Free-riding in alliances: testing an old theory with a new method

Free-Riding in Alliances

Testing an Old Theory with a New Method

Published in:

Conflict Management and Peace Science (forthcoming)

Thomas Plümper[a] and Eric Neumayer[b]

[a] Department of Government, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK, tpluem@essex.ac.uk

[b] Department of Geography and Environment, London School of Economics and Political Science (LSE), London WC2A 2AE, UK, e.neumayer@lse.ac.uk

Abstract

We revisit the old and well-established theory of free-riding in military alliances. Existing empirical evidence infers free-riding from the larger military expenditures per gross domestic product (GDP) of countries of larger GDP. Yet, larger countries have broader military and geostrategic interests that result in larger defense burdens, thus creating an identification problem for existing tests of free-riding behavior. We therefore develop alternative predictions that ignore differences in the level of military spending and instead relate to growth in spending over time. The safety level of smaller members of an alliance is affected, simultaneously, by changes to military spending of the largest alliance member as well as by spending changes of the main enemy. Using the North Atlantic Treaty Organization (NATO) as test case, we estimate country-specific response functions of the smaller alliance members to growth in United States (US) military spending on the one hand and to growth of Soviet spending (if in excess of US growth) on the other hand, covering the period 1956 to 1988. Results from our quasi-spatial approach corroborate one part of the theory in that we find the vast majority of the smaller NATO allies to be free-riders. However, our empirical evidence flatly contradicts the other part of the free-riding theory: the extent of free-riding is not a function of country size. Smaller allies free-ride, but the relatively larger of the smaller allies do not free-ride any less than the relatively even smaller alliance partners.

Key words: military alliances, free-riding, burden sharing, NATO, defense spending

1. Introduction

Smaller members of a military alliance are suspected of free-riding on the defense burden covered by the largest alliance member and the more so the smaller they are (Olson 1965; Olson and Zeckhauser 1966). The theory of free-riding in alliances appears to be theoretically plausible, but the empirical evidence that has been brought forward to support the theory suffers from a serious identification problem. In short, the mere fact that larger North Atlantic Treaty Organization (NATO) allies have a larger share of military spending to their gross domestic product (GDP) or gross national product (GNP)[1] than smaller allies, that the correlation between country size and the share of military spending to GDP is positive, or that country size has a positive estimated effect on military spending as share of GDP do not provide convincing evidence for the free-riding in alliance hypothesis because larger countries have broader military and geostrategic interests than smaller members.

Compare, for example, the United States to Denmark. On the one hand, we have a small European country whose neighbors are NATO members, that had colonies in Faroe Islands and Greenland, and that has no geostrategic interest beyond the Northern Atlantic and North-West Europe. On the other hand, the USA is a superpower with troops stationed in 150 countries of the world, that entertains roughly 80 percent of the global fleet of air craft carriers, and that after the Second World War was involved in militarized conflicts on all continents with the exception of Australia and Antarctica. Political scientists should therefore not be surprised that the USA spends a larger share of her gross domestic income on defense than all other NATO countries.

[1] Some studies look at spending per GDP, others at spending per GNP. As this makes practically no difference, we use the GDP terminology throughout this paper.

Existing differences in geostrategic interests render simple tests of correlation of country size with the size of military burden or the effect of country size on the military burden implausible. Multivariate regression analysis could in principle deal with this problem if researchers managed to fully control for the influence of geostrategic interests on military spending. This is unlikely, however. Equally importantly, there are many other reasons why countries differ in the absolute size of their military spending or their spending relative to GDP, such as differences in military recruitment (conscription or professional army) or their historical legacy (e.g., West Germany). Unless one can adequately control for these other influences, which again seems implausible, multivariate regression in levels of military spending or spending to GDP will suffer from an identification problem.

This leaves us with a plausible theory – the theory of free-riding in an alliance – but implausible evidence supporting this theory. We therefore suggest alternative predictions from the free-riding theory which lead to different tests that do not focus on the allocation of the total defense burden among NATO allies but on the smaller NATO allies' responsiveness to growth in military spending of the USA and the Soviet Union, respectively. We do not analyze variation in *levels* of military expenditure between the NATO members, but variation in *growth rates* of military spending. Growth rates in military spending are far less subject to the sources for cross-country heterogeneity enlisted above and can therefore provide a cleaner test for the free-riding hypothesis than estimations in levels of military spending could – given the difficulty of controlling for this heterogeneity. To be sure, our new method is not entirely free of the identification problem described above, but growth rates are much less subject to the impact of heterogeneous geostrategic interests and other causes of

heterogeneity than levels of military spending.[2] In fact, if the change in geostrategic interests was uncorrelated with the level of geostrategic interests, our estimates would be free of bias from this source.

