, and Military Spending

Joshua Alley*

Abstract

How does alliance participation affect military spending? Some argue that alliance membership increases military expenditures, while others contend that it produces spending cuts. I argue that deep formal defense cooperation modifies the impact of alliance participation on military expenditures. When security-seeking non-major powers join deep alliances they usually decrease military spending, because these treaties are more credible. Joining shallow alliances often increases non-major power military spending, however. I test the argument by creating a latent measure of alliance treaty depth and using it to predict differences in how alliance participation affects military spending. The research design generates new empirical evidence linking alliance participation and percentage changes in state military spending from 1816 to 2007. I find that deeper alliance treaties tend to decrease non-major power military spending, and shallow alliances often increase military spending. These results help scholars and policymakers better understand a central question about alliance politics that has been debated in scholarship for decades.

*Graduate Student, Department of Political Science, Texas A&M University.

1 Introduction

Scholars of international relations have long acknowledged that there are two ways for states to increase their security. They can invest in indigenous military capability or form alliances (Morgenthau, 1948; Altfield, 1984; Morrow, 1993). Because both policies provide security, broadly defined, alliance participation should change how states invest in military capability. But exactly how alliances influence military spending remains unclear.

Existing scholarship contains contradictory theoretical predictions and evidence on the question of alliance participation and military spending. One view expects that alliance participation will reduce military spending e.g., (Barnett and Levy, 1991; Morrow, 1993; Conybeare, 1994). The other predicts that alliance participants will spend more on defense e.g., (Diehl, 1994; Morgan and Palmer, 2006; Quiroz Flores, 2011). This paper addresses the divide by using alliance treaty design to explain when alliance participation leads to more or less defense spending. In doing so, it helps clarify a longstanding debate about alliance politics.

I use variation in alliance design and membership to predict how alliance participation affects military spending. Scholars have engaged in extensive study of the sources and consequences of differences in alliance treaty design and membership (Mattes, 2012; Benson, 2012; Poast, 2019; Morrow, 1991; Leeds, 2003*b*; Leeds and Anac, 2005; Fordham, 2010; Mattes, 2012; Poast, 2013; Johnson, Leeds and Wu, 2015). Despite the importance of different alliance treaty designs for outcomes like conflict (Leeds, 2003*b*; Benson, 2012) and trade (Long, 2003; Long and Leeds, 2006) the debate about alliance participation and military spending has largely treated alliances as homogeneous.[1] But alliance participation could plausibly increase or decrease defense expenditures, under different alliance treaties. In particular, I emphasize how treaty depth modifies the impact of alliance participation on military spending. Deep alliances formalize extensive defense cooperation between members. In addition to commitments of military support, deep treaties require

[1] See DiGiuseppe and Poast (2016) for an important exception.

defense coordination and cooperation among alliance members.

To explore the consequences of deep and shallow alliances, I examine a particular set of states: non-major powers. I focus on non-major powers because these states clearly show a tradeoff between reassurance and allied military spending in deep alliances. Participation in deep alliances allows non-major powers to reduce military spending due to greater treaty credibility and reduced allied leverage on defense spending. Joining a shallow alliance often increases non-major power military spending because realizing foreign policy gains from alliance participation depends on defense spending when members fear abandonment.

I employ a novel research design to test my argument. First, I develop a latent measure of alliance treaty depth. I then incorporate that measure into a multilevel model which estimates how alliance characteristics modify the impact of total allied defense expenditures on annual percentage changes in military spending. Because allied capability is the ultimate source of security in alliances, I use total allied capability to measure alliance participation. Allied capability is a useful proxy for alliance participation because it combines the impact of joining an alliance and changing allied capability during treaty membership, both of which change the security benefits of alliance participation. Multilevel modeling matches my conditional argument and captures heterogeneous effects of alliance participation across individual treaties. I fit the model on a sample of non-major power states from 1816 to 2007 and find that while deep alliances decrease percentage changes in non-major power military spending, participation in shallow alliances often increases spending.

The argument and findings illuminate a salient debate in US foreign policy about the costs and benefits of alliances. Advocates of deep engagement (Brooks, Ikenberry and Wohlforth, 2013) and restraint (Posen, 2014) in grand strategy have different views of alliances. Proponents of restraint argue that the United States should withdraw from many alliances, because allies spend too little on defense, which then increases US defense spending (Preble, 2009). The deep engagement school argues that the benefits of alliances exceed the costs and believe that the problem of low allied military spending is overstated (Brands and Feaver, 2017). My argument and findings suggest

that policymakers can adjust alliance commitments to reassure partners or encourage higher allied military spending, but will struggle to do both. The United States often uses deep alliances to reassure partners, but this may encourage lower allied military spending. This tradeoff suggests that there is little room for compromise between deep engagement and restraint in grand strategy.

The paper proceeds as follows. First, I summarize competing claims on alliance participation and military spending. Then I describe my argument in more detail. After the argument, I present the research design and results. The final section concludes with a discussion of the results and implications for scholarship and policy.

2 Do Alliances Increase or Decrease Military Spending?

Scholarship on alliance participation and military spending is divided between two views. One expects that alliance participation will decrease military spending, while the other anticipates a positive relationship between alliance participation and defense expenditures. Each view predicts a different average effect of alliance participation by emphasizing one aspect of alliance politics.

Two types of arguments predict a negative association between alliance participation and defense spending. First, the economic theory of alliances (Olson and Zeckhauser, 1966) claims that alliances are subject to a collective action problem because security from an alliance is a public good. Because alliance security is neither rivalrous nor excludable, members contribute inadequate resources to collective defense. Alliance members can "free-ride" and smaller states exploit larger partners. Lower spending allows alliance members to consume more non-defense goods, but the alliance provides suboptimal security.[2] Second, substitution arguments recognize that states employ one policy in place of another (Most and Starr, 1989). Alliances provide security without requiring additional military spending (Morrow, 1993; Conybeare, 1994). Given extra security, states rely on their allies and and reallocate military spending to other goods. Both the substitution

[2]Sandler and Forbes (1980), Oneal (1990) and Sandler and Hartley (2001) all modify the public goods logic while relying on Olson and Zeckhauser's core intuition.

and public goods models expect that alliance participation reduces military spending due to the opportunity costs of military expenditures. States want to rely on their allies for security because higher defense expenditures leave fewer resources for other goods (Fordham, 1998; Fearon, 2018).

A contradictory perspective asserts that alliance participation increases military expenditures. Several arguments predict higher military spending by alliance members, using a shared intuition that states increase military spending to support their alliance commitments. Diehl (1994) argues that alliances create new foreign policy obligations, necessitating extra military spending. Because alliances expand what a state can achieve in international relations, states might increase military spending to pursue other foreign policy goals (Morgan and Palmer, 2006). For example, buffer states use conscription to make themselves a more attractive alliance partner (Horowitz, Poast and Stam, 2017). Others assert that cooperation within alliances generates higher defense spending (Palmer, 1990; Quiroz Flores, 2011). These predictions of a positive correlation between alliance participation and military spending contradict expectations of lower military spending by alliance members.[3]

2.1 Mixed Evidence

Mixed findings reinforce theoretical divisions between the contradictory views of alliances. Some studies find a positive association between alliance participation and military spending. Others find a negative relationship.[4]

Scholars have studied the connection between alliance participation and military spending in two ways. General studies of military spending and alliances compare many states through dummy indicators of alliance participation, which collapse alliances into a state-level measure. This design compares states with at least one alliance to those with none. By contrast, specific research designs

[3] Senese and Vasquez (2008) argue that military spending and alliances are part of a conflict spiral of simultaneous growth in military expenditures and alliance participation, which suggests that conflict behavior drives any correlation between alliances and military spending.

[4] Because tests of the public goods model use military spending as a share of GDP as the their outcome of interest, I do not include most of those results in this summary.

examine defense expenditures within a single alliance. Much of this work seeks to explain variation in spending among NATO members, focusing on how states respond to increases in contributions to collective defense by the United States.

Table 1 summarizes previous results on the issue of alliance participation and military spending. Both specific and general research designs produce mixed findings. There is one negative, three positive and two null estimates of the correlation between alliance participation and spending in general studies. Specific studies turn up five negative and two positive results.

Research Design	Study	Decrease	Increase	Null
General	Most and Siverson (1987)			X
	Conybeare (1994)	X		
	Diehl (1994)		X	
	Goldsmith (2003)			X
	Morgan and Palmer (2006)		X	
	Quiroz Flores (2011)		X	
Specific	Conybeare and Sandler (1990)		X	
	Barnett and Levy (1991)	X		
	Morrow (1993)	X		
	Sorokin (1994)	X		
	Chen, Feng and Masroori (1996)		X	
	Plümper and Neumayer (2015)	X		
	George and Sandler (2017)	X		

Table 1: Findings of the association between alliance participation and military spending. The top block details results from general studies, which compare states with at least one alliance to states without any alliances. The bottom block shows results from specific studies, which examine how national military spending changes in response to shifting allied capability.

The mixed empirical results reflect a theoretical problem. Both perspectives make unconditional claims about the average effect of alliance participation on military spending. With one exception (DiGiuseppe and Poast, 2016), scholarship on alliance participation and military spending ignores differences between alliances. Treaty obligations and membership vary widely across alliances, however (Leeds et al., 2002). I focus on a key difference between alliances that can help us understand their heterogeneous effects on military spending: the depth of military cooperation

in the treaty.

3 Argument

My argument explores how deep military cooperation in an alliance treaty modifies the impact of alliance participation on non-major power military spending. Because non-major powers focus on security, alliances protect them from external threats. Treaty depth shapes whether the security benefits of alliance participation depend on military spending. Even in the face of significant threats, non-major powers can gain more security with less military spending if they form deep alliances. Conversely, security gains in shallow alliances often depend on non-major powers maintaining or even increasing their military spending to manage abandonment concerns.

