

# Alliance Participation, Treaty Depth and Military Spending

Joshua Alley\*

September 5, 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 growth in military expenditures depends on the depth of military cooperation in the treaty. Close military cooperation modifies the impact of alliance participation on military spending by constraining free-riding through treaty value and allied leverage. Security-seeking non-major powers often free-ride on their allies by reducing defense spending, but deeper alliances increase their military spending growth. 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 growth in state military spending from 1816 to 2007. I find that deep alliances increase growth in military spending from alliance participation in non-major powers 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. States 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 contrary predictions and evidence on the question of alliance participation and military spending. One view, which I call the *force multiplier perspective*, expects alliance participation will reduce military spending (Morrow, 1993; Conybeare, 1994; DiGiuseppe and Poast, 2016). The other view, which I label the *foreign entanglement perspective*, 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 force multiplier and foreign entanglement perspectives largely ignores heterogeneity among alliances,<sup>1</sup> which is essential to alliance politics scholarship (Morrow, 1991; Leeds, 2003; 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 partners. Rather than make arms-length commitments, these treaties make unconditional promises of military support, promise aid, and obligate members to coordinate defense policy. For example, a 1963 alliance between Jordan and Iraq commits to broad military support as well as military aid, bases and institutions to govern military coordination.

---

<sup>1</sup>See DiGiuseppe and Poast (2016) for an important exception.

Adding treaty depth addresses a salient threat to cooperation in alliances— free-riding. Non-major powers are prone to free-ride as they focus on security and have high opportunity costs of military spending. Because alliance participation provides extra security, it often reduces non-major power growth in military spending. Greater treaty depth attenuates that negative correlation between alliance participation and military spending among non-major powers by increasing allied leverage and treaty value.

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 growth in military spending. I fit the model on a sample of non-major power states from 1816 to 2007. I find that deeper alliances increase the impact of alliance participation on growth in non-major power military spending, relative to shallow treaties.

For non-major powers, deep treaties provide more benefits at the cost of freedom to reduce military spending. 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 take different views of alliances. Advocates of restraint argue that US allies spend too little on defense, which increases US defense spending (Preble, 2009) so the US should withdraw many of its alliances. 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 the limited formal depth of most US treaties. While my argument gives a partial accounting of alliance costs and benefits, it suggests adding or maintaining deep military cooperation could help check free-riding. US allies spend less on defense in part because the United States usually offers shallow alliances. While limits on treaty depth limit entanglement abroad, they also constrain US leverage.

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. The final section concludes with a discussion of the results and implications for scholarship and policy.

## **2 Force Multiplier or Foreign Entanglement?**

Scholarship on alliance participation and military spending is divided between force multiplier and foreign entanglement 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 Force Multiplier**

Force multiplier 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 military spending. 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.<sup>2</sup>

Another force multiplier argument 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 more security without requiring additional military spending (Morrow, 1993; Conybeare, 1994). Given extra security, states rely on their allies and reallocate military spending to other goods.

---

<sup>2</sup>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, the foreign entanglement perspective asserts that alliance participation increases military expenditures.

## **2.2 Foreign Entanglement**

The foreign entanglement view contains several arguments. 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 the force multiplier perspective.<sup>3</sup>

---

<sup>3</sup>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 force multiplier and foreign entanglement 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.<sup>4</sup>

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 are two negative, four 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	
DiGiuseppe and Poast (2016)	X		
Horowitz, Poast and Stam (2017)		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 support for force multiplier arguments comes from specific designs (Barnett and Levy, 1991; Morrow, 1993; Sorokin, 1994; Plümper and Neumayer, 2015). Other specific studies find increased spending by alliance members, however (Conybeare and Sandler, 1990; Chen, Feng and Masroori, 1996).

<sup>4</sup>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 have heterogeneous effects. Conflict (Leeds, 2003; 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 undermine 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 alliance credibility from 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 reduces non-major powers' tendency to free-ride by increasing allied leverage and making the treaty more valuable. Non-major powers have less freedom to reduce defense spending under greater influence from their partners. As for treaty value, inadequate defense spending weakens alliances and could antagonize partners into reducing or eliminating a valuable commitment. Therefore, non-major power growth in military spending from alliance participation is higher in deep treaties, relative to 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 growth in non-major power military spending.

### **3.1 Free-Riding 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).

Free-riding through low defense spending is one form of opportunism in alliances.<sup>5</sup> 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). Reduced defense spending can be a form of shirking that leaves allied states bear a greater military burden.

Enforcing cooperation around free-riding is difficult.<sup>6</sup> 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 substantial security and can reallocate resources to other

---

<sup>5</sup>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.

<sup>6</sup>See Poast (2012, 2013) and Long and Leeds (2006) on how issue linkages address abandonment, which is another salient form of opportunism in alliances.



priorities. The ability to reduce defense spending and spend more on other social goods sometimes motivates states to form alliances (Kimball, 2010; Allen and DiGiuseppe, 2013).

Deep alliances provide two mechanisms that address incentives to free-ride. Stipulating deep cooperation increases both allied influence and the value of the alliance. As a result, members have less freedom and motivation to reduce defense spending.

### **3.2 Depth of Military Cooperation**

Alliance depth is the extent of defense cooperation formalized in the treaty. There are two sources of depth. The first is what support the alliance promises and when it applies. Formal peacetime defense cooperation is the second source of depth. In general, deep alliances stipulate close cooperation between members. By contrast, shallow alliances promise more arms-length cooperation.

The core obligations of the alliance add depth by stipulating how and when the treaty applies. Less than half of alliances promise military support. Commitments of non-aggression, neutrality or consultation are minimal, arms-length obligations. Defensive or offensive commitments of military intervention, by contrast, require closer cooperation to manage the alliance (Snyder, 1997). Promises of military support are a key source of depth.

