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 and can explain increases and decreases in spending by small alliance members. Security-seeking junior members of deep alliances usually decrease military spending, because these treaties are more credible. Joining shallow alliances often increases junior alliance member 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 1919 to 2007. I find that deeper alliance treaties tend to decrease military spending by junior alliance members, 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.

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.

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 alliance treaty design for outcomes like conflict (Leeds, 2003*b*; Benson, 2012) and trade (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 treaties. I focus on how treaty depth modifies the impact of alliance participation on military spending. Deep alliances formalize costly defense coordination and cooperation among alliance members to bolster commitments of military support.

To explore the consequences of deep and shallow alliances, I focus my argument and empirical

---

[1]See DiGiuseppe and Poast (2018) for an important exception.

analysis on junior alliance partners, who have limited capability relative to their allies. I focus on small alliance members because these states show a tradeoff between reassurance and military spending in alliance treaty design. As small states seek security, participation in deep alliances allows them to reduce military spending because the sunk costs of cooperation increase credibility and reduce allied leverage over their defense spending. Joining a shallow alliance often increases junior alliance member military spending because the foreign policy benefits of alliance participation depend 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 the depth 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. 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 affect the security benefits of alliance participation. Multilevel modeling matches my conditional argument and captures heterogeneous effects of participation in individual treaties. I fit the model on a sample of all states from 1919 to 2007 and find that while deep alliances decrease percentage changes in junior alliance member 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 will struggle to both reassure alliance partners and encourage higher military spending. The United States often uses deep alliances to reassure partners, but this may encourage

lower allied military spending, which leaves little room for compromise between deep engagement and restraint in grand strategy.

2   Do Alliances Increase or Decrease Military Spending?

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

Two types of arguments expect 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 can 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 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

---

[2]Sandler and Forbes (1980), Oneal (1990) and Sandler and Hartley (2001) build on the public goods model.

that alliances generate foreign policy obligations that necessitate 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 allied cooperation 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.[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.

Scholars have assessed 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 examine defense expenditures within a single alliance, most often the North Atlantic Treaty Organization (NATO).

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

The mixed empirical results reflect a theoretical problem. Both perspectives make unconditional claims about the average effect of alliance participation on military spending. With the ex-

---

[3]Senese and Vasquez (2008) argue that conflict spirals drive simultaneous growth in military expenditures and alliance participation, which suggests that conflict behavior drives the correlation between alliances and military spending.

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	
	Oneal and Whatley (1996)	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		
	Olson and Zeckhauser (1966)	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.

ception of DiGiuseppe and Poast (2018), 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 alliance treaties.

3 Argument

Deep military cooperation modifies the impact of alliance participation on junior alliance members' military spending. For security-seeking junior alliance partners, alliance participation generates a dilemma, where military spending that makes a state a more valuable ally also has opportunity costs. Alliance members must invest in defense, but they can also rely on allied capability to

reduce defense burdens. Treaty depth shapes how alliance participants manage this tension. Even under significant threats, junior members of deep alliances can gain security without increasing military spending, because depth generates credible commitments through sunk and installment costs in cooperation. Conversely, for junior partners to retain shallow alliances, they often must maintain or increase their military spending to manage abandonment concerns.

Alliances are a form of international cooperation. Promising military support in a treaty facilitates credible intervention commitments through hands-tying signals (Fearon, 1997; Morrow, 2000). Allied support then helps members achieve foreign policy goals like deterrence or winning wars (Walt, 1990; Snyder, 1997).

Partners' military capability makes alliances valuable (Fordham and Poast, 2014). An alliance treaty formalizes when a state can expect military intervention (Morrow, 2000), but the treaty itself does not provide security. Promises of intervention are only meaningful when allies have sufficient military capability to affect the expected outcome of conflicts. Therefore, greater allied capability increases alliance value, all else equal (Johnson, Leeds and Wu, 2015).

Because treaty value depends on allied capability, alliance participation and military spending are related. Forming and maintaining alliances creates a dilemma in military spending decisions. On the one hand, military spending makes a state a more valuable ally and can provide a costly indicator of commitment. At the same time, defense spending has domestic opportunity costs— funds spent on the military could be allocated to other goods. Alliance members must invest enough in defense to maintain the foreign policy benefits of alliance participation while managing opportunity costs.

How states address the military spending dilemma in alliances depends on their relative military capability, which structures relations between alliance members. 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 states with large capability differences, where large members increase their foreign policy influ-

7

ence and smaller partners gain security (Morrow, 1991). As a result, alliance participation affects defense spending by leading and junior treaty members in different ways. Small states face the tension between demonstrating commitment and value through military spending while managing opportunity costs. Large alliance members provide substantial capability, so they are less subject to the spending tension than small members, though they must also consider the opportunity costs of spending.

As with all cooperation, alliance members must account for opportunism. Even as states commit to an alliance, they could defect and take advantage of allied cooperation. Sometimes the perceived benefits of defection outweigh the long-run benefits of cooperation, which creates an enforcement problem (Fearon, 1998; Koremenos, Lipson and Snidal, 2001).

