Alliance Participation, Treaty Depth, and Military Spending

Joshua Alley*

September 25, 2019

Abstract

How does alliance participation affect military spending? Some argue that joining an alliance increases defense expenditures, while others contend that it produces spending cuts. I argue that how alliance participation affects military expenditures depends on the depth of military cooperation in the treaty. Alliance depth reveals a tradeoff between reassurance and free-riding, as security-seeking non-major powers free-ride more on their allies in deeper alliances. 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 deep alliances reduce the impact of alliance participation on non-major power military spending relative to shallow treaties. This result helps 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 produces contradictory predictions and evidence on the question of alliance participation and military spending. One view expects alliance participation will reduce military spending (Morrow, 1993; Conybeare, 1994; DiGiuseppe and Poast, 2016). The other predicts alliance participants will spend more on defense (Diehl, 1994; Morgan and Palmer, 2006). This paper addresses the divide by explaining 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 between treaties and states, alliance participation could plausibly increase or decrease defense expenditures. I emphasize the role of treaty depth to explain how alliance treaty participation impacts military spending.

Deep alliances formalize extensive defense cooperation between members, which reassures allies and helps coordinate policy. Rather than make arms-length commitments, these treaties make promise aid, create formal institutions and obligate members to coordinate defense policy. For example, a 1963 alliance between Jordan and Iraq commits to military support as well as military aid, bases and institutions to govern military coordination.

Additional treaty depth increases free-riding by non-major power alliance participants. Non-

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

major powers are prone to free-ride by reducing military spending, as they emphasize immediate security and have high opportunity costs of military spending. Because alliance participation provides extra security, it often reduces non-major power military spending. Reassurance and credibility from a deep alliance, as well as efficiency gains, allow non-major powers to reduce military spending. As a result, deep alliances reduce non-major power military spending relative to shallow alliances. Shallow alliances are more likely to increase military spending because they involve non-major powers in international disputes with less assurance allies will honor their commitments.

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 treaty characteristics modify the impact of alliance participation on percentage changes in military spending. Multilevel modeling corresponds to the argument and generates useful comparisons between alliances. I fit the model on a sample of non-major power states from 1816 to 2007. I find that deep alliances reduce the impact of alliance participation on percentage changes in non-major power military spending, relative to shallow treaties.

Though increasing treaty depth reassures allied states, it increases the prevalence of free-riding. This trade off illuminates a salient debate in US foreign policy about the costs and benefits of US alliances. Advocates of deep engagement (Brooks, Ikenberry and Wohlforth, 2013) and restraint (Posen, 2014) in grand strategy have different views of alliances. Advocates of restraint argue that the United States should withdraw from many alliances, as allied states spend too little on defense, which increases US defense spending (Preble, 2009). Proponents of continued deep engagement argue that the benefits of alliances exceed the costs, and the extent of allied free-riding is overstated (Brands and Feaver, 2017). Debates about how to respond to "free-riding" by US allies should consider that efforts to reassure may increase the extent of free-riding. Many US alliances make commitments of deep military cooperation, which is likely to exacerbate free-riding.

The paper proceeds as follows. First, I summarize competing claims on alliance participation and military spending. Then I describe my argument in more detail. The third and fourth sections

present the research design and results. As part of the results, I show how US alliances are often below average in formal depth and their impact on allied military spending. 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 or unconditional effect of alliance participation by emphasizing one aspect of alliance politics. I start with the substitution and public goods logics connecting alliance participation with reduced defense spending.

2.1 Why Alliances Could Decrease Military Spending

These arguments begin with the premise that alliances and military spending both provide security. The first such model treats security from an alliance as a public good. Olson and Zeckhauser argue that alliances are subject to a collective action problem (Olson and Zeckhauser, 1966). 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. Spending less allows alliance members to consume more non-defense goods, but the alliance provides suboptimal security.[2]

Another argument that predicts reduced defense spending focuses on substitution between foreign policy instruments. Substitution arguments recognize that states employ one policy in place of the other (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.

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

Under substitution, allied military capability replaces other members' defense expenditures. DiGiuseppe and Poast refine the substitution logic by arguing that states will only reduce spending if the alliance is credible (DiGiuseppe and Poast, 2016). Because democracies make more credible commitments, they assert that defense pacts with democracies lower defense spending. This conditional argument is a useful step towards bridging the theoretical debate.

Both the substitution and public goods models expect that alliance participation reduces military spending due to opportunity costs of military expenditures. States have incentives to rely on their allies for security because spending more on the military leaves less for other goods (Fordham, 1998; Fearon, 2018). On the other hand, a contradictory perspective asserts that alliance participation increases military expenditures.

2.2 Why Alliances Could Increase Military Spending

Several arguments predict higher military spending by alliance members. All share an intuition that states increase military spending to support their alliance commitments. In these models, investing in the military secures foreign policy gains from alliance participation.

Diehl argues that alliances increase foreign policy obligations, necessitating extra military spending (Diehl, 1994). Because alliances expand what a state can achieve in international relations, states increase military spending to pursue other foreign policy goals Morgan and Palmer (2006). For example, buffer states increase defense effort to make themselves a more attractive alliance partner (Horowitz, Poast and Stam, 2017). Others assert that alliances generate cooperation, leading to 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 allied states.[3]

[3]Senese and Vasquez (2008) argue that military spending and alliances are part of a conflict spiral of simultaneous growth in military expenditures and alliance participation. This argument suggests that any correlation between alliances and military spending is driven by conflict behavior, not treaty participation.

2.3 Mixed Evidence

Debate between the contradictory views of alliances could be settled by a consistent set of average effects, but mixed results reinforce the theoretical division. Some studies find a positive association between alliance participation and military spending. Others find a negative relationship.[4]

General studies of military spending and alliances compare many states through dummy indicators of alliance participation, which combine 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.

	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 Association Between Alliance Participation and Military Spending.

Unlike general studies, specific research designs examine individual treaties and estimate responses to military spending of key allies. Most evidence of reduced military spending by alliance members comes from alliance-specific research designs (Barnett and Levy, 1991; Morrow, 1993; Sorokin, 1994; Plümper and Neumayer, 2015; George and Sandler, 2017). Other specific studies find increased spending by alliance members, however (Conybeare and Sandler, 1990; Chen, Feng and Masroori, 1996).

[4]Because tests of the public goods model use military spending as a share of GDP as the their outcome of interest, I ignore most of those results.

