Testing the Public Goods Theory of Alliances

Joshua Alley*

November 29, 2018

1 Introduction

Olson and Zeckhauser (1966) argue that international alliances are subject to a collective action problem. Smaller alliance participants "free ride" on the contributions of larger members. This claim is intuitive and influential.

As of November 2018, Olson and Zeckhauser's article has been cited 1670 times. Furthermore, they claim alliances are exemplars of collective action in international organizations more generally. Given its salience and broad implications, this public goods theory of alliances merits careful scrutiny.

But even after 52 years, we have limited evidence for or against free-riding in alliances. The lack of solid evidence is the result of two research design issues. First, many empirical models are unidentified. Second, the overwhelming majority of tests focus on NATO.

Olson and Zeckhauser's approach to measuring the defense burden and state size creates an identification problem. They estimated a correlation between military spending as a share of national income on national income, and has been widely emulated. This approach places GDP on both sides of the equation, so the models are unidentified.

Ratio dependent variables such as military expenditures as a share of GDP are likely to produce spurious results in general (Kronmal, 1993). In Olson and Zeckhauser's test, changes in GDP im-

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

pact the independent and dependent variable. Changes in income will change the ratio of military expenditures to GDP.

Plümper and Neumayer (2015) address the identification problem by examining changes in spending among NATO members. In their framework, unresponsiveness to US and Soviet military spending is evidence of free riding among non-US NATO members. They find no correlation between the size of a NATO member and free-riding, which they argue contradicts Olson and Zeckhauser. However, they do not include the United States in their sample, which omits a crucial data point for testing the size argument.[1]

Despite the canonical status of Olson and Zeckhauser (1966), their predictions have not been subjected to a general, well-identified test. The majority of free-riding tests focus on NATO, and we can conclude NATO members spend less on the military thanks to allied capability (Plümper and Neumayer, 2015; George and Sandler, 2017). Although NATO is salient, it is only one case.

We do not know whether the public goods theory of alliances applies beyond NATO. My survey of the literature on alliance participation and military spending found six tests of the public goods theory of alliances outside of NATO. All six of those studies include GDP in the independent and dependent variable, creating an identification problem.

Due to identification problems and emphasis on NATO, we have no evidence for the generalizability of the public goods theory of alliances. Understanding NATO is worthwhile. But it is insufficient to assess the overall explanatory power of the public goods theory of alliances.

Failure to test the public goods theory of alliances well has two important consequences. First, it hinders the accumulation of knowledge. Without a valid and comprehensive test, the value of the public goods theory of alliances is unclear.

Second, the public goods theory of alliances has an outsized impact on policy debates. The idea of free-riding is ubiquitous in popular and policy debates. If the public goods theory of alliances has limited explanatory power, charges of free-riding should be reassessed.

[1]Given the focus on responses to US spending, this decision is understandable.

Other scholars have questioned the public goods theory of alliances on theoretical grounds. Sandler and Hartley (2001) summarize several extensions of the argument. But such theoretical revisions are premature without knowledge that the more parsimonious public goods model is inappropriate. This paper provides the first comprehensive test of the public goods theory of alliances.

Establishing the empirical validity of the public goods theory of alliances is crucial for academic theory and policy debates. Below, I outline two possible solutions to this challenge. Both examine the role of alliance participant size in multiple alliances from 1816 to 2007.

The first approach uses a conventional panel data research design, with an aggregate measure of alliance participation. In this model, I examine whether small states reduce military expenditures as a result of increasing allied spending, and that effect is diminishing in state size. The second design employs a Bayesian model to estimate the association between treaty contribution and member's military spending for every alliance. Neither design finds strong evidence for Olson and Zeckhauser's predictions.

The paper proceeds as follows. First, I summarize the public goods theory of alliances and generate two testable predictions. Then, I describe the results of a panel-data test of the first prediction. The third section tests the second prediction with a multilevel model. The final section assesses aggregate support for the public goods logic, as well as implications for scholarship and policy.

2 The Public Goods Theory of Alliances

Why are alliances subject to a public goods problem? The aggregate capability of an alliance provides security for members. Because a treaty cannot exclude members without undermining its purpose, security is a public good.

Individual members gain security from treaty participation, regardless of their individual contribution. Olson and Zeckhauser expect that larger members of the alliance with bear a higher defense burden, because these states value defense from the alliance more. Smaller alliance members free-ride on the contributions of larger partners.

