When military alliances are expensive, they naturally raise distributional issues. This article considers two theories to explain how much a state will voluntarily contribute to the economic burdens of defense. Empirical work has relied largely on data from the twentieth century. This article provides an out-of-sample test to evaluate the models. Using data on the Quintuple Alliance, the results are more consistent with the predictions of the joint products model than the pure public goods model. Due to credible commitment problems, and intra-alliance cleavages, I argue that we should not expect substantial free riding in most conventional military alliances.

Keywords: Alliances; Public Goods; Defense Economics; Free-riding; Burden-sharing; Europe

JEL Codes: D70; D71; D74

At least since 404 BCE, when the Spartan alliance defeated Athens to end the Peloponnesian War, military alliances have been recognized for their role in deterring conflict, and in deciding victory when deterrence fails. Although their function may seem fairly straightforward, alliances can be expensive. How the costs are distributed between members is intensely debated in the scholarly literature and in policy circles (see Sandler and Hartley, 2002, for a recent review). Three specific questions have motivated the debate: (1) why would one ally contribute more of its scarce resources to the collective undertaking than another? (2) Can the distribution of allied benefits explain the distribution of the allied costs? (3) Is free riding a serious problem in military alliances?

This article considers two prominent theories to explain how much a state will voluntarily contribute to the economic burdens of defense – the pure public goods model and the joint products model. Empirically, most work on these questions has relied on data from the twentieth century. The North Atlantic Treaty Organization (NATO), in particular, has been the focus of most studies. Although NATO is an exceptionally important alliance, it is also a highly unusual alliance because of its size (in terms of world GDP and the number of members, today 28) and because it was a nuclear alliance with a very credible and secure second-strike capability. As a result, NATO was able to issue a credible commitment to all its members that the bigger members would act in concert in the event of an attack. NATO’s principle was ‘an attack against one is an attack against all’. The assurance was credible in NATO’s case: it had boots on the ground, nuclear weapons in hardened silos and fighter jets permanently in the sky. Most military alliances offer significantly

*Email: david.siroky@asu.edu

ISSN 1024-2694 print: ISSN 1476-826X online © 2011 Taylor & Francis
http://dx.doi.org/10.1080/10242694.2011.596654
fewer guarantees, however, so states have rarely relaxed their efforts on account of other’s increased efforts. The argument emphasized in this article is that most alliances historically and in much of the modern world behave more in line with a different principle: ‘each man for himself’.

This study provides an out-of-sample test to evaluate this claim. An assessment of the two principal models cannot be adequate if the empirical testing is limited largely to nuclear alliances like NATO and the Warsaw Pact. Using data on the Quintuple Alliance, this paper shows that the results are more consistent with the predictions of the joint products than the pure public goods model, which should be viewed as a limiting case when excludable benefits and rivalry are low, but credible deterrence is high. The article argues that in most alliance settings we should not observe substantial free riding due to commitment problems and intra-alliance cleavages. The next section discusses the two theoretical models, and then tests them on data from a nineteenth-century alliance.

**THE PURE PUBLIC GOOD MODEL**

Zeckhauser and Olson (1966: 266) proposed a public good model ‘to explain the workings of international organizations and to test that model against existing international institutions’. The model was parsimonious, falsifiable and politically relevant. NATO, they argued, was an international institution that produced a public good – deterrence. It should be possible therefore to test NATO’s empirical distribution of costs against the theory’s predictive distribution. If deterrence is a public good (the consumption of deterrence is non-rival and non-excludable), then alliances ought to be characterized by free riding, because of non-excludability, and by sub-optimality, since marginal costs are not equated with marginal benefits (Samuelson, 1954; Kreps, 1990: 168–9). They tested it against NATO data from 1964 and found that small NATO allies were free riding, and that the total amount of security provided was therefore sub-optimal.