2.4 The Theoretical Challenge

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 vary widely across alliances (Leeds et al., 2002), which leads to heterogeneous effects. Conflict (Leeds, 2003*b*; Benson, 2012) and trade (Long, 2003; Long and Leeds, 2006) are two domains where alliance design shapes the impact of treaty participation.

Differences in alliance treaty design should lead us to question binary measures of alliance participation in general studies and limit the generalizability of inferences from specific studies. Alliance participation could increase or decrease military spending, depending on treaty obligations. While DiGiuseppe and Poast (2016) focus on the conditionality of alliance credibility based on democratic membership, I look at the depth of cooperation in the treaty.

My argument focuses on treaty depth as an important source of heterogeneity in how alliance participation impacts military spending. Formal treaty depth alters the consequences of alliance participation for non-major powers by limiting their ability to reduce defense spending. The next section summarizes the argument in more detail.

3 Argument

Deep military cooperation in an alliance treaty reduces the impact of alliance participation on non-major power military spending. Treaty depth increases non-major powers' tendency to free-ride through reassurance and efficiency gains. Given greater treaty credibility, non-major powers have more freedom to reduce defense spending. Efficiency gains allow non-major powers to depend on their allies for crucial military capabilities. Therefore, deep alliances have a lower impact on percentage changes in non-major power military spending than shallow ones.

I start the argument by describing problems of opportunism and enforcement in alliances. Then I describe the role of deep formal military cooperation. Last, I show how alliance depth attenuates the tendency of alliance participation to reduce non-major power military spending.

3.1 Opportunism in Alliances

Alliances are a form of international cooperation. By promising to aid one another in war, states make a credible commitment of intervention (Fearon, 1997; Morrow, 2000). Allied support helps members achieve crucial foreign policy goals like deterrence or success in war (Walt, 1990; Snyder, 1997).

Like all cooperation, alliances can suffer from 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).

There are two especially salient forms of opportunism in alliances: abandonment and free-riding.[5] Abandonment, where states violate their alliance commitments, is a common problem (Berkemeier and Fuhrmann, 2018). As a result, alliance members must convey the credibility of their commitments. Free-riding through low defense spending is another common form of opportunism in alliances.[6] Though states contribute 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).

As DiGiuseppe and Poast (2016) observe, some states have fewer credibility concerns due to

[5]Some argue that entrapment is another form of opportunism (Snyder, 1984), but some evidence suggests entrapment is rare (Kim, 2011; Beckley, 2015)

[6]Though the public goods model of alliances has serious theoretical and empirical flaws, it is common practice to describe low defense spending in an alliance as free-riding.

inherent characteristics like their political regime. Credibility is multifaceted, however. Alliance members often attempt to increase the perceived credibility of their commitments through treaty design. For example, placing few conditions of military support indicates greater credibility (Benson, 2012; Chiba, Johnson and Leeds, 2015). Issue linkages also increase alliance credibility (Long and Leeds, 2006; Poast, 2012, 2013). Here, I focus on treaty depth as a key source of credibility.

In general, alliance members increase the credibility of their commitments by making costly commitments. These costs allow alliance members and potential adversaries to infer the credibility of the alliance (Leeds, 2003*b*; Fuhrmann and Sechser, 2014). Reassuring allies through costly cooperation reduces the risk of abandonment, but it does not address free-riding.

Enforcing cooperation around free-riding is difficult. Normative appeals to common interests rarely work. Though verbal communication or "cheap talk" can have value in international politics (Trager, 2010), it is unlikely to overcome incentives to free-ride. Even when they spend less on the military, 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).

States need sufficient leverage to address allied free-riding. Leverage requires either a credible threat of abandonment in response to free-riding or control over allied policies. Residual control of allied military spending decisions is more common when the treaty reflects hierarchical relationships like an informal empire (Lake, 1996). Without such direct influence, states must possess a credible threat of abandonment in response to free-riding. Otherwise, free-riding allies will dismiss weaker signals and threats.

Greater alliance credibility reduces the credibility of threatening to abandon free-riding allies. Put differently, states cannot simultaneously reassure their allies and maximize leverage over free-riding. As alliance members make costly commitments, partners can take the security benefits and reduce defense spending more. In less credible treaties, such as alliances between erstwhile rivals, members are less likely to reduce defense spending (Niou and Zeigler, 2019).

Deep alliances highlight this tradeoff between reassurance and free-riding. Stipulating deep cooperation reassures partners. Increasing treaty depth also promotes specialization and efficiency from coordinated defense effort. Compared to shallow alliances, deep alliances will lead to lower percentage changes in military spending.

3.2 Depth of Military Cooperation

Alliance depth is the extent of defense cooperation formalized in the treaty. To be more precise, depth is the extent of policy coordination and military cooperation required by formal alliance treaty obligations. Formal defense cooperation is the primary source of depth. Deep alliances stipulate close cooperation between members. By contrast, shallow alliances promise more arms-length cooperation.

Defense cooperation in a deep alliance can take many forms. Allies can form an integrated military command, provide military aid, commit to a common defense policy, provide basing rights, set up a formal organization to manage the alliance 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. These commitments also impose monetary and policy costs on members.

Increasing military coordination adds ties between alliance members beyond a promise of military support, which adds substantial depth. 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 alliance between Greece and Bulgaria which only commits to mutual defense and consultation if either state is attacked by Turkey.

Alliance depth increases the extent of free-riding in two ways. First, adding depth to an alliance reassures partners. Second, deep alliances provide efficiency gains for members.

Deep alliances are more credible because defense cooperation is costly. Making costly com-

mitments of foreign bases, policy coordination, or aid reassures allies. Depth is especially useful because alliance members face a time inconsistency problem. Alliance treaty fulfillment depends largely on shared foreign policy interests (Morrow, 2000; Leeds, 2003*a*). Changing foreign policy interests are therefore a serious threat to alliance fulfillment (Leeds and Savun, 2007). A deep alliance makes a series of repeated transfers, and states can signal commitment by maintaining those transfers.

Shallow alliances still have some credibility. Promises of military support generate some credibility through hands-tying signals (Fearon, 1997), as well as audience Morrow (2000) and reputational costs (Gibler, 2008; Crescenzi et al., 2012). These more limited commitments will be less likely to reduce members military spending, however.

Policy coordination in a deep alliance is also likely to reduce military spending. Given a reliable alliance, members can exploit economies of scale by focusing their defense efforts (Leeds, 2003*a*). Reducing redundant military capabilities may allow some alliance participants to reduce their defense spending. In a deep alliance, members attain more foreign policy goods with less defense effort.

