

Reassessing the Public Goods Theory of Alliances

Joshua Alley*

February 4, 2020

Abstract

The public goods theory of alliances exerts substantial influence on scholarship and policy, but it has not received sufficient empirical scrutiny. The public goods model claims that smaller alliance participants can free-ride on larger partners because alliances face collective action problems. Prior statistical tests of free-riding suffer from model specification and generalizability problems, so this study addresses those limitations. Using data on 285 alliances from 1816 to 2007, I examine how often states with little economic weight in an alliance decrease military spending, and states with high economic weight increase military spending. I find little evidence to support this alternative expression of the free-riding hypothesis. Therefore, the argument that alliances provide a public good and generate free-riding is a poor foundation for public debate and theoretical development.

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

1 Introduction

Olson and Zeckhauser (1966) argue that international alliances generate a collective action problem. According to their theory, security from an alliance is a public good, so smaller alliance participants “free ride” on the contributions of larger members. Free-riding is reflected by disproportionate allocations of resources to defense, where smaller alliance members spend a lower share of their national income on the military. This view of alliances and defense effort plays an important role in scholarly debates about alliances (Walt, 1990; Mearsheimer, 1994; Goldstein, 1995; Sandler and Hartley, 2001; Garfinkel, 2004; Walt, 2009; Norrlof, 2010; Barrett, 2010; Plümper and Neumayer, 2015). Some theories of alliances like the joint product model (Sandler, 1993) also use the public goods logic as a starting point.

In addition, policy and popular discussions of alliances employ collective action ideas. Pundits and American policymakers often refer to allied “free-riding.” Treating alliances as a public good and beset by collective action problems generates concern that the United States is “being taken advantage of” by junior partners. US policymakers often use free-riding and disproportionate contributions to the common good to criticize lackluster allied defense expenditures. For example, Barack Obama complained in 2016 that “Free riders aggravate me” and US allies “have to pay your fair share.” Donald Trump has implied the United States would not protect allies who spend too little on defense. Such complaints and exhortations go back as far as the Eisenhower administration (Lanoszka, 2015).

Olson and Zeckhauser’s argument has shaped a generation of policy and scholarly debate about alliances. Increasing great power competition makes assessing this theory even more important. Competing visions of US grand strategy hinge in part on the explanatory power of the public goods model. Advocates of retrenchment and “restraint” use the public goods logic to claim that US allies free-ride, so the United States should withdraw from many alliances (Preble, 2009; Posen, 2014). Others assert that alliances do not provide a public good and the benefits of alliance participation

outweigh the costs (Brooks, Ikenberry and Wohlforth, 2013; Brands and Feaver, 2017).

In this paper, I reassess the public goods model by testing the extent of free-riding across 285 alliances. The academic and policy salience of alliances and military spending make revisiting Olson and Zeckhauser's classic argument worthwhile. Although academic theory has progressed substantially since 1966, the public goods model retains an important place in academic and policy discourse, despite a major problem. 53 years after the publication of "An Economic Theory of Alliances," there is little reliable and general evidence for or against Olson and Zeckhauser's prediction that small alliance members free-ride on larger partners.

Existing tests of the public goods logic suffer from two key limitations. First, many empirical estimates of free-riding within alliances are misspecified. Olson and Zeckhauser measure defense burdens using military spending as a share of GDP and state size with GDP. This approach is widely emulated, but it creates a problem because GDP is present on both sides of the equation. There is a deterministic component in the relationship between GDP and defense burdens— changes in GDP shift the defense burden.¹ Moreover, ratio dependent variables often generate spurious results (Kronmal, 1993).

One notable paper addresses the specification problem, but may not produce generalizable findings. Plümper and Neumayer (2015) examine how growth in military spending by North Atlantic Treaty Organization (NATO) members responds to changing US and Soviet spending. They demonstrate that small NATO members are unresponsive to US and Soviet military spending, and present this as evidence of free riding. Plümper and Neumayer (2015) find no correlation between NATO member size and the extent of free-riding, however, which they argue contradicts Olson and Zeckhauser. The emphasis on NATO in this paper brings me to the second limitation: a lack of generalizability.

NATO is the epicenter of free-riding discussions. Following Olson and Zeckhauser's emphasis

¹When GDP shifts, the defense burden remains constant only if military spending also changes in such a way that defense spending's share of GDP remains the same. Such changes are highly unlikely.

on military spending as a share of GDP, accusations of free-riding emphasize that NATO members have lower defense burdens than the United States. Scholars, pundits and policymakers have spent decades arguing over whether the public goods model applies to NATO e.g., (Sandler and Forbes, 1980; Palmer, 1990; Boyer, 1993; Gates and Terasawa, 1992; Sandler and Hartley, 2001; Lanoszka, 2015; Plümper and Neumayer, 2015).

