

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. Treaty depth reveals a tradeoff between reassurance and greater allied military spending. 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 measure of alliance treaty depth and employing it in a multilevel model. 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 alliance participation will reduce military spending e.g., (Morrow, 1993; Conybeare, 1994). The other predicts alliance participants will spend more on defense e.g., (Diehl, 1994; Morgan and Palmer, 2006). 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.

Debate between the two perspectives largely ignores heterogeneity among alliances,¹ which is essential to alliance politics scholarship (Morrow, 1991; Leeds, 2003*b*; Leeds and Anac, 2005; Fordham, 2010; Mattes, 2012; Benson, 2012; Poast, 2013; Johnson, Leeds and Wu, 2015). Given differences in alliance design and membership, alliance participation could plausibly increase or decrease defense expenditures. In this paper, I use variation in alliance design and membership to predict how alliance participation affects military spending.

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 defense coordination and cooperation among alliance members.

Deep and shallow alliances have different effects on non-major power military spending. Par-

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

ticipation 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, as states face a higher risk of abandonment and allies use credible threats of abandonment to demand investment in military capability. The argument focuses on non-major powers because these states clearly show the tradeoff between reassurance and military spending.²

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 percentage changes in military spending. Allied capability is a useful proxy for alliance participation because it combines the effects of joining an alliance and changing allied capability during treaty membership, both of which shape the impact of alliances on military spending. Multilevel modeling matches my conditional argument and estimates heterogeneous effects of alliance participation that are specific to 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, shallow alliances increase 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). Advocates of continued deep engagement argue that the benefits of alliances exceed the costs and believe that the problem of low allied military spending is overstated (Brands and Feaver, 2017). If efforts to reassure allies through treaty depth lower allied military spending, it will be difficult to compromise between

²I explore the process connecting alliance participation and military spending for major powers in a separate paper.

these two positions.

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. Each predicts a different average effect of alliance participation by emphasizing one aspect of alliance politics.

Substitution and public goods arguments predict that alliances reduce defense spending as states can replace security from military spending with security from alliances. Olson and Zeckhauser (1966) argue 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.³ 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 reallocate military spending to other goods. Both the substitution 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).

³Sandler 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.

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.⁴

2.1 Mixed Evidence

Debate between the contradictory views of alliances could be settled by a consistent set of results, but mixed findings reinforce the theoretical division. Some studies find a positive association between alliance participation and military spending. Others find a negative relationship.⁵

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 an alliance to those without. Table 1 summarizes previous results from general models of alliance participation and military spending. There is one negative, three positive and two null estimates of the correlation between alliance participation and spending.

Unlike general studies, specific research designs estimate how states respond to military spending by a few key allies. If states reduce their own military spending as allied spending rises, specific

⁴Senese 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 correlations between alliances and military spending.

⁵Because 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.

	Decrease	Increase	Null
Most and Siverson (1987)			X
Conybeare (1994)	X		
Diehl (1994)		X	
Goldsmith (2003)			X
Morgan and Palmer (2006)		X	
Quiroz Flores (2011)		X	

Table 1: General findings of the association between alliance participation and military spending.

studies conclude alliances decrease military spending. Most evidence of reduced military spending by alliance members comes from alliance-specific designs (Barnett and Levy, 1991; Morrow, 1993; Sorokin, 1994; Plümper and Neumayer, 2015; George and Sandler, 2017). Other specific studies find states increase their military spending as allied spending rises, however (Conybeare and Sandler, 1990; Chen, Feng and Masroori, 1996).

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). Conflict (Leeds, 2003*b*; Benson, 2012) and trade (Long, 2003; Long and Leeds, 2006) are two domains where alliance design shapes the consequences of treaty participation. Building on this work, I focus on one 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

Deep military cooperation in an alliance treaty modifies the impact of alliance participation on non-major power military spending. Given greater treaty credibility in a deep alliance, joining these alliances often reduces non-major power defense spending. Conversely, fear of abandonment

in shallow alliances means non-major power participants in these alliances often increase defense spending.

I start the argument by describing a general framework for non-major power alliances. Then I discuss the role of deep formal military cooperation. Last, I show 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 to draw on allied capability (Fordham and Poast, 2014) and thereby advance their foreign policy aims.