If deep alliances offer efficiency gains, why are many alliances shallow? Depth in an alliance comes at the cost of foreign entanglement. In general, alliance members sacrifice some of their freedom of action in foreign policy in return for the benefits of alliance participation (Altfield, 1984; Snyder, 1997). Deeper military cooperation sharpens this trade off between benefits and foreign policy freedom. Shallow alliances provide fewer direct benefits, but they also have less foreign entanglement.[7] Domestic politics (Davis, 2004), democracy (Chiba, Johnson and Leeds, 2015) and a state's international reputation (Mattes, 2012) may all lead to more arms-length cooperation and shallow treaties.

In summary, deep alliances increase free-riding among alliance members. This is especially

[7]Also deep alliances do not solve every problem, as more institutionalized alliances may not be more reliable (Leeds and Anac, 2005).

relevant for non-major powers. These states are more likely to use alliances to reduce defense spending.

3.3 Deep Cooperation and Non-Major Powers

Non-major powers have limited military capabilities and less status in international relations. As a result, they usually focus on ensuring their immediate security. The emphasis on security creates opportunities for exchanges where non-major powers trade foreign policy autonomy for allied protection (Altfield, 1984; Morrow, 1991). By giving non-major powers more security, alliances allow them to reduce defense spending.

Besides their security goals, non-major powers face higher opportunity costs of military spending. The marginal cost per taxpayer of increases in defense spending is decreasing in the number of taxpayers (Dudley and Montmarquette, 1981), and non-major powers have fewer taxpayers. Small states also have limited economies of scale in military spending (Moravcsik, 1991; Kapstein, 1991; Anderton, 1995; Devore, 2013). These economic factors and non-major powers' foreign policy goals encourage free-riding.

Greater alliance treaty depth makes free-riding more likely. For security-conscious non-major powers, abandonment is a serious concern. Depth reassures non-major powers and allows them to focus their defense efforts.

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 a coordinated occupation of the Rhineland to defensive obligations. Given this additional security, the Franco-Belgian alliance likely reduced Belgian defense expenditures. A more limited treaty between France and Poland that only included military support (ATOPID 2135) had no effect on Polish spending.

These brief examples and the argument suggest that treaty depth modifies the impact of alliance participation on non-major power military spending. Alliance participation is expressed through

allied capability, not just the presence of a treaty. Non-major powers free-ride on allied capability, so my measure of alliance participation uses capability to capture the effect of participation. Greater depth increases the usual negative impact of alliance participation on non-major power military spending. Deeper alliances will lead to higher percentage changes in non-major power military spending than shallow treaties.

> HYPOTHESIS 1: THE IMPACT OF ALLIANCE PARTICIPATION ON PERCENTAGE CHANGES IN NON-MAJOR POWER MILITARY SPENDING WILL BE LOWER IN DEEP ALLIANCES, RELATIVE TO SHALLOW ALLIANCES.

Hypothesis 1 predicts how percentage changes in non-major power military spending differ between deep and shallow alliances. Percentage changes in military spending take changes in spending as a share of the previous year's defense budget. This variable is an appropriate outcome of interest. Percentage changes express the opportunity costs of military spending. All else equal, a large increase in spending relative to the previous year's defense budget imposes great constraints on other goods. Moreover, percentage changes are easier to compare across states and years.

Because my argument focuses on differences between deep and shallow treaties, the research design must measure alliance treaty depth and compare alliances. I use a measurement model to infer treaty depth from formal content, then connect alliances to military spending with a multilevel model. The next section describes the research design in more detail.

4 Research Design

The research design makes two contributions. First, I develop a latent measure of alliance treaty depth. Second, I employ that measure in a multilevel model, connecting alliance-level variation with state-level outcomes. 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

Observed alliance commitments reflect the underlying depth of the treaty. Deep treaties promise more military cooperation. Therefore, I use observed alliance characteristics to infer treaty depth.

Using observed treaty conditions as indicators of underlying depth could produce two measures. One possible measure is an additive index of treaty depth, where treaties with multiple commitments have higher index values. This assumes each indicator is equally important, which is unlikely. Instead, I employ latent variable modeling, which is a more flexible way to use observable characteristics to infer an underlying trait. This approach allows different variables to contribute more or less to depth, and includes ordinal variables. The measurement model estimates the correlations between alliance treaty content and the underlying formal depth to predict the depth of each treaty.

Measurement models have a rich history in political science (Clinton, Jackman and Rivers, 2004; Treier and Jackman, 2008; Fariss, 2014). Benson and Clinton use a mixed factor analysis model to measure alliance scope, depth and capability (Benson and Clinton, 2016; Quinn, 2004). I emulate Benson and Clinton's approach, but employ different indicators of depth and another estimator.

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 model breaks the dependence between the latent factors and marginal distributions by using copulas to encode the dependence among the latent variables.[8] Beyond the semiparametric aspect, this measurement model is a standard mixed factor analysis.

I estimated the measurement model using observed data from 289 alliances with military sup-

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

port in the alliance-level ATOP data (Leeds et al., 2002). I examine the 289 alliances with military support because prior studies of alliance participation and military spending emphasize these treaties. Indicators of treaty depth include military aid, bases, international organization formation, integrated military command, defense policy coordination and commitments to form companion military agreements. The argument suggests there is a single factor underlying variation in all six indicators, so I fit 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. Both the factor loadings and the latent factors match the theoretical argument.

Figure 1 summarizes the posterior distribution of the factor loading for each variable. All six indicators increase treaty depth, but the strength of the association varies. There is some variation in how strongly different factors are associated with latent depth. Policy coordination and integrated military command are the largest positive correlates of latent depth, and formal organization is only a little weaker. Bases, aid and companion military agreements contribute less to depth, but still increase the depth of an alliance.

The factor loadings in Figure 1 predict the latent depth of each alliance. A linear combination of these variables and loading generates a unique posterior distribution of latent depth for each alliance. I use the mean of that posterior to measure treaty depth, so each alliance has its own depth value.