I focus on non-major powers to maintain theoretical and empirical parsimony. As I explain below, major and non-major powers use alliances to achieve different foreign policy goals, which changes how alliance participation affects their military spending. My argument also provides novel insights about non-major powers. Some scholarship and much popular discourse assumes that non-major powers regularly reduce military spending in alliances. I challenge this assertion by showing that alliance participation sometimes requires non-major powers to increase defense expenditures.

I start the argument by describing a general framework where alliances are institutionalized military cooperation between states. Then I discuss how deep formal military cooperation affects alliance credibility. Last, I explain how alliance depth affects the connection between alliance participation and non-major power military spending.

3.1 Cooperation in Alliances

Alliances are a form of international cooperation. Promising military support through a treaty generates a credible commitment of intervention (Fearon, 1997; Morrow, 2000). Allied support

then helps members achieve crucial foreign policy goals like deterrence or winning wars (Walt, 1990; Snyder, 1997). States form alliances so they can use other states' military capabilities to back their foreign policy aims.

Because allied capability gives a treaty foreign policy value, alliance participation is inseparable from allied capability (Fordham and Poast, 2014). The presence of an alliance treaty formalizes when a state can expect military intervention (Morrow, 2000), but the treaty itself does not provide security. Allied capability is the ultimate source of security in an alliance. Greater allied capability increases the value of an alliance, all else equal (Johnson, Leeds and Wu, 2015), so I conceptualize alliance participation in terms of allied capability.[5]

Alliance treaties and the capability they draw on can support many foreign policy aims, which often facilitates exchanges between alliance participants. One common exchange occurs in asymmetric alliances between major and non-major powers (Morrow, 1991). Large states form asymmetric alliances to increase their foreign policy influence, while smaller partners gain protection from external threats. Not all alliances are asymmetric,[6] but the divergent motives of major and non-major powers in these treaties reflect general tendencies in alliance politics. Major powers often use alliances to address the global balance of power and increase their influence. Smaller non-major powers tend to emphasize immediate security and regional concerns in their alliances. As a result, there are distinct processes behind non-major and major power alliance participation and these states respond to allied capability and treaty design in different ways.

Non-major powers use allied support to provide protection from external threats. But as with all cooperation, alliance members must account for opportunism, or "behavior with guile" (Williamson, 1985). Even as states commit to an alliance, they can also benefit from defecting and taking advantage of allied cooperation. Sometimes the perceived benefits of defection outweigh the long-run benefits of cooperation, so alliance members face an enforcement problem (Fearon,

[5] A binary conceptualization of alliance participation assumes all alliances are equally valuable.

[6] 130 of 289 ATOP alliances with offensive or defensive obligations are asymmetric pacts with at least one major and one non-major power, but a further 122 alliances are symmetric treaties between non-major powers.

1998; Koremenos, Lipson and Snidal, 2001).

Non-major powers are especially concerned with abandonment, which is the most common form of opportunism in alliances. One estimate suggests that the rate of compliance with military intervention obligations is only 50% (Berkemeier and Fuhrmann, 2018). For security-focused non-major powers, these alliance violations threaten their main goal. Concern with abandonment then affects military spending decisions.

Abandonment and military spending are related because greater alliance credibility allows states to lower military expenditures.[7] Though states can augment the collective military capability of an alliance through investing in military capability, they can also reduce defense spending and rely on their partners (Olson and Zeckhauser, 1966; Morrow, 1993; Conybeare, 1994; Sandler and Hartley, 2001). Under credible alliances, such reductions in military expenditures are more likely, because states have less fear of abandonment.

As DiGiuseppe and Poast (2016) observe, some alliances have fewer credibility concerns due to members' political regime type. They show that defense pacts with democracies lower defense spending, as democracies make more credible commitments. This insight about conditional credibility is a useful starting point because credibility is multifaceted. In treaty design, depth, unconditional military support (Benson, 2012; Chiba, Johnson and Leeds, 2015) and issue linkages (Long and Leeds, 2006; Poast, 2012, 2013) all increase credibility.[8] In this paper, I focus on treaty depth, which is a costly promise that allows alliance members and potential adversaries to infer the credibility of the alliance (Leeds, 2003*b*; Fuhrmann and Sechser, 2014).

If alliance participation can reduce non-major power military spending, is that a problem for their allies? This argument assumes that allied states usually prefer higher non-major power military spending because it adds capability to the alliance.[9] Allies also encourage higher military

[7]The public goods model of alliances calls reduced military spending in alliances free-riding. I do not use this language because it has come to imply that reduced defense spending is problematic. But as Olson and Zeckhauser (1966, pg. 278) themselves observe, lower defense spending is not normatively problematic.

[8]Though the argument emphasizes depth, the research design accounts for multiple sources of alliance credibility.

[9]This preference is somewhat attenuated by concern that substantial changes in capability will provide too much

spending so their partners can manage internal threats. This preference for higher defense spending is not a driving force in alliance treaty design, however, because states prioritize abandonment and entrapment concerns in treaty design. States cannot form an alliance without addressing abandonment and entrapment, but they can leave bargaining over military spending to alliance maintenance (Snyder, 1997). Also, some evidence suggests that alliance negotiations emphasize when and how states provide military support (Poast, 2019).

Assuming that allies of non-major powers usually prefer higher military spending is reasonable. States often sell or transfer weapons to allies in order to augment allied capability (Yarhi-Milo, Lanoszka and Cooper, 2016; Pamp, Dendorfer and Thurner, 2018). Even in alliances between superpowers and non-major powers, where highly asymmetric capability and interests might reduce concerns about non-major power defense expenditures, superpowers often encourage higher military spending. US demands for higher military spending by NATO members are well-known (Schuessler and Shifrinson, 2019). The Soviet Union also regularly pressured Warsaw Pact members to increase defense spending (Simon, 1985).

When allies want non-major powers to spend more, enforcing such cooperation is difficult. Normative appeals to common interests are ineffective. Though verbal communication or "cheap talk" has value in international politics (Trager, 2010), it is unlikely to overcome the opportunity costs of defense spending. Even after reducing defense expenditures, alliance members retain foreign policy benefits and can reallocate resources to other priorities. The ability to reduce defense spending and spend more on other goods sometimes motivates states to form alliances (Kimball, 2010; Allen and DiGiuseppe, 2013).

Therefore, if alliance members want to encourage greater allied defense spending, they need leverage. States gain leverage either from a credible threat to abandon low-spending partners or direct control over allied policies. Policy control of allied spending decisions occurs when the

autonomy. For example, the United States has inhibited nuclear acquisition by some allies (Gavin, 2015; Lanoszka, 2015).

alliance reflects hierarchical relationships (Lake, 1996). Without such direct influence, states must possess a credible threat to leave the alliance over low defense spending. Otherwise, allies will dismiss weaker signals and threats due to uncertainty and incomplete information.

Reassuring allies reduces the credibility of threats to abandon states that spend too little on defense. States cannot simultaneously reassure their allies and maximize leverage on defense spending. As alliance members use costly commitments to reassure, partners have more freedom to reduce defense spending.

Under less credible alliances, low military expenditures could increase the risk of abandonment, so members have less freedom to reduce defense spending. Moreover, although states in less credible alliances increase their foreign obligations, they also face the risk of having to meet foreign contingencies without allied support. As treaty credibility falls, the foreign policy benefits of alliance participation become contingent on military spending.[10] This occurs because alliance members must hedge against abandonment and partners can leverage that concern to encourage higher defense spending. Thus, alliances with less credibility will be more likely to increase military spending.

Treaty depth highlights this tradeoff between reassurance and military spending. Where deep alliances often reduce non-major power military spending, participation in shallow alliances often increases it. Stipulating deep cooperation reassures partners and reduces leverage over allied military expenditures. Credibility from treaty depth also promotes efficiency gains from coordinated defense effort. I now describe the role of treaty depth in more detail.

3.2 Alliance Treaty Depth

I focus on depth because it provides theoretical leverage to predict when alliance participation increases and decreases military spending, which reveals a trade off between reassurance and de-

[10]For example, (Niou and Zeigler, 2019) use a formal model to show that in alliances between rivals, states are less likely to reduce military spending, because their foreign policy gains from alliance participation depend on military spending.

fense spending. Moreover, treaty depth is a common policy choice.[11] Though states probably do not change their political regime to reassure allies, they often form deep alliance treaties. Over half of defensive or offensive ATOP alliances have some depth.

Alliance depth is the extent of defense cooperation formalized in the treaty. Deep alliances require additional policy coordination and military cooperation beyond a promise of military support. While shallow alliances stipulate more arms-length cooperation between members, deep treaties lead to closer cooperation. Defense cooperation in a deep alliance takes many forms. Allies can form an integrated military command, provide military aid, commit to a common defense policy, provide basing rights, set up an international organization or undertake companion military agreements. All of these obligations move alliance members away from an arms-length partnership towards close cooperation via policy coordination and regular interaction, while imposing monetary and policy autonomy costs.[12]

One example of a deep alliance is a 1948 defense pact between the United Kingdom and Jordan. This alliance included basing rights, military aid, official military contact, and an Anglo-Transjordan Joint Defense Board. The basing, aid, and cooperation added more depth than a 1912 treaty between Greece and Bulgaria which only commits to mutual defense and consultation if either state is attacked by Turkey.