Due to their emphasis on security, small states are especially concerned with abandonment, which is the most common form of opportunism in alliances. The rate of compliance with military intervention obligations may be as low as 50% (Berkemeier and Fuhrmann, 2018). Abandonment threatens the security goals of junior alliance members, and thus affects military spending decisions.

Abandonment and military spending are related because fear of abandonment shapes how much small states must spend on defense to maintain an alliance. Security-seeking alliance members will only lower military expenditures if the alliance is credible.[4] Though states can augment the collective military capability of an alliance with military capability investments, they can also reduce defense spending, rely on their partners and so mitigate the opportunity costs of defense spending (Olson and Zeckhauser, 1966; Morrow, 1993; Conybeare, 1994; Sandler and Hartley, 2001).

Credible alliances make lower military expenditures more likely by reducing the fear of abandonment. As DiGiuseppe and Poast (2018) observe, some alliances have fewer credibility concerns

---

[4]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 a problem. But as Olson and Zeckhauser (1966, pg. 278) observe, lower defense spending is not normatively problematic.

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. In general, alliance members can establish credible commitments through costly signals (Fearon, 1997; Quek, 2021). All alliance treaties have a hands-tying component, though the strength of this signal varies. I focus on treaty depth, which employs sunk and installment cost promises to help alliance members and potential adversaries infer the credibility of the alliance (Leeds, 2003*b*; Fuhrmann and Sechser, 2014).

Although credible alliances might encourage lower spending by junior members, other alliance members prefer higher allied military spending because it adds capability to the alliance and helps manage their own military spending opportunity costs.[5] Allies also encourage higher military spending so their partners can manage internal threats.[6] This makes lower allied spending a side effect of large states' efforts to establish credible commitment.

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

---

[5]This preference for higher spending has limits, as substantial capability changes will increase allied autonomy. For example, the United States has inhibited allied nuclear programs (Gavin, 2015; Lanoszka, 2015).

[6]This preference for higher defense spending is not a driving force in alliance treaty design, however. 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).

Enforcing greater allied spending 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. Freedom to reduce defense spending and spend more on other goods sometimes motivates alliance formation (Kimball, 2010; Allen and DiGiuseppe, 2013), so lower spending can be a goal of small states.

Given the opportunity costs of military spending, alliance members need leverage to encourage greater allied defense spending. Leverage comes either from a credible threat to abandon low-spending partners or direct control over allied policies. Policy control of allied spending decisions holds when the 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.

Reassurance reduces the credibility of threats to abandon low-spending allies. States cannot reassure their allies and maximize defense spending leverage. Partners can reduce expenditures when alliance members sink costs to reassure.

Under less credible alliances, low military expenditures increase the risk of abandonment, so members have less freedom to reduce defense spending. Moreover, although alliance members increase their foreign obligations, states with credibility concerns risk facing conflicts without allied support. As treaty credibility falls, the foreign policy benefits of alliance participation become contingent on military spending.[7] 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.

---

[7]For example, (Niou and Zeigler, 2019) claim that states are less likely to reduce military spending in alliances between rivals because the benefits of alliance participation depend on military spending.

Treaty depth highlights the tradeoff between reassurance and military spending in alliance politics. Costly cooperation generates credible commitments in deep alliances. While deep alliances often reduce junior members' military spending, participation in shallow alliances often increases it. Treaty depth reassures and reduces leverage over allied military expenditures. Credibility from treaty depth also promotes efficiency gains from coordinated defense effort.

3.1   Alliance Treaty Depth

I focus on depth because it is a common policy choice. Over half of defensive or offensive ATOP alliances have some depth. Moreover, the sunk and installment costs of deep treaties generate a trade off between reassuring small alliance members and defense spending.

Alliance depth is the extent of defense cooperation in the alliance treaty. Deep alliances require additional policy coordination and military cooperation beyond a promise of military support. Defense cooperation in a deep alliance can take many forms, including an integrated military command, military aid, common defense policies, basing rights, international organizations, specific contribution requirements or companion military agreements. All of these obligations move alliance members towards close cooperation via policy coordination and regular interaction, while imposing monetary and policy autonomy costs.[8]

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 rights, 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 junior alliance member military spending in two ways. First, depth reassures partners and reduces leverage on defense spending. Deep

---

[8]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 junior partners rarely use basing rights in deep alliances to deploy their troops abroad.

alliances are more credible because defense cooperation is a sunk cost signal of commitment.[9] Making costly commitments of bases, policy coordination, or aid reassures allies, as peacetime coordination facilitates credible commitment in alliances (Morrow, 1994). When alliance members sustain costly cooperation, they inform allies and adversaries alike that they are resolved to uphold their treaty commitments. Sunk costs cooperation under deep alliances is especially informative because alliance members face a time inconsistency problem. Alliance treaty fulfillment depends on common interests (Morrow, 2000; Leeds, 2003a), so changing foreign policy interests undermine commitment (Leeds and Savun, 2007). Under a deep alliance, members can signal ongoing commitment by continuing to sink costs.

Second, credibility in deep alliances can facilitate more efficient defense spending, as alliance members coordinate investment and provide specific capabilities. Specialization allows members of deep alliances to spend less on the military and retain adequate security. Efficiency gains depend on the credibility of the alliance, because states will only specialize if they believe the alliance is reliable (Leeds, 2003a).