Figure 2 describes the latent measure for ATOP alliances with defensive or offensive commitments from 1815 to 2016. The posterior mean captures the expected depth of an alliance treaty, conditional on its formal promises. There is substantial variation in the depth of alliance treaties. The top panel of Figure 2 is a histogram of mean treaty depth for alliances promising military support. Many treaties make no promises of deep military cooperation, and are clustered on around -0.8. 171 alliances have a depth score higher than -0.6 as they have at least one source of depth

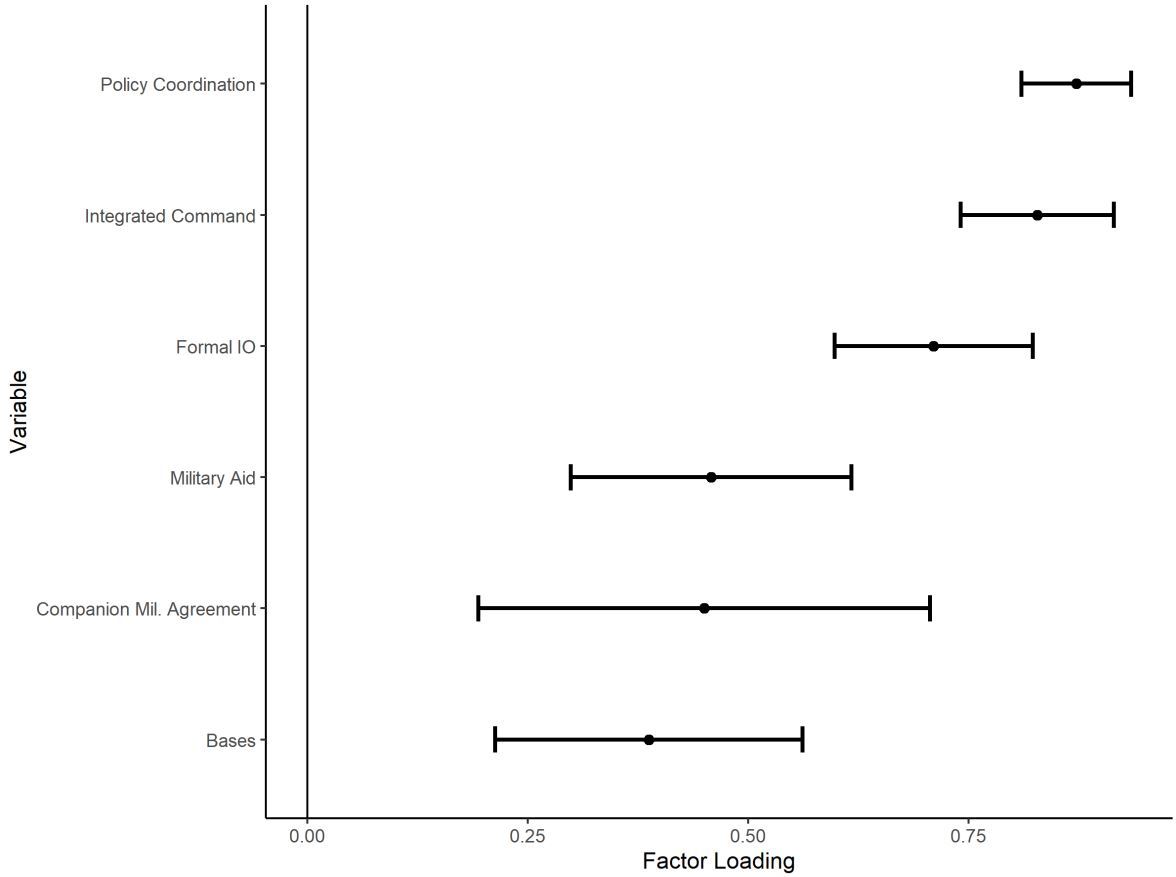

Figure 1: Summary of factor loadings on latent alliance treaty depth for all ATOP alliances from 1816 to 2016. Points mark the posterior mean and the error bars summarize the 90% credible interval. The loading captures the direction and strength of the correlation between each variable and latent depth.

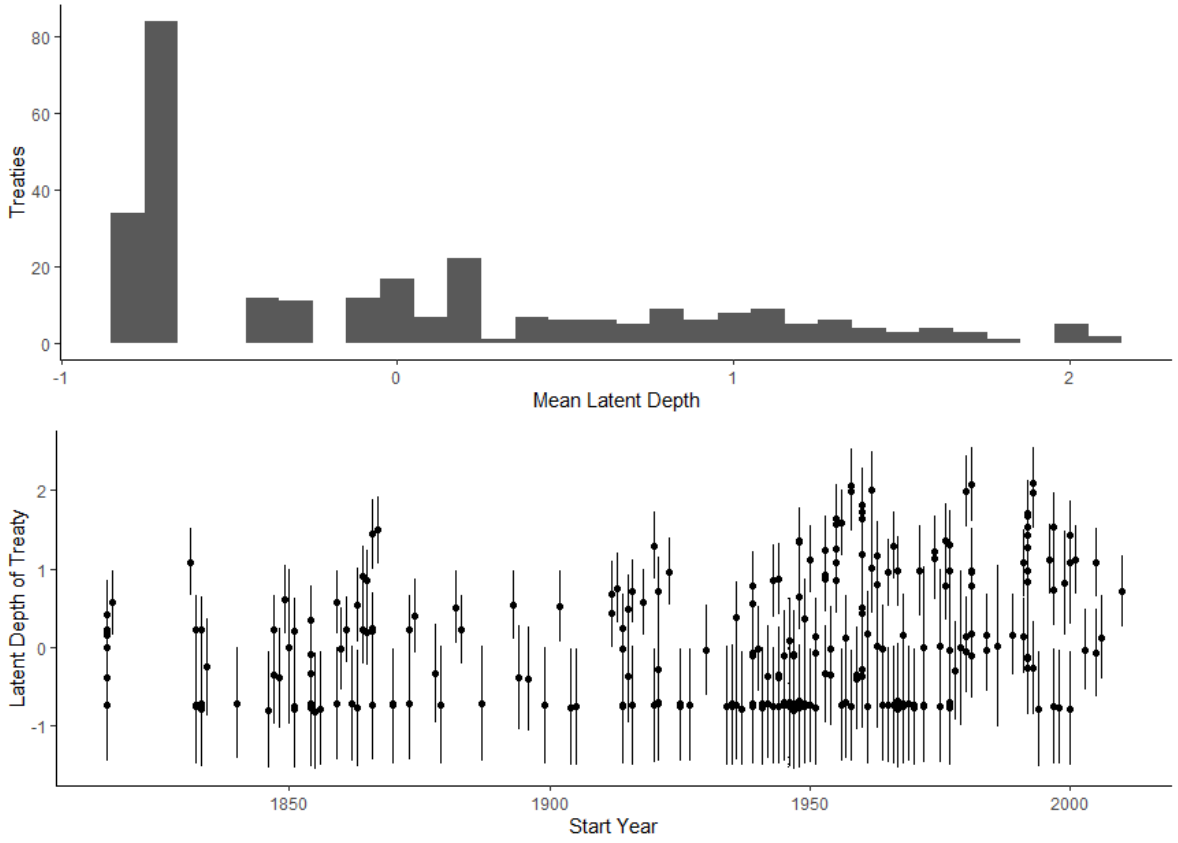

Figure 2: Summary of latent measure of alliance treaty depth for 289 alliances promising military support from 1816 to 2016. The top panel is a histogram of the expected of alliance treaty depth. The bottom panel plots mean treaty depth (points) and the standard deviation (error bars) against the start year of the treaty.