Besides depth from military support, depth comes from conditions on primary obligations. Conditional alliances limit military intervention to specific circumstances. For example, an alliance may only promise defensive intervention if the fighting was unprovoked or involves a specific state. Alliance members use conditions to address moral hazard problems in alliances (Benson, 2012; Benson, Meirowitz and Ramsay, 2014) and manage their own obligations (Chiba, Johnson and Leeds, 2015). Conditional obligations reduce the depth of the alliance by limiting when members must cooperate. Unconditional alliances apply to many circumstances, increasing the extent of defense cooperation between members.

Beyond the core obligations and conditions, states add depth to alliances by incorporating

promises of peacetime defense cooperation. This cooperation has 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 through policy coordination and regular interactions.

Increasing military coordination adds ties between alliance members beyond a promise of military support, which adds substantial depth to the alliance. 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 far 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 is a mechanism to reduce free-riding. Depth restricts the ability of members to free-ride in two ways. Adding depth to an alliance gives partners more leverage and increases the value of alliance participation.

Partners have more influence in a deep alliance because deep defense cooperation provides additional leverage. Common defense policies, military aid, international organizations, aid and bases provide multiple avenues to check free-riding. Depth allows allies to work out a division of labor that limits free-riding.

Deep cooperation also aligns defense policy across issues beyond formal treaty content. As depth increases, the nature of the relationship the alliance codifies changes. Arms-length relationships are more anarchic, while close cooperation is more hierarchical. In hierarchical relationships, partners give up control on issues beyond those the treaty formally addresses. Lake (1996) calls this a loss of residual control, where one partner gives up control over issues outside the formal agreement. Deep alliances often reflect past territorial control, attempts to form federations, or wartime coalitions. all of which are more hierarchical than arms-length agreements among equals.

As a result, deep treaties give alliance partners more influence over issues like military spend-

ing. Inasmuch as allies understand this arrangement, direct coercion over free-riding will be rare in deep alliances. Additional leverage is not the only check on free-riding, however.

Members of deep alliances also lose some motivation to reduce defense spending because deep cooperation is more valuable than arms-length alliances. Besides military support, members benefit from policy coordination and transfers. Depth increases the foreign policy benefits of alliance participation by helping members coordinate policy stances and defense efforts. Deep cooperation is also a clear signal of commitment.

Because deep alliances are more valuable, members are less likely to free-ride. Insufficient defense spending could antagonize allies and lead them to abandon the alliance. Free-riding in a deep alliance endangers more than a promise of military support, it imperils close defense cooperation.

If deep alliances are a valuable source of influence over allied behavior, 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.<sup>7</sup> 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 restrict free-riding among alliance members. This is especially relevant for non-major powers. These states are more likely to free-ride.

### **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 immediate security. The emphasis on security creates

---

<sup>7</sup>Also deep alliances do not solve every problem, as more institutionalized alliances may not be more reliable (Leeds and Anac, 2005).

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 emphasis on security, encourage free-riding.

Alliance participation often reduces non-major power defense spending. Greater alliance treaty depth attenuates this negative correlation, however. Allied leverage and treaty value can overcome non-major power incentives to free-ride.

To illustrate the logic, consider two related alliances from the inter-war period. In 1925, France and Poland formed a defense pact promising military support in the event of an unprovoked German attack (ATOPID 2135). The treaty contained no other obligations and provided enough security for Poland to reduce growth in their defense budget. On the other hand, a 1920 treaty between France and Belgium (ATOPID 2055) added commitments of military aid and a coordinated occupation of the Rhineland to defensive obligations. The Franco-Belgian treaty increased growth in Belgian military spending.

These brief examples and the argument suggest that treaty depth modifies the impact of alliance participation on non-major power military spending. Greater depth increases the usual negative impact of alliance participation on non-major power military spending. Deeper alliances will lead to higher growth in non-major power military spending than shallow treaties.

HYPOTHESIS 1: AS ALLIANCE TREATY DEPTH INCREASES, GROWTH IN NON-MAJOR POWER MILITARY SPENDING FROM ALLIANCE PARTICIPATION WILL INCREASE.

### 3.4 Interpreting Predictions about Spending Growth

When treaty depth increases growth in non-major power spending, the level of military spending may only change in a counterfactual sense. Higher expenditure growth need not produce larger defense budgets. Only positive growth in spending increases the level of military expenditures. When increasing growth in military spending remains negative, spending levels will fall.

For example, a non-major power whose military budget goes from 10 million to 9\$ million due to an alliance experiences -10% growth in spending. If the same alliance had greater depth and spending decreased from 10 million to 9.5\$ million, growth falls to -5%, but the level still falls. Therefore, my arguments about growth apply to the level of military spending through counterfactual reasoning Fearon (1991). Increasing growth implies a state would have spent less on the military, all else equal.

Figure 1 illustrates a possible implication of treaty depth about military spending by plotting a hypothetical link between treaty depth and the impact of alliance participation on military spending. The shape and level of the curve are not essential to the argument. Figure 1 only shows the need to be careful moving from changes in spending growth to changes in levels.

For non-major powers, higher growth in spending in a broad treaty depth could remain negative. Deep treaties attenuate the negative correlation between alliance participation and military spending in non-major powers. But so long as growth from alliance participation remains negative, the level of military spending will fall.

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.