On the other hand, shallow alliances have fewer sunk cost signals of reliability, so participation in these treaties will often increase military spending. Shallow alliances have some baseline credibility from hands-tying signals (Fearon, 1997), as well as the audience (Morrow, 2000) and reputation (Gibler, 2008; Crescenzi et al., 2012) costs of violation. Therefore, shallow alliances make junior alliance members more secure, even if they lack the influence to secure a deep commitment. Even so, threats to abandon low-spending allies under a shallow treaty 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 shallow alliance participation thus requires extra defense spending, because low military spending might endanger the treaty or make

---

[9]In the appendix I reassess the finding of Leeds and Anac (2005) that institutionalized alliances are less reliable, and show that these inferences depend in part on the measure of treaty depth.

states vulnerable 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.[10]

The experience of junior NATO members illustrates the argument. European NATO members balanced their desire to substitute defense for social spending with investing enough in the military to avoid antagonizing the United States (Sayle, 2019, pg. 182-3). At the same time, joint defense planning and troop deployments tied the United States to Europe, even in the face of Congressional skepticism (Sayle, 2019, pg 175).

Taken together, the example and argument suggest that treaty depth modifies the impact of alliance participation on small alliance member 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 junior alliance member 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.[11]

> HYPOTHESIS 1: ON AVERAGE, PARTICIPATION IN SHALLOW ALLIANCES WILL IN-
> CREASE PERCENTAGE CHANGES IN SMALL ALLIANCE MEMBER MILITARY SPEND-
> ING.

> HYPOTHESIS 2: ON AVERAGE, PARTICIPATION IN DEEP ALLIANCES WILL DE-
> CREASE PERCENTAGE CHANGES IN SMALL ALLIANCE MEMBER MILITARY SPEND-
> ING.

---

[10]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.

[11]Hypothesis 3 follows from Hypotheses 1 and 2.

HYPOTHESIS 3: AS ALLIANCE TREATY DEPTH INCREASES, THE IMPACT OF AL-
LIANCE PARTICIPATION ON PERCENTAGE CHANGES IN SMALL ALLIANCE MEMBER
MILITARY SPENDING WILL DECREASE.

Before detailing how I test these hypotheses, I must briefly justify two conceptual and mea-
surement choices. First, the three hypotheses predict how percentage changes in junior alliance
member military spending differ under deep and shallow alliances. Percentage changes in mili-
tary 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,[12] a larger increase in spending relative to the previous year's
defense budget 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 allied capability supports alliance participants' foreign policy
goals. States do not respond to a treaty per se, rather they respond to expectations that allies will
fight on their behalf. This makes alliances with more capable states are more valuable. 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 and creates a unified understanding of how
alliances affect military spending.

---

[12]Especially holding economic growth constant.

14

4 Research Design

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 of military spending by all states between 1919 and 2007. Throughout the analysis, I examine alliances with military support because prior studies of alliance participation and military spending focus on these treaties. The next section describes the measure of alliance treaty depth.

4.1 Measuring Alliance Treaty Depth

Formal treaty commitments reflect alliance depth,[13] so I use observed alliance treaty characteristics to infer depth. Before describing my measure, I discuss related ordinal and latent measures in previous work.[14] Due to the limitations of extant 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 alliance institutionalization that assigns 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 gives alliance obligations different theoretical weights. 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 two sets of depth variables and does not aggregate multiple sources of depth. For example, it

---

[13]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.

[14]Neither measure is a simple additive index. Adding the presence of different sources of depth assumes that all factors are equally important, which creates theoretical and empirical problems with comparing alliances (Benson and Clinton, 2016, pg. 876).

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 to measure alliance scope, depth and capability. Measurement models have a rich history in political science (Clinton, Jackman and Rivers, 2004; Fariss, 2014). Latent variable modeling uses data to estimate the weights of different observable variables in some unobserved 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. 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 assump-

tions. 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.[15] 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). Indicators of treaty depth include military aid, a dummy indicator of basing rights, international organization formation, integrated military command, defense policy coordination, subordination of forces in wartime, specific contribution requirements, and commitments to form companion military agreements. While these obligations may impact defense spending, they do not mandate specific spending levels. 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.

---

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

These factor loadings match the argument. First, all the variables have strong positive factor loadings, as they stipulate cooperation between alliance members. Second, I define depth as the extent of cooperation a treaty requires, so it makes sense that an integrated command, military policy coordination, and formal organizations have the highest factor loadings, as these three variables are key sources of peacetime cooperation.[16] Third, defining depth in terms of coordination and cooperation may also explain why basing rights enters the second tier of factor loadings. Although basing rights in an alliance treaty clearly increase treaty depth, they may not add as much depth as other factors because basing has cooperative and hierarchical elements. Providing basing rights and managing their daily interactions requires costly cooperation, 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 can be 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 of -0.6 or higher 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 with military support alone. 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