Utilizing NATO data from 1964, Olson and Zeckhauser (1966: 274–277) demonstrated that the share of defense costs paid by each ally was positively correlated with its economic capabilities. This meant that the smaller members did not pay the same ‘share’ of their GNP (defense expenditures/gross national product) as bigger members, who paid more than their share. For instance, the US, which was the biggest member of NATO, devoted many times more of its GNP to defense than Luxembourg, which was the smallest member (1966: 273).\(^1\) The policy implication was that NATO’s smaller allies needed to pay more and stop free riding because the result was sub-optimal deterrence for all alliance members. If alliances are institutions that produce a public good – deterrence – the model implied that it should be impracticable to exclude alliance members who do not fully share the cost of providing security. At the margin, new members can be added without significantly reducing the level of security available. Larger alliance members would be forced to bear a disproportionately large share of the security costs (Olson and Zeckhauser, 1966: 268–276; Olson, 1971: 871; Thies, 1987: 300).

---

\(^1\) In a more realistic setting, the ability of the most powerful to shift the burden to smaller powers may depend upon the extent to which they perceive a threat. Britain, for example, did not regard Russian troops, and Germany did not regard Austro-Hungarian troops, as 1-for-1 substitutes in World War I. Substitutability is more realistic to assume for lesser assignments. For an excellent paper that takes the role of external threat seriously, see Oneal and Whatley, 1996.
In its simplest form, the model involves \( n \) allies each of which allocates its GNP or income, \( I \), between a pure public good, \( q \), and a private good, \( y \). A unitary actor maximizes a well-behaved, strictly quasi-concave utility function\(^3\): \( U_i = U^i(y^i, q^i + QT) \), where \( Q \) is the additive sum of the other allies’ defense expenditure and \( T \) is the exogenous threat, such that: \( Q = \sum_{j \neq i} q^j \). Utility increases when the total level of defense expenditure \( Q \) and the production of the excludable good \( y \) increase. Since \( Q \) is additive rather than multiplicative or exponential, defense expenditures are perfect substitutes.\(^4\) To bound the model’s predictions, each ally possesses a linear budget constraint, \( I_i = h y_i + q_i \) where the price \( h \) of one unit of defense \( q \) and one unit of the excludable good \( y \) is one. The Pareto-optimal contribution to the alliance, \( Q^o \), results when every ally elects to contribute so that the relative price of defense is equated to the sum of the marginal rate of substitution for all allies. Since the marginal benefits obtained from other allies’ contributions are not accounted for (because decisions are simultaneous), the public good, deterrence, will be underprovided, which is sub-optimal.

This important prediction, however, rests on five basic assumptions (Sandler and Hartley, 2002). First, a unitary actor decides the level of defense expenditure.\(^5\) Second, all decisions are made simultaneously. Third, expenditure decisions produce a public good, which is shared among allies. Fourth, marginal defense costs per unit of defense are equivalent for all allies. Fifth, allied defense contributions are perfectly substitutable (Sandler and Hartley, 2002, p. 871).\(^6\) Several of these assumptions are relaxed in the joint products model.

**THE JOINT PRODUCTS MODEL**

Only one year after the original paper, Van Ypersele de Strihou (1967) presented an alternative model in which defense expenditures yield ally-specific benefits that were excludable. Building on this effort, Sandler developed ‘the joint-products model,’ which allowed defense expenditures to produce excludable benefits for individual allies. If alliances produce significant excludable benefits, then individual allies have incentives to pull their weight in proportion to benefits obtained, and we should observe little disproportionality in costs (Sandler and Cauley, 1975; Sandler and Hartley, 2002).\(^7\) Since the credibility of an ally coming to the aid of a threatened member varies with the geographic location of the ally, the weapons system deployed, intra-alliance cleavages, and the reputation of the guarantor, the basic conditions for non-excludable benefits are regularly violated.