Alliance participation is inseparable from allied capability. The presence of an alliance treaty formalizes when a state can expect military intervention (Morrow, 2000), but the treaty paper alone lacks value. An alliance is only as useful if partners provide meaningful military capability. Greater allied capability increases the value of an alliance, all else equal (Johnson, Leeds and Wu, 2015).⁶

Alliance treaties and the capability they aggregate can support many foreign policy aims. This flexibility often facilitates exchange between alliance participants with different goals. 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,⁷ but the divergent motives of major and non-major powers reflect general tendencies in alliance politics. Major powers often use alliances to address the global

⁶A binary conceptualization of alliance participation assumes all alliances are equally valuable.

⁷130 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.

balance of power and issues of influence. Smaller non-major powers tend to be less ambitious in their alliances and emphasize immediate security.

As a result, there are distinct processes behind non-major and major power alliance participation and the consequences of their alliances. MAJOR and non-major powers respond to allied capability and treaty design in different ways. Here I focus on non-major powers to ensure theoretical and empirical parsimony. 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 assumption by showing that realizing alliance participation often requires additional military spending by non-major powers.

As 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. Non-major powers can gain increased security with less military spending, even in the face of greater threats, if they form deep alliances. Conversely, security gains in shallow alliances may depend on non-major powers maintaining or even increasing their military spending to offset fears of abandonment.

Given non-major power security concerns, allied support and cooperation is extremely valuable. 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, 1998; Koremenos, Lipson and Snidal, 2001).

Non-major powers are especially concerned with abandonment, which is the most common form of opportunism in alliances. States often violate their alliance obligations and abandon their partners (Berkemeier and Fuhrmann, 2018). For security-focused non-major powers, abandonment is a major threat. The risk of abandonment means alliance members must maintain the perceived credibility of their commitment.

Abandonment and reduced military spending are related because greater treaty credibility allows states to lower military expenditures.⁸ Though states can add to the collective military capability of an alliance through their military spending, they can also reduce defense spending and rely on their partners (Olson and Zeckhauser, 1966; Morrow, 1993; Conybeare, 1994; Sandler and Hartley, 2001). Such reductions in defense spending by non-major powers are not always problematic, especially if the alliance reflects an exchange. But there are circumstances where states would prefer their allies spend more on the military.

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. To give three examples, depth, unconditional military support (Benson, 2012; Chiba, Johnson and Leeds, 2015) and issue linkages (Long and Leeds, 2006; Poast, 2012, 2013) are also sources of credibility.⁹ 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 defense spending.

Moreover, treaty depth is a common policy choice.¹⁰ 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, so my argument clarifies how a common source of credibility shapes alliance politics. Because deep alliances contain costly commitments, depth reduces the perceived risk of abandonment. Costly promises allow alliance members and potential adversaries to infer the credibility of the alliance (Leeds, 2003*b*; Fuhrmann and Sechser, 2014).

⁸The public goods model of alliances calls reduced military spending free-riding. I do not use this language because it implies reduced defense spending is problematic. But exchange in asymmetric alliances, or the reduction of threat in general, may make reduced spending acceptable.

⁹Though the argument emphasizes depth, the research design accounts for multiple sources of alliance credibility.

¹⁰In a separate 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, unconditional military support, and the number of treaty members all increase treaty depth, while the presence of issue linkages reduces depth.

Greater alliance depth and credibility does not lead states to increase military spending, however. Enforcing higher defense effort 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).

Alliance members must have leverage if they want to ensure increases in allied defense spending. Leverage either comes from a credible threat to abandon states who spend too little or control over allied policies. Policy control of allied spending decisions holds when the alliance reflects hierarchical relationships like an informal empire (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 can reduce defense spending.

In less credible treaties, such as alliances between erstwhile rivals (Niou and Zeigler, 2019), failure to spend on the military 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 a greater risk of abandonment. Being left with foreign obligations and without allied support is a concern for security-conscious non-major powers. As treaty credibility falls, the foreign policy benefits of alliance participation become contingent on military spending because alliance members hedge against abandonment and partners can use that concern as leverage. Thus, alliances with less credibility will tend 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 to check reductions in military expenditures. Credibility from treaty depth also promotes efficiency gains from coordinated defense effort.

3.2 Alliance Treaty 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.¹¹

One example of a deep alliance is a 1948 defense pact between the United Kingdom and Jordan, which includes unconditional military support, basing rights, military aid, official military contact, and an Anglo-Transjordan Joint Defense Board. This is a deeper alliance than a 1912 treaty between Greece and Bulgaria which only commits to mutual defense and consultation if either state is attacked by Turkey. Increasing military coordination adds ties between alliance members beyond a promise of military support.