Deep alliances like the Anglo-Jordanian pact reduce non-major power military spending in two ways. First, depth reassures partners and reduces leverage on defense spending. Deep alliances are more credible because defense cooperation is costly. Making costly commitments of bases, policy coordination, or aid reassures allies (Morrow, 1994). Depth is especially useful because alliance members face a time inconsistency problem. Alliance treaty fulfillment depends on shared foreign

[11] In a related paper, I explore the sources of alliance treaty depth. All the sources of depth that paper identifies are included as control variables in the empirical analysis. Average democracy and threat at the time of alliance formation are the two largest sources of higher treaty depth.

[12] Although depth can have monetary costs, reductions in military spending due to greater treaty credibility will often outweigh those costs. Alliance contributions are a small part of most defense budgets, and non-major powers rarely use basing rights in deep alliances to deploy their troops abroad.

policy interests (Morrow, 2000; Leeds, 2003*a*), so changing foreign policy interests threaten alliance fulfillment (Leeds and Savun, 2007). A deep alliance makes a series of repeated transfers, and states can signal commitment by maintaining those transfers.[13] Second, credibility in deep alliances can facilitate more efficient defense spending, as alliance members to provide specific capabilities. Specialization means members of deep alliances spend less on the military but retain adequate security. These efficiency gains depend on the credibility of the alliance, because states will only specialize if they believe the alliance is reliable (Leeds, 2003*a*).

On the other hand, shallow alliances are less credible, so participation in these treaties is more likely to increase military spending. Shallow alliances have some basic credibility from hands-tying signals (Fearon, 1997), as well as the audience (Morrow, 2000) and reputational (Gibler, 2008; Crescenzi et al., 2012) costs of violation. Even so, threats to abandon low-spending allies are more credible than in a deep alliance where partners have taken pains to reassure their partners. In a shallow alliance, members must hedge against abandonment, which partners can use as leverage to discourage low defense spending. Maintaining the benefits of alliance participation then requires defense spending, because low military spending might endanger the treaty or expose states to adverse consequences if they are abandoned. Shallow alliances are also less likely to facilitate military specialization due to the fear of abandonment. As a result, participation in shallow alliances often increases military spending.[14]

To illustrate the logic, consider two related alliances from the inter-war period. A 1920 treaty between France and Belgium (ATOPID 2055) added commitments of military aid and policy coordination to defensive obligations. Given this depth, the Franco-Belgian alliance reduced Belgian defense expenditures, even while Belgium helped occupy the Ruhr. A more limited treaty with only military support between France, Belgium, the United Kingdom, Italy and Germany

[13] Conversely, eliminating or reducing planned transfers reduces the credibility of the whole alliance.

[14] One objection to this argument is that deep alliances are more valuable to members, which augments allied influence on defense spending. Although alliance value adds some leverage, it cannot offset reducing the credibility of threats to abandon low-spending allies. Value increases leverage because states fear their allies will abrogate a valuable alliance, and deep alliances counteract this essential concern.

(ATOPID 2130) increased Belgian spending, on the other hand.

Taken together, the examples and argument suggest that treaty depth modifies the impact of alliance participation on non-major power military spending. Shallow alliances often increase military spending, and deep alliances usually reduce spending. Therefore, with a continuous conceptualization of depth, there should be a negative correlation between treaty depth and the impact of alliance participation on non-major power military spending as the positive effects of shallow treaties turn into negative effects in deep alliances. This implies three separate hypotheses, one about shallow alliances, another about deep alliances, and the third about how changes in treaty depth modify the association between alliance participation and military spending.[15]

> HYPOTHESIS 1: ON AVERAGE, PARTICIPATION IN SHALLOW ALLIANCES WILL INCREASE PERCENTAGE CHANGES IN NON-MAJOR POWER MILITARY SPENDING.

> HYPOTHESIS 2: ON AVERAGE, PARTICIPATION IN DEEP ALLIANCES WILL DECREASE PERCENTAGE CHANGES IN NON-MAJOR POWER MILITARY SPENDING.

> HYPOTHESIS 3: AS ALLIANCE TREATY DEPTH INCREASES, THE IMPACT OF ALLIANCE PARTICIPATION ON PERCENTAGE CHANGES IN NON-MAJOR POWER MILITARY SPENDING WILL DECREASE.

Before detailing how I test these hypotheses, I must briefly justify two conceptual and measurement choices. First, the three hypotheses predict how percentage changes in non-major power military spending differ under deep and shallow alliances. Percentage changes in military spending express changes in spending as a share of the previous year's defense budget. This variable is an appropriate outcome of interest, in part because it expresses the opportunity costs of military spending. All else equal,[16] a larger increase in spending relative to the previous year's defense budget

[15] Hypothesis 3 follows from Hypotheses 1 and 2.
[16] Especially holding economic growth constant.

imposes more constraints on other goods. Using percentage changes also facilitates comparisons across diverse states and years.

Second, I use allied capability to express the consequences of alliance participation for military spending, rather than a dichotomous indicator of participation. Measuring participation using allied capability matches my argument and encompasses the division in previous research designs. My argument starts with the premise that states form alliances so that allied capability supports their foreign policy goals. States do not respond to a treaty per se, rather they respond to expectations that allies will employ military capability on their behalf. This makes alliances with more capable states are more valuable, all else equal. If allied capability is the main source of security, alliance participation affects military spending through joining an alliance and changes in allied capability after the treaty forms. In previous scholarship, general research designs address the implications of joining an alliance, while specific designs focus on the ramifications of changes in allied capability. Conceptualizing alliance participation in terms of allied capability encapsulates both designs, creating a unified approach to understanding how alliances affect military spending.

Because my argument focuses on differences between deep and shallow treaties, the research design must measure alliance treaty depth and show how depth modifies the impact of allied capability on military spending. I use a measurement model to infer depth from treaty content, then connect alliance characteristics to military spending with a multilevel model. The next section describes the research design in more detail.

4 Research Design

The research design involves two steps. First, I develop a latent measure of treaty depth for alliances with military support. Second, I employ that measure in a multilevel model to estimate how treaty depth modifies the impact of alliance participation on military spending and apply the multilevel model to a sample of non-major powers from 1816 to 2007. The next section describes

the measure of alliance treaty depth.

4.1 Measuring Alliance Treaty Depth

Formal treaty commitments reflect alliance depth,[17] so I use observed alliance treaty characteristics to infer depth. This approach could produce ordinal or continuous measures of treaty depth. Before describing my measure, I discuss two related measures in previous work.[18] Due to the limitations of these measures for my purposes in this paper, I create a new measure of treaty depth in offensive and defensive ATOP alliances using a semiparametric factor analysis.

In an ordinal index of treaty depth, researchers theoretically assign depth values to alliances. (Leeds and Anac, 2005) develop an ordinal measure of military institutionalization alliance that includes scores of zero, one or two based on the extent of investment in joint action required by the alliance treaty. This measure is not an additive index of treaty depth, as it weights different alliance obligations. The resulting measure roughly captures treaty depth, but it understates variation in this crucial concept. Leed's and Anac's measurement strategy imposes equal weights on different sets of depth variables and does not aggregate multiple sources of depth. For example, it treats an integrated military command and military bases as equivalent sources of depth, and does not aggregate the two promises if both are present. Therefore an alliance with an integrated command and bases is just as institutionalized as an alliance with only an integrated command. Given these limitations, I employ a more flexible measurement strategy.

My particular measure builds on work by Benson and Clinton (2016), who use a latent variable model (Quinn, 2004) to measure alliance scope, depth and capability. Measurement models have a rich history in political science e.g. (Clinton, Jackman and Rivers, 2004; Fariss, 2014). Latent variable modeling uses data to estimate the weights of different observable variables in some unob-

[17] Formal treaty obligations may not be fully implemented, but formal depth should produce practical depth as states try to uphold the credibility of their commitments.

[18] Neither of these measures uses an additive index of the different sources of treaty depth. Adding together the presence of different sources of depth assumes that every source is equally important, which creates theoretical and empirical problems with making comparisons between alliances (Benson and Clinton, 2016, pg. 876).

served concept. In this case, the measurement model uses correlations between observable alliance treaty content and unobserved latent depth to predict the depth of each treaty. With this approach, theory identifies the relevant correlates of treaty depth, but the data drives how much depth each observed variable adds to the alliance.

Given the advantages of latent variable modeling, I emulate Benson and Clinton's approach, but use a different concept, sample of alliances and estimator. Conceptually, Benson and Clinton (2016) define depth as the general costliness of the alliance, so they include measures of economic issue linkages and secrecy in their observed data. My definition of depth emphasizes military cooperation between alliance members, and I view issue linkages as a separate source of credibility. Given their broad definition of depth, Benson and Clinton also include neutrality pacts in their sample of alliances. I am only interested in offensive and defensive alliances, however. As for the estimator, in Benson and Clinton's parametric latent variable models, the latent variables influence the form of the dependence structure and the marginal distributions of the latent value estimates. These correlations can produce misleading inferences about latent scores, so I employ a semiparametric estimator.

I use a Bayesian Gaussian Copula Factor Model (Murray et al., 2013) to measure alliance treaty depth in offensive and defensive alliances. Murray et al's model improves inferences from mixed factor analysis for continuous, ordinal, and binary observed data by relaxing distributional assumptions. Given discrete observed variables and non-Gaussian latent variables, the dependence among the latent variables and their marginal distributions are both influenced by the latent variables. This approach breaks the dependence between the latent factors and marginal distributions by using copulas to encode the dependence among the latent variables.[19] Beyond the semiparametric aspect, this measurement model is a standard ordinal factor analysis.