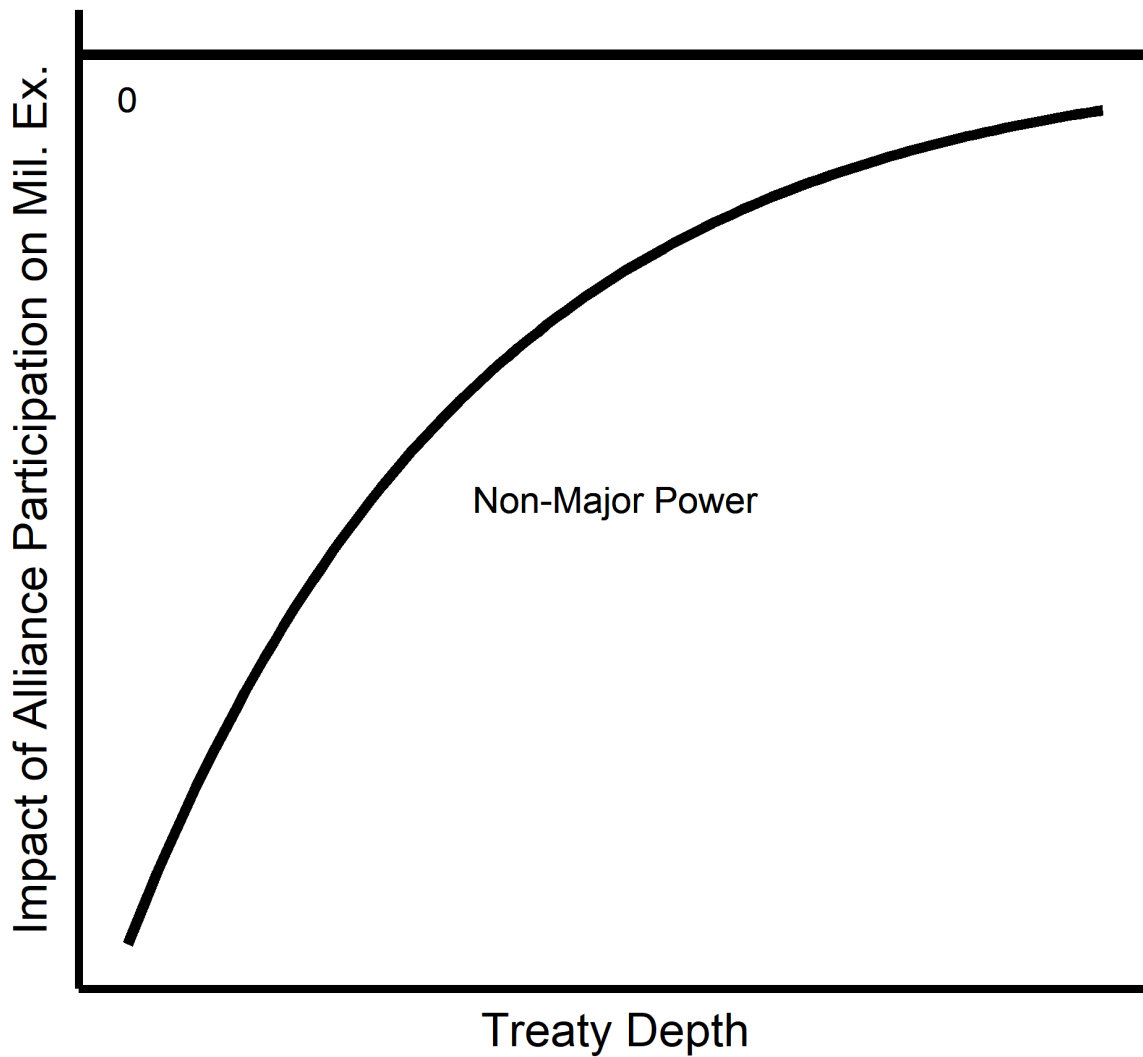


Figure 1: Abstract illustration of the link between treaty depth and the impact of alliance participation on military spending. The curve represents a possible impact of alliance participation on growth in non-major power military spending across the range of treaty depth.

## 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. 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.<sup>8</sup>

Beyond the semiparametric aspect, this measurement model is a standard mixed factor analysis. Factor analysis estimates the association between observed variables and some latent factor. Each observed variable has a factor loading— the association between the observed variable and the latent variable. Like standardized regression coefficients, factor loadings range from -1 to 1, so observed variables are positively or negatively correlated with the latent factor. For each observation, a linear combination of observed alliance characteristics predicts latent treaty depth, like a regression with an unobserved outcome.

I estimated the model using observed data from all 745 alliances in the alliance-level ATOP data (Leeds et al., 2002). Indicators of treaty depth are divided between the primary obligations and additional cooperation. Primary obligations include promises of defensive support, offensive support, neutrality, consultation, non-aggression and whether military support is unconditional. I created the indicator of unconditional military support using ATOP data on whether offensive or defensive promises in the alliance are conditional on specific circumstances. Indicators of defense cooperation 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 latent factor underlying variation in all 11 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. Because treaty depth is the quantity of interest, I focus on the posteriors of the latent factor.

Each alliance has a unique posterior distribution of its latent depth. I use the mean of that

---

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



posterior to measure treaty depth. Figure 2 describes the latent measure for ATOP alliances with defensive or offensive commitments from 1815 to 2016. I examine the 289 alliances with military support because prior studies of alliance participation and military spending emphasize these treaties.<sup>9</sup> 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. Most treaties are concentrated between .5 and 1.5 on the latent scale, but approximately 40 treaties fall outside this range. The bottom panel of Figure 2 plots the posterior means and uncertainty in those 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 mean of treaty depth is 0.92, and the median is 0.89. The median treaty is a 1967 agreement between Czechoslovakia and East Germany (ATOP ID 3550). Two of the most shallow treaties are twin 1870 neutrality and offense pacts, one between France and Britain (ATOPID 1300), and the other between Prussian and Britain (ATOPID 1295).<sup>10</sup> Britain used these alliances to protect Belgian neutrality during the Franco-Prussian war. Both treaties have only offensive and neutrality promises, conditional on France and Prussia respecting Belgian neutrality.