Alliance 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. Depth is especially useful because alliance members face a

¹¹Although depth can have monetary costs, reductions in military spending due to greater treaty credibility will often outweigh those costs. Such an effect is plausible, as contributions to alliances are a small part of defense spending, and non-major powers rarely use basing rights in deep alliances to deploy their troops abroad.

time inconsistency problem. Alliance treaty fulfillment depends on shared foreign policy interests (Morrow, 2000; Leeds, 2003a), 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.¹²

Credibility in deep alliances may also facilitate more efficient defense spending. States often use alliances to formulate joint war plans (Poast, 2019), which allows alliance members to provide specific capabilities. Specialization means members of deep alliances spend less on the military but retain foreign policy benefits. Alliance credibility and efficiency gains are inseparable. States will only specialize if they believe the alliance is reliable (Leeds, 2003a).

On the other hand, shallow alliances are less credible. These treaties 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 increased 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 see coordination and specialization due to the fear of abandonment. As a result, participation in shallow alliances often increases military spending.¹³

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 they participated in occupying the Ruhr. A more limited treaty

¹²Conversely, eliminating these transfers reduces the credibility of the whole alliance.

¹³One objection to this argument is that deep alliances are more valuable to members, which augments allies 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.

with only military support between France, Belgium, the United Kingdom, Italy and Germany (ATOPID 2130) increased Belgian spending, on the other hand.

These brief examples and the 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. If we think about depth as a continuous variable, 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 towards negative effects in the deepest 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.¹⁴

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.

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,¹⁵ a larger increase in spending relative to the previous year's defense budget imposes

¹⁴Hypothesis 3 follows from Hypotheses 1 and 2.

¹⁵Especially holding economic growth constant.

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

To understand the consequences of alliance participation for military spending, I focus on allied capability rather than a simple dichotomous indicator of participation. 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 the treaty creates about allied intervention. Therefore, allied capability is more meaningful than treaty presence and alliance participation affects military spending through joining an alliance as well as changing allied capability after the treaty forms. In previous scholarship, general research designs address the first path, while specific designs focus on the latter. 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 treaty depth from formal 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. I estimate the multilevel model in 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.¹⁶ Therefore, I use observed alliance treaty characteristics to infer depth, which could produce two types of measures. The first measure is an additive index of treaty depth, where treaties with multiple commitments have higher index values.

(Leeds and Anac, 2005) develop a similar ordinal measure by assigning alliances military institutionalization scores of zero, one or two based on the extent of investment in joint action required by the alliance treaty. The result is a coarse measure of alliance depth which understates variation in treaty depth. This measurement strategy imposes equal weights on different depth sources 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 add the two sources together if both are present. I assess these theoretical assumptions with a more flexible measurement strategy.

I employ latent variable modeling to create a continuous measure of treaty depth that makes more nuanced distinctions between alliances. The specific 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 variable adds to the alliance.

Measurement models have a rich history in political science (Clinton, Jackman and Rivers, 2004; Treier and Jackman, 2008; Fariss, 2014). 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. I emulate Benson and Clinton's approach, but use a different concept, sample of alliances and estimator. First, Benson and Clinton (2016) define depth as the costliness of the alliance in general, so they include measures of economic issue linkages and secrecy. My definition of depth only includes military cooperation, because I view issue linkages as a separate source of credibility. Given their broad definition of depth, Benson and Clinton also include neutrality pacts

¹⁶Of course, formal treaty obligations may not be fully implemented, but formal depth often leads to practical depth.

in their sample of alliances. I am only interested in offensive and defensive alliances, however. As for the estimator, in some latent variable models, the latent variables influence the form of the dependence structure and the form of the marginal distributions in the latent value estimates. I use a different estimator that does not have this limitation (Murray et al., 2013).

Due to the limits of ordinal measures and key conceptual differences with Benson and Clinton's latent measure, existing measures of treaty depth do not fit my purposes in this paper. See the appendix for more detailed comparisons and evidence that Leeds and Anac (2005)'s measure of institutionalization produces similar inferences. I create a new measure of treaty depth in offensive and defensive ATOP alliances using a semiparametric factor analysis.

I use a Bayesian Gaussian Copula Factor Model (Murray et al., 2013) to measure alliance treaty depth. 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.¹⁷ Beyond the semiparametric aspect, this measurement model is a standard ordinal factor analysis.