We test our alternative predictions of the "free-riding in alliances" theory over the period 1956 to 1988, just before the end of the Cold War, by using a quasi-spatial approach. We estimate the reaction functions of the non-US members of NATO to the growth rate in US and Soviet military spending. Specifically, we argue that free-riding is a function of the responsiveness of NATO allies to both growth in US spending and growth in Soviet spending (if in excess of US spending growth) taken together. We define a country as "free-rider" if the sum of both response parameters is smaller than 1.0. Based on this definition, we estimate country-specific degrees of free-riding. As we show, there is evidence for free-riding by the vast majority of smaller NATO allies, but the degree of free-riding is not a simple function of country size as measured by GDP. Instead, once Portugal is removed as an outlier we find the degree of free-riding to be negatively correlated with sharing a land border with a Warsaw Pact country or being separated from it by sea by less than 150 miles and to be positively correlated with increasing distance between the ally's capital and Moscow.

2. The Theory of Free-Riding in Alliances: A Brief Review of the Literature

This section reviews the theoretical and empirical literature on free-riding in military alliances. We focus on NATO as the most stable and important military alliance in the world. In one of his last speeches as departing US Defense Secretary, Robert M. Gates warned of a

[2] Moreover, in one robustness test we further allow for time-invariant heterogeneity and an ally-specific trend in growth rates to further eliminate such contamination of the testing of the free-riding hypothesis.

growing divide within the NATO "between those willing and able to pay the price and bear the burden of commitments, and those who enjoy the benefits of NATO membership but don't want to share the risks and costs" (International Herald Tribune 2011). Gates predicted that future political leaders of the USA, "those for whom the Cold War was not the formative experience" as it was for him, may be less inclined to accept an unequal burden sharing between the US and her European allies. This American political discussion about burden sharing suggests that the by far largest ally, the US, feels exploited by all other allies ("the Europeans") and in the absence of an obvious distinction, by each one of the European allies. In this paper, we will focus on the hypothesis of "the exploitation of the largest ally by all others".

2.1. The Theory of Free Riding in Alliances

Countries join military alliances to pool their resources against a common threat. However, complaints about the "free-riding" of European allies at the expense of the USA are (almost) as old as the NATO. To political scientists, there is little new in the suggestion that NATO offers incentives for free-riding behavior by the smaller allies. The theory of free-riding in alliances, as first developed by Olson and Zeckhauser (1966), argues that defense in an alliance is a pure public good: the benefits associated with defense spending are non-excludable and non-rival in consumption among allies (Sandler and Hartley 1999: 29). Two core predictions follow from this theory: First, the largest ally – call her the hegemon – bears a disproportionately large share of the aggregated defense burden of the alliance. And second, the smaller an ally is, which the extant literature typically defines in terms of GDP, the more it free-rides. The reason for the second prediction is that the smaller an ally is, the easier it is for this ally to under-supply its contribution to the public good since it can reduce its military spending with no noticeable effect on the alliance's overall ability to defend its borders, while potential cuts in defense spending of larger countries would have a more noticeable effect on

the alliance's military capabilities (Olson 1965; Olson and Zeckhauser 1966; Palmer 1990a; Sandler 1993). If, for example, a small NATO member, say Denmark, halved her military expenditures, the total defense spending of NATO would decline by less than 0.5 percentage points. The United States, the NATO's biggest defense spender, could increase her military spending by a little less than 0.5 percent to compensate for Denmark's blatant attempt to free ride.

This does not mean that middle-sized allies such as Italy or even one of the larger ones like West Germany or the United Kingdom do not free-ride. If one single country strongly dominates an alliance as is clearly the case with the USA in NATO, every single one of the smaller allies has an incentive to under-invest in defense. To put it bluntly: Since the defense spending of each single one of the non-US NATO allies is virtually irrelevant for the alliance's joint military strength, the theory predicts that small countries under-contribute to the joint effort or make no own defense spending contribution at all (as Iceland does). Why does the hegemon put up with free-riding then? The answer is that the dominant country in alliances can accept some extent of free-riding of smaller alliance members because despite free-riding the alliance reduces the cost of an arms race with other large rival powers and their alliances. While in relative terms they contribute less than the dominant country in the alliance (Diehl 1994), in absolute terms, even free-riders contribute to the fight against the enemy and the alliance is more powerful than the dominant nation alone.[3] For example, in the