NATO (ATOPID 3180) has a mean latent depth of 0.69, placing it just above the first quartile of military alliances. NATO's two costly provisions are the conditional defense obligations in Article 5 and establishing 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

---

<sup>9</sup>I estimated the measurement model on all alliances, including those without military support to capture the contributions of primary obligations to treaty depth.

<sup>10</sup>Sampling variation produces small differences in the scores. These alliances do not enter the later test, as they only involve major powers.

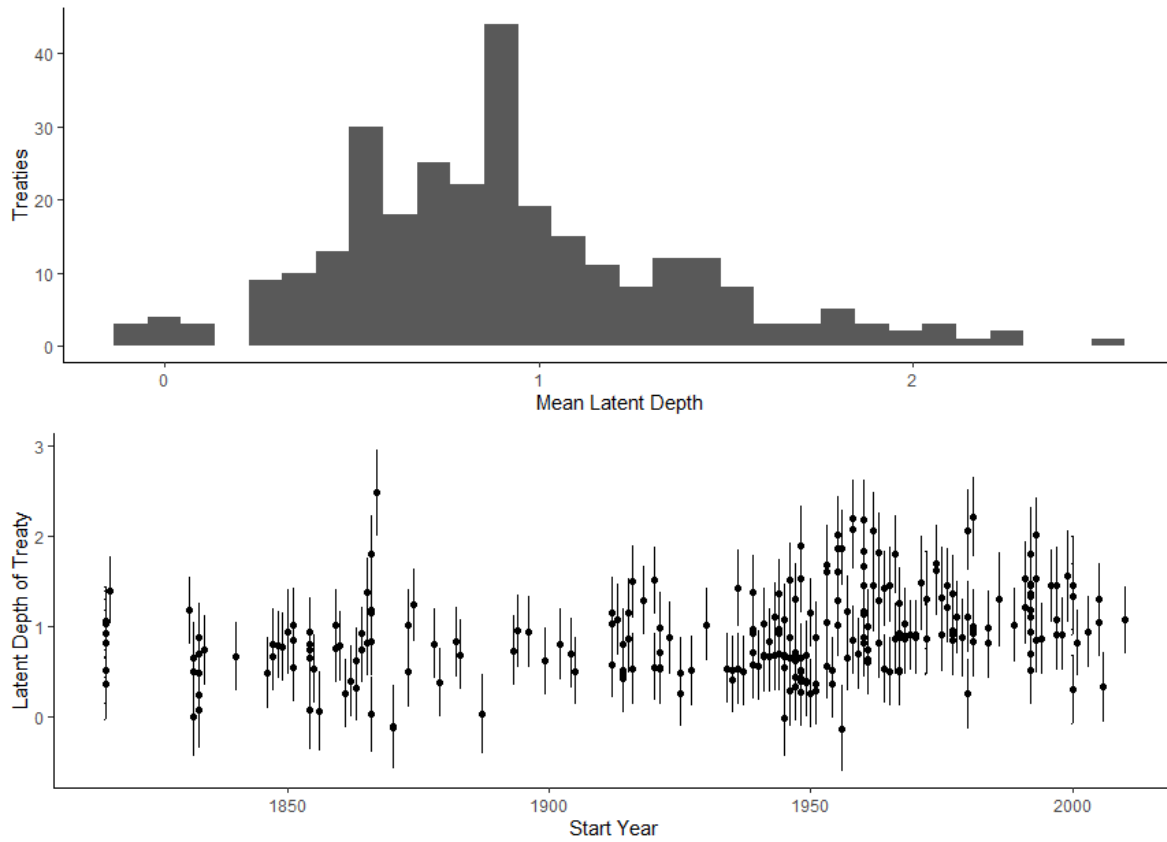


Figure 2: Summary of latent measure of alliance treaty depth for 289 alliances promising military support from 1816 to 2016. The top panel is a histogram of the expected of alliance treaty depth. The bottom panel plots mean treaty depth (points) and the standard deviation (error bars) against the start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

do not qualify as separate alliances, nor are they part of the overall NATO structure.” The NATO treaty is below average in formal depth in part because additional commitments fall outside the formal treaty.

The three deepest treaties are an 1867 alliance between Prussia and Hesse (ATOPID 1290), a 1958 alliance between the UAE and Yemen (ATOPID 3345), and a 1981 pact between Gambia and Senegal. All these alliances include unconditional military support and extensive defense cooperation. The Prussia–Hessian treaty formed on the eve of war with Austria and includes military aid, integrated military command and basing agreements in addition to offensive and defensive promises. The other two treaties attempted to establish a federation among their members through defensive support, basing, and defense policy coordination. Both federations were eventually undone by policy disputes, but that does not diminish the breadth of cooperation they sought.

The latent measure has some face, concept, and discriminant validity. For face validity, the Gambia-Senegal federation requires deeper cooperation than a promise to respect Belgian neutrality. The most shallow treaties promise little beyond conditional military support, matching my conceptualization of treaty depth. Last, Figure 2 shows that this measure can distinguish between broad and narrow commitments.

My argument uses variation in treaty depth between alliances to explain growth in military spending. Differences in depth at the alliance level modify the impact of alliance participation on growth 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 is a natural way to bridge levels of analysis. My model estimates heterogeneous effects of alliance participation on military spending and uses treaty characteristics to explain variation across alliances. I make simultaneous inferences about the specific impact of

individual treaties and the general role of treaty characteristics like depth. 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 adds complexity to a traditional panel data model. But the additional components generate novel inferences to connect the argument and research design. The hypothesis emphasizes differences in treaty depth, so the model contains a corresponding coefficient.

Most panel models use a state-level proxy for alliance characteristics, which compares states rather than alliances. Aggregating depth 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. Connecting the alliance and state level of analysis generates inferences of alliance-level variation impacts annual growth in state military expenditures.