I estimated the measurement model using observed data from 289 alliances with offensive or defensive obligations in the alliance-level ATOP data (Leeds et al., 2002). 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.¹⁸ The argument suggests there is a single factor underlying variation in all eight indicators, so I fit

¹⁷Copulas are a distribution function on $[0, 1]^p$ where each univariate marginal distribution is uniform on $[0, 1]$.

¹⁸These are the variables Leeds and Anac (2005) use, with the addition of a companion military agreements dummy.

the model with one latent factor.

I used Parameter expanded Gibbs sampling, the default generalized double Pareto (GDP) prior, 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.

The posterior mean of the latent factor for each alliance is my summary of treaty depth, 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 the correlations between these observed variables and the latent factor. Policy coordination, integrated military command, and formal organizations are the three largest correlates of depth. The other five factors have roughly comparable associations with latent treaty depth.

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 deep military cooperation, 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, 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),

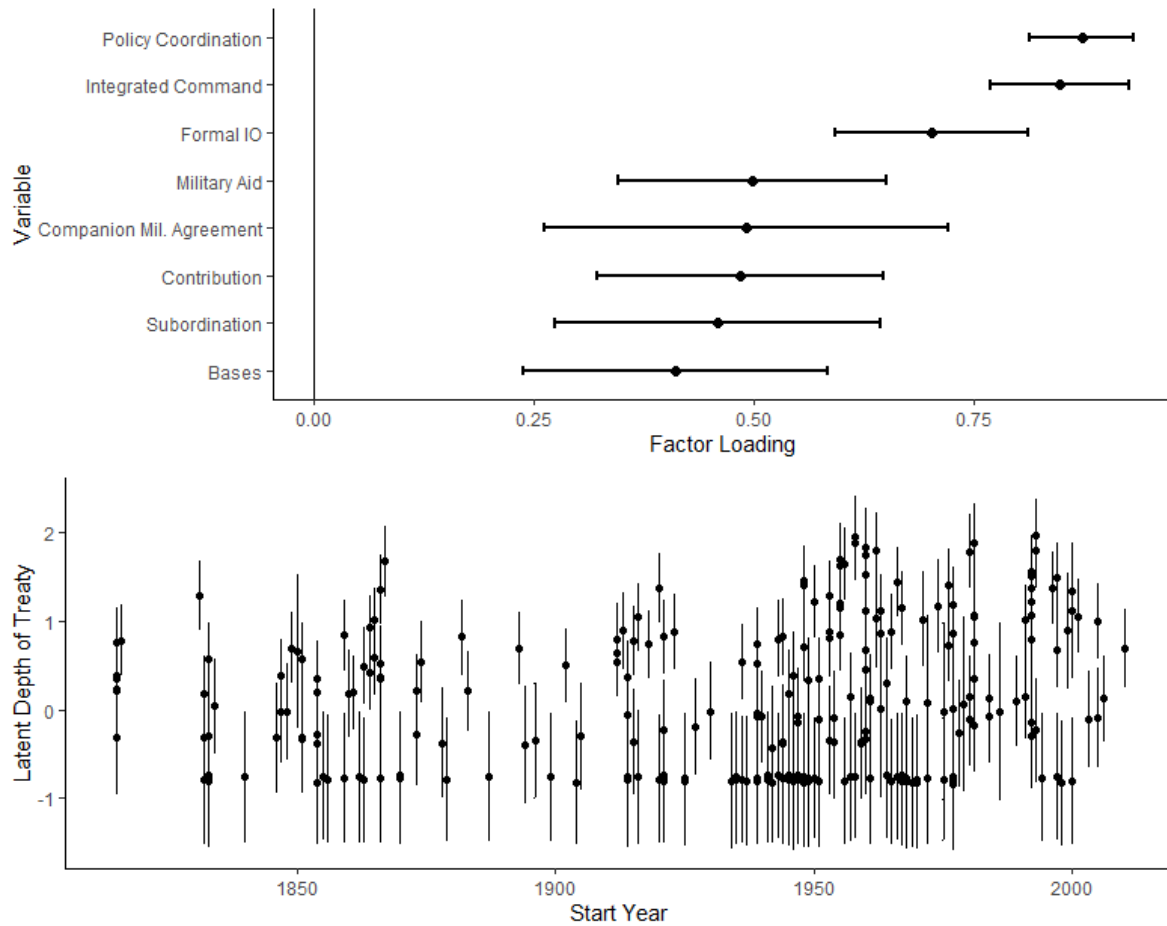


Figure 1: Summary of the latent measure of alliance treaty depth for 289 defensive or offensive alliances from 1816 to 2016. The top panel plots the factor loadings with 90% credible intervals. The bottom panel plots mean treaty depth (points) and the standard deviation (error bars) against the start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

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.