There are two implications of this argument. First, large and small states should respond differently to changes in allied military spending. If Olson and Zeckhauser's argument is correct, smaller states should decrease military spending in response to greater allied military spending. Larger states will be less responsive to increased allied spending. This implies a conditional relationship between allied spending, state size, and state military spending.

> HYPOTHESIS 1: The marginal effect of allied spending on state military spending will be negative for small states, and increasing in state size.

The other implication deals with differences within individual alliances. If larger states bear a disproportionate share of the alliance burden, then within each alliance, increasing contribution will lead to higher military spending. States that contribute more to the alliance will increase military expenditures to maintain the output from the treaty.

> HYPOTHESIS 2: Within alliances, increasing treaty contribution will be positively associated with military spending.

Alliances where we observe a positive correlation between treaty contribution and military spending conform to Olson and Zeckhauser's expectations. No correlation implies larger members are not increasing military expenditures. A negative correlation between treaty contribution and military spending implies larger members decrease spending.

I now test these two predictions from the public goods theory of alliances in a sample of all states except microstates from 1816 to 2007.[2] The next section examines Hypothesis 1 by interacting changes in total allied spending and GDP. This approach aggregates multiple treaties into a single measure of alliance participation.

[2]Including states with really small GDP values makes estimating the interactions in Section 2 difficult.

3 Testing Hypothesis 1

To examine Hypothesis 1, I rely on conventional measures of state size and alliance participation. Olson and Zeckhauser use GDP as a proxy for state size, and I employ the Maddison project's data (Bolt et al., 2018). Total allied capability is a common measure of alliance participation (Sorokin, 1994; Morgan and Palmer, 2003). I measure total allied capability as the military spending of all defensive or offensive alliance partners of a state (Leeds et al., 2002). The dependent variable is changes in the natural log of military spending (Singer, 1988).[3]

Because Hypothesis 1 implies a conditional relationship, I interact allied spending and GDP in the following regression model:

$$y = \beta_1 + \beta_2 \text{GDP} + \beta_3 \text{Ally Spend} + \beta_4 \text{GDP} \times \text{Ally Spend} + \beta X \tag{1}$$

β_2 and β_3 are the constituent terms and X is a matrix of control variables with coefficient vector β. Hypothesis 1 implies β_2 should be positive, and β_3 should be negative. The interaction term β_4 should be positive.

I estimate this model with a robust regression estimator. Even after the log transformation, military spending produces heavy-tailed residuals, making OLS inefficient (Rainey and Baissa, 2018). Robust regression re-weights observations according to the size of the residual, making it less sensitive to large deviations than OLS.

In the robust regression, I control for several correlates of alliance participation and military spending. First, I include a lagged dependent variable to address temporal autocorrelation.[4]. I control for international war (Reiter, Stam and Horowitz, 2016), civil war participation (Sarkees and Wayman, 2010), and a count of annual MIDs (Gibler, Miller and Little, 2016). Other controls address salient alliance characteristics, including average democracy and number of members across a state's alliance treaties. Last, I control for regime type, external threat (Leeds and Savun, 2007),

[3]The log-level of military spending is non-stationary.

[4]Therefore, the model functionally estimates changes in spending

and Cold War years.

3.1 Results

Estimates from this model do not match the expectations of Hypothesis 1. Table 1 summarizes results from the robust regression estimator. As expected, GDP is positively correlated with military spending. Changes in allied spending have no effect on military spending, however. The interaction term of changes in allied spending and GDP is also statistically insignificant.

	Value	Std. Error	t value
Change Allied Spending	-0.020	0.044	-0.453
ln(GDP)	0.012	0.002	6.161
Change Allied Spending × ln(GDP)	0.001	0.002	0.805
Lag ln(Military Spend.)	0.992	0.001	805.703
Avg Alliance Size	0.001	0.000	4.370
Avg Alliance Democracy	-0.001	0.001	-1.547
International War	0.090	0.010	9.396
Civil War Participation	0.005	0.007	0.717
Polity	-0.000	0.000	-0.777
International Threat	0.026	0.011	2.437
Cold War	0.046	0.004	11.472
Constant	-0.180	0.038	-4.690

Table 1: Robust Regression Estimates: Allied Spending, GDP, and military spending 1816-2007.

Interactions cannot be interpreted using the coefficients alone (Brambor, Clark and Golder, 2006). Table 3.1 plots the marginal effect of allied military spending across the distribution of GDP. The marginal effect of allied military spending is never negative. There is no conditional relationship between state size and changes in allied spending.