Besides connecting alliance and state level variation, the multilevel model generates novel comparisons between alliances by estimating the specific impact of each alliance on members' military expenditures. Partial pooling of these alliance-specific parameters generates reasonable estimates for every treaty, 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 is similar to a random effects panel data regression. A second alliance-level regression modifies parameters in the first regression, 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— growth in military spending. I model spending growth using a t-distribution with degrees of freedom  $\nu$  to address heavy tails.<sup>11</sup>  $\sigma$  is analogous to the error term in a frequentist regression as it captures unexplained variation in spending growth.  $\mu$ , the mean of the outcome, depends on several factors.

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

Growth in spending is a function of an overall intercept  $\alpha$ , state and year varying intercepts  $\alpha^{st}$  and  $\alpha^{yr}$  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.

I use total allied spending in the alliance participation matrix because more capable alliances provide extra benefits. Increasing allied capability makes promises of military support more valuable (Johnson, Leeds and Wu, 2015).  $\mathbf{Z}$  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.

$\lambda$  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  $\lambda$  parameters capture alliance members' responsiveness to allied capability. Each alliance has a unique  $\lambda$ . The  $\lambda$  parameters have shared distribution, so I assume alliances are similar but different in how they impact growth in military spending.

---

<sup>11</sup>I estimate  $\nu$  directly.

The second part of the multilevel model uses alliance characteristics to predict how alliance participation is associated with growth in military spending. The  $\lambda$  parameters are the outcome in an alliance-level regression. As a result, the impact of alliance participation on members' military spending depends on treaty characteristics, including depth. In this second-level regression:

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

and

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

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

Treaty depth impacts military spending by changing the consequences of alliance participation. Changing treaty depth shifts  $\lambda$ , which in turn affects growth in military spending.  $\beta_1$  compares deep and shallow treaties. Hypothesis 1 predicts  $\beta_1$  will be positive for non-major powers.

Consider one observation as an example of how the model works. Growth in Argentina's military spending in 1955 depends on Argentina's economic growth, political regime, conflict participation, and rival military spending. Argentine participation in the Rio Pact and OAS also

changes growth in spending through allied capability.

$$\begin{aligned} \text{Argentina 1955: Spending Growth} = & \text{Overall mean} + \text{Argentine Intercept} + 1955 \text{ Intercept} + \text{Argentine Character} \\ & + \lambda_{OAS} * \text{OAS Expenditure} + \lambda_{Rio} * \text{Rio Pact Expenditure} \end{aligned} \quad (5)$$

$\lambda_{OAS}$  and  $\lambda_{Rio}$  capture the impact of participating in each alliance and are a function of the alliance level regression. Characteristics of the OAS and Rio Pact alter their respective  $\lambda$  parameters. Other alliances have no impact on growth in Argentine military spending.

In this model, the  $\beta$  parameters capture the general association between key alliance characteristics and military spending. The  $\lambda$  parameters express the impact of participation in each alliance, permitting heterogeneous effects of different treaties. 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 growth in military spending. Growth in military expenditures is calculated as:

$$\text{Growth Mil. Expend} = \frac{\text{Change Mil. Expend}_t}{\text{Mil. Expend}_{t-1}} \quad (6)$$

I used the Correlates of War Project's data on military spending to measure growth (Singer, 1988).

Growth 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 extreme outliers, I apply the inverse hyperbolic sine transformation to the growth variable.<sup>12</sup>

Using military expenditure growth 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 growth 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. 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 hierarchical security relationships, I include a count of foreign policy concessions in the alliance including stipulations on competing alliances, not aiding

---

<sup>12</sup>This transformation applies to positive, negative and zero values. It has little impact on most growth 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.



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.<sup>13</sup>

Because I use Bayesian modeling to estimate the association between treaty depth and growth 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 coefficients to assess Hypotheses 1.

The preponderance of evidence matches the predictions of Hypotheses 1. Although the 90% credible barely overlaps zero, there is a 92% chance treaty depth is positively associated with growth in military spending for non-major powers. Treaty depth also has a substantively important effect. For non-major powers, the mean of the treaty depth coefficient is 0.02, and median growth in military expenditures is 0.06. In expectation, greater treaty depth adds about a third to typical