---

[16]Policy coordination is an ordinal indicator of ties between allied armed forces for developing military plans, and includes none, wartime coordination, peacetime coordination and full defense policy coordination.
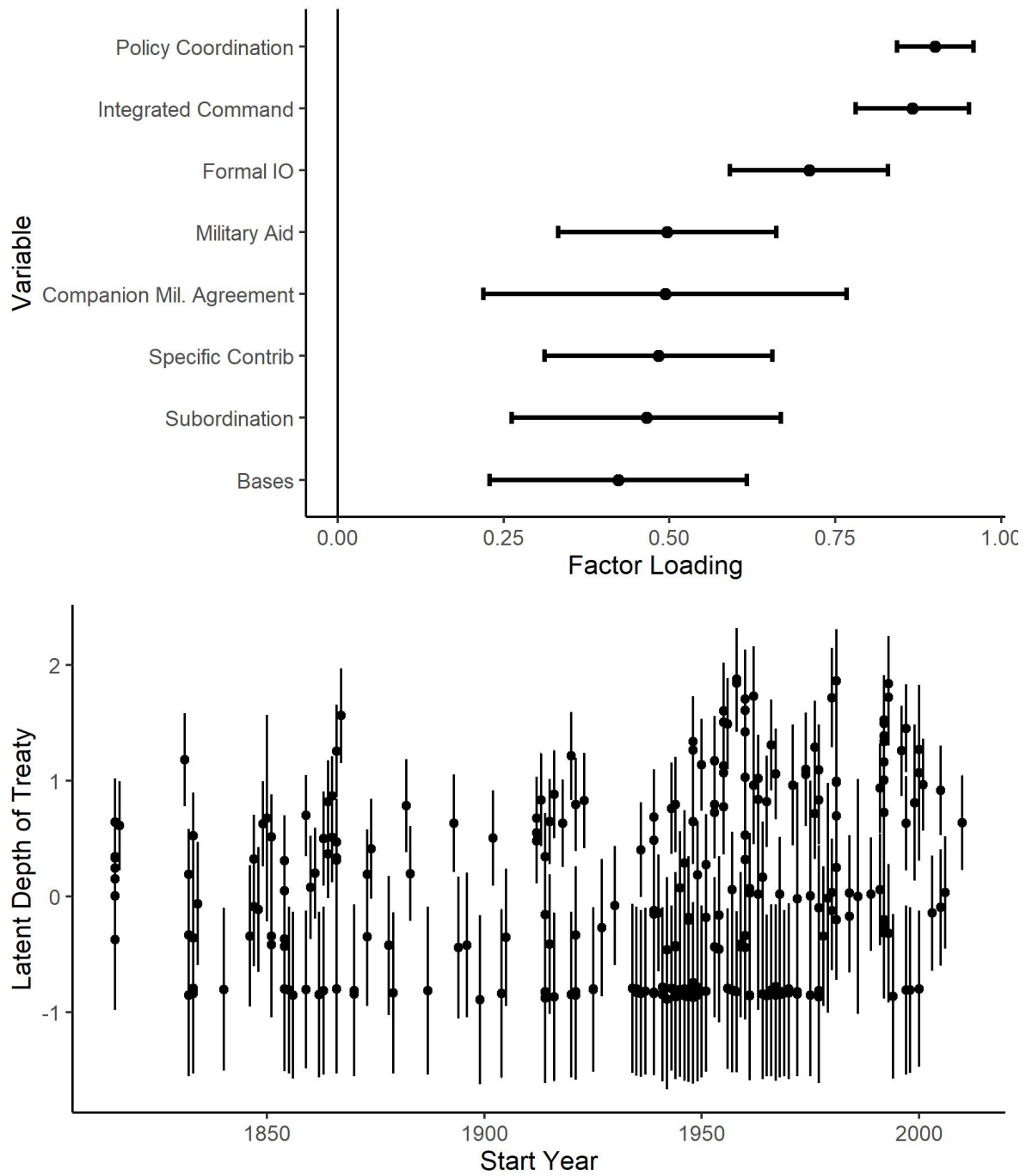
Figure 1: Summary of the latent measure of alliance treaty depth for 289 defensive or offensive alliances from 1812 to 2016. The top panel plots the factor loadings with 90% credible intervals. The bottom panel plots mean treaty depth (points) and the standard deviation (error bars) against the start year of the treaty.

alliance 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.

Alliance	Depth Score	Year	Depth Sources
France-USSR	-.85	1935	None
NATO	.2	1949	Formal Organization, Basing rights
France-Russia	.63	1893	Moderate Policy Coord., Specific Contribution
Brazil-Argentina-Uruguay	.89	1865	Integrated Command, Military Aid, Low Policy Coord., Subordination.
Warsaw Pact	1.1	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.

Table 2 provides five other examples of deep and shallow alliance treaty obligations. France's 1935 alliance with the USSR (ATOPID 2360) has a minimal depth score, as it did not include any cooperation besides defensive and consultation obligations. An 1893 Franco-Russian alliance (ATOPID 1385) called for policy coordination and specified the number of troops each partner would provide in war. The 1865 Triple Alliance between Brazil, Argentina and Uruguay (ATOPID 1255) included extensive cooperation in a war against Paraguay. NATO (ATOPID 3180) has above average latent treaty depth, but its formal treaty obligations are not extraordinarily deep.[17] Because the Warsaw Pact (ATOPID 3285) obligates members to join a formal organization, coordinate defense policy, and participate in an integrated command, it has greater formal depth than NATO.