These results hold if I replace robust regression with OLS estimator or employ kernel estimation of the interactive relationship (Hainmueller, Mummolo and Xu, 2017).[5] I also estimated a model that treats states with an alliance as the estimation sample, and used the Heckman two-stage estimator to address non-random selection into alliances.

[5] See the Supplementary files.

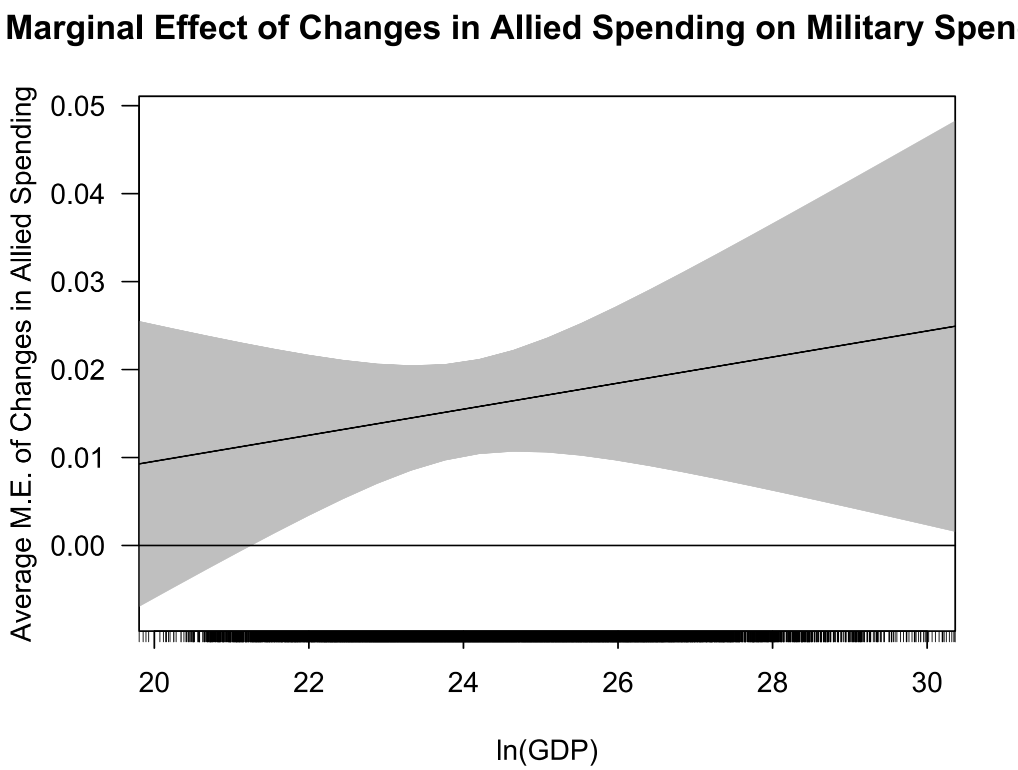

Figure 1: Average Marginal Effect of increasing allied military spending on a state's military spending, across the range of GDP.

It is possible that by aggregating multiple alliance treaties into a single measure, this regression does not capture heterogeneous effects of individual treaties. Small states can make a large contribution to an alliance if their partner is even smaller. So the next section examines how many alliances match the expectations of the public goods theory.

4 Testing Hypothesis 2

Testing Hypothesis 2 requires estimating the association between alliance contribution and military spending for each alliance. For each of the 285 alliances that promise military support,[6] I estimate a unique parameter. Bayesian estimation regularizes estimates with hundreds of parameters, so I fit the following model using STAN (Carpenter et al., 2016).

The full model is starts with state-year changes in the natural log of military spending y_{it}. I model the DV with a t-distribution to account for heavy tails. ν is the degrees of freedom parameter— as ν increases, the t distribution becomes more like a normal distribution.

$$y_{it} \sim student_t(\nu, \mu, \sigma) \tag{2}$$

Most of this model follows standard panel data designs. The expected value of the outcome μ_{it} depends on a constant α, state and year varying intercepts α^{st} and α^{yr}, a lagged DV y_{t-1}, and control variables $X_{it}\beta$. In this specification, I include all controls besides the alliance portfolio averages from the first robust regression in X.