For the simplest case where alliance defense expenditures produce two goods, or products, \( x \) and \( z \), one private \((x^i = \alpha q^i)\) and one public \((z^i = \beta q^i)\), the total security produced by the alliance is \( Z \). Following Sandler and Hartley (2002: 877), I take \( Z \), as the total

---

\(^2\)The exposition of these two models follows Sandler and Hartley, 2002, 872–877.

\(^3\)If both the private and public goods are normal, that is demand increases with income, then there exists a unique Nash equilibrium. See Sandler and Murdoch, 1990; Cornes, Hartley and Sandler, 1999.

\(^4\)For other means of aggregating the individual contributions, see Hirschleifer, 1983.


\(^6\)Also transaction costs are assumed to be zero.

\(^7\)Olson and Zeckhauser later relaxed the fourth assumption on comparative advantage and allocative efficiency within international organizations by allowing differences in the marginal cost of defense for different allies. Olson and Zeckhauser, 1967. Allocative efficiency means that burden sharing would be driven not by direct proportionality of ability-to-pay, but by comparative advantage.
security from all allies minus the one: \( Z = Z^i + Z_{-i} \). The entire alliance security level can be written: \( Z = \beta(q^i + Q_{-i}) \) and the utility function for an ally \( i \) can be written: \( U^i = (y^i, x^i, Z, T) \), where \( T \) is the threat and \( y^i \) is the private good. The non-cooperative (Nash) maximization problem for each ally is therefore: \( \arg_{y^i,q^i} \{ U^i[y^i, x^i, \beta(q^i + Q_{-i}) T] \} \), subject to a linear budget constraint: \( I^i = y^i + pq^i \). This implies that \( x\psi_{xy}^i + \beta\psi_{zy}^i = \theta \) where \( \psi_{xy}^i \) is the willingness to pay for the excludable security benefits and \( \psi_{zy}^i \) is for the alliance-wide security benefits. If \( \alpha = 1 \) and \( \beta = 0 \), then security is entirely excludable within the alliance, whereas if \( \alpha = 0 \) and \( \beta = 1 \), then the alliance is producing a pure public good.

The joint products model implies that alliance members usually possess incentives to supplement their security beyond that which the alliance provides, and thus predicts little if any disproportionality once we account for expected benefits. If the consumption of security is non-rival, there is no efficiency reason to exclude additional members: benefits are not attenuated, costs are just split into smaller and smaller shares. In conventional warfare settings, however, many alliance benefits are arguably rival (Sandler and Hartley, 2002: 876). If so, then the joint products model is likely to perform better in such settings than pure public goods models. Within the joint products model, the level of provision is thus unrelated to the number of members when excludable benefits to total benefits ratio is sufficiently large, whereas in the public goods model larger groups are predicted to be less efficient and to exhibit more free-riding (Cornes and Sandler, 1984). In the joint products model, free riding is unlikely to be a problem as long as the proportion of excludable to total benefits is sufficiently high, since marginal benefits and marginal costs are equated (Sandler and Sargent, 1995).

The observable implications of the two theories are therefore quite distinct about the extent of free riding we should expect. As the paper shows in the empirical section below, alliance members displayed very limited free-riding behavior, and generally bore costs proportional to the benefits they expected to receive from the alliance. In the final section, we point to several theoretical and practical implications of this result for economic theories of alliances.

THE HOLY ALLIANCE

The Holy Alliance (subsequently the Quintuple Alliance) was perhaps the first military alliance in the modern state system (Cresson 1922). It was initially an alliance of Russia, Austria and Prussia, created on the initiative of Tsar Alexander I in November 1815, and soon added Great Britain and France. Although initially religious and quite vague in its practical implications, Prince Metternich of Austria soon modified the alliance into a more practical instrument of statecraft (Kissinger 1957). In particular, the alliance originally was founded as a pact against secularism, but became a tool to suppress democratic revolutions and to defend monarchies. In short, it was a collective security pact in Olson’s original sense, designed to uphold the settlement of the Napoleonic wars, and therefore represents a valid test of the theory.