---

<sup>13</sup>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>

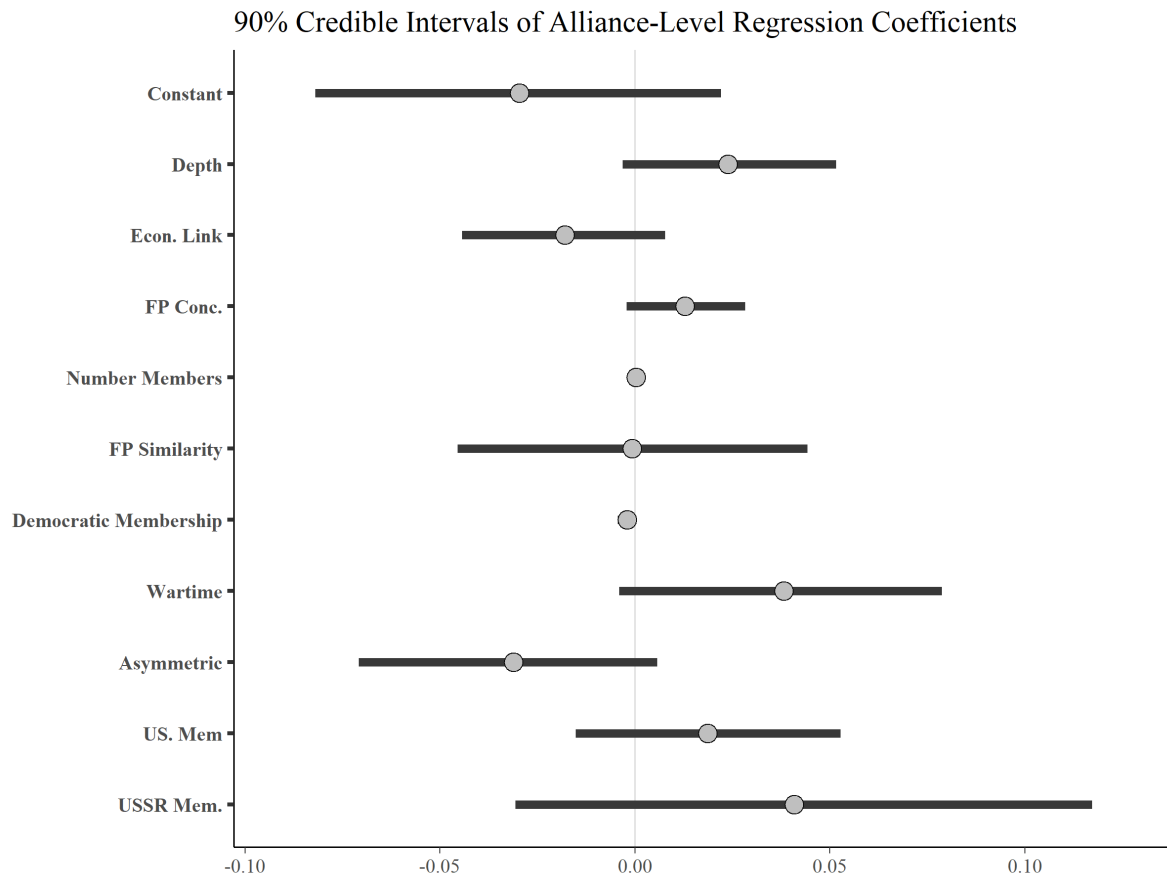


Figure 3: Posterior densities of coefficients in the alliance-level regression of the multilevel model.

growth in minor power military expenditures.<sup>14</sup> This change in spending is a plausible effect—5.1% of the 2018 US defense budget was spent directly on NATO.<sup>15</sup>

I also assess substantive importance by looking at patterns in the  $\lambda$  parameters. Each  $\lambda$  measures the impact of treaty participation. If treaty depth has a large influence on alliance participation, it will appear in the  $\lambda$  estimates. There should be a positive trend in the expected value of  $\lambda$  as treaty depth increases in non-major power alliances. On average, shallow alliances should have a more negative effect on members' growth in military spending than broad alliances.

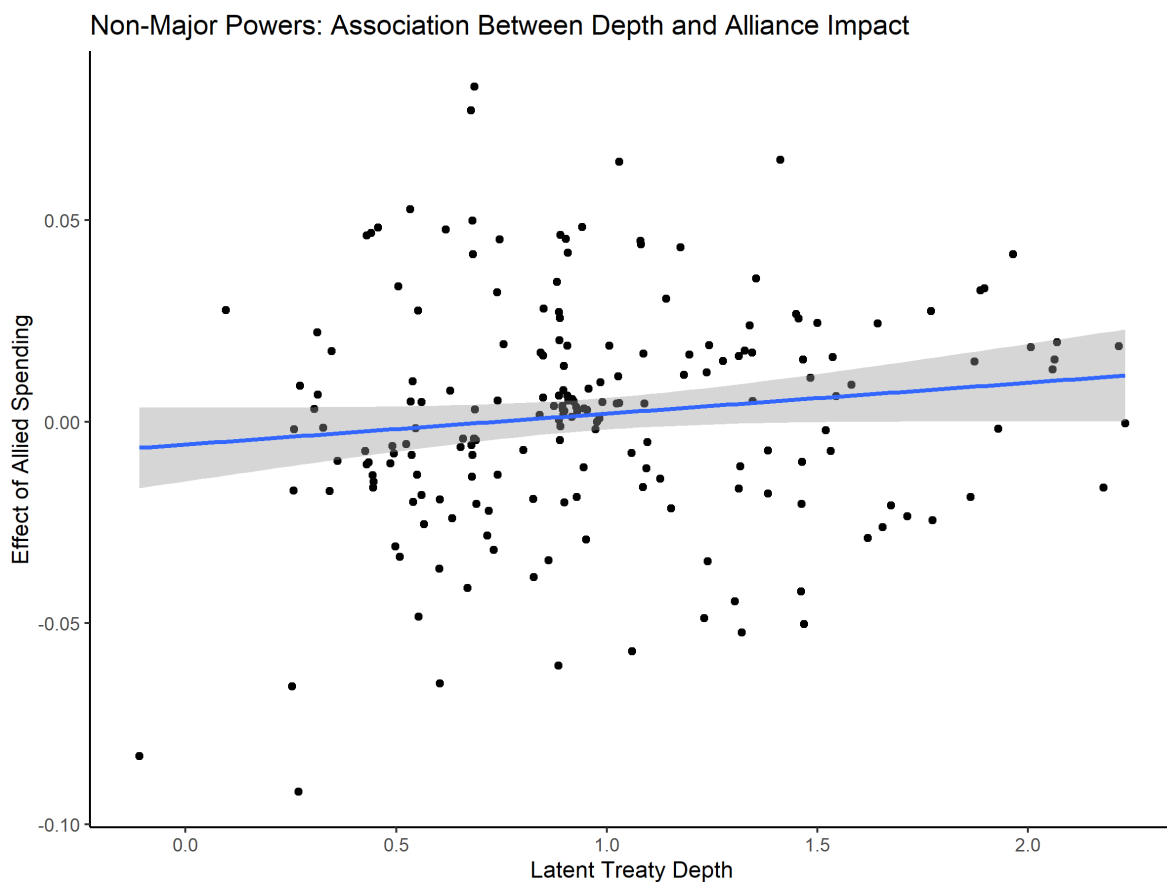


Figure 4: Scatter plots of trends in mean  $\lambda$  parameters and treaty depth.  $\lambda$  is the total impact of alliance participation on growth in military spending. For non-major powers, there is a slight but perceptible increase in  $\lambda$  as treaty depth rises. Trend line estimated using linear regression.