$$\mu_{it} = \alpha + \alpha^{st} + \alpha^{yr} + \eta y_{it-1} + X_{it}\beta + Z_{it}\lambda \tag{3}$$

The $Z_{it}\lambda$ term captures the impact of multiple alliances. Z is a matrix of state participation in alliances— columns are alliances, rows are state-year observations. If a state is not in an alliance,

[6] ATOP offensive and defensive treaties.

the corresponding element of the matrix is equal to zero. If a state is part of an alliance, the corresponding element of the matrix is equal to a state's military spending as a share of allied spending. The alliance contribution elements of the matrix range from 0 to 1, because some states have no military spending.[7]

λ is a vector of 285 alliance-specific parameters. These parameters capture the association between treaty contribution and military spending. A positive λ implies that as contribution to the alliance increases, members spend more.

When a state is not in an alliance, the corresponding λ is multiplied by zero, and has no impact. Alliance participation only affects military expenditures if a state contributes. Each alliance a state is a member of has a separate impact on military spending. Thus, each alliance estimate holds the impact of other treaties constant.

4.1 Results

If the public goods theory of alliances is correct, we should observe many positive λ parameters. Because I estimated this model in the Bayesian framework, each λ has a posterior distribution.[8] I focus interpretation on the posterior mean and 90% credible intervals.[9] The posterior mean is the expected value of λ, while the credible intervals capture uncertainty around that estimate.

Figure 2 summarizes the posterior mean of the 285 alliance coefficients. Most posterior means are close to zero. There are large positive values, but also a good number of large negative values.

141 of 285 alliances have a positive posterior mean. 144 of 285 alliances have a negative posterior mean. The large number of negative values is some evidence against the prediction of Hypothesis 2. However, the posterior means do not convey whether the impact of increasing alliance contribution can be distinguished from zero.

Figure 3 plots the λ parameter for each alliance against the start year of the treaty. Points mark

[7] Costa Rica is the best-known example of this.
[8] See the appendix for a full summary of priors, convergence and model fit.
[9] I use the 90% credible intervals because inferences around the 95% intervals can be less stable.