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.

My argument uses variation in treaty depth between alliances to explain percentage changes in military spending. 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. Because alliances and states are separate levels of analysis, I use a multilevel model to estimate the association between treaty depth and military spending. The next section summarizes this estimation strategy.

4.2 Multilevel Model

Multilevel modeling bridges levels of analysis (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. I 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.¹⁹ 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. This means the research design should compare the impact of different alliances on military spending. The multilevel model explicitly compares the impact of participation in deep and shallow alliances by estimating how the changes in treaty depth modify the consequences of alliance participation.

Standard research designs 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). Averaging or otherwise aggregating alliance characteristics at a different level of analysis changes the mean and variance of key independent variables, which then affects inferences. Multilevel modeling avoids this problem by retaining the structure of the data.

Multilevel modeling incorporates several key characteristics of alliance data. First, states can participate in more than one alliance and I expect that alliances have heterogeneous effects on military spending. The multilevel model estimates the specific impact of each alliance on members' military expenditures, which reveals differences between individual treaties. Aggregating multiple alliances into a few state level indicators will mask any heterogeneous effects of individual treaties.²⁰

Furthermore, multiple alliance characteristics modify the consequences of alliance participation. The multilevel model captures multiple sources of heterogeneity in how alliances impact military spending. In a panel estimator with state-level proxies for alliance characteristics, accounting for correlated aspects of alliance design is difficult. Treaty depth is correlated with other aspects of alliance membership and design, so this step is important.²¹ Panel estimates that account for one or two alliance characteristics can only do this by averaging different parts of a state's alliance portfo-

¹⁹See the appendix for results from several models with state-level indicators of alliance depth.

²⁰Partial pooling of the alliance-specific parameters generates reasonable estimates for each alliance.

²¹For example, I show in another paper that democratic alliance membership and treaty depth are correlated.

lio or subsetting 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. The next section details the model specification.

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:

$$y \sim student_t(\nu, \mu, \sigma) \quad (1)$$

y is the dependent variable— percentage changes in military spending. I model the outcome using a t-distribution with degrees of freedom ν to address heavy tails.²² 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.

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

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.

²²I estimate ν directly.

\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. If a state is not in the alliance, the corresponding cell of the matrix 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 in the alliance participation matrix 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 at once, so observations are not neatly nested. This specification allows each alliance to have a unique impact on military spending as 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 Z are allied spending, the λ parameters capture alliance members' response to allied capability. Each alliance has a unique λ . The λ parameters have 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 λ 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:

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

and

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

²³Normalization keeps the parameters on similar scales, which is important for modeling. I selected normalization theoretically and corroborated this choice by comparing models fit with different ways of expressing allied capability. See the appendix for details.

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 β s 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 the impact of participation in each alliance, permitting heterogeneous effects of individual treaties. The β parameters capture how alliance characteristics modify the impact of alliance participation on military spending. Again, using alliance characteristics to modify 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. 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 8,668 state-year observations and 192 alliances.

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

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

I used the Correlates of War Project's data on military spending to measure percentage changes in spending (Singer, 1988).²⁴ 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 benchmarked to previous spending levels. To facilitate model fitting, I apply the inverse hyperbolic sine transformation to this variable.²⁵

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. This variable enters the model in the alliance-level regression and I expect it will have a negative coefficient. I also include several 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) 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

²⁴Estimating the model on different military spending data produces similar results: see the appendix for details.

²⁵This 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.

alliance promises any kind of 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. With the full research design in hand, I now turn to results.

5 Results

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.²⁶

I use Bayesian modeling to estimate the association between treaty depth and percentage changes in military spending, so each coefficient has a posterior distribution— the likely values of the coefficient 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 3.²⁷ Figure 2 summarizes

²⁶See the appendix for details on convergence and other robustness checks.

²⁷I use 90% intervals because inferences about 95% intervals are sensitive to simulation variance in Bayesian anal-

the results, including the estimated substantive effect of a large increase in treaty depth.