---

8Benefits in the joint products model are those ‘objects of value that are protected by conventional and/or nuclear forces’, including (1) vulnerable borders, (2) industries and (3) population centers. In its original form, expected benefits was taken as the average of each ally’s share of NATO’s GDP (proxy for industrial base), its share of NATO’s population and its share of NATO’s exposed borders. Other work has considered additional proxies, e.g., Khanna and Sandler, 1996, 1997; Sandler and Murdoch, 2000, and Khanna et al., 1998; Sandler, 2005. Sander and Forbes, 1980, estimate a proxy for the share of benefits that each ally receives from membership in the security institution and equate it to the share that each ally spends.
Using data from the Quintuple Alliance, I show that disproportionality in burden sharing was negligible and that each state relied primarily upon its own efforts. Data from this period are of course sparser and possibly less reliable than post World War II data. Given these issues, 1820 was chosen as the year with the most complete information needed to test the theories. Table I draws on historical datasets to provide the relative rankings for the members of the alliance in terms akin to Olson and Zeckhauser’s measures of ‘size’ and measures of ‘benefits’. Russia was the ‘largest’ member in terms of overall GDP, military manpower and raw population size. It was the goliath of the alliance. These various measures are provided to show that, regardless of the specific measures used, two patterns emerge clearly. Olson-Zeckhauser’s model predicts expenditure according to ‘size’: Russia > UK > France > Prussia > Austria. The UK was the alliance member with the greater expected benefits in terms of GNP per capita, iron and steel production and overall share of the world system’s capabilities. The joint products model therefore leads us to expect: UK > Russia > France > Austria > Prussia. Table II compares these predictive distributions to the observed distributions, and calculates the divergence.

The data structure is fortunately quite straightforward and does not require elaborate pattern recognition strategies. A simple measure of model performance is the sum of rank errors ($M_k = \sum_i (\hat{\phi} - \phi_i)$) for each model $k = 1,2$ and the average error in ranks $M_k/N$.

For model 1 (the pure public goods model), $M_1 = 6$, which implies an average error of $\frac{M_1}{N} = 1.2$. For model 2 (the joint products model), $M_1 = 2$, which implies an average error of $\frac{M_1}{N} = 0.4$. The ratio of errors from the two models provides a direct indicator of their relative performance: $\frac{M_1}{M_2} = 3.0$. The pure public goods model is three times more prone to error than the joint products model on these data. While neither model is perfect, I interpret the evidence as a reasonably strong endorsement of the joint products model over the public goods model.

It is not clear that any additional statistical machinery is warranted, but I report the results of the non-parametric Spearman rank test because it has been frequently reported in this literature. It is important to use exact rather than asymptotic p-values for these tests, however (Valz and Thompson 1994). In general, nonparametric methods are believed to be more appropriate when the sample is small, the data are ranked, and the analyst is uncertain about the parameters or the distribution of the variable of interest in the population. The statistical results reinforce the calculations above. The non-parametric Spearman

---

9The choice between parametric and non-parametric testing is difficult to make when the samples are small because it is not known whether the data come from a Gaussian population: non-parametric tests lack statistical power, and tend to inflate p-values when the data are in fact gaussian, but parametric tests are not robust and will produce incorrect p-values if the data are non-gaussian.