Most studies of the public goods model focus on NATO, but NATO is a difficult case for making general conclusions. NATO is exceptionally large, durable and capable. There are only seven tests of the public goods model outside of NATO, all of which examine a few alliances (Russett, 1970; Starr, 1974; Reisinger, 1983; Thies, 1987; Conybeare and Sandler, 1990; Oneal and Whatley, 1996; Siroky, 2012). Six of these studies include GDP in the independent and dependent variable, so they suffer from the aforementioned specification problem. The seventh study uses case comparisons. There leaves a need for a general examination of the public goods logic.

Here, I address the specification and generalizability issues in a test of the public goods model. I use the public goods logic to predict that states with high economic weight in an alliance will increase percentage changes in military spending, and states with low economic weight will decrease military spending. Then I employ a Bayesian model to estimate the association between economic weight and percentage changes in military spending within 285 alliances. I find few alliances with clear evidence for the predictions of the public goods model.

My findings will not surprise the many skeptics of the public goods model of alliances (Palmer, 1990; Gates and Terasawa, 1992; Sandler and Hartley, 2001; Norrlof, 2010; Niou and Zeigler, 2019), many of whom are skeptical due to inconsistent correlations between GDP and defense burdens. Such misspecified models provide unreliable evidence about the public goods model, however. Without a reliable and general test, theoretical revisions of a parsimonious public goods model may be premature. Before dismissing or modifying the public goods model, we should firmly establish that it lacks explanatory power. Therefore, this paper advances knowledge by checking an important point in academic debate and scrutinizing an argument with substantial

popular appeal.

The paper proceeds as follows. First, I summarize the public goods theory of alliances and use it to derive observable implications of free-riding. Then, I describe the model and results. The final section assesses implications for scholarship and policy.

2 Free-Riding in Alliances

To identify observable implications of free-riding, I use the public goods argument, but make slightly different predictions. This theoretical exercise does not critique the public goods model. Instead, I take the public goods logic as given and generate alternative implications of free-riding by small states. Following Plümper and Neumayer (2015) I use percentage changes in military spending to conceptualize defense effort. I then assess the role of relative state size using economic weight.

Why might alliances suffer from a collective action problem? The aggregate military capability of an alliance provides security for members and states contribute by investing in their military. Because an alliance cannot exclude members without undermining its purpose, alliance security is a public good. Alliance members receive security regardless of their individual contribution.² Thus, states have incentives to rely on their alliance partners and reduce their own military expenditures.

Olson and Zeckhauser expect that larger members of the alliance will bear a higher defense burden, because these states value security from the alliance more. Small alliance members rely on the contributions of larger partners for security and bear a lower defense burden. As a result, smaller states free-ride on larger alliance participants. Moreover, smaller states have greater bargaining leverage, because a large state cannot credibly threaten to reduce their contribution and has “relatively less to gain than its small ally from driving a hard bargain” (Olson and Zeckhauser, 1966, pg. 274).

²The marginal costs and benefits of participation depend in part on group size, but Olson and Zeckhauser’s model shows free-riding even in a bilateral alliance.

Olson and Zeckhauser compare alliance members using economic resources, specifically GDP. Economic weight within an alliance, or each state's share of total allied GDP, is a related way to conceptualize differences in state size.³ Using economic weight facilitates comparisons of state size across diverse alliances. A greater share of total economic resources in an alliance gives a state more economic weight and increases their potential contribution of military spending.

Because Olson and Zeckhauser expect that large states contribute disproportionately to an alliance, economic weight shapes defense expenditures. Inasmuch as larger states value security from an alliance more, alliance participation will increase their investment in military capability. Smaller members can free-ride and lower military spending. Therefore, alliance participation will increase the military spending of large states, and decrease military spending by small states.

HYPOTHESIS 1: For states with a large share of the total GDP in an alliance, alliance participation will increase annual percentage changes in military spending.

HYPOTHESIS 2: For states with a small share of the total GDP in an alliance, alliance participation will decrease annual percentage changes in military spending.