[3] An alliance may also be stronger than the sum of its parts by generating scale economies in government procurement of military products. Alliance members are more likely to agree on joint production of military goods and they may even agree to specialize in their production. Thus, the existence of an alliance is likely to be beneficial for the dominant alliance member even if smaller or all other countries partly free-ride on the military expenditures of the dominant ally. Under these circumstances, alliances

case of NATO the smaller alliance members contributed approximately 32 percent to the alliance's total defense spending over the period 1955 to 1988, even if no single one of the smaller allies produced more then 7.4 percent (the United Kingdom).

The theory of free-riding in military alliances clearly predicts that smaller allies exploit the hegemon. Yet, to what extent the smaller allies are allowed to free-ride is less clear. As Palmer (1990a, 1990b) shows, if a Cournot-type model were to adequately describe the behavior of alliance members, then one would see a strong form of free-riding by the smaller allies: smaller allies negatively respond to an increase in the defense effort by the US by lowering their own effort. Such a Cournot-type model is based on short-term utility maximizing behavior, which in respect to defense policies appears to be an unrealistic and implausible assumption. Palmer shows that a bargaining-type model in which the alliance members play an iterative game over time and can bargain over their relative contributions results in a much milder form of free-riding in that the smaller allies now positively respond to an increase in the defense effort by the US, but do not necessarily increase their own effort by the same proportional amount. Such a bargaining-type model also neatly supports the second core prediction of the free-riding in alliances theory. Since the marginal absolute effect of spending increases by the smaller of the small allies is smaller than the effect of spending increases by the larger of the small allies, it follows that the larger of the small allies "will respond more to changes in the defense effort of the hegemon (the U.S.)" (Palmer 1990: 156) than will the smaller of the small allies.

Another strand of reasoning has also provided arguments why the extent of free-riding by the smaller allies and thus the exploitation by the hegemon may be severely limited. Sandler and

---

can be remarkably stable despite free-riding, moderate conflict of interest, and the absence of a plausible military threat.

his co-authors have questioned the extent to which a military alliance like NATO really provides a public good (Sandler 1977, Sandler and Forbes 1980; Murdoch and Sandler 1984; Sandler and Harting 2001). As mentioned at the start of this section, the theory of free-riding in military alliances is based on the assumption that the alliance produces a public good that is non-rival in consumption and from which none of the allies can be excluded.

But is defense non-rival in consumption among alliance members? On a trivial level, military goods are rival in consumption. If the US Army fires a bullet on an enemy, another NATO army cannot fire the same bullet on another enemy. Yet, on a non-trivial level the case is less clear. If defense spending aims at deterrence, then deterrence by the US military of its enemies will be non-rival in consumption by other alliance members since it will also deter other NATO members' enemies if these enemies believe the principle laid out in article 5 of the NATO treaty. Especially after NATO abandoned the strategy of mutually assured destruction and implemented the strategy of flexible response in 1967, smaller allies may have wondered, however, whether they are important enough to guarantee a NATO intervention in case of an open military dispute, which puts doubt on the credibility of deterrence. At the same time, the US begun to develop more "protective" than "deterrent" weaponry (Sandler 1977, Sandler and Forbes 1980). Sandler and his co-authors thus augment Olson and Zeckhauser's pure public good theory of military deterrence to account for "changes in NATO's military strategy and the development of new weapon systems" (Sandler and Forbes 1980: 426). To theoretically account for these changes, their 'joint product' model distinguishes between "deterrent", "protective", and "defense" weapons. In their view, only deterrent weapons provide pure public goods, while the other types are excludable. This approach perceives all weapons as being based on a continuum with protective and deterrent weapons marking the two extremes (Sandler and Forbes 1980: 427). According to this line of reasoning, the extent to which the smaller allies can exploit the

hegemon becomes a function of the extent to which weapons are protective and defensive rather than deterrent: "…a high ratio of excludable benefits – ally-specific and damage-limiting protection – to total benefits means that an ally must support its own defense, regardless of its size, if it is going to be protected. As this ratio approaches one, the exploitation hypothesis is anticipated to lose its relevancy, so that the disproportionality between allies' GDP and their share of GDP devoted to defense is expected to dissipate" (Sandler and Harting 2001: 878).