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.

---

[17]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.

The latent treaty depth measure is my key explanatory variable. 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.

4.2   Estimation: Multilevel Model

Because alliances and state-year observations are separate levels of analysis, I use a multilevel model to bridge the two data levels and estimate the association between treaty depth and military spending. My model estimates heterogeneous effects of alliance participation on military spending as a function of alliance characteristics to infer 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 (2018) use, but the multilevel components add substantial value by connecting the argument and research design. I argue that treaty depth modifies the impact of alliance participation on growth in military spending. Therefore, 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 through alliance-level regression coefficients.

Unlike the multilevel model, standard panel models employ state-level proxies for alliance characteristics, which compare states rather than alliances. Summarizing alliance characteristics at a different level of analysis changes the mean and variance of key independent variables. This practice of aggregating alliances at the state-year level of analysis may produce misleading inferences (McElreath, 2016, pg. 356). Multilevel modeling avoids the 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. Aggregating multiple alliances into state level indicators would mask the heterogeneous effects of individual treaties.[18]

Furthermore, multiple alliance characteristics affect the consequences of alliance participation. The multilevel model captures multiple sources of heterogeneity in how alliances impact military spending. Different aspects of alliance membership and design are correlated (Chiba, Johnson and Leeds, 2015), so this step is important.[19] 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 spurious findings through multiple comparisons. In a multilevel model, I account for multiple alliance characteristics 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 adjusts for key alliance and state characteristics.

4.2.1   Model Specification

This multilevel model connects regression models at the state and alliance levels of analysis. The base is a state-year-level regression, which includes the impact of alliance participation when a state is a large or small alliance member. Two alliance-level regressions then modify the effect of alliance participation on military spending for small and large alliance participants, like an interaction. Because alliance participation has different consequences for large and small alliance members, distinct models capture the impact of alliance participation on large states with substantial capability, and small states with a limited share of total allied capability.

---

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

[19]Average democracy, threat and alliance size at the time of alliance formation are the most important sources of treaty depth.

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

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

y is the dependent variable— percentage changes in military spending. I model the outcome using a Cauchy distribution address heavy tails. σ is analogous to the error term in a frequentist regression as it captures unexplained variation. μ, the outcome mean, depends on several factors.

$$\mu = \alpha + \alpha^{st} + \alpha^{yr} + \mathbf{W}_{n \times k}\gamma_{k \times 1} + \mathbf{Z}^{sm}_{n \times a_{sm}}\lambda^{sm}_{a_{sm} \times 1} + \mathbf{Z}^{lg}_{n \times a_{lg}}\lambda^{lg}_{a_{lg} \times 1} \tag{2}$$

Percentage changes in spending depend on 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 $\mathbf{Z}\lambda$ terms incorporate alliance participation as a large and small treaty member. $\mathbf{Z}_{lg}$ is the alliance participation matrix when a state is a large member, while $\mathbf{Z}_{sm}$ captures alliance participation when a state is a junior partner. A state can be a large or small member of different alliances, depending on relative capability. In both alliance participation matrices, columns correspond to alliances and rows to state-year observations. $\mathbf{Z}_{lg}$ has 195 columns and 7845 rows and $\mathbf{Z}_{sm}$ has 159 columns and the same number of rows, as some alliances have limited capability divergence, making all members large states. If a state is part of an alliance, a corresponding element of the large or small membership matrix contains the log of total allied military spending, normalized by year. If a state is not in a particular alliance the matrix cells in those alliance columns are zero in both membership matrices.

I use total allied spending to express the effect of alliance participation as it matches the theoretical emphasis on allied capability. The $\mathbf{Z}$ matrices encode quasi-spatial indicators of alliance participation for the a alliances in the data. States can be members of multiple treaties, so state-year observations are not nested within alliances. In this specification, each alliance has a unique impact on military spending, even when states participate in multiple treaties. Furthermore, every

23

alliance member experiences temporal variation in allied capability that starts when the alliance forms, continues with fluctuations in allied spending while the alliance is active and ends with alliance termination.

λ^{sm} and λ^{lg} are vectors of parameters which estimate the impact of participation in specific alliances on military spending for large and small members. Because the non-zero elements of $\mathbf{Z}$ are allied spending, the λ parameters capture alliance members' response to allied capability. These parameters apply to all states of the same size, so alliance formation and membership have a common effect on multiple states when there are multiple large or small alliance members. There are 195 large and 159 small alliance participation parameters, leaving 354 unique alliance participation parameters. Both sets of λ parameters have their own distribution, so I assume alliances are similar but different in how they impact military spending.[20]

The second part of the multilevel model uses alliance characteristics to predict how alliance participation is associated with percentage changes in military spending. The large and small member alliance participation parameters are the outcome in these two alliance-level regressions, where the impact of alliance participation on military spending depends on treaty characteristics, including depth. Given separate regressions for large and small states, alliance characteristics like depth have different consequences for small and large alliance members. For the two alliance-level regressions where large and small alliance participation is indexed by j:

$$\lambda_{aj} \sim N(\theta_{aj}, \sigma_{all}^{j}) \tag{3}$$

and

$$\theta_{aj} = \alpha_{all}^{j} + \beta_j \text{treaty depth} + \mathbf{X}_{aj \times l}^{j} \beta_j \tag{4}$$