present. The bottom panel of Figure 2 plots the posterior means and uncertainty of the depth estimates against the start year of the treaty. 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 -0.11, and the mean is 0.02. The median treaty is the Organization of American States (OAS), which includes an Organ for Consultation (ATOP ID 3075). There are many shallow treaties, all of which only include military support. One such alliance is an 1855 pact between France, the UK and Sweden (ATOPID 1190) which only includes defense and consultation.

NATO (ATOPID 3180) has a mean latent depth of 0.37, placing it solidly in the third quartile. The main formal source of depth is a commitment to establish the Atlantic Council. According to the ATOP coding sheet for NATO, "There are numerous bilateral agreements among NATO members re: military aid, bases, etc. but they do not qualify as separate alliances, nor are they part of the overall NATO structure." The NATO treaty is still above average in formal depth, however.

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

The latent measure has some face, concept, and discriminant validity. For face validity, the Gambia-Senegal federation requires deeper cooperation than arms-length commitments of military support. The most shallow treaties promise little beyond military support, matching my conceptualization of treaty depth. Last, Figure 2 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 at the alliance level modify the impact of alliance participation on percentage changes in military spending at the state-year level. Therefore I use a multilevel model to estimate the association between treaty depth and military spending. The next section summarizes the multilevel model.

4.2 Multilevel Model

Multilevel modeling bridges levels of analysis (Steenbergen and Jones, 2002; Gelman and Hill, 2007). My specific model estimates heterogeneous effects of alliance participation on military spending, which are a function of alliance characteristics. I make inferences about the specific

impact of individual alliances and how alliance characteristics like formal depth modify the impact of alliance participation 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 traditional panel data model. But the additional components add substantial value, first by connecting the argument and research design.I argue that treaty depth modifies the impact of alliance participation on growth in military spending. Therefore, the multilevel model examines that prediction using an alliance-level coefficient. The theory compares alliances, and so does this key coefficient estimate.

Standard panel models use a state-level proxy for alliance characteristics, which compares states rather than alliances. This practice of aggregating alliances at the state-year level of analysis may produce misleading inferences (McElreath, 2016). Multilevel modeling retains the structure of the data, where states are members of multiple alliances, and each alliance has multiple salient characteristics. Depth is not the only possible source of differences in how alliance participation impacts military spending: allied democracy is at least one other factor (DiGiuseppe and Poast, 2016). Accounting for how multiple alliance characteristics chance the consequences of alliance participation is straightforward in a multilevel model.

Besides connecting alliance and state level variation, the multilevel model generates useful comparisons between alliances by estimating the specific impact of each alliance on members' military expenditures. By contrast, aggregating multiple alliances at the state level could mask different effects of individual treaties. Partial pooling of these alliance-specific parameters generates reasonable estimates for each alliance, which can be used to compare treaties. 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:

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

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.[9] σ is analogous to the error term in a frequentist regression as it captures unexplained variation. μ, the mean of the outcome, depends on several factors.

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

Percentage changes in spending are a function of an overall intercept α, state and year varying intercepts α^{st} and α^{yr} and a matrix of state-level control variables $\mathbf{W}$. These components comprise a standard random effects model. The $\mathbf{Z}\lambda$ term incorporates alliance participation.

$\mathbf{Z}$ is a matrix of state participation in alliances. Columns correspond to each of the a alliances in the data, and rows to state-year observations. 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.[10]

I use total allied spending in the alliance participation matrix because more capable alliances are more valuable (Johnson, Leeds and Wu, 2015). Major capability changes within alliances (the

[9] I estimate ν directly.

[10] Normalization keeps the parameters on similar scales, which is important for modeling. I selected normalization by comparing model fit with different ways of expressing allied capability.

columns of $\mathbf{Z}$) are driven by changes in alliance membership. $\mathbf{Z}$ 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' responsiveness 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:

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

and

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

In the alliance-level regression, $\mathbf{X}$ is a matrix of the l alliance-level control variables and α_{all} is the constant. Adding σ_{all} means predictions of λ are not deterministic— the alliance level regression contains an error term. A larger σ_{all} indicates more variation in how alliance participation impacts military spending. The second-level regression includes treaty depth, and each β parameter modifies the impact of alliance participation on percentage changes in military spending. The βs are like marginal effects in an interaction.

Treaty depth impacts military spending by changing the consequences of alliance participation.

Changing treaty depth shifts λ, which in turn affects military spending. β_1 compares deep and shallow treaties. Hypothesis 1 predicts β_1 will be negative for non-major powers, which implies deep alliances have a more negative effect on military spending.

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

4.3 Sample and Covariates

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 observations and 192 alliances.

The dependent variable is percent changes in military spending. Percentage changes in military expenditures is calculated as:

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

I used the Correlates of War Project's data on military spending to measure percentage changes in spending (Singer, 1988). The percentage change in spending is equal to changes in spending as a share of the previous year's military spending, so changes are relative to previous levels of spending. To address outliers, I apply the inverse hyperbolic sine transformation to this variable.[11]

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

Using percent 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 Hypothesis 1 predicts 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 variables that are correlated with 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 of GDP because GDP levels are non-stationary, and economic growth shapes the opportunity costs of military spending (Kimball, 2010; Zielinski, Fordham and Schilde, 2017).

The alliance-level regression contains the mean of the latent treaty depth— the key independent variable. 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 also control for issue linkages by creating a dummy indicator of whether the alliance promises any kind of economic cooperation (Poast, 2013; Long and Leeds, 2006). As another possible indicator of hierachical security relationships, I include a count of foreign policy concessions in the alliance including stipulations on competing alliances, not aiding enemies, third party ties, how to divide gains, and domestic intervention. 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.