---

**TABLE I** Measures of “‘Benefits” and “‘Size’” for the Quintuple Alliance in 1820

<table>
<thead>
<tr>
<th>Alliance Member</th>
<th>Benefits</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GNP/cap</td>
<td>Iron and Steel</td>
</tr>
<tr>
<td>Prussia</td>
<td>1077 (4)</td>
<td>50 (5)</td>
</tr>
<tr>
<td>Russia</td>
<td>688 (5)</td>
<td>135 (3)</td>
</tr>
<tr>
<td>Austria</td>
<td>1218 (2)</td>
<td>70 (4)</td>
</tr>
<tr>
<td>UK</td>
<td>1706 (1)</td>
<td>320 (1)</td>
</tr>
<tr>
<td>France</td>
<td>1135 (3)</td>
<td>140 (2)</td>
</tr>
</tbody>
</table>

---

9The choice between parametric and non-parametric testing is difficult to make when the samples are small because it is not known whether the data come from a Gaussian population: non-parametric tests lack statistical power, and tend to inflate p-values when the data are in fact gaussian, but parametric tests are not robust and will produce incorrect p-values if the data are non-gaussian.
and Kendall correlation in ranks for the joint products model are quite high \( \theta = 0.90 \) and \( \theta = 0.80 \), respectively and are significant, whereas the same statistics for the pure public goods model are small and insignificant.\(^\text{10}\) This result makes sense because, as the discussion below illustrates, the alliance produced excludable benefits to its members, displayed imperfect substitutability and intra-alliance cleavages.

Certain shared security interests – say in preventing revolution in Europe – need not imply mutual interest on other questions – such as preventing it in the colonies. The Quintuple Alliance provides several examples of unsuccessful efforts to act collectively that call into question its internal coherence (Schenk, 1947). One typical failure followed from France and Russia’s proposal to aid Spain reconquer her colonies in America. Great Britain opposed, historians suggest, less on principle and more because it had begun benefiting from fewer restrictions on trade with Latin America (Bushnell, 1994). The alliance remained formally intact for several years, and is typically not considered to have become defunct until Alexander’s death in December 1825, but the credibility of its extended deterrence guarantees was never substantial (Artz, 1934; Spahn, 1910). When it was proposed that the alliance intervene in Spain to influence its constitutional struggle, Britain refused. When Russia and Prussia proposed to intervene against the Barbary Corsairs to limit piracy, Britain again baulked because it did not want a Russian fleet in the Mediterranean.

The alliance was largely unable to act as a collective due to such internal divides, especially rifts over intervening in other states. Britain, the largest alliance member, uniformly objected to Russian, Prussian and Austrian proposals to intervene abroad. Alliance members often appeared to fear each other as much if not more than they were concerned with threats to the alliance as a whole (Schroeder, 1994). If Britain felt that it could rely on France’s or Russia’s army, and they in turn could rely on Britain’s navy, this substitutability may have resulted in some disproportionality. Although they all shared a common interest in preventing instability within their borders, the alliance’s commitment to mutual defense was hardly credible enough to allow individual allies to free ride on the efforts of larger members.

The joint products model is more appropriate in these settings because it does not assume that extended deterrence will be fully credible, or that the alliance has one guiding purpose. Free-riding will be limited because ‘each man is for himself.’ The Quintuple Alliance displayed little if any disproportionality in burden sharing – marginal costs and benefits were roughly in line (Table II). The alliance did serve complementary interests, though not necessarily collective ones.

\(^{10}\)P-values are exact and not based on the Student’s t. Kendall’s tau for pure public goods model is 0.40 with p-value of 0.46.
The analysis of this seminal alliance in the European state system calls into question some aspects of the classically realistic view of alliances as mechanisms to aggregate individual efforts against external threats. The alliance did this of course to a degree, but it also and perhaps more importantly served an alternative function, which is emphasized in institutional accounts of alliances. It served as a ‘security management institution’ to monitor ‘allies,’ to deal with conflicts among them and to influence their behavior (Haftendorn, Keohane and Wallender, 1999). Without denying that alliances function to aggregate individual efforts and provide deterrence, the analysis indicates that their ability to make credible commitments varies considerably, and depends on the type of deterrence (nuclear or conventional), the size of the alliance (many members or few) and intra-alliances cleavages.

CONCLUSIONS

This study has examined the conditions under which the two main economic theories of alliances provide useful explanations for actual alliance behavior. While they may not be fully predictive models, they can provide crucial guidance in our effort to explain economic behavior in alliances.