In the alliance-level regressions, $\mathbf{X}$ is a matrix of alliance-level control variables and α_{all} is the constant. Adding a variance parameter σ_{all} means predictions of λ are not deterministic—

---

[20]I assume that the alliance-level regression parameters are uncorrelated. A model that relaxes this assumption with a varying slopes specification produces similar results. Results on file with the author.

the alliance level regressions contain 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 parameters 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 that β_1 will be negative in the junior state alliance equation, which is the focus of the results in this paper.

Each λ expresses the effects of participation in an alliance for small or large member states, which is a function of changes in allied capability throughout alliance formation and maintenance. The β parameters estimate how alliance characteristics modify the impact of alliance participation on military spending, so they compare the impact of alliances with different characteristics like depth. Therefore, treaty depth indirectly affects military spending by modifying the impact of allied capability. 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 all states from 1919 to 2007, including state-year observations without alliances. Alliance participation data comes from the version 4 of the ATOP data (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 7,845 state-year observations and 195 alliances.

I divide alliance members into the large and small participation matrices based on relative capability within an alliance. If a state has the largest share of total alliance capability, is part of a symmetric alliance between major powers, has more than 40% of total capability in a bilateral alliance, or is a major power in an asymmetric multilateral alliance without a superpower member,

they are a large alliance member. Small states thus have less than 40% of total allied capability in a bilateral alliance, or are non-major powers with less than the largest capability share in multilateral alliances. Superpowers like the United States and USSR are always the lone large alliance member under this classification. I employ these classification rules because the distributions of relative capability are very different in bilateral and multilateral alliances. Again, a state can be a large member of one alliance, and a small member elsewhere. For example, Germany is a junior member of NATO, but a large member of the European Defense Pact.

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 create this variable (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.[21] 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 and 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

---

[21]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.

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).

The other alliance level variables are correlates of treaty depth and military spending, including the number of members and average polity score of treaty members at time of alliance formation. 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 with a dummy variable. 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). Dummy indicators of wartime alliances, asymmetric obligations (Leeds et al., 2002) and superpower membership complete the alliance-level regression specification.[22] Having described the measure of treaty depth, multilevel model, and covariates, I now turn to the results.

5 Results

This section summarizes inferences about treaty depth and junior alliance members from the multilevel model.[23] 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 small alliance member military spending are negatively correlated.

---

[22]Though I discuss these variables as controls, many of them are theoretically interesting as well.

[23]Alliance-level regression results for large states, which show a different connection between depth and defense spending, are in the appendix.

Results are based on 3,000 samples from four chains, with 1,500 warm-up iterations. To facilitate model fitting, I employed a non-centered parameterization of the varying intercepts and sparse matrix representations of the $\mathbf{Z}$ matrices. Standard convergence diagnostics indicate the chains adequately explored the posterior.

Because I used Bayesian modeling to estimate the association between treaty depth and percentage changes in military spending, each coefficient has a posterior distribution of likely parameter values conditional on the priors and observed data. There are no indicators of statistical significance. I use 90% credible intervals and calculate the negative posterior probability for the treaty depth coefficient to assess Hypothesis 3.[24]

Figure 2 summarizes the coefficient estimates from the alliance-level regression and a simulated substantive effect of increasing 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 small alliance members. As treaty depth rises, the effect of alliance participation on junior alliance member 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 participation parameter λ by .1 in expectation, which then lowers military spending growth. 90% intervals summarize 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 median allied capability, the difference in spending growth between the shallow and deep treaty has a mean of -.04. The 90% credible interval of this predicted fall in military spending due to increasing treaty depth