I estimated the measurement model using observed data from all 289 alliances with offensive or defensive obligations in the alliance-level ATOP data (Leeds et al., 2002). This includes major

[19]Copulas are a distribution function on $[0, 1]^p$ where each univariate marginal distribution is uniform on [0, 1].

power alliances, non-major power alliances, and asymmetric alliances. I examine alliances with military support because prior studies of alliance participation and military spending focus on these treaties. Indicators of treaty depth include military aid, bases, international organization formation, integrated military command, defense policy coordination, subordination of forces in wartime, specific contribution requirements, and commitments to form companion military agreements. The argument suggests there is a single factor underlying variation in all eight binary and ordinal indicators, so I fit the model with one latent factor. To estimate the model, I used Parameter expanded Gibbs sampling, a normal prior on the factor loadings, 20,000 burn-in iterations of the MCMC chain, and 30,000 samples thinned every 30 observations to ensure convergence. The estimates include posterior distributions for the factor loadings and the latent factor.

To summarize treaty depth, I use the posterior mean of the latent factor for each alliance, so each alliance has its own depth value. The posterior mean captures the expected depth of an alliance treaty, conditional on its formal promises. Figure 1 describes the latent depth of ATOP alliances with defensive or offensive commitments from 1815 to 2016. There is substantial variation in alliance treaty depth, which has several sources. The top panel in Figure 1 shows the factor loadings from the latent variable model, which are essentially correlations between the observed variables and latent factor. Military policy coordination, integrated military command, and formal organizations are the three largest correlates of depth. The other five factors have smaller but still strong and positive associations with latent depth.

The factor loadings behind this model match the argument. First, all the variables have strong positive factor loadings, as all require additional cooperation between alliance members. Second, I define depth as the extent of cooperation a treaty requires, so it makes sense that an integrated command, policy coordination requirements, and formal organizations have the highest factor loadings. All three of these variables are key sources of peacetime cooperation. Third, thinking about depth in terms of coordination and cooperation may also explain why basing rights falls into the second tier of factor loadings, which is a perhaps the least intuitive estimate. Though basing rights

in an alliance treaty clearly increase treaty depth, they may not add as much depth as some other factors because basing mixes cooperative and hierarchical elements. Providing basing rights and managing their daily interactions is cooperative, of course. But bases can also operate as enclaves and focus more on power projection or control than alliance-building. Relative to the top sources of treaty depth, basing rights are easier to separate from alliance cooperation.

The bottom panel of Figure 1 plots the posterior means and uncertainty of the depth estimates against the start year of the treaty. Many treaties have no depth, and are clustered on around -0.8. 171 alliances have a depth score higher than -0.6 because at least one source of depth is present. Even after accounting for uncertainty, it is possible to distinguish between some alliances.

Although the values of the latent measure are not intrinsically meaningful and can therefore be positive or negative, differences between treaties on the latent scale are informative. The median of treaty depth is -0.09, and the mean is 0.05. The median treaty is the Southeast Asian Treaty Organization (SEATO), which includes a formal international organization (ATOP ID 3260). There are many shallow treaties that only include military support. One such alliance is an 1855 pact between France, the UK and Sweden (ATOPID 1190) which promises defense and consultation.

Three of the deepest treaties are a 1993 alliance between Russia and Tajikistan (ATOPID 4470), a 1958 alliance between the UAE and Yemen (ATOPID 3345), and a 1981 pact between Gambia and Senegal (ATOPID 3930). All these alliances stipulate extensive defense cooperation. The alliance between Russia and Tajikistan includes military aid, bases, a companion military agreement, and integrated military command. The other two treaties attempted to establish a federation through military support, international organizations, basing, and defense policy coordination.

Table 2 provides five other examples of deep and shallow alliance treaty obbligations. The 1865 Triple Alliance between Brazil, Argentina and Uruguay included extensive cooperation for their war against Paraguay. An 1893 Franco-Russian alliance called for policy coordination and specified the number of troops each partner would provide in war. France's 1935 alliance with Poland

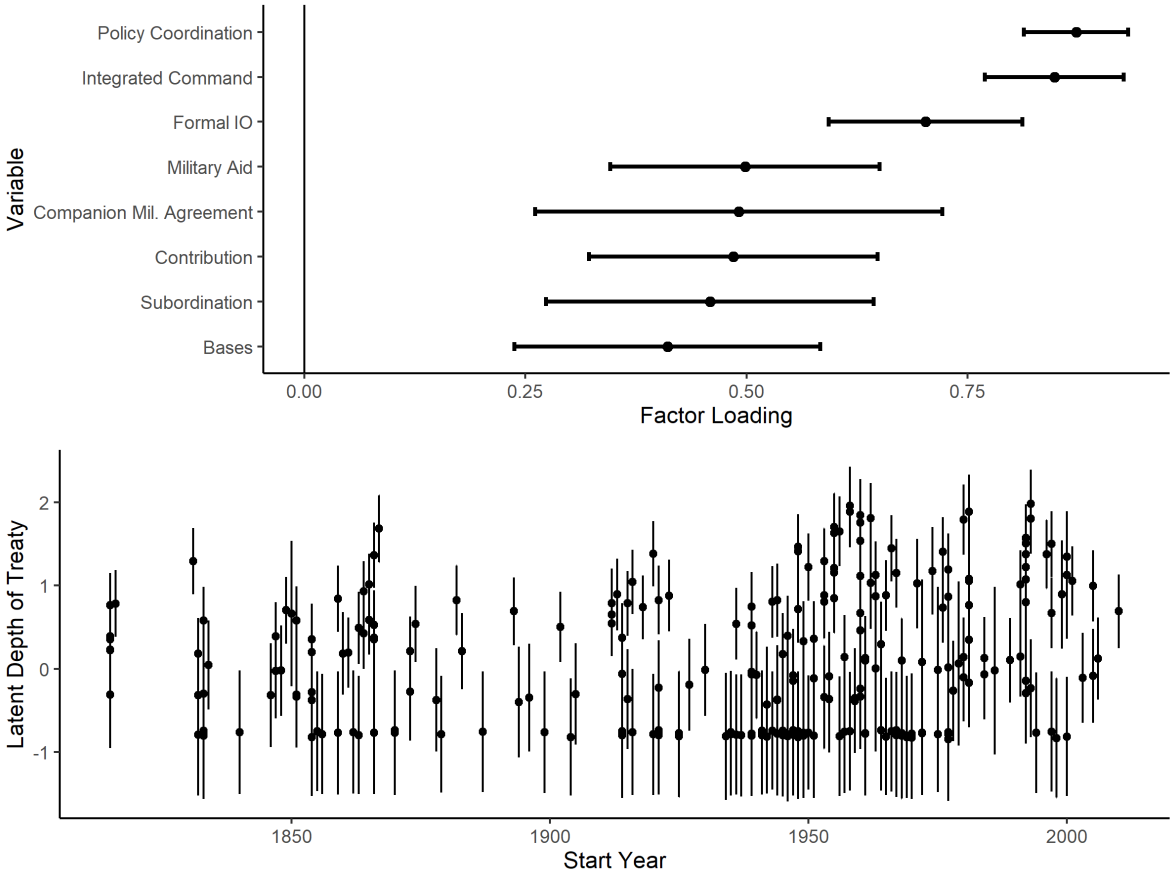

Figure 1: Summary of the latent measure of alliance treaty depth for 289 defensive or offensive alliances from 1816 to 2016. The top panel plots the factor loadings with 90% credible intervals. The bottom panel plots mean treaty depth (points) and the standard deviation (error bars) against the start year of the treaty.

Alliance	Depth Score	Year	Depth Sources
Brazil-Argentina-Uruguay	1.01	1865	Integrated Command, Military Aid, Low Policy Coord., Subordination.
France-Russia	.69	1893	Moderate Policy Coord., Specific Contribution
France-Poland	-.75	1935	None
NATO	.32	1951	Formal Organization, Basing rights
Warsaw Pact	1.2	1955	Formal IO, High Policy Coord., Integrated Command

Table 2: Table of treaty depth estimates, year of formation and sources of depth for five alliances.

has a minimal depth score, as it did not include any cooperation besides defensive and consultation obligations. As a result of ATOP's decision to omit the integrated command from formal treaty obligations, NATO scores above average on latent treaty depth, but it is not an overwhelmingly deep alliance in its formal treaty obligations.[20] Because the Warsaw Pact obligates members to join a formal organization, coordinate defense policy, and participate in an integrated command, it has greater formal depth than NATO. NATO has less depth because states like France and Greece could leave the integrated command structure without violating the alliance.

The latent measure has some face, concept, and discriminant validity. As an example of face validity, the Gambia-Senegal federation requires deeper cooperation than arms-length commitments of military support. Shallow treaties promise little beyond military support, matching my conceptualization of treaty depth. Last, Figure 1 shows that this measure can distinguish between deep and shallow commitments.

This latent treaty depth measure is the key explanatory variable in my empirical analysis. My argument claims that differences in depth among alliances modify the impact of alliance participation on percentage changes in military spending by alliance members at the state-year level of analysis. I use a multilevel model to capture this relationship, and summarize the estimation strategy in the next section.

[20]Participation in the integrated command is not required by the NATO treaty, so ATOP does not count this as a treaty obligation. Basing rights are part of a 1951 additional protocol.

4.2 Estimation: Multilevel Model

Because alliances and state-year observations are separate levels of analysis, I use a multilevel model to estimate the association between treaty depth and military spending and bridge the two data levels (Steenbergen and Jones, 2002; Gelman and Hill, 2007). My model estimates heterogeneous effects of alliance participation on military spending as a function of alliance characteristics, in order to make inferences about how alliance characteristics like formal depth modify the impact of individual alliances on military spending. To facilitate computation and interpretation, I fit the model using Bayesian estimation in STAN (Carpenter et al., 2016). See the appendix for details of the weakly informative prior distributions and evidence the chains converged.