Yet, a move toward "protective" weaponry does not itself weaken the credibility of deterrence. All other things equal, countries that own protective weaponry might even increase their commitment to protect smaller allies as their costs of doing so would decline, given that they are now more protected from the devastating consequences of an attack by the enemy. In other words, the credibility of the American commitment to protect smaller allies partly positively depends on the effectiveness of the US's protective weaponry. Such weaponry thus does not necessarily produce exclusively private benefits.

Alignment of interests and the consequent credibility of the hegemon's commitment to defend the smaller allies thus trumps any distinction of type of weaponry. The same is true for the non-excludability of alliance members, the second requirement of a public good.[4] The North Atlantic Treaty that established the NATO promises non-excludability in article 5: "The Parties agree that an armed attack against one or more of them in Europe or North America shall be considered an attack against them all and consequently they agree that, if such an armed attack occurs, each of them (…) will assist the Party or Parties so attacked

[4] Non-alliance members can of course always be excluded, making defence in an alliance a club good rather than a pure public good. We nevertheless use the language of public good, keeping in mind that the relevant population are only countries within, not outside an alliance.

(…).” If this rule were obeyed under all circumstance, then military spending would produce a non-excludable good. However, the solemn promise of article 5 has never been tested – the only time Article 5 was invoked was after 9/11 when the biggest member of the alliance was under attack, not by a foreign state, however, but by international terrorists. It is therefore unknown whether some governments of NATO countries would decide not to defend a smaller ally in violation of the wording of the treaty. In fact, in some sense article 5 already contains a potential loophole in that it states that the alliance “will assist the Party or Parties so attacked by taking forthwith, individually and in concert with the other Parties, such action *as it deems necessary*, including the use of armed force, to restore and maintain the security of the North Atlantic area” (emphasis added). For the smaller allies a particular problem would arise if the US commitment to article 5 were in doubt. However, incentives for smaller allies to free ride continue to exist despite the largest ally’s commitment to protect them being uncertain. In fact, the only way the dominant ally could eliminate free-riding is by credibly committing herself not to honor her commitment towards smaller alliance members that free-ride. In the presence of at least partly aligned interests, however, such a statement could never become credible. The largest alliance member could only deter being exploited by her smaller allies via a credible threat of exclusion if the protection of these smaller allies from a potential enemy is not in the interest of the hegemon.

In sum, the extent to which defense in an alliance is a public good and therefore invites free-riding by the smaller alliance members is a function of the homogeneity of interests within the alliance (Gates and Terasawa 2003) rather than determined by the existence of an alliance per se or a matter of military technology. Unless the interests of alliance members are independent, the existence of NATO allows the smaller allies to free-ride to some extent. To what extent is essentially an empirical rather than theoretical question and we next turn to the available empirical evidence.

2.2. Existing Empirical Evidence for Free-Riding in Alliances

Theoretically, free-riding in the provision of a public good can be clearly defined as receiving larger benefits from the public good than contributing to the costs of providing the public good. Empirically, in the context of free-riding in military alliances the problem is that the benefits an individual alliance member receives from the public good are very difficult to estimate. Most empirical tests have therefore focused on the contribution to the cost side. Here as well scholars encounter problems, however, since not all military spending contributes to the public good of the alliance, but contributes to the private interests of a country.

Empirically, free-riding is a matter of definition. Whether it is detected for an alliance member depends on the definition and on the identification strategy employed. The simplest definition of "free-riding" merely compares the share of military spending to a country's GDP across alliance members, as Olson and Zeckhauser (1966) in their seminal contribution have done, finding support for both parts of the free-riding hypothesis. This is not surprising, defined in this way, free-riding becomes immediately obvious. The US's military spending as share of its gross domestic product (GDP) was almost 7.5 percent over the period 1956 to 1988 while all other NATO members on average spent 2.1 percent of their GDP for defense, with the smaller of non-US NATO allies on the whole below and the larger of allies above this average (data from Whitten and Williams 2010).

However, this definition invites an obvious counterargument. It simply assumes that the NATO members have identical geostrategic interests outside the NATO area. This is implausible. The USA has global military interests that other NATO members either do not have or have to a far lesser extent. One should thus expect that the USA devotes a larger

share of her GDP to military spending. Comparing defense spending to GDP ratios does not make sense if one tries to identify free-riding.

A very similar and thus not any more plausible second test uses correlations: the NATO members' military expenditures as percentage of their GDP is correlated with their