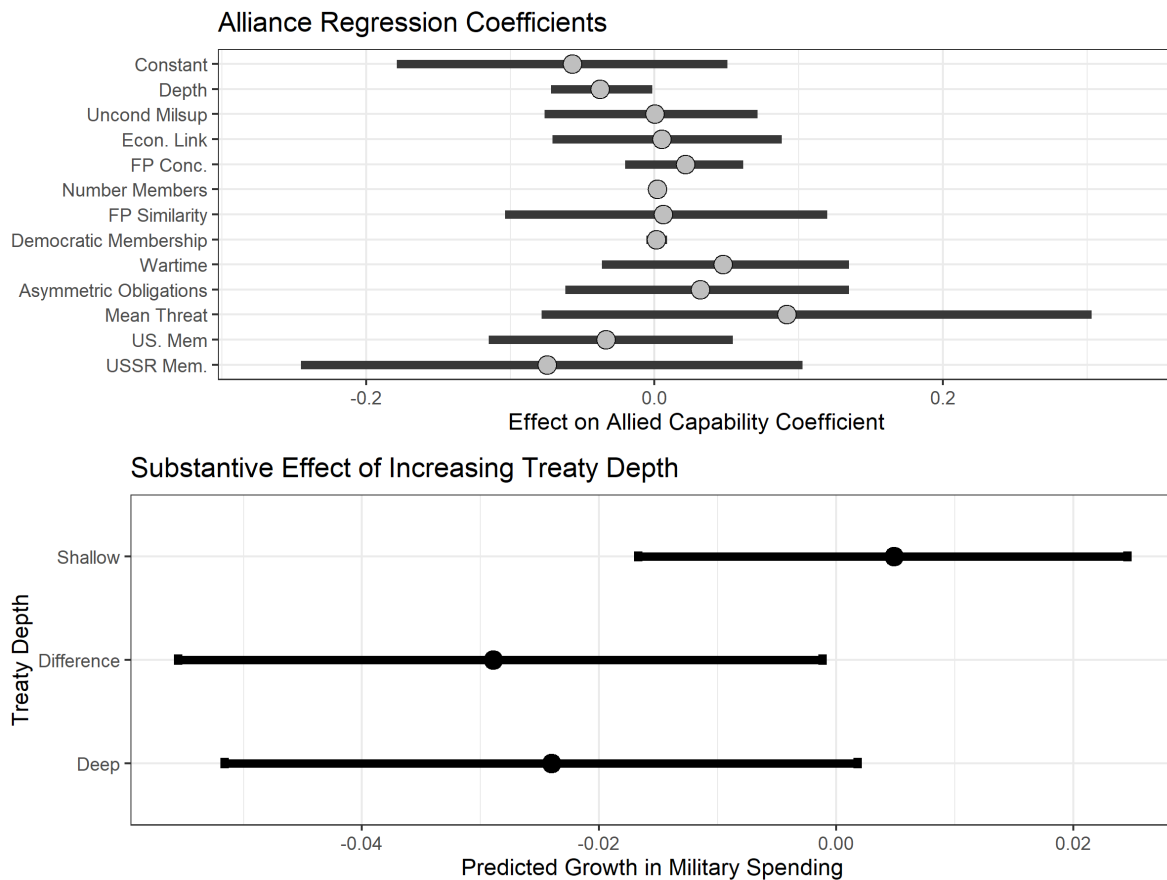


Figure 2: Summary of alliance-level regression results from the multilevel model. The top panel shows the 90% credible intervals to summarize the posterior densities of coefficients in the alliance-level regression. The bottom panel plots the estimated effect of participation in an alliance with average 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.

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. Also, the 90% credible interval for treaty depth does not include zero.

This alliance-level depth coefficient has a substantively important effect on growth in military

spending. I assess this substantive effect by simulating the effect of changing treaty depth from the minimum value of -0.8 to 1.5, which is in the fourth quartile. Holding other alliance covariates at their modes or medians, this increase in depth reduces a hypothetical λ by .08 in expectation. Shifting λ then produces a substantial difference in military spending growth, as shown in the bottom panel of Figure 2. That panel plots the 90% intervals for predicted growth in military spending in the shallow alliance, predicted spending growth in a 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. There is a meaningful difference between non-major power military spending growth in deep and shallow alliances, all else equal.

The substantive importance of treaty depth is also evident in patterns in the λ parameters, which broadly correspond with Hypotheses 1 and 2. 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. Therefore, there should be a negative trend in the expected value of λ as treaty depth increases.

The top panel of Figure 3 plots the expected value of λ across the range of treaty depth. As expected, shallow treaties often have positive λ values for non-major powers,²⁸ which corresponds to Hypothesis 1. Most of the deepest treaties have a negative λ , which matches the prediction of Hypothesis 2. Because other treaty characteristics and unmeasured factors also influence the λ estimates, there is tremendous variation in how alliance participation impacts non-major power military spending.