Adjusting for all of these covariates helps address systemic differences between states and alliances from strategic selection into alliances. Regime type and external threat are especially important in that endeavor. The next section describes the results from the major and non-major samples.

5 Results

Results are based on 2,000 total 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 density.[12]

Because I use Bayesian modeling to estimate the association between treaty depth and percent changes in military spending, 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, Figure 3 summarizes the 90% credible intervals of the parameters, and I calculate the positive posterior probability for the treaty depth coefficient to assess Hypothesis 1.

The preponderance of evidence matches the predictions of Hypothesis 1. There is a 96% chance treaty depth is negatively associated with percent changes in military spending for non-major powers. Moreover, the the 90% credible interval for treaty depth does not include zero, though it is close. This is one important indicator that participation in a deep alliance increases non-major power military spending, relative to a shallow alliance.

Treaty depth also has a substantively important effect, which I assessed by simulating the effect of changing treaty depth from the minimum value of -0.8 to 1, which is in the third quartile. Holding other alliance covariates at their means, this change in depth reduces the hypothetical λ by .06. As a result of this change in λ, alliance members' military spending would fall by .02,

[12] See the appendix for more details on convergence and other robustness checks. An up to date appendix can be found here: https://github.com/joshuaalley/arms-allies/blob/master/appendix/appendix.pdf

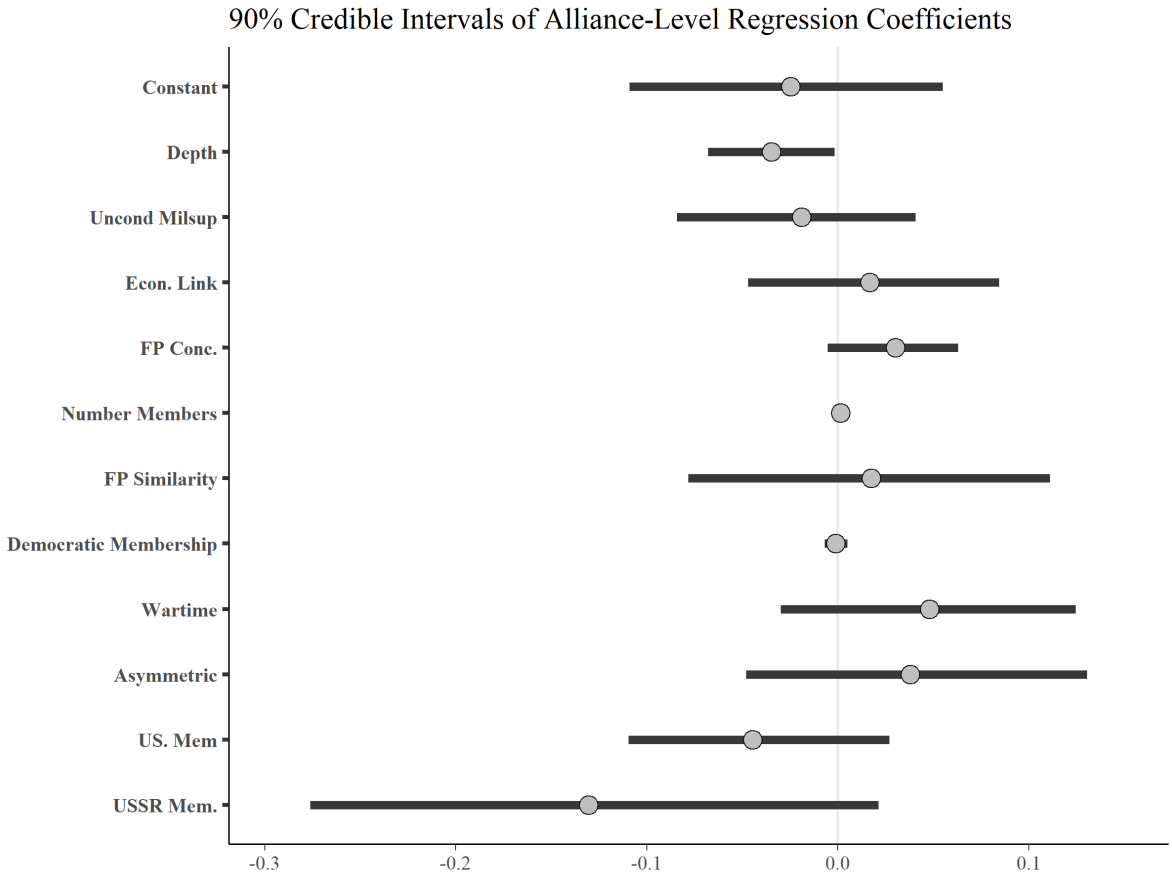

Figure 3: 90% credible intervals to summarize the posterior densities of coefficients in the alliance-level regression. Points mark the posterior mean, and the bars encapsulate the width of the credible interval.

assuming median allied capability values. The 90% credible interval for this change in depth ranges from -0.04 to 0.001.

The substantive importance of treaty depth is reflected by patterns in the λ parameters. Each λ measures the impact of treaty participation. If treaty depth has a large influence on alliance participation, it will appear in the λ estimates. There should be a negative trend in the expected value of λ as treaty depth increases in non-major power alliances. On average, deep alliances should have a more negative effect on members' percent changes in military spending than shallow alliances.