---

[24]I use 90% intervals because inferences about 95% intervals are sensitive to simulation variance in Bayesian analysis.
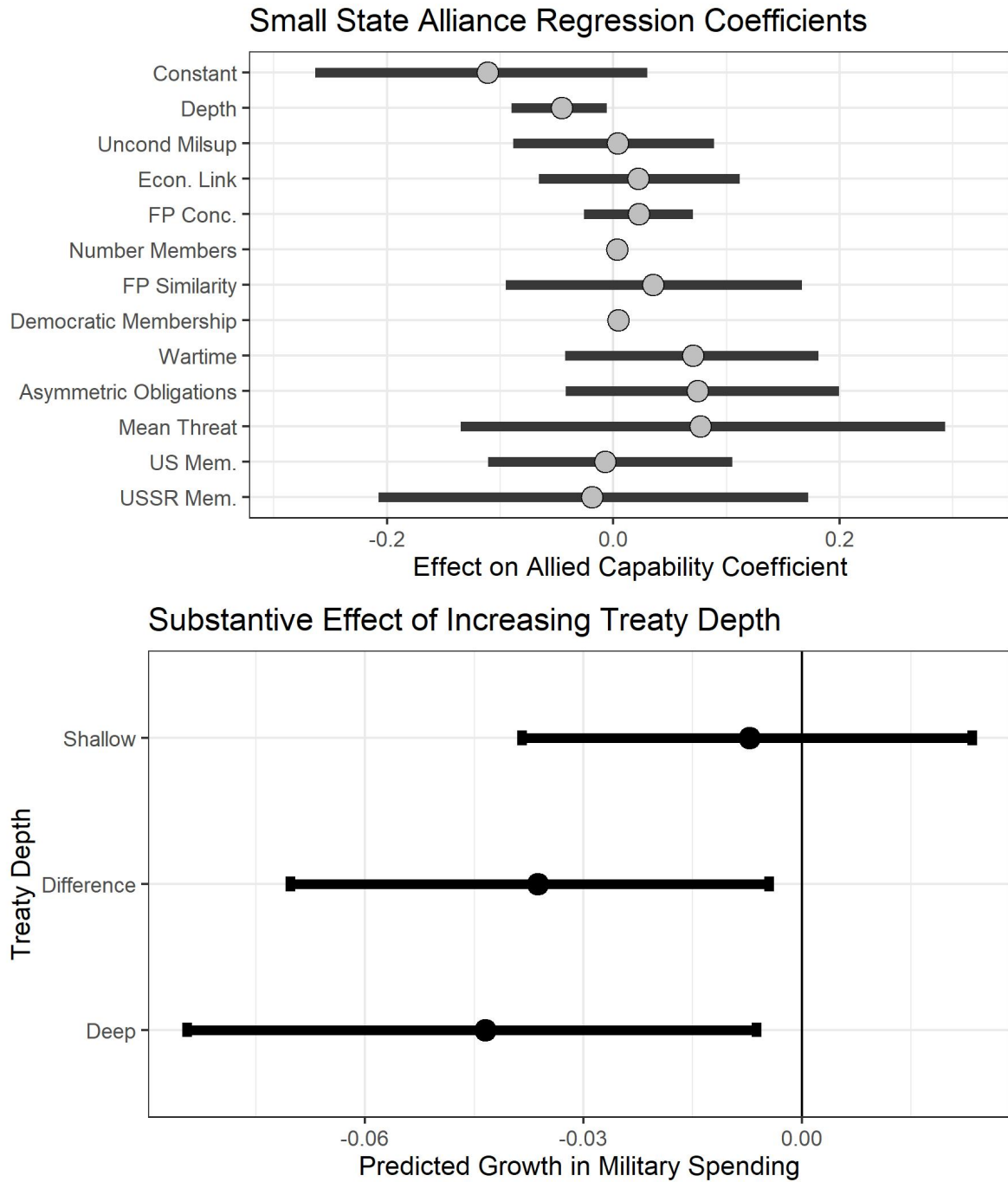
Figure 2: Summary of alliance-level regression results from the multilevel model of alliance participation and military spending. The top panel shows the 90% credible intervals for coefficients in the alliance-level regression. The bottom panel plots the estimated effect of participation in an alliance 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.

ranges from -0.07 to -0.002. Although the range of estimated substantive effects includes small values, there is a perceptible difference between junior alliance member 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 alters the consequences of alliance participation, it will appear in the λ estimates. On average, participation in deep alliances should have a negative effect on percentage changes in military spending and shallow alliances should have a positive effect. If this is true, there will also be a negative trend in average λ values 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, most shallow treaties have positive λ parameters, 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 junior alliance member 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 members' defense spending, I multiplied the alliance membership matrix $\mathbf{Z}$ by the λ parameters to generate 7,936 non-zero state-alliance-year predictions.[25] The scatter plot in the bottom panel of Figure 3 shows the distribution of predicted changes in military spending from alliance participation over the range of treaty depth. Each point marks the mean estimated effect of an alliance on growth in military spending for that state-year. To avoid overplotting the 7,936 point estimates, I combined them into hexagons. Darker hexagons mark areas with more points and alliance estimates usually concentrate within hexagons. Figure 3 shows the indirect consequences of treaty depth, which modifies