I assess the two hypotheses by estimating the association between economic weight and military spending across 285 alliances. I code economic weight such that positive alliance parameters imply more military spending for large members and less spending for small members. A negative correlation between economic weight and military spending implies that larger members spend less, and small members spend more. To assess the generalizability of the public goods model, I examine how many alliances have a positive correlation between shares of total allied GDP and military spending.

The hypotheses approximate Olson and Zeckhauser's prediction that smaller alliance members will spend a smaller share of their economic resources on the military. Though my predictions are

³Different states may be large or small depending on the alliance.

different, they facilitate a reliable and general empirical test. I now test the hypotheses in a sample of all non-microstates from 1816 to 2007. State-year observations are the unit of analysis.

3 Testing the Public Goods Logic

For each of the 285 alliances that promise military support,⁴ I estimate a parameter measuring the association between economic weight and military expenditures. This model examines alliance free-riding using evidence from many treaties.⁵ Bayesian estimation regularizes estimates with many parameters, so I fit the following model using STAN (Carpenter et al., 2016).

The model starts with state-year percentage changes in military spending y_{it} , transformed with an inverse hyperbolic sine. I model this variable using a t-distribution with degrees of freedom ν to account for heavy tails. The expected value of military spending μ_{it} depends on a constant α , state and year varying intercepts, and control variables $\mathbf{X}_{it}\beta$.⁶

$$y_{it} \sim student_t(\nu, \mu_{it}, \sigma) \quad (1)$$

$$\mu_{it} = \alpha + \alpha^{st} + \alpha^{yr} + \mathbf{X}_{it}\beta + \mathbf{Z}_{it}\gamma \quad (2)$$

The $\mathbf{Z}_{it}\gamma$ term captures the impact of economic weight in alliances. \mathbf{Z} is a matrix of state participation in alliances— columns are alliances, rows are state-year observations. If a state is part of an alliance, the corresponding value in \mathbf{Z} depends on their share of total GDP in the alliance. Small alliance members are assigned a value of negative 1 if their economic weight is less than the median of .5 in bilateral alliances, or less than the median of .125 in multilateral alliances.⁷

⁴ATOP offensive and defensive treaties (Leeds et al., 2002).

⁵I also regress percentage changes in military spending on state's average weight in their alliances, and make similar inferences, which I report in the appendix.

⁶See the appendix for a full description of all the variables in the model.

⁷The distribution of economic weight is very different in bilateral and multilateral alliances, which requires different thresholds.

Large alliance members have a value of positive one in \mathbf{Z} if their economic weight is above those thresholds. If a state is not part of the alliance, the corresponding matrix element is zero, so the positive and negative one values are relative to the absence of the alliance. Multiplying a positive γ by negative one for small states and positive one for large states will lead to negative military spending growth for small states and positive growth for large states.

\mathbf{Z} is a quasi-spatial approach to capturing the impact of participation in multiple alliances. Alliance participation affects military spending through economic weight in this model. The γ parameters capture the correlation between economic weight in an alliance and military spending.

γ is a vector of 285 alliance-specific parameters. Because \mathbf{Z} contains each state's share of allied GDP, these coefficients estimate the association between economic weight and military spending.⁸ When a state is not in an alliance, the corresponding γ is multiplied by zero, and has no impact.

Each alliance has a separate impact on military spending. While these alliance parameters are distinct, they have a common prior distribution. Partial pooling estimates the dispersion of the alliance parameters from the data, so the prior for γ is normally distributed with mean θ and variance σ_{all} . θ is the mean hyperparameter of the alliance coefficients and each γ deviates from θ based on the variance hyperparameter σ_{all} . Every alliance estimate holds the impact of other treaties constant. A positive γ implies that alliance members with more economic weight have higher percentage changes in military spending, as Hypothesis 1 predicts.

3.1 Results

The public goods model would expect many positive γ parameters. Because I employed Bayesian modeling, each γ has a posterior distribution.⁹ I focus interpretation on the posterior

⁸This assumes symmetric effects across small and large states, which I relax somewhat in the appendix by using a weighted coding of \mathbf{Z}

⁹See the appendix for a full summary of priors, convergence and model fit. I also show that the model can recover known parameters from simulated data.

mean and 90% credible intervals.¹⁰ The posterior mean is the expected value of γ , while the credible intervals capture uncertainty around that estimate.

Figure 1 plots the γ parameter for each alliance against the start year of the treaty. Points mark the posterior mean. The error bars encapsulate the 90% credible interval.