<sup>14</sup>Of course, uncertainty in the posterior estimates includes larger and smaller effects.

<sup>15</sup>See the following IISS blog post: <https://www.iiss.org/blogs/military-balance/2018/07/us-and-nato-allies-costs-and-value>.

Figure 4 plots the expected value of  $\lambda$  across the range of treaty depth. For non-major powers, the trend is positive. The correlation of .14 between mean  $\lambda$  and treaty depth is statistically significant, but of small substantive magnitude. Still, the trend matches the underlying logic of Hypotheses 1. Narrow treaties often reduce growth in military spending among non-major powers, while most broad treaties have a less negative effect. Because other treaty characteristics and unmeasured factors also influence the  $\lambda$  estimates, Figure 4 shows tremendous variation in how alliance participation impacts non-major power military spending. Treaty depth is only one source of heterogeneity in how alliance participation impacts military spending.

The statistical model also allows us to learn about individual alliances. Each  $\lambda$  parameter captures the impact of an individual alliance. Next, I summarize the impact of some US alliances on non-major power defense spending.

## 5.1 US Alliances

Figure 5 summarizes the impact of 18 post World War II alliances with US-membership on growth in non-major power military spending. The figure includes both the posterior mean of the  $\lambda$  parameter for each alliance and the 90% credible interval. Thirteen of the 18 alliances have a negative mean  $\lambda$  estimate. Among US alliances, shallow obligations are usually associated with more evidence for a negative impact of the alliance.

Moreover, most US alliances have limited depth, compared to other treaties. Only three US alliances have above average depth, as Figure 5 shows. Ten of eighteen US alliances with non-major powers after 1945 have a below average  $\lambda$  as well.

Calculating the relative positive or negative posterior probability of each  $\lambda$  is an alternative way to assess how US allies respond to allied capability. As with regression coefficients, 90% positive or negative probability is a common threshold for a clear effect. One US alliances has a positive  $\lambda$  parameter with greater than 90% positive posterior mass. An alliance between the US and other belligerents in World War II (ATOPID 3010) has a mean  $\lambda$  of 0.0506. The positive effect of this

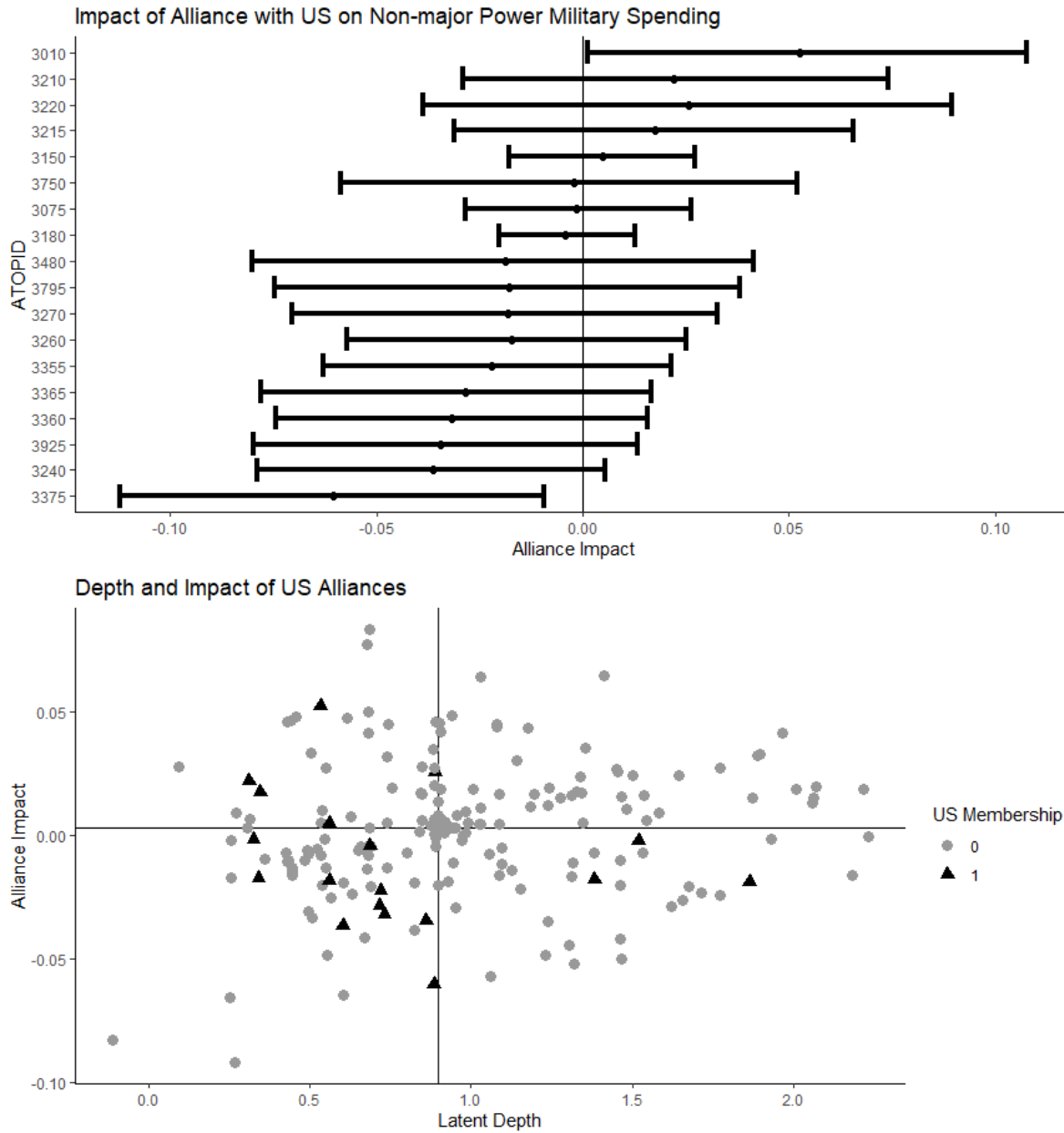


Figure 5: Impact of an alliance with the US on non-major power military spending, 1945-2007. In the top figure, the point summarizes the mean  $\lambda$  parameters and the error bars capture the 90% credible interval. The bottom figure plots the mean  $\lambda$  value against the depth of the treaty and highlights US alliances. The vertical and horizontal lines mark the median depth of non-major power alliances and median  $\lambda$ , respectively.

alliance is driven by the wartime nature of the treaty and a below average depth score of 0.56.