---

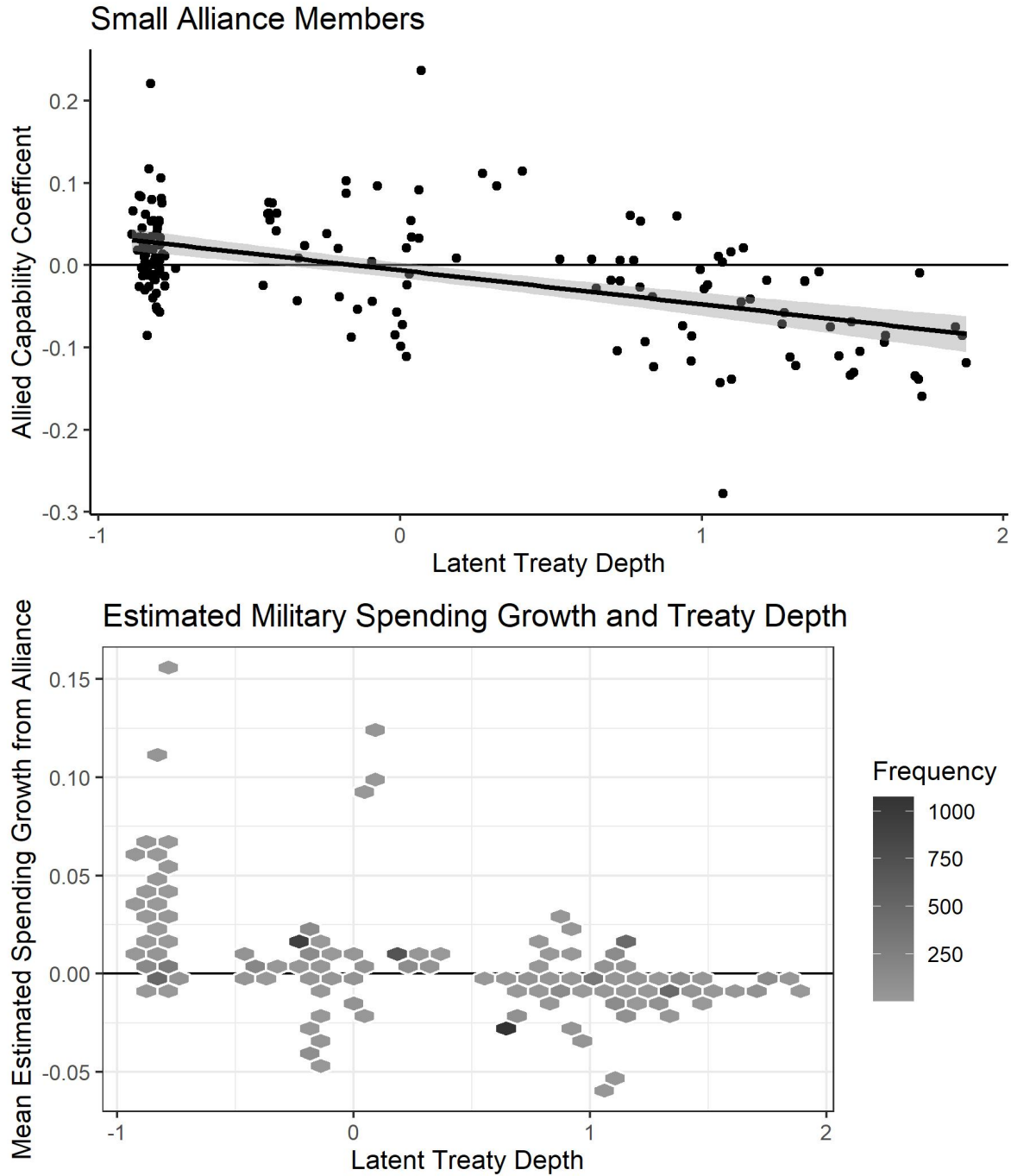[25]These estimates hold all state-level variables constant.

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 1919 to 2007. The top panel plots the mean of each λ parameter by treaty depth. The bottom panel plots 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.

the impact of alliance participation. As Hypothesis 1 predicts, participation in shallow alliances regularly increases military spending. Many shallow alliances have little effect on military spending, however. Participation in deep alliances is more likely to reduce military spending growth, with the exception of wartime alliances. Some of the deepest alliances have a negligible effect on military spending despite negative λ parameters due to limited allied capability.

I report several checks of these results in the appendix. First, I fit the multilevel model with the ordinal measure of (Leeds and Anac, 2005), and find that military institutionalization also reduces the impact of alliance participation on small alliance member 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 with average and dummy summaries of states' alliance portfolios.

In summary, treaty depth modifies the impact of alliance participation on military spending. Deep alliance participation often reduces military spending, while membership of shallow alliances 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 that 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, thereby building on other conditional arguments (DiGiuseppe and Poast, 2018). 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 junior members' military spending, which contradicts expectations that these states are inveterate free-riders. To gain security from alliance participation, junior partners sometimes need to increase their defense bud-

get (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 widespread credibility concerns.

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 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 military spending growth. Even with this effort, selection into different alliances could still produce unobserved differences between

alliances I cannot account 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 more or less military spending by junior alliance members, 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 in international institutions. 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 military spending? Many theories expect a close connection between alliances and spending, 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. This creates a tradeoff between credible commitments and greater allied military spending.

The United States currently faces the credibility-military spending tradeoff. Washington has often complained that allies 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 would generate credibility problems, however. Low

allied defense spending may be the price of credible commitments. Therefore, this paper is not an unconditional call for shallow 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. 2018. "Arms versus Democratic Allies." *British Journal of Political Science* 48(4):981–1003.

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.

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 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.

Oneal, John R and Hugh Carter Whatley. 1996. "The effect of alliance membership on national defense burdens, 1953–88: A test of Mancur Olson's theory of collective action." *International Interactions* 22(2):105–122.

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.

Quek, Kai. 2021. "Four Costly Signaling Mechanisms." *American Political Science Review* 115(2):537–549.

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.

Sayle, Timothy Andrews. 2019. *Enduring Alliance: A History of NATO and the Postwar Global Order*. Ithaca NY: Cornell University 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.

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.

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.