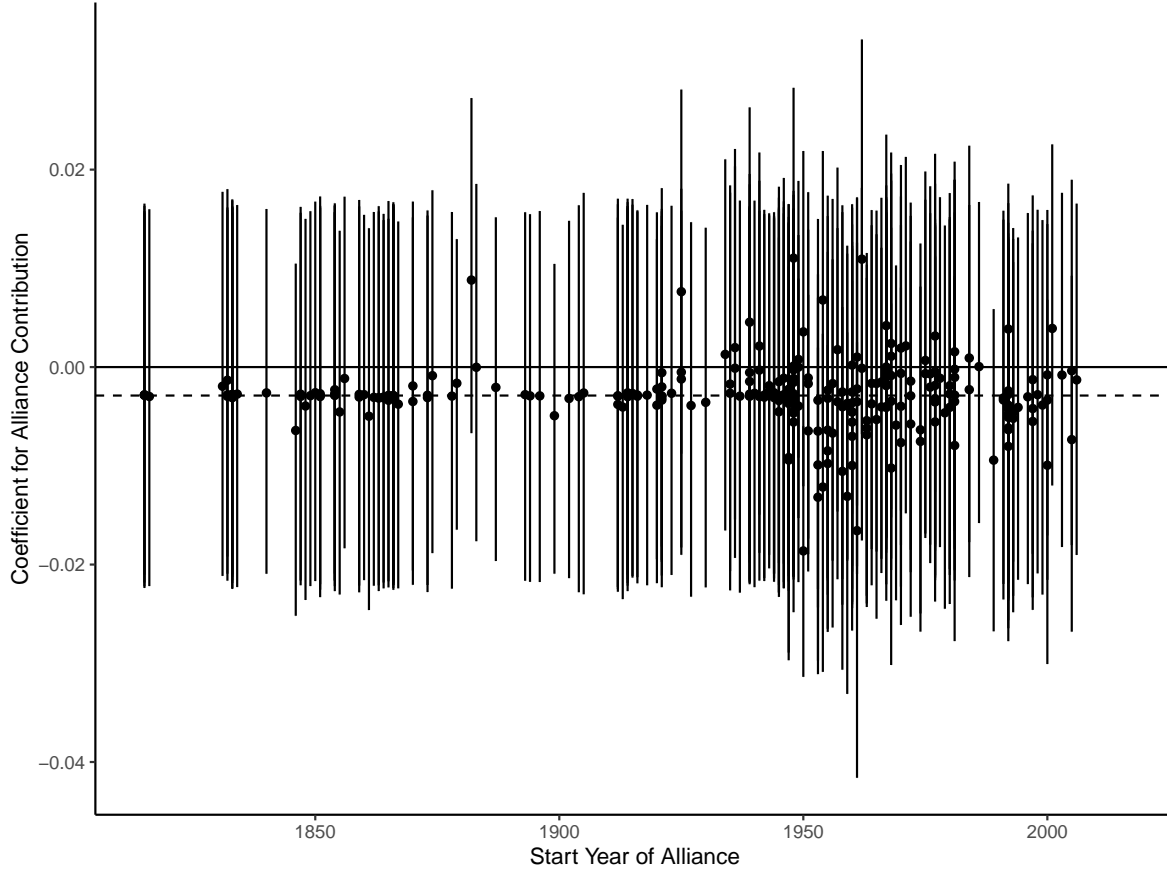


Figure 1: Estimated association between share of total allied GDP and defense spending in 285 defensive and offensive alliances from 1816 to 2007. Points represent the posterior mean and the error bars cover the 90% credible interval. The dashed line marks the posterior mean of the θ parameter, which represents the average association between economic weight and percentage changes in military spending.

No alliances have a uniformly positive 90% credible interval. Most credible intervals are consistent with effects ranging from -0.02 to 0.015. Only 28 of the 285 alliances have a positive

¹⁰I use 90% credible intervals because inferences around 95% intervals are less stable. The 90% credible interval falls between the 5% and 95% quantiles of the posterior.

posterior mean.

As Figure 1 shows, partial pooling of the γ parameters regularizes many estimates. θ has a posterior mean of -0.002. The effects of increasing economic weight γ are tightly clustered around the mean hyperparameter. Under these circumstances, it is difficult to distinguish most estimates from zero, even as partial pooling shares information across coefficients and reduces uncertainty.

The limited connection between economic weight and greater military spending is further reinforced by a simulation. To assess whether increasing a state's share of allied GDP leads to higher defense spending, I simulated the effect of changing economic weight on percentage changes in military spending. In the simulated data, I used the full posteriors of the intercept α , all the β coefficients, and one γ parameter. I selected the γ parameter with the most positive posterior mass, so this is the *best case alliance* for the public goods predictions. I set the state-level variables at their median or modal value and changed economic weight from -1 to 1.