The public goods theorists begin from the premise that security institutions serve the common interests of its members. It is typically assumed that defense expenditures are perfectly substitutable, meaning that more is always better, since it spreads costs over more actors while not adversely affecting benefits. If allied contributions were not assumed to be perfectly substitutable, then the collective good would not be pure. For instance, if its function was not to serve collective security interests but only complementary interests, or if increases in others’ expenditures were not viewed as necessarily beneficial, which were both the case in the empirical study of the Holy Alliance, then the public good model of alliances would in general not work very well. The pure public goods model also assumes that forces are mobile and expenditures are collectively transferable. These assumptions may be reasonable in certain modern alliances, especially those based on nuclear second strike capabilities. They seem less likely to hold in historical European alliances, and in alliances today in much of the developing world. Without credible commitments to shift (or ‘substitute’) forces, allies in conventional alliances can seldom be relied upon to defend the homeland. In terms of the public goods literature, defense in pre-World War II alliances is an impure public good, subject to intense rivalry, exclusion and imperfect substitution. The joint products model accounts for the impurity of the public good by having power projection and spatial rivalry influence the ratio of excludable to total benefits (Arce and Sandler, 2001). States rarely relax efforts on account of other’s increased efforts.

Alliances in the nineteenth century and arguably in much of the developing world today are as much about aggregating individual efforts against external threats as they are about controlling internal threats by keeping them close (Schroeder, 1994). As far as NATO and the Warsaw Pact are concerned – that is, for relatively large, alliances – the pure public goods assumptions may be more reasonable. The deterrence produced by most alliances is to some extent always rival –maybe less so in nuclear alliances and more so in conventionally armed ones –and this implies that the credibility of the alliance’s commitment to shifting forces shapes patterns of expenditure.

11If increases in allies’ expenditures meant an arms race, then the function of an alliance might be to monitor other nations more closely (e.g., Italy and Austria-Hungary).
The Warsaw Pact and NATO were relatively compact compared to the several pre-World War II alliances, including the Franco-Czech alliance, the Triple Alliances or the Quintuple Alliance. Allied forces in such alliances were not fully substitutable and power projection capabilities were truly limited, even for major powers. Substitutability is partly determined of course by geography, by transportation infrastructures and by intra-alliance rivalries. Again, one could speculate that these factors are less relevant in nuclear alliances. Most alliances are conventionally armed, and therefore more likely to exhibit behavior that is more consistent with the joint products model. The claim is not that the single case study disproves the pure public goods model, but rather that in combination with studies of other pre-World War II alliances (Thies 1987) it casts doubt on its applicability to a large subset of alliances. If economic models are to provide a more general explanation for variation in the distribution of costs among allied states, then the joint products model should be favored, since the pure public goods model can be derived from it as an important but limiting case.

Future research would benefit from further examining the issue of substitutability between forces and the implications of taking alliances as more than capability aggregators but also as mechanisms of control. More substitutability means that smaller allies can rely upon their larger allies and those with more intense preferences can bear more of the alliance costs. The same is true of power projection capabilities, which has historically limited the extent to which smaller allies could rely upon larger ones. Pre-World War II alliances generally worked against smaller allies, who were often left to fend for themselves when the alliance was called upon to serve its purpose. If substitutability and power projection capabilities have increased over time, albeit unevenly across the world, smaller allies should have more opportunities to free ride, so we should observe a secular increase in disproportionality. These are important issues for future research to investigate.

ACKNOWLEDGEMENTS

I wish to thank Daniel Arce, David Banks, Charles Becker, Lenka Bustikova, Michael Hechter, Donald Horowitz, Robert Keohane, Todd Sandler, Carolyn Warner, colleagues at Duke, Yale and Arizona State universities and reviewers for valuable comments and conversations, which have improved this article. All errors are mine.

References