This research design is more complicated than a panel data model like the estimator DiGiuseppe and Poast (2016) use, but the multilevel components add substantial value, especially by connecting the argument and research design. I argue that treaty depth modifies the impact of alliance participation on growth in military spending. Differences in how alliance participation impacts military spending are the outcome of interest. The multilevel model explicitly compares the impact of participation in deep and shallow alliances by estimating how changes in treaty depth modify the consequences of alliance participation.

Unlike the multilevel model, standard panel models employ state-level proxies for alliance characteristics, which compare states rather than alliances. This practice of aggregating alliances at the state-year level of analysis may produce misleading inferences (McElreath, 2016, pg. 356). Summarizing alliance characteristics at a different level of analysis changes the mean and variance of key independent variables, which could affect inferences. Multilevel modeling avoids this aggregation problem by retaining the structure of alliance data, where states participate in multiple alliances and alliances could have heterogeneous effects on military spending. The multilevel model estimates the specific impact of each alliance on members' military expenditures, thereby revealing differences between individual treaties. On the other hand, aggregating multiple alliances

into state level indicators will mask any heterogeneous effects of individual treaties.[21]

Furthermore, multiple alliance characteristics affect the consequences of alliance participation. The multilevel model captures multiple sources of heterogeneity in how alliances impact military spending. Treaty depth is correlated with other aspects of alliance membership and design, so this step is important.[22] In a panel estimator with state-level proxies for alliance characteristics, accounting for correlated alliance characteristics is difficult. To account for several alliance characteristics, panel estimates must average different parts of a state's alliance portfolio or subset the data. Averaging reduces theoretically interesting alliance-level variation, and analysis of multiple subsets risks generating spurious findings through multiple comparisons. In a multilevel model, I can account for how multiple alliance characteristics change the consequences of alliance participation by including other variables besides treaty depth in an alliance level regression. Therefore, my estimate of how treaty depth modifies the impact of alliance participation on military spending holds key alliance and state characteristics constant. I now describe the model specification in more detail.

4.2.1 Model Specification

This multilevel model connects two distinct regressions. The base is a state-year-level regression, which includes the impact of alliance participation. A second alliance-level regression modifies the effect of alliance participation on military spending, like an interaction.

The state-year-level regression starts with a distribution for the outcome:

$$y \sim student_t(\nu, \mu, \sigma) \tag{1}$$

y is the dependent variable— percentage changes in military spending. I model the outcome

[21] Partial pooling of the alliance-specific parameters generates reasonable estimates for each alliance.

[22] For example, I show in another paper that democratic alliance membership is positively correlated with treaty depth.

using a t-distribution with degrees of freedom ν to address heavy tails and estimate ν directly. σ is analogous to the error term in a frequentist regression as it captures unexplained variation. μ, the mean of the outcome, depends on several factors.

$$\mu = \alpha + \alpha^{st} + \alpha^{yr} + \mathbf{W}_{n\times k}\gamma_{k\times 1} + \mathbf{Z}_{n\times a}\lambda_{a\times 1} \tag{2}$$

Percentage changes in spending are a function of an overall intercept α, state and year varying intercepts α^{st} and α^{yr} and a matrix of state-level control variables $\mathbf{W}$. The $\mathbf{Z}\lambda$ term incorporates alliance participation. $\mathbf{Z}$ is a matrix of state participation in alliances. Columns correspond to each of the a alliances in the data, and rows to state-year observations. For this sample of non-major powers, $\mathbf{Z}$ has 190 columns and 8,290 rows. If a state is not in a particular alliance, the corresponding matrix cell in that alliance column is zero. If a state is part of the alliance in a given year, the matrix element contains the log of total allied military spending, which is normalized by year.[23] I use total allied spending to express the effect of alliance participation to match the theoretical emphasis on allied capability. $\mathbf{Z}$ encodes a quasi-spatial indicator of alliance participation for all a alliances in the data. States can be members of multiple treaties, so state-year observations are not neatly nested within alliances. This specification means each alliance has a unique impact on military spending, even when states participate in multiple treaties.

λ is a vector of parameters which estimate the impact of participation in specific alliances on military spending. Because the non-zero elements of $\mathbf{Z}$ are allied spending, the λ parameters capture alliance members' response to allied capability. Each alliance has a unique λ, so there are 190 alliance participation parameters. The λ parameters have a shared distribution, so I assume alliances are similar but different in how they impact military spending.

The second part of the multilevel model uses alliance characteristics to predict how alliance participation is associated with percentage changes in military spending. The alliance participation

[23]Normalization keeps the parameters on similar scales, which is important for modeling. The normalization is based of all 289 ATOP alliances, not just alliances with non-major powers.

parameters are the outcome in an alliance-level regression. As a result, the impact of alliance participation on members' military spending depends on treaty characteristics, including depth. In this second-level regression:

$$\lambda_a \sim N(\theta_a, \sigma_{all}) \tag{3}$$

and

$$\theta_a = \alpha_{all} + \beta_1 \text{treaty depth} + \mathbf{X}_{a \times l}\beta \tag{4}$$

In the alliance-level regression, $\mathbf{X}$ is a matrix of the l alliance-level control variables and α_{all} is the constant. Adding σ_{all} means predictions of λ are not deterministic— the alliance level regression contains an error term. A larger σ_{all} indicates more variation in how alliance participation impacts military spending.

The second-level regression includes treaty depth, and each β parameter modifies the impact of alliance participation on percentage changes in military spending. The βs are like marginal effects in an interaction. Treaty depth impacts military spending by modifying the consequences of alliance participation. Changing treaty depth shifts λ, which in turn affects military spending. Hypothesis 3 predicts β_1 will be negative for non-major powers.

In this model, the λ parameters express heterogeneous effects of participation in individual alliances. The β parameters estimate how alliance characteristics modify the impact of alliance participation on military spending. Again, using alliance characteristics to predict the impact of alliance participation matches my conditional argument. I now describe the sample and key variables in the analysis.

4.3 Sample and Key Variables

I estimate the multilevel model on a sample of non-major power states from 1816 to 2007, including state-year observations with no alliances. I identify non-major powers using a measure of

major power status from the Correlates of War Project. Alliance participation data comes from the ATOP project (Leeds et al., 2002). I focus on participation in defensive and offensive treaties, because prior studies of alliances and military spending examine these treaties. The sample contains data from 8,280 state-year observations and 190 alliances.

The dependent variable is percent changes in military spending, which is calculated as:

$$\% \text{ Change Mil. Expend} = \frac{\text{Change Mil. Expend}_t}{\text{Mil. Expend}_{t-1}} \tag{5}$$

I used the Correlates of War Project's data on military spending to measure percentage changes in spending (Singer, 1988).The annual percentage change in spending equals that year's change in spending as a share of the previous year's military spending. Thus, annual changes are bench marked to previous spending levels. To facilitate model fitting, I apply the inverse hyperbolic sine transformation to this variable.[24] Using percentage changes in military expenditures as the dependent variable helps the research design. The level of military spending is not stationary for most states, especially in longer panels. Thus, using percentage changes in spending reduces the risk of spurious inferences. Benchmarking changes to prior expenditures also facilitates comparisons across states and over time.

The key independent variable is the mean latent depth of each alliance, based on the measurement model. This variable enters the model in the alliance-level regression and I expect it will have a negative coefficient. I also include a series of state and alliance-level controls.

In the state-level regression, I adjust for several correlates of alliance participation and military spending. State-level covariates include GDP growth (Bolt et al., 2018) regime type, international war (Reiter, Stam and Horowitz, 2016), civil war participation (Sarkees and Wayman, 2010), annual MIDs (Gibler, Miller and Little, 2016), rival military spending (Thompson and Dreyer, 2012)

[24]This transformation applies to positive, negative and zero values. It has minimal impact on values between -1 and 1, but pulls in large positive values, which range as high as 140. Inferences about treaty depth and other alliance characteristics are comparable with and without the transformation.

and a dummy for Cold War years. Conflict participation, alliances, and military spending are all correlated (Senese and Vasquez, 2008). I include growth in GDP instead of levels because GDP levels are non-stationary and economic growth shapes the opportunity costs of military spending (Kimball, 2010; Zielinski, Fordham and Schilde, 2017).

Other alliance level variables are correlates of treaty design and military spending, including the number of members and share of democracies in a treaty at time of formation (Chiba, Johnson and Leeds, 2015). I control for issue linkages by creating a dummy indicator of whether the alliance promises economic cooperation (Poast, 2013; Long and Leeds, 2006). As an indicator of hierarchical security relationships, I include a count of foreign policy concessions in the alliance. I also mark the presence of unconditional military support using a dummy variable I constructed using existing indicators of conditional support in the ATOP data. Because threat may drive states to form deeper alliances and affect subsequent military spending, I control for the average threat of alliance members at the time of alliance formation using the threat measure of Leeds and Savun (2007). I adjust for superpower membership— whether the United States or Soviet Union participated in a treaty during the Cold War. Two dummy indicators of wartime alliances and asymmetric obligations (Leeds et al., 2002) complete the alliance-level regression specification. Though I discuss these variables as controls, many of them are theoretically interesting in their own right. Having described the measure of treaty depth, multilevel model, and covariates, I now turn to the results of the analysis.