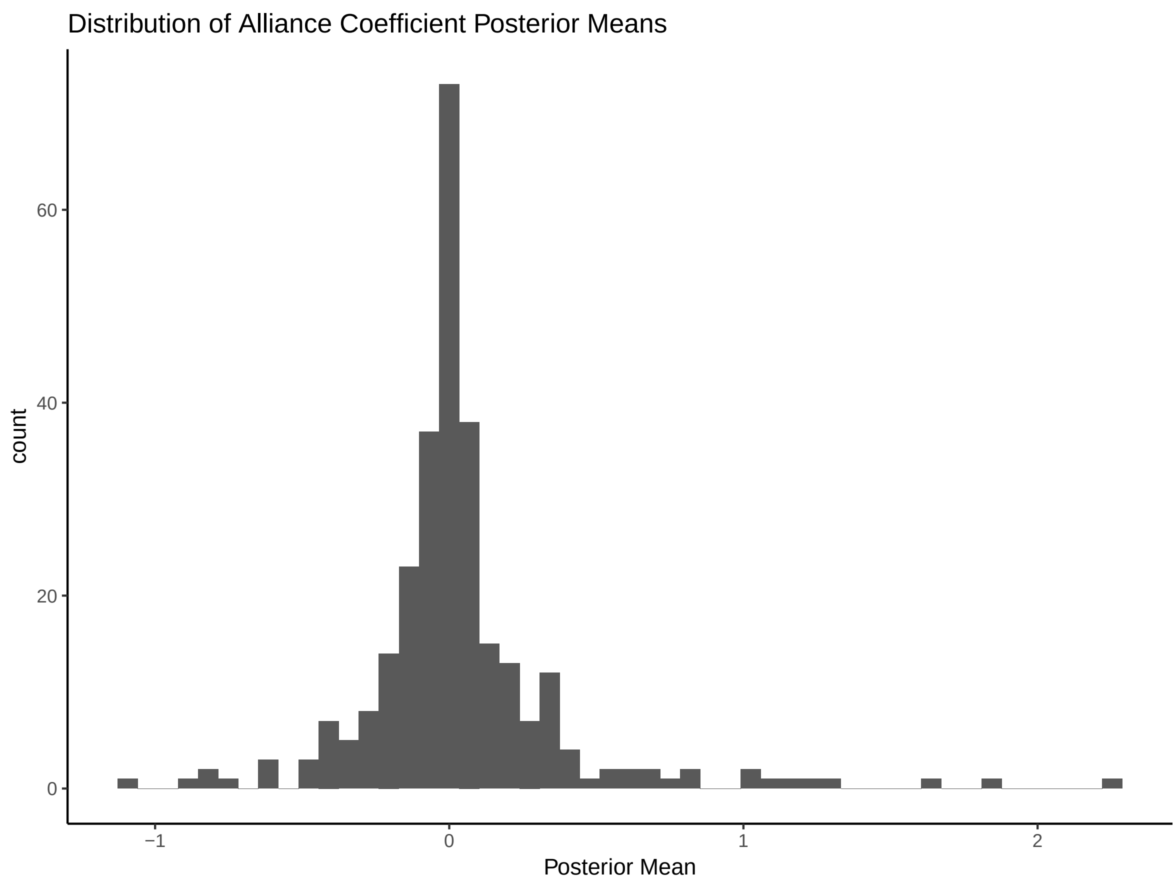

Figure 2: Posterior mean of association between alliance contribution and military spending in 285 defensive and offensive alliances from 1816 to 2007.

the posterior mean. The error bars encapsulate the 90% credible interval.[10]

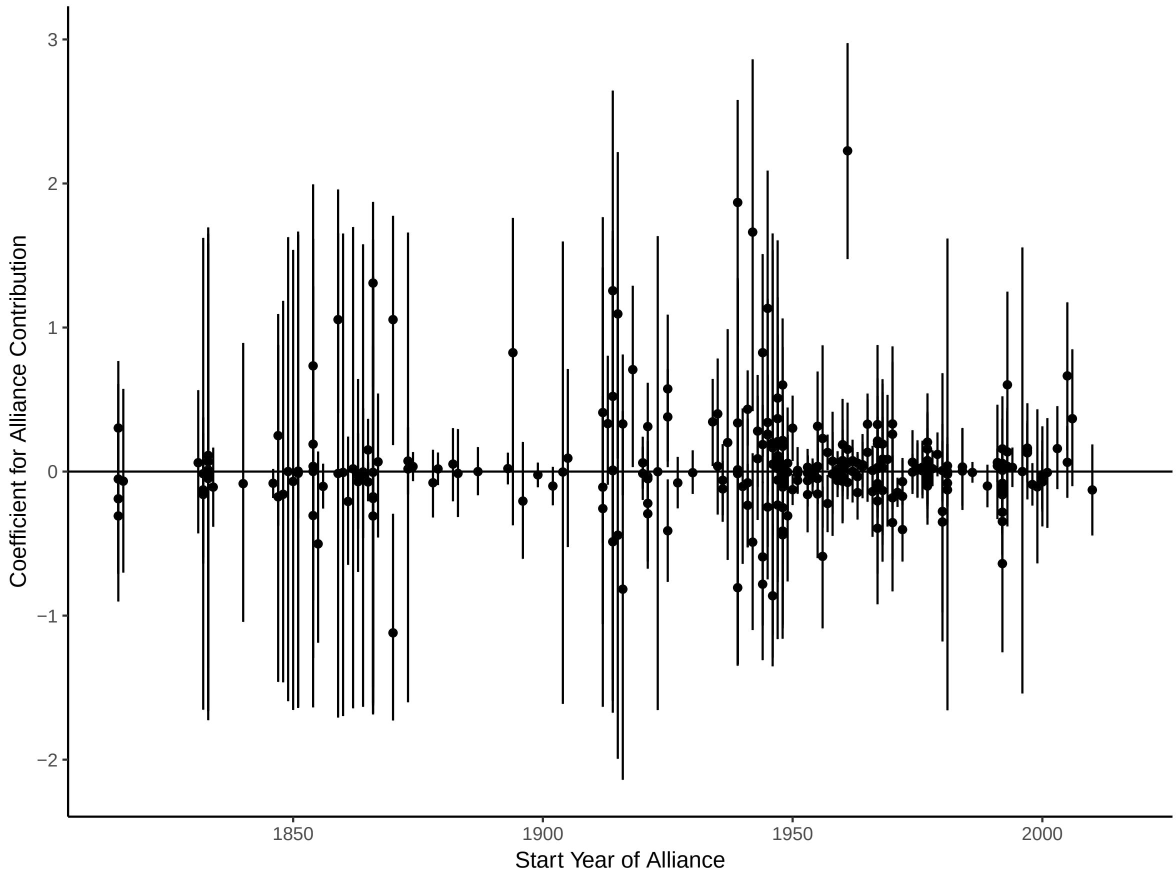

Figure 3: Estimated association between alliance contribution and defense spending in 285 defensive and offensive alliances from 1816 to 2007. Points represent the posterior mean and the error bars cover the 90% credible interval.