There are three alliances with more than 90% negative posterior mass in  $\lambda$ . The US alliance with Japan (ATOPID 3375) is close to the median alliance in its depth as it scores .89 on the latent measure. A formal alliance between the US and Israel from 1981 to 1991 is also close to median depth. Last, the US alliance with South Korea (ATOPID 3240) has a depth of .63, as it offers conditional military support and provides basing for US troops.

Several other alliances have between 75 and 80% negative posterior mass and below average depth. US alliances with Turkey (ATOPID 3360) and Iran (ATOPID 3365) score around .74 on depth. The US alliance with Taiwan (ATOPID 3270) has .58 latent depth, as it includes only a conditional defense obligation.

NATO, the most well-known US alliance, is below average in formal depth. The expected impact allied capability on growth in defense spending among non-major powers in NATO is only -0.004, however.<sup>16</sup> The small negative effect of NATO may reflect extensive defense cooperation outside the formal treaty obligations. Defense contacts, regular exercises and bases among NATO members may have limited the negative impact of the alliance on European defense spending. Many of these arrangements are outside the formal scope of NATO.

Limited formal depth is standard for US alliances. Only one US alliance offers unconditional military support— a 1963 defense pact with Spain (ATOPID 3480). Moreover, few US alliances promise additional defense cooperation. US alliances are exemplars of anarchic, arms-length military cooperation (Lake, 1996).

Most defense cooperation with the US takes place through separate agreements from the formal alliance. While constraining the depth of US treaties limits formal entanglement abroad, it also weakens the connection between defense cooperation and the alliance. When defense cooperation is distinct from alliance obligations, non-major power US allies have more freedom to reduce

---

<sup>16</sup>One possible explanation for this estimate is that it does not include major powers like the UK, France and post-unification Germany that also rely of US protection.

defense expenditures.

## 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. Treaty depth increases non-major power expenditure growth from alliance participation.<sup>17</sup> 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<sup>18</sup> and clarifies where the force multiplier and foreign entanglement perspectives best apply. The force multiplier perspective covers non-major powers well. But non-major powers cannot reduce growth in military spending in all alliances, because deep treaties constrain. In some cases, alliance participation clearly increases non-major power military spending, and depth contributes to those positive correlations.

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 itself the subject of a rich literature (Whitten

---

<sup>17</sup>See the appendix for a robustness check of the multilevel model results with a single-level regression.

<sup>18</sup>DiGiuseppe and Poast (2016)

and Williams, 2011; Alptekin and Levine, 2012). Furthermore, domestic politics shapes how states define 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 growth in military spending among non-major powers, addressing debate between the force multiplier and foreign entanglement 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. Though alliance participation often reduces non-major power defense spending, spending growth is increasing in alliance treaty depth. 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 account-



ing for heterogeneity among alliances. Alliances have heterogeneous effects on the risk of war, trade and military spending (Leeds, 2003; 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 growth in military spending, the design of international alliances alters the domestic political economy of member states. The domestic 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. States may be able to check allied free-riding through increasing the depth of their alliance commitments. Greater influence requires deeper engagement abroad, however.

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 shallow formal commitments. “Foreign entanglements” may provide more influence to curb falling allied defense spending.

Therefore, growing institutionalization of NATO, including the agreement for all allies to spend at least 2% of GDP on defense, should help. More defense cooperation within NATO ties the United States more closely to Europe, but are a better check on allied free-riding than verbal exhortations. Threatening to leave the alliance undermines whatever benefits the United States receives from its alliances.

This is not an unconditional call for greater depth in US alliance commitments, however. Adjusting existing treaties may be more difficult than designing new alliances. The full consequences of treaty depth and attempting to change it 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.
- Benson, Brett V. 2012. *Constructing International Security: Alliances, Deterrence, and Moral Hazard*. Cambridge University Press.
- Benson, Brett V, Adam Meirowitz and Kristopher W Ramsay. 2014. "Inducing Deterrence through Moral Hazard in Alliance Contracts." *Journal of Conflict Resolution* 58(2):307–335.
- Benson, Brett V and Joshua D Clinton. 2016. "Assessing the Variation of Formal Military Alliances." *Journal of Conflict Resolution* 60(5):866–898.
- 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.
- 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. 1991. "Counterfactuals and Hypothesis Testing in Political Science." *World Politics* 43(02):169–195.
- 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.
- 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.
- 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. "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 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.
- 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 interstate, extra-state, intra-state, and non-state wars, 1816-2007*. Washington, DC: CQ Press.
- Senese, Paul D and John A Vasquez. 2008. *The steps to war: An empirical study*. Princeton University Press.
- Singer, J David. 1988. "Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985." *International Interactions* 14(2):115–132.
- Snyder, Glenn H. 1997. *Alliance Politics*. Ithaca: Cornell University Press.
- Sorokin, Gerald L. 1994. "Arms, Alliances, and Security Tradeoffs in Enduring Rivalries." *International Studies Quarterly* 38(3):421–446.
- 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.