5 Results

This section summarizes inferences from the multilevel model. I find support for all three hypotheses. Because shallow alliances tend to increase military spending and deep alliances often decrease spending, treaty depth and the effect of alliance participation on non-major power military spending are negatively correlated. Results are based on 2,000 samples from four chains, with

1,000 warm-up iterations. To facilitate model fitting, I employed a non-centered parameterization of the varying intercepts and a sparse matrix representation of $\mathbf{Z}$. Standard convergence diagnostics indicate the chains adequately explored the posterior.[25]

Because I used Bayesian modeling to estimate the association between treaty depth and percentage changes in military spending, each coefficient has a posterior distribution— the likely parameter values conditional on the priors and observed data. There are no indicators of statistical significance. Instead, I use the 90% credible intervals of the parameters and calculate the negative posterior probability for the treaty depth coefficient to assess Hypothesis 3.[26]

Figure 2 summarizes the coefficient estimates from the alliance-level regression and the simulated substantive effect of greater treaty depth. The preponderance of evidence matches Hypothesis 3, as shown in the top panel of Figure 2. There is a 96% chance treaty depth is negatively correlated with the impact of alliance participation on percent changes in military spending for non-major powers. As treaty depth rises, the effect of alliance participation on military spending falls.

I plot the substantive effect of the alliance-level depth coefficient in the bottom panel of Figure 2. I assess this substantive effect by simulating the effect of changing treaty depth from the minimum value of -0.8 to 1.5. Holding other alliance covariates at their modes or medians, this increase in depth reduces a hypothetical alliance parameter λ by .08 in expectation, which then lowers military spending growth. The results of the the simulation are summarized by 90% intervals for predicted growth in military spending in the hypothetical shallow alliance, predicted spending growth in the hypothetical deep alliance, and the difference between those two scenarios. Assuming the hypothetical alliance has median capability, the difference in spending growth between the shallow and deep treaty has a mean of -.03. The 90% credible interval of this predicted fall in military spending due to increasing treaty depth ranges from -0.056 to -0.001. Although the range of estimated substantive effects includes small values, there is a perceptible difference between

[25] See the appendix for details on convergence and other robustness checks.

[26] I use 90% intervals because inferences about 95% intervals are sensitive to simulation variance in Bayesian analysis.

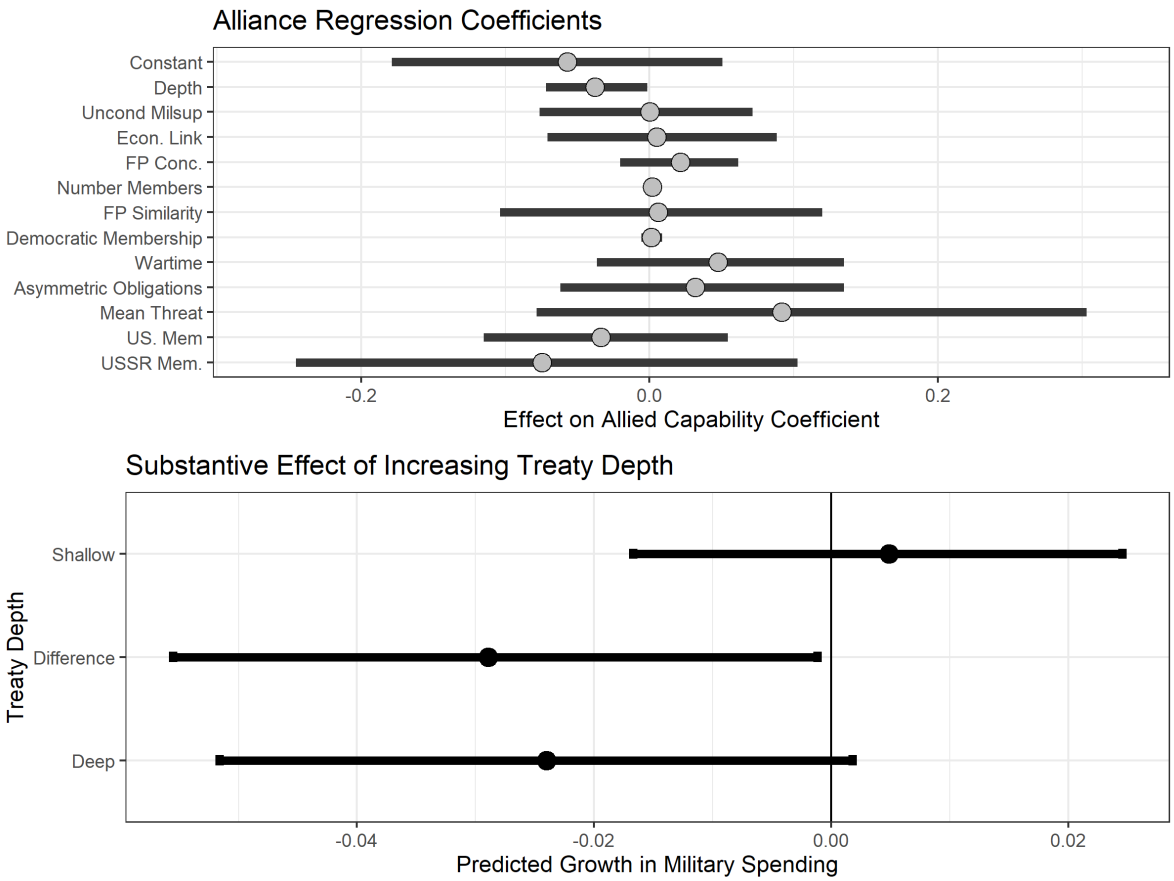

Figure 2: Summary of alliance-level regression results from the multilevel model of alliance participation and military spending. The top panel shows the 90% credible intervals for coefficients in the alliance-level regression. The bottom panel plots the estimated effect of participation in an alliance with average capability on growth in military spending for a deep and shallow treaty, as well as the difference between deep and shallow alliances. In both panels, points mark the posterior mean, and the bars encapsulate the width of the 90% credible interval.

non-major power military spending growth in deep and shallow alliances, all else equal.

To assess Hypotheses 1 and 2, I examine patterns in the alliance participation parameters across the range of treaty depth. Each λ measures the impact of treaty participation, so if treaty depth has a large influence on alliance participation, it will appear in the λ estimates. On average, participation in deep alliances should have a negative effect on members' percent changes in military spending and shallow alliances should have a positive effect. If this is true, there will also be a negative trend in the expected value of λ as treaty depth increases.

The preponderance of evidence in Figure 3 matches the predictions of Hypotheses 1 and 2. The top panel of Figure 3 plots the expected value of the alliance participation parameter across the range of treaty depth. As expected, shallow treaties often have positive λ values,[27] which corresponds to Hypothesis 1. Most of the deepest treaties have a negative λ, which matches Hypothesis 2. Because other treaty characteristics and chance also influence the λ estimates, there is substantial variation in how alliance participation impacts non-major power military spending.

I then used the λ posteriors to estimate the impact of alliance participation on state-year military spending growth. To assess the annual impact of alliances on their members, I multiplied the alliance membership matrix $\mathbf{Z}$ by the λ parameters to generate 9,124 non-zero state-alliance-year predictions.[28] The scatter plot in the bottom panel of Figure 3 shows the distribution of predicted changes in military spending from alliance participation. Each point marks the mean estimated effect of an alliance on growth in military spending for that state-year. To avoid overplotting the 9,124 point estimates, I combined them into hexagons. Darker hexagons mark areas with more points. As Hypothesis 1 predicts, participation in shallow alliances regularly increases military spending. Many alliances with shallow depth have little effect on military spending, however. As treaty depth increases, alliance participation is more likely to reduce military spending growth.[29]

[27] All the negative λ estimates in alliances with minimal depth are treaties between the Soviet Union and Eastern European states during the Cold War.

[28] These estimates hold all state-level variables constant.

[29] Wartime alliances are the main exception to this trend.