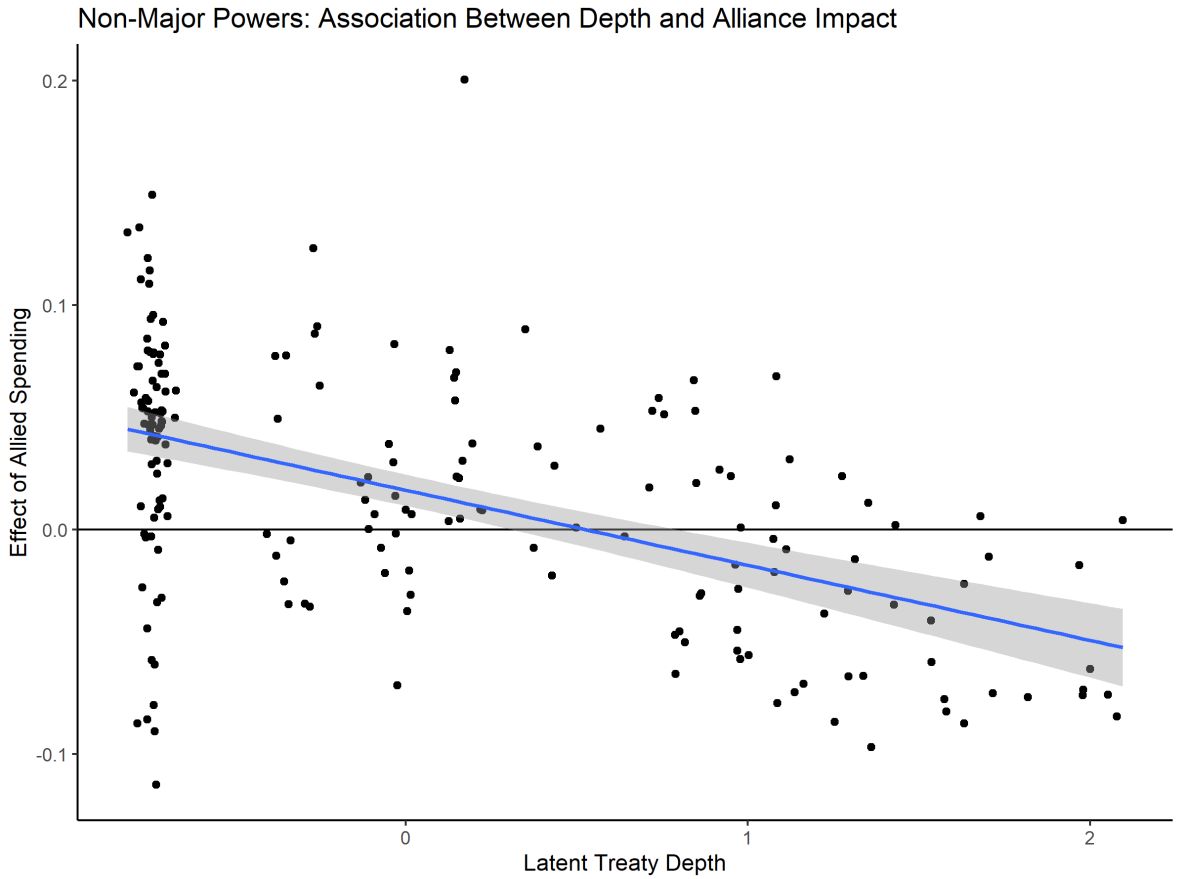

Figure 4: Scatter plots of trends in mean λ parameters and treaty depth in peacetime alliances. λ is the total impact of alliance participation on percent changes in military spending. For non-major powers, there is clear decrease in λ as treaty depth rises. Trend line estimated using linear regression.

Figure 4 plots the expected value of λ across the range of treaty depth. For non-major powers, the trend is negative. The correlation of .5 between mean λ and treaty depth matches the underlying logic of Hypotheses 1. Shallow treaties often increase percentage change in military spending among non-major powers, while most deep treaties have a negative effect. Because other treaty characteristics and unmeasured factors also influence the λ estimates, Figure 4 shows tremendous variation in how alliance participation impacts non-major power military spending.

Treaty depth is a key source of heterogeneity in how alliance participation impacts military spending. This has important consequences for our understanding of US alliances. Next, I summarize the impact of some US alliances on non-major power defense spending.

5.1 US Alliances

?? summarizes the impact of 18 post World War II alliances with US participation on percentage changes in non-major power military spending. Sixteen of the eighteen alliances have a negative mean λ estimate. Among US alliances, deep obligations are usually associated with more evidence for a negative impact of alliance participation on non-major power spending.

Most US alliances have above average depth, compared to other treaties. Eleven US alliances have greater than median depth, as Figure 5 shows. All but one of those US alliances with non-major powers after 1945 has a below average λ as well.

Deep promises of military cooperation are common in US alliances. If anything, formal treaty content understates the depth of US cooperation, because other depth is often established through separate bilateral arrangements The depth of US alliances limits leverage to address allied free-riding. By using bases, exercises and defense cooperation to reassure allies, the US loses leverage to demand allies spend more on defense. When coupled with substantial US military capability, US alliances are designed in a way that is likely to lead to free-riding.

In summary, I find that treaty depth decreases the impact of alliance participation on military spending. Deep alliances have a more negative impact on members military spending than shallow