Because many of the posterior means are close to zero, 251 credible intervals include zero. There are 20 alliances where the credible interval includes only positive values. That leaves 14 treaties where the credible interval covers only negative values.

20 of 285, or 7%, is a dismal prediction success rate for Hypothesis 2. There is no association between treaty contribution and military spending in most alliances. The 14 negative λ parameters imply larger members of those alliances *spend less* on the military, which the public goods theory cannot explain.

[10]The credible intervals summarize the 5% and 95% quantiles of the posterior.

The estimated λ for NATO offers no support for the public goods theory of alliances. The posterior mean is -0.004, and the credible interval ranges from -.26 to .26. Greater contribution to NATO is unassociated with military spending. This finding corroborates the results of Plümper and Neumayer (2015). The key difference is that this model controls for membership in other alliances.

This second set of estimates provides little support for predictions from the public goods theory of alliances. In 88% of alliances, there is no association between treaty contribution and military spending. There is little support for Hypothesis 2.

5 Conclusion

Taking the results of testing Hypothesis 1 and 2 together, there is little evidence for the predictions of the public goods theory of alliances. There is no conditional relationship between GDP and changes in allied spending. Moreover, there is no association between treaty contribution and military spending in most alliances.

These findings should increase our skepticism about the public goods model of alliances. Although Olson and Zeckhauser's model is simple and intuitive, it lacks explanatory power. Better identified empirical models and examination of multiple treaties shows little evidence small states leave larger counterparts to bear a higher burden.

My results reinforce extant theoretical skepticism about the public goods model. Palmer (1990) and Sandler and Hartley (2001) provide two theoretical critiques that merit further scrutiny. Constructing parsimonious models of alliances and defense effort is the next challenge for scholars. Components of the public goods model and these alternatives may provide a useful starting point.

Some of the results raise additional questions for theoretical work. Why do we observe 14 alliances where increasing alliance contributions are associated with reduced spending? Why is the effect of greater allied spending mostly positive in Model 1? Derivatives of the public goods theory may struggle to address these questions.

Furthermore, scholars should reassess the idea of "free-riding" in alliance politics. Free-riding is inextricable from a public goods understanding of alliances. But if key predictions of the public goods model have little empirical support, free-riding is an inaccurate description of reduced defense effort by alliance participants. Charges of free-riding are normatively loaded and may mask significant exchanges between alliance members (Lanoszka, 2015).

We should not abandon the public goods model in international politics more generally. The public goods model may apply to other international organizations. Inferring collective action problems in other international organizations from alliances is inappropriate, however.

The public goods model of alliances should have a less salient role in discussions of alliance participation and defense effort. I find little evidence that reduced military spending by alliance members is the result of a collective action problem. Public goods models are useful in many contexts, but they have little explanatory power in alliance politics.

References

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* .

Brambor, Thomas, William Roberts Clark and Matt Golder. 2006. "Understanding interaction models: Improving empirical analyses." *Political Analysis* 14(1):63–82.

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.

George, Justin and Todd Sandler. 2017. "Demand for military spending in NATO, 19682015: A spatial panel approach." *European Journal of Political Economy* pp. 1–15.

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.

Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2017. "How Much Should We Trust Estimates from Multiplicative Interaction Models?: Simple Tools to Improve Empirical Practice.".

Kronmal, Richard A. 1993. "Spurious Correlation and the Fallacy of the Ratio Standard Revisited." *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 156(3):379–92.

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

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

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

Morgan, T Clifton and Glenn Palmer. 2003. "To protect and to serve: Alliances and foreign policy portfolios." *Journal of Conflict Resolution* 47(2):180–203.

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

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.

Rainey, Carlisle and Daniel K. Baissa. 2018. "When BLUE Is Not Best: Non-Normal Errors and the Linear Model." *Political Science Research & Methods* .

Reiter, Dan, Allan C. Stam and Michael C. Horowitz. 2016. "A Revised Look at Interstate Wars, 18162007." *Journal of Conflict Resolution* 60(5):956–976.

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.

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

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