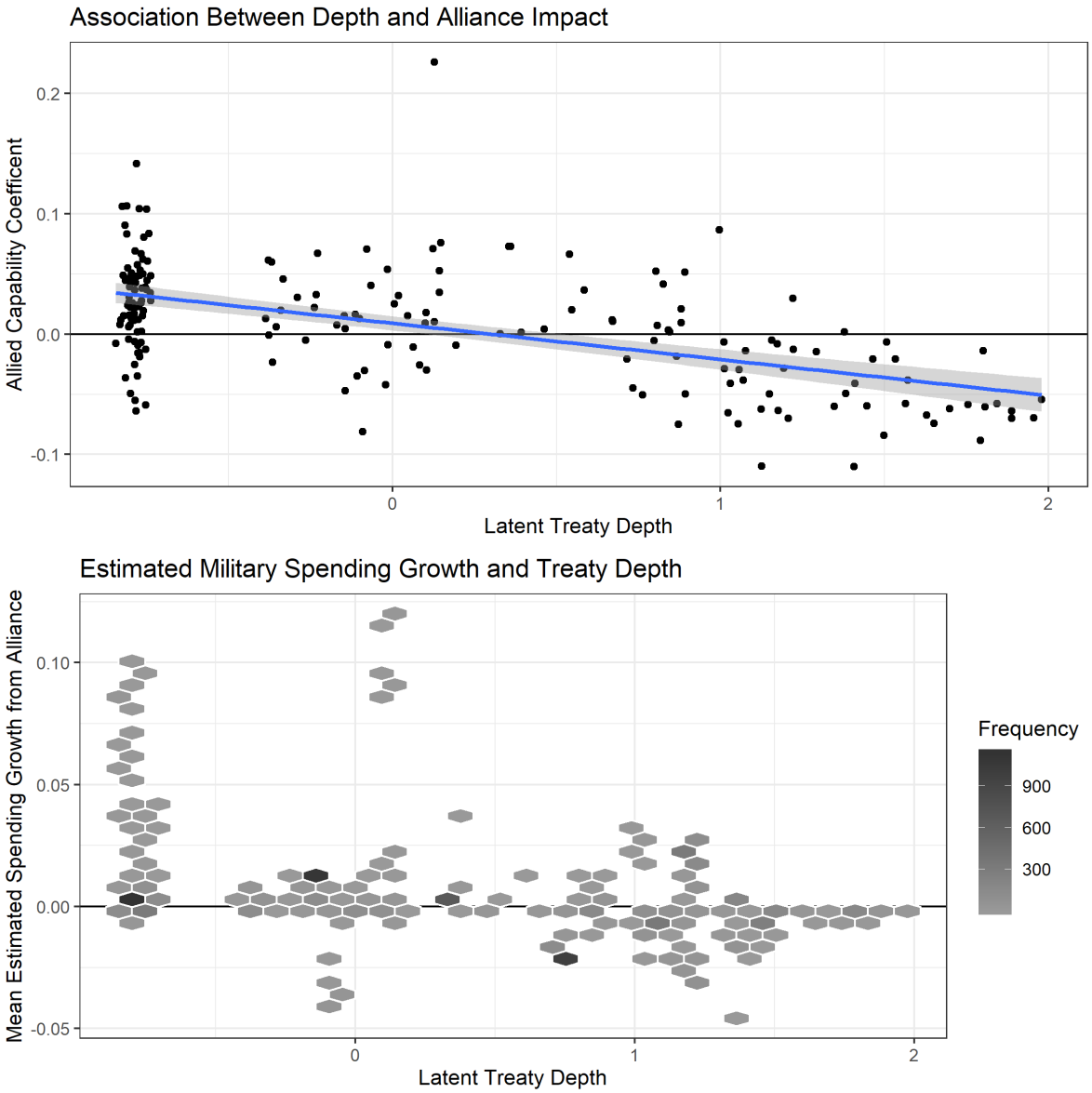

Figure 3: Summary of the predicted effect of alliance participation on growth in military spending across the observed values of treaty depth for 190 alliances from 1816 to 2007. The top panel plots the mean of each λ parameter by treaty depth. The bottom panel plots 9,128 state-alliance-year estimates of how participation in individual alliances affects growth in military spending, which are the product of the λ for the alliance and allied capability value for that state-year observation. Darker shading indicates more data points in the hexagon.

Some of the deepest alliances have a negligible effect on military spending despite negative λ values due to limited capability.

The appendix contains a series of robustness checks for this specification. First, I fit models with the three-level ordinal measure of (Leeds and Anac, 2005), and find that their military institutionalization measure also reduces the impact of alliance participation on non-major power military spending. I also fit a model that shifted the distribution of latent depth to include only zero and positive values. Last, I include single level panel regressions that average or summarize states' alliance portfolio in several ways.

In summary, I find that treaty depth modifies the impact of alliance participation on military spending. Participating in deep alliances often reduces military spending, while being part of a shallow alliance often increases spending. I now discuss the implications of the argument and results.

6 Discussion and Conclusion

This paper addresses a longstanding debate over whether alliance participation increases or decreases military spending. Claims alliance participation only increases or decreases military spending are incomplete. My argument shows how treaty depth modifies the impact of alliance participation on military spending, which builds on other conditional arguments (DiGiuseppe and Poast, 2016). I show that whether alliance participation increases or decreases military spending depends on treaty depth. Compared to no alliance at all, joining a shallow treaty usually increases military expenditures, while participation in a deep alliance often lowers defense spending.

There are two other noteworthy findings. First, many alliances increase non-major power military spending, which cuts against expectations that non-major powers are inveterate free-riders. To gain security from alliance participation, non-major powers sometimes need to increase their defense budget (Horowitz, Poast and Stam, 2017). Second, the military spending growth predictions

in Figure 3 suggest that many alliances have little effect on military spending, which is puzzling given widespread expectations that alliances and military spending are related. These negligible effects could reflect offsetting effects from different parts of alliance treaty design and membership, or concerns about the credibility of many alliance treaties.

Given my departure from previous research designs, how should we compare these results to prior evidence on alliance participation and military spending? Connecting my results with earlier evidence requires renewed attention to specific and general research designs. Recall that general studies compare states in an alliance to those without one in a global sample and specific studies estimate responses to allied military spending in a few alliances. The results encompass specific and general research designs by using allied capability to measure alliance participation. Each alliance participation parameter includes the effect of joining an alliance and changes in allied capability. I then use the alliance-level regression to understand how the impact of alliance participation varies with treaty design and membership.

Although my research design advances the debate on alliance participation and military spending, it has two limitations. First, my findings only address formal treaty depth. The measure of treaty depth only includes formal promises, in part because informal depth is harder to observe. As a result, my test of alliance depth may be conservative— it does not capture phenomena that should have a similar effect. It could instead overstate the findings if formal depth is not implemented, however. Strategic alliance design is the second possible weakness of the test. Domestic politics can affect alliance obligations, for example (Davis, 2004; Chiba, Johnson and Leeds, 2015). To address this issue, I controlled for correlates of alliance participation and treaty depth at each level of the model, with a particular focus on factors like democracy, alliance size, external threat, and other sources of credibility. At the state level, I adjusted for threat, economic growth, and regime type, all of which are possible correlates of treaty depth and growth in military spending. Even with this effort, selection into different alliances could still produce unobserved differences between alliances I did not adjust for.

Despite these limitations, the argument and results generate valuable insights about alliance participation and military spending. I explain when alliance participation is associated with increases or decreases in military spending among non-major powers, which addresses a debate between contradictory views of alliances. I provide evidence that how alliance participation impacts military spending depends on state capability and alliance treaty depth using a new measure of alliance treaty depth and a multilevel model.

There are several implications for scholarship. First, my argument and findings reinforce the importance of accounting for heterogeneity among alliances and institutional design more generally. Second, my research design could apply to other international institutions where institutional design shapes the consequences of participation. Besides these general implications, the results raise interesting questions for future research. For example, why do many alliances have a negligible effect on non-major power military spending? Because arms and alliances both provide security, existing arguments expect they are connected, but I find that many alliances have little substantive effect. The domestic economic and political consequences of alliance participation also remain relatively unexplored. Research on this topic could explore how changes in allied military spending affect public support for alliance participation and the economic consequences of changes in military spending.

Besides their scholarly value, the argument and evidence help inform policy debates about military spending. My argument claims that reassurance from deep alliances leads to lower defense spending. States can use deep cooperation to increase alliance credibility, but allied military spending may fall in response. Therefore, there is a tradeoff between treaty credibility and allied military spending.

The United States is currently wrestling with the credibility-military spending tradeoff. Washington has often decried allies who provide too little for their own defense (Lanoszka, 2015). But allies are able to maintain low military spending partly because the United States makes deep commitments. Reducing the depth of US alliances could generate credibility problems, however. Low

allied defense spending may be the price of credible commitments. Therefore, this paper is not an unconditional call to reduce the depth of US alliances. Adjusting existing treaties may be more difficult than designing new alliances and could have other ramifications. The full consequences of shifting treaty depth require additional scrutiny.

References

Allen, Michael A and Matthew DiGiuseppe. 2013. "Tightening the Belt: Sovereign Debt and Alliance Formation." *International Studies Quarterly* 57(4):647–659.

Altfield, Michael F. 1984. "The Decision to Ally: A Theory and Test." *Western Political Quarterly* 37(4):523–44.

Barnett, Michael N and Jack S Levy. 1991. "Domestic Sources of Alliances and Alignments: The Case of Egypt, 1962-73." *International Organization* 45(3):369–395.

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

Benson, Brett V and Joshua D Clinton. 2016. "Assessing the Variation of Formal Military Alliances." *Journal of Conflict Resolution* 60(5):866–898.

Berkemeier, Molly and Matthew Fuhrmann. 2018. "Reassessing the fulfillment of alliance commitments in war." *Research & Politics* April-June:1–5.

Bolt, J, R Inklaar, H de Jong and JL van Zanden. 2018. "Maddison Project Database, Version 2018." *Rebasing 'Maddison': new income comparisons and the shape of long-run economic development* .

Brands, Hal and Peter D. Feaver. 2017. "What are America's Alliances Good for?" *Parameters* 47(2):15–30.

Brooks, Stephen G, G John Ikenberry and William C Wohlforth. 2013. "Don't come home, America: the case against retrenchment." *International Security* 37(3):7–51.

Carpenter, Bob, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li and Allen Riddell. 2016. "Stan: A probabilistic programming language." *Journal of Statistical Software* 20:1–37.

Chen, Baizhu, Yi Feng and Cyrus Masroori. 1996. "Collective action in the Middle East? A study of free-ride in defense spending." *Journal of Peace Research* 33(3):323–339.

Chiba, Daina, Jesse C Johnson and Brett Ashley Leeds. 2015. "Careful Commitments: Democratic States and Alliance Design." *The Journal of Politics* 77(4):968–982.

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2):355–370.

Conybeare, John AC. 1994. "Arms versus Alliances: The Capital Structure of Military Enterprise." *Journal of Conflict Resolution* 38(2):215–235.

Conybeare, John AC and Todd Sandler. 1990. "The Triple Entente and the Triple Alliance 1880–1914: A Collective Goods Approach." *American Political Science Review* 84(4):1197–1206.

Crescenzi, Mark JC, Jacob D Kathman, Katja B Kleinberg and Reed M Wood. 2012. "Reliability, Reputation, and Alliance Formation." *International Studies Quarterly* 56(2):259–274.

Davis, Christina L. 2004. "International Institutions and Issue Linkage: Building Support for Agricultural Trade Liberalization." *American Political Science Review* 98(1):153–169.

Diehl, Paul F. 1994. "Substitutes or Complements?: The Effects of Alliances on Military Spending in Major Power Rivalries." *International Interactions* 19(3):159–176.

DiGiuseppe, Matthew and Paul Poast. 2016. "Arms versus Democratic Allies." *British Journal of Political Science* pp. 1–23.