In Figure 2, I summarize predicted changes in military spending at different levels of economic weight. In this figure, the point marks the mean and the error bars summarize the 90% credible interval. There is limited evidence that moving from small to large economic weight increases military spending.

The γ estimates also produce inferences about individual alliances. The estimated γ for NATO offers no support for the public goods theory of alliances. The posterior mean of this parameter is close to zero, and it has roughly half positive and half negative posterior mass. A greater share of NATO's total GDP has no clear association with percentage changes in military spending. This finding corroborates the size result of Plümper and Neumayer (2015), but NATO members may still spend less on the military thanks to allied capability (George and Sandler, 2017).

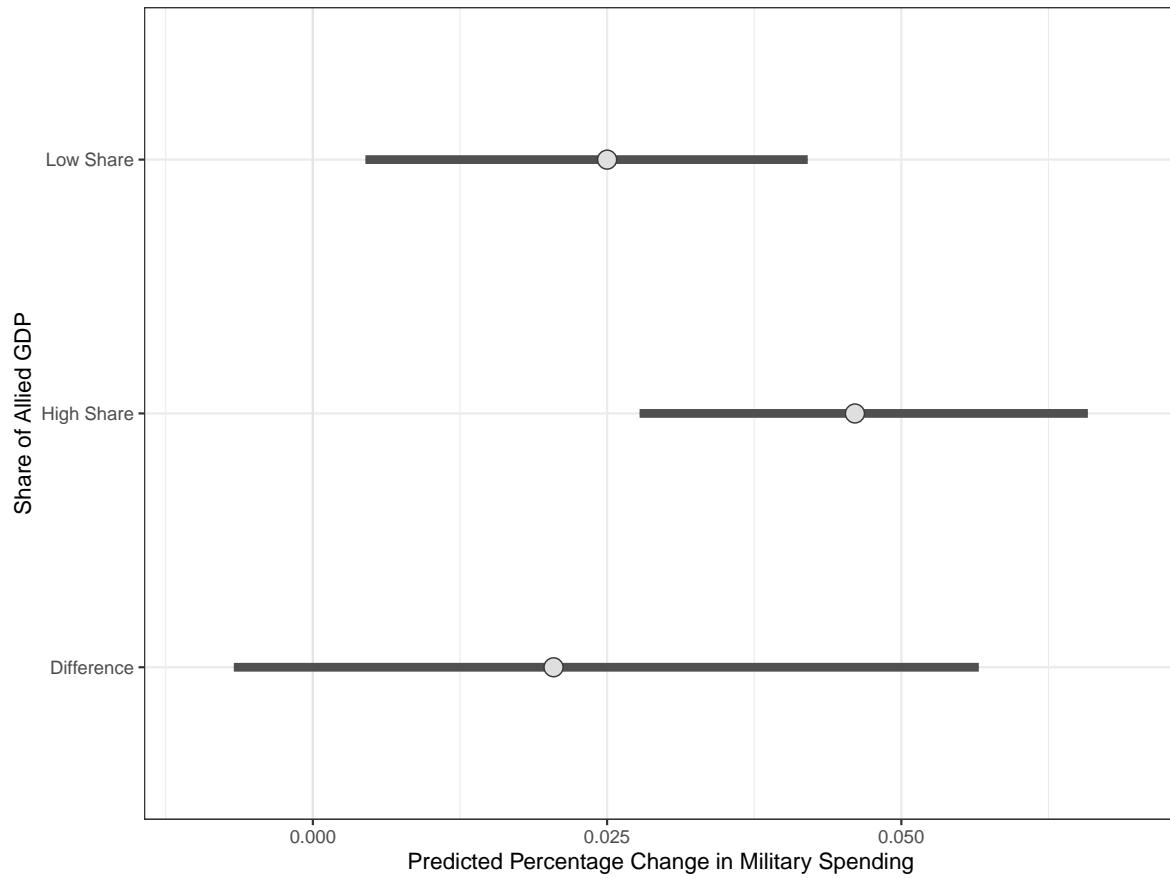


Figure 2: Predicted percentage changes in military spending for a simulated state with low or high shares of total allied GDP. Points mark the median value, and the error bars summarize the 90% credible interval. The difference bar captures the effect of moving from having low to high weight.