Because they reflect the impact of allied capability, the λ values are not direct predictions of

²⁸All the negative λ estimates in alliances with minimal depth are treaties between the Soviet Union and Eastern European states during the Cold War.

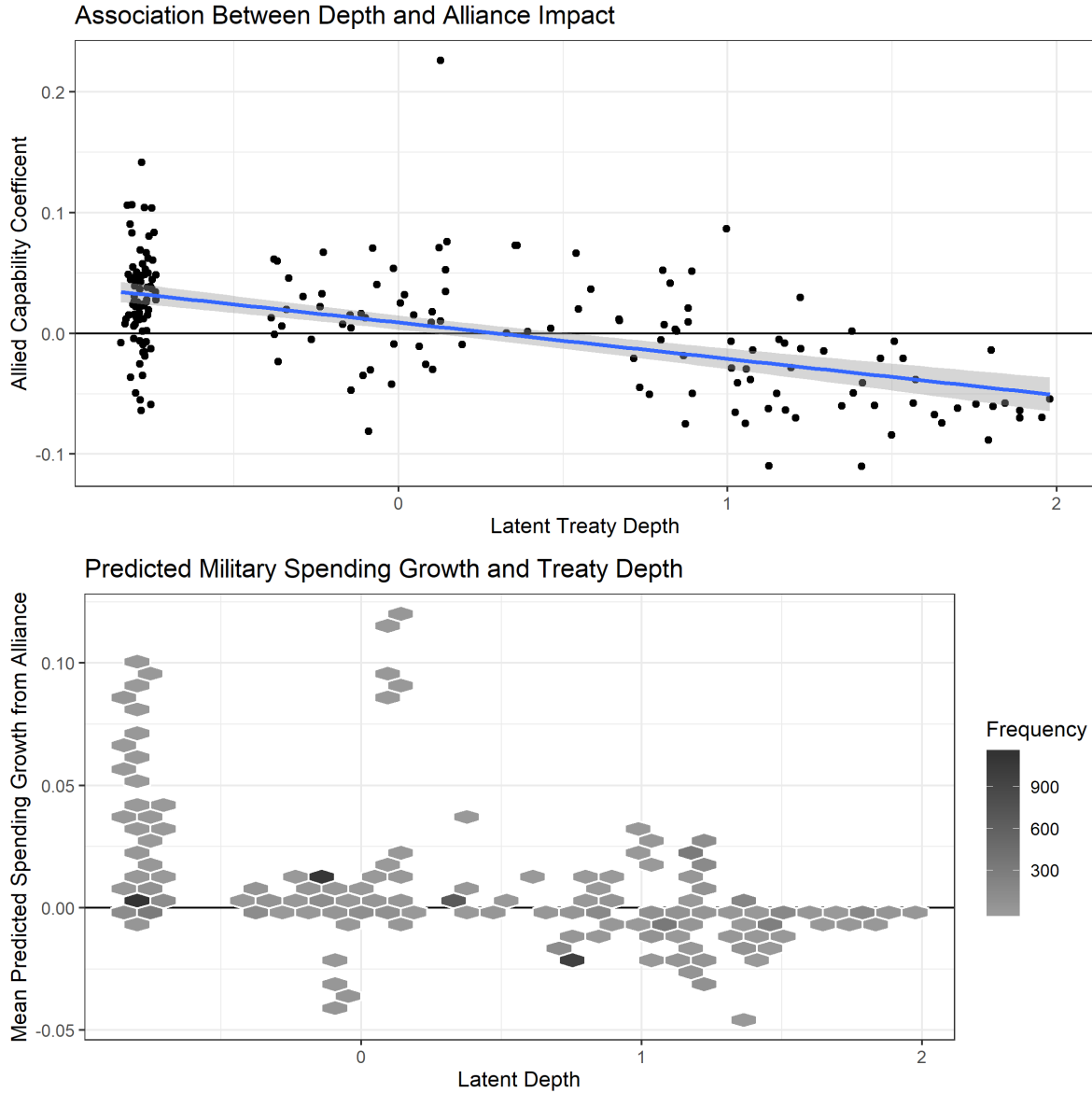


Figure 3: Summary of the predicted effect of alliance participation across the observed values of treaty depth for 192 alliances from 1816 to 2007. The top panel plots the mean of each λ parameter by treaty depth. The bottom panel plots 9,128 state-alliance-year predictions of growth in military spending from alliance participation, which combines the λ value for the alliance and allied capability value for that state-year observation. Darker colors indicate more data points in the particular hexagon.

how alliance participation affects military spending. Therefore, I multiplied the capability values in the alliance membership matrix by the λ parameters to generate 9,124 state-alliance-year predictions.²⁹ The bottom panel of Figure 3 shows the distribution of these predicted effects. This figure is a scatter plot, where each point marks the mean estimated effect of participation in a particular alliance on growth in military spending for that year. I combined the points into hexagons to address overplotting. Darker hexagons mark areas with more points.