Fariss, Christopher J. 2014. "Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability." *American Political Science Review* 108(2):297–318.

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

Fearon, James D. 1998. "Bargaining, enforcement, and international cooperation." *International Organization* 52(2):269–305.

Fearon, James D. 2018. "Cooperation, Conflict, and the Costs of Anarchy." *International Organization* pp. 1–37.

Fordham, Benjamin. 1998. *Building the Cold War Consensus: The Political Economy of US National Security Policy, 1949-51*. Ann Arbor: University of Michigan Press.

Fordham, Benjamin O. 2010. "Trade and asymmetric alliances." *Journal of Peace Research* 47(6):685–696.

Fordham, Benjamin and Paul Poast. 2014. "All Alliances Are Multilateral Rethinking Alliance Formation." *Journal of Conflict Resolution* 58(27):1–26.

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):919–935.

Gavin, Francis J. 2015. "Strategies of inhibition: US grand strategy, the nuclear revolution, and nonproliferation." *International Security* 40(1):9–46.

Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Vol. 1 Cambridge University Press New York, NY, USA.

George, Justin and Todd Sandler. 2017. "Demand for military spending in NATO, 1968–2015: A spatial panel approach." *European Journal of Political Economy* 53:222–236.

Gibler, Douglas M. 2008. "The Costs of Reneging: Reputation and Alliance Formation." *Journal of Conflict Resolution* 52(3):426–454.

Gibler, Douglas M, Steven V Miller and Erin K Little. 2016. "An Analysis of the Militarized Interstate Dispute (MID) Dataset, 1816–2001." *International Studies Quarterly* 60(4):719–730.

Goldsmith, Benjamin E. 2003. "Bearing the Defense Burden, 1886-1989: Why Spend More?" *Journal of Conflict Resolution* 47(5):551–573.

Horowitz, Michael C, Paul Poast and Allan C Stam. 2017. "Domestic Signaling of Commitment Credibility: Military Recruitment and Alliance Formation." *Journal of Conflict Resolution* 61(8):1682–1710.

Johnson, Jesse C., Brett Ashley Leeds and Ahra Wu. 2015. "Capability, Credibility, and Extended General Deterrence." *International Interactions* 41(2):309–336.

Kimball, Anessa L. 2010. "Political Survival, Policy Distribution, and Alliance Formation." *Journal of Peace Research* 47(4):407–419.

Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55(04):761–799.

Lake, David A. 1996. "Anarchy, Hierarchy and the Variety of International Relations." *International Organization* 50(1):1–33.

Lanoszka, Alexander. 2015. "Do Allies Really Free Ride?" *Survival* 57(3):133–152.

Leeds, Brett Ashley. 2003*a*. "Alliance Reliability in Times of War: Explaining State Decisions to Violate Treaties." *International Organization* 57(4):801–827.

Leeds, Brett Ashley. 2003*b*. "Do Alliances Deter Aggression? The Influence of Military Alliances on the Initiation of Militarized Interstate Disputes." *American Journal of Political Science* 47(3):427–439.

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

Leeds, Brett Ashley and Sezi Anac. 2005. "Alliance Institutionalization and Alliance Performance." *International Interactions* 31(3):183–202.

Leeds, Brett, Jeffrey Ritter, Sara Mitchell and Andrew Long. 2002. "Alliance Treaty Obligations and Provisions, 1815-1944." *International Interactions* 28(3):237–260.

Long, Andrew G. 2003. "Defense Pacts and International Trade." *Journal of Peace Research* 40(5):537–552.

Long, Andrew G and Brett Ashley Leeds. 2006. "Trading for Security: Military Alliances and Economic Agreements." *Journal of Peace Research* 43(4):433–451.

Mattes, Michaela. 2012. "Reputation, Symmetry, and Alliance Design." *International Organization* 66(4):679–707.

McElreath, Richard. 2016. *Statistical Rethinking: A Bayesian course with examples in R and Stan*. Vol. 122 CRC Press.

Morgan, T Clifton and Glenn Palmer. 2006. *A Theory of Foreign Policy*. Princeton: Princeton University Press.

Morgenthau, Hans. 1948. *Politics Among Nations: The Struggle for Power and Peace*. New York: Alfred A. Knopf.

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. 1993. "Arms versus allies: trade-offs in the search for security." *International Organization* 47(2):207–233.

Morrow, James D. 1994. "Alliances, Credibility, and Peacetime Costs." *Journal of Conflict Resolution* 38(2):270–297.

Morrow, James D. 2000. "Alliances: Why Write Them Down?" *Annual Review of Political Science* 3:63–83.

Most, Benjamin A and Harvey Starr. 1989. *Inquiry, Logic and International Politics*. Columbia: University of South Carolina Press.

Most, Benjamin A. and Randolph M. Siverson. 1987. Substituting Arms and Alliances 1870–1914: an Exploration in Comparative Foreign Policy. In *New Directions in the Study of Foreign Policy*, ed. Charles F Hermann, Charles W. Kegley Jr and James N. Rosenau. Boston: Allen & Unwin pp. 131–160.

Murray, Jared S, David B Dunson, Lawrence Carin and Joseph E Lucas. 2013. "Bayesian Gaussian Copula Factor Models for Mixed Data." *Journal of the American Statistical Association* 108(502):656–665.

Niou, Emerson MS and Sean M Zeigler. 2019. "External Threat, Internal Rivalry, and Alliance Formation." *The Journal of Politics* 81(2):571–584.

Olson, Mancur and Richard Zeckhauser. 1966. "An Economic Theory of Alliances." *The Review of Economics and Statistics* 48(3):266–279.

Oneal, John R. 1990. "The theory of collective action and burden sharing in NATO." *International Organization* 44(3):379–402.

Palmer, Glenn. 1990. "Alliance Politics and Issue Areas: Determinants of Defense Spending." *American Journal of Political Science* 34(1):190–211.

Pamp, Oliver, Florian Dendorfer and Paul W. Thurner. 2018. "Arm your friends and save on defense? The impact of arms exports on military expenditures." *Public Choice* 177(1):165–87.

Plümper, Thomas and Eric Neumayer. 2015. "Free-riding in alliances: Testing an old theory with a new method." *Conflict Management and Peace Science* 32(3):247–268.

Poast, P. 2012. "Does Issue Linkage Work? Evidence from European Alliance Negotiations, 1860 to 1945." *International Organization* 66(1):277–310.

Poast, Paul. 2013. "Can Issue Linkage Improve Treaty Credibility? Buffer State Alliances as a "Hard Case"." *Journal of Conflict Resolution* 57(5):739–764.

Poast, Paul. 2019. *Arguing about Alliances: The Art of Agreement in Military-Pact Negotiations*. Cornell University Press.

Posen, Barry R. 2014. *Restraint: A New Foundation for U.S. Grand Strategy*. Cornell University Press.

Preble, Christopher A. 2009. *The Power Problem: How American Military Dominance Makes Us Less Safe, Less Prosperous, and Less Free*. Ithaca: Cornell University Press.

Quinn, Kevin M. 2004. "Bayesian Factor Analysis for Mixed Ordinal and Continuous Responses." *Political Analysis* 12(4):338–353.

Quiroz Flores, Alejandro. 2011. "Alliances as Contiguity in Spatial Models of Military Expenditures." *Conflict Management and Peace Science* 28(4):402–418.

Reiter, Dan, Allan C. Stam and Michael C. Horowitz. 2016. "A Revised Look at Interstate Wars, 1816–2007." *Journal of Conflict Resolution* 60(5):956–976.

Sandler, Todd and John F Forbes. 1980. "Burden sharing, strategy, and the design of NATO." *Economic Inquiry* 18(3):425–444.

Sandler, Todd and Keith Hartley. 2001. "Economics of alliances: The lessons for collective action." *Journal of Economic Literature* 39(3):869–896.

Sarkees, Meredith Reid and Frank Whelon Wayman. 2010. *Resort to War: A Data Guide to Interstate, Extra-state, Intra-state, and Non-state Wars, 1816-2007*. Washington, DC: CQ Press.

Schuessler, John M and Joshua R Shifrinson. 2019. "The Shadow of Exit from NATO." *Strategic Studies Quarterly* 13(3):38–51.

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

Simon, Jeffrey. 1985. *Warsaw Pact Forces: Problems of Command and Control*. Boulder, CO: Westview.

Singer, J David. 1988. "Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985." *International Interactions* 14(2):115–132.

Snyder, Glenn H. 1997. *Alliance Politics*. Ithaca: Cornell University Press.

Sorokin, Gerald L. 1994. "Arms, Alliances, and Security Tradeoffs in Enduring Rivalries." *International Studies Quarterly* 38(3):421–446.

Steenbergen, Marco R and Bradford S Jones. 2002. "Modeling Multilevel Data Structures." *American Journal of Political Science* pp. 218–237.

Thompson, William R. and David R. Dreyer. 2012. *Handbook of International Rivalries 1494–2010*. Washington, DC: CQ Press.

Trager, Robert F. 2010. "Diplomatic Calculus in Anarchy: How Communication Matters." *American Political Science Review* 104(02):347–368.

Walt, Stephen M. 1990. *The Origins of Alliance*. Cornell University Press.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

Yarhi-Milo, Keren, Alexander Lanoszka and Zack Cooper. 2016. "To Arm or to Ally? The Patron's Dilemma and the Strategic Logic of Arms Transfers and Alliances." *International Security* 41(2):90–139.

Zielinski, Rosella Cappella, Benjamin O Fordham and Kaija E Schilde. 2017. "What goes up, must come down? The asymmetric effects of economic growth and international threat on military spending." *Journal of Peace Research* 54(6):791–805.