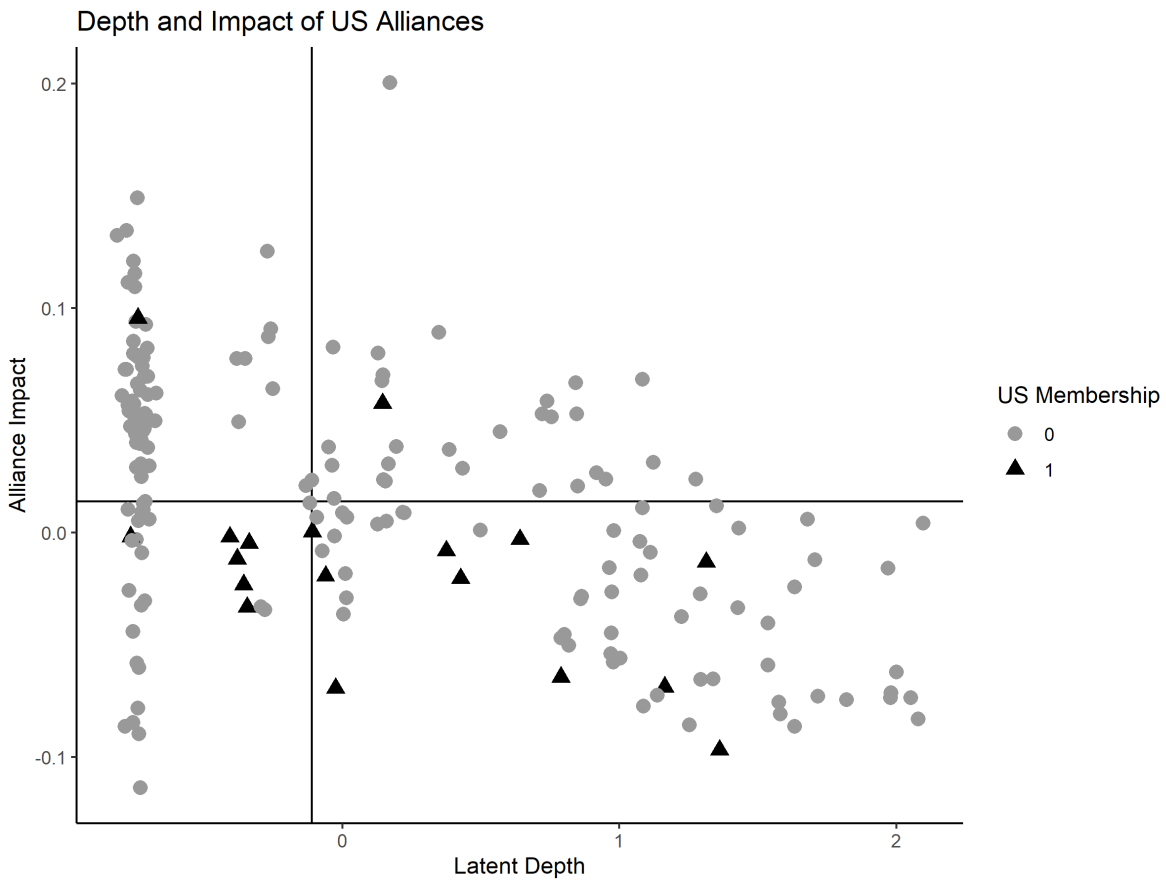

Figure 5: Impact of an alliance with the US on non-major power military spending, 1945-2007. This figure modifies Figure 4 by highlighting US alliances. The vertical and horizontal lines mark the median depth of non-major power alliances and median λ, respectively.

alliances. This has important consequences for our understanding 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 inaccurate. Rather, the impact of alliance participation on defense expenditures is lower in deep alliances, relative to shallow alliances. Due to treaty depth and other factors, alliance participation has a wide range of effects on military spending.

My argument builds on other conditional arguments about alliance participation and military spending (DiGiuseppe and Poast, 2016). Non-major powers are more likely to reduce percentage changes in military spending in deep treaties. In some cases, alliance participation clearly increases non-major power military spending, especially in shallow alliances.

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 responsiveness to allied military spending in a few treaties.

The results encompass specific and general studies, as I estimate both the impact of individual treaties and general differences between treaties. My research design emulates specific studies by estimating the unique impact of participation in individual treaties. The alliance-level coefficients compare treaties to capture the general consequences of alliance characteristics.

This paper has several limitations. First, the argument offers a cursory treatment of the domestic political economy of military spending, which is the subject of a rich literature (Whitten and Williams, 2011; Alptekin and Levine, 2012). Furthermore, domestic politics shapes how states de-

fine their foreign policy interests and the tools they use to pursue those interests (Fordham, 1998, 2011; Narizny, 2007). At the moment, my argument treats foreign policy interests as given.

My findings also 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 treaty depth may be conservative— it does not capture phenomena my argument expects should have a similar effect.

Strategic alliance design is another possible weakness of the test. Non-random selection into different alliances could produce systematic differences between members that are not captured for in my statistical model. I attempted to control for correlates of alliance treaty depth, but oversights are possible.

Despite these limitations, the argument and results provide valuable insights about alliance participation and military spending. I explain when alliance participation is associated with more or less 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. Deep alliance treaties reduce the impact of alliance participation on non-major power defense expenditures. Alliance participation has a wide range of effects on defense expenditures by non-major powers.

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).

Another implication is the distributional consequences of changes in military spending within states and among alliance members. By altering military spending, alliance design changes the domestic political economy of member states. The economic consequences of alliance participation are a possible subject for future research.

Besides their scholarly value, the argument and evidence help inform policy debates. Tradeoffs in alliance treaty design can guide our understanding of why some treaties lead to "free-riding" and possible policy responses. Because deep alliances reassure, they increase free-riding. States can use deep cooperation to reassure their allies, but they lose leverage to check free-riding as a result.

The United States is currently wrestling with the implications of treaty depth. Washington has often decried "free-riding" by allies who provide too little for their own defense (Lanoszka, 2015). But allies are able to free-ride partly because the United States makes deep formal commitments. Therefore, attempts to reassure NATO allies may outweigh the agreement for all allies to spend at least 2% of GDP on defense.

Reducing the depth of US alliances may generate credibility problems, however. Free-riding may be the price of credible commitments. Therefore, this is not an unconditional call to reduce the depth of US alliance commitments. Adjusting existing treaties may be more difficult than designing new alliances and it has other consequences. The full consequences of attempting to change 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.

Alptekin, Aynur and Paul Levine. 2012. "Military expenditure and economic growth: A meta-analysis." *European Journal of Political Economy* 28(4):636–650.

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

Anderton, Charles H. 1995. Economics of the Arms Trade. In *Handbook of Defense Economics*, ed. Keith Hartley and Todd Sandler. Vol. 1 Elsevier pp. 523–61.

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.

Beckley, Michael. 2015. "The Myth of Entangling Alliances: Reassessing the Security Risks of U.S. Defense Pacts." *International Security* 39(4):7–48.

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.

Devore, Marc R. 2013. "Arms Production in the Global Village: Options for Adapting to Defense-Industrial Globalization." *Security studies* 22(3):532–572.

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.

Dudley, Leonard and Claude Montmarquette. 1981. "The demand for military expenditures: an international comparison." *Public Choice* 37(1):5–31.

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 O. 2011. "Who wants to be a major power? Explaining the expansion of foreign policy ambition." *Journal of Peace Research* 48(5):587–603.

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 Multilevel/Hierarchical Models*. Vol. 1 Cambridge University Press New York, NY, USA.

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

Kapstein, Ethan B. 1991. "International Collaboration in Armaments Production: A Second-Best Solution." *Political Science Quarterly* 106(4):657–675.

Kim, Tongfi. 2011. "Why Alliances Entangle But Seldom Entrap States." *Security Studies* 20(3):350–377.

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

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

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

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

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

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

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

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

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

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

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

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

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

Moravcsik, Andrew. 1991. "Arms and Autarky in Modern European History." *Daedalus* 120(4):23–45.

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.

Narizny, Kevin. 2007. *The Political Economy of Grand Strategy*. Ithaca: Cornell University Press.

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.

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 inter-state, 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. 1984. "The Security Dilemma in Alliance Politics." *World Politics* 36(04):461–495.

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.

Whitten, Guy D. and Laron K. Williams. 2011. "Buttery Guns and Welfare Hawks: The Politics of Defense Spending in Advanced Industrial Democracies." *American Journal of Political Science* 55(1):117–134.

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.