These predictions show how each λ translates into military spending growth. As Hypothesis 1 predicts, participation in shallow alliances regularly increases military spending. Many state year alliances with shallow alliances see little effect on military spending, however. As treaty depth increases, more alliances reduce military spending growth among their members.³⁰ Some deep alliances have a negligible effect on military spending despite large λ values because they have comparatively little capability. The same pattern holds in the top and bottom panels of Figure 3, which corresponds to the expectations of Hypotheses 1 and 2.

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 it. This has important consequences for our view of alliance participation and military spending.

6 Discussion

My findings add to our understanding of alliance participation and military spending and address debates 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 and builds on other conditional arguments (DiGiuseppe and Poast, 2016).

²⁹These estimates hold all state-level covariates like threat and regime type constant.

³⁰Wartime alliances are the main exception to this trend.

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. The impact of alliance participation on non-major power military spending tends to decrease with treaty depth. There are many alliances that increase non-major power military spending, which cuts against popular expectations that non-major powers are likely to free-ride at every opportunity.

The military spending growth predictions in Figure 3 also raise a point existing literature and my own argument largely ignore— many alliances have little effect on military spending. Many of the predicted effects of alliance participation are quite close to zero, which implies that allied capability does very little for non-major power military spending.

How do the findings compare to prior evidence on alliance participation and military spending? Connecting my results with earlier evidence requires renewed attention to specific and general research designs. General studies compare states in an alliance to those without one. Specific studies estimate responses to allied military spending in a few treaties.

The results encompass specific and general studies through my use of allied capability to measure alliance participation and estimates from both levels of the multilevel model. My research design emulates specific studies by estimating the unique impact of participation in individual treaties. The alliance-level coefficients compare treaties to assess the general role of alliance characteristics.

There are a couple limitations to the above results. 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 may be also overstate the findings if formal depth is not implemented, however.

Strategic alliance design is another possible weakness of the test. Domestic politics can affect alliances (Davis, 2004; Chiba, Johnson and Leeds, 2015), for example. 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, 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 cannot adjust for.

Despite these limitations, the argument and results provide 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, addressing 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. The argument and findings have implications for scholars and policymakers.

7 Conclusion

Alliance participation does not uniformly increase or decrease military spending. Rather, participation in deep alliances often decreases military spending, while participation in shallow alliances often increases spending. This relationship reflects a tradeoff between reassurance and defense spending in alliances.

There are several implications of my findings. First, they reinforce the importance of accounting for heterogeneity among alliances. Alliances have heterogeneous effects on the risk of war, trade and military spending (Leeds, 2003*b*; Long and Leeds, 2006; Benson, 2012; DiGiuseppe and Poast, 2016).

The distributional consequences of changes in military spending within states are another salient implication. By altering military spending, treaty design affects the domestic political economy of alliance members. The economic consequences of alliance participation are a pos-

sible subject for future research.

Besides their scholarly value, the argument and evidence help inform policy debates. Tradeoffs in alliance treaty design add to our understanding of why some alliances reduce military spending and possible policy responses. Reassurance from deep alliances leads to lower defense spending. States can use deep cooperation to increase alliance credibility, but they lose influence over allied military spending as a result.

The United States is currently wrestling with the implications of treaty depth. 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 reveals a tradeoff—it is not an unconditional call to reduce the depth of US alliances. Adjusting existing treaties may be more difficult than designing new alliances and will 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* .
- 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.
- Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multi-level/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* pp. 1–15.
- 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. 2003a. "Alliance Reliability in Times of War: Explaining State Decisions to Violate Treaties." *International Organization* 57(4):801–827.
- Leeds, Brett Ashley. 2003b. "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. 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.
- 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.
- Senese, Paul D and John A Vasquez. 2008. *The steps to war: An empirical study*. Princeton University Press.
- 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.

- Treier, Shawn and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1):201–217.
- Walt, Stephen M. 1990. *The Origins of Alliance*. Cornell University Press.
- Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.
- 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.