4 Conclusion

Few alliances see large differences in military spending by economic weight. Although Olson and Zeckhauser's model is parsimonious, it barely applies to most alliances.

Scholars have already developed a wide range of arguments with better explanatory power in alliance politics. The best theoretical alternative to the public goods approach emphasizes exchange between alliance members and intra-alliance bargaining (Morrow, 1991; Norrlof, 2010; Brooks, Ikenberry and Wohlforth, 2013; Johnson, 2015; Kim, 2016). Other arguments with roots in the public goods approach (Sandler and Hartley, 2001) may also merit further empirical and theoretical development.

Policymakers and pundits should reassess the way they use “free-riding” to describe alliance politics. Free-riding is inextricable from a public goods understanding of alliances. But if the public goods model has little empirical support, free-riding is an inaccurate description of reduced defense effort by alliance participants.

Scholars should not abandon the public goods model in international politics, however. Using alliances as exemplars of collective action problems in other international organizations is inappropriate, but collective action can apply to other international organizations. Thinking about alliances using exchange is more fruitful than the collective action problem envisaged by Olson and Zeckhauser, so policymakers and scholars should exercise caution in using the public goods model to understand alliance politics.

References

- Barrett, Scott. 2010. *Why Cooperate?: The Incentive to Supply Global Public Goods*. Oxford University Press.
- Boyer, Mark A. 1993. *International Cooperation and Public Goods: Opportunities for the Western Alliance*. Baltimore: Johns Hopkins University Press.

- 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.
- 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.
- Garfinkel, Michelle R. 2004. "Stable alliance formation in distributional conflict." *European Journal of Political Economy* 20(4):829–852.
- Gates, William R and Katsuaki L Terasawa. 1992. "Commitment, Threat Perceptions, and Expenditures in a Defense Alliance." *International Studies Quarterly* 36(1):101–118.
- George, Justin and Todd Sandler. 2017. "Demand for military spending in NATO, 1968–2015: A spatial panel approach." *European Journal of Political Economy* pp. 1–15.
- Goldstein, Avery. 1995. "Discounting the free ride: alliances and security in the postwar world." *International Organization* 49(1):39–71.
- Johnson, Jesse C. 2015. "The cost of security: Foreign policy concessions and military alliances." *Journal of Peace Research* 52(5):665–679.
- Kim, Tongfi. 2016. *The Supply Side of Security : A Market Theory of Military Alliances*. Stanford University Press.
- 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, Jeffrey Ritter, Sara Mitchell and Andrew Long. 2002. "Alliance Treaty Obligations and Provisions, 1815–1944." *International Interactions* 28(3):237–260.
- Mearsheimer, John J. 1994. "The False Promise of International Institutions." *International security* 19(3):5–49.
- 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.
- Niou, Emerson MS and Sean M Zeigler. 2019. "External Threat, Internal Rivalry, and Alliance Formation." *The Journal of Politics* 81(2):571–584.

- Norrlof, Carla. 2010. *America's global advantage: US hegemony and international cooperation*. Cambridge 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 and Hugh Carter Whatley. 1996. "The effect of alliance membership on national defense burdens, 1953–88: A test of Mancur Olson's theory of collective action." *International Interactions* 22(2):105–122.
- Palmer, Glenn. 1990. "Alliance Politics and Issue Areas: Determinants of Defense Spending." *American Journal of Political Science* 34(1):190–211.
- 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.
- 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.
- Reisinger, William M. 1983. "East European military expenditures in the 1970s: collective good or bargaining offer?" *International Organization* 37(1):143–155.
- Russett, Bruce. 1970. *What Price Vigilance?: The Burdens of National Defense*. New Haven: Yale University Press.
- Sandler, Todd. 1993. "The economic theory of alliances: a survey." *Journal of Conflict Resolution* 37(3):446–483.
- 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.
- Siroky, David S. 2012. "Each Man For Himself? Rival Theories of Alliance Economics in The Early State System." *Defence and Peace Economics* 23(4):321–330.
- Starr, Harvey. 1974. "A Collective Goods Analysis of the Warsaw Pact after Czechoslovakia." *International Organization* 28(3):521–532.
- Thies, Wallace J. 1987. "Alliances and Collective Goods: A Reappraisal." *Journal of Conflict Resolution* 31(2):298–332.
- Walt, Stephen M. 1990. *The Origins of Alliance*. Cornell University Press.
- Walt, Stephen M. 2009. "Alliances in a Unipolar World." *World Politics* 61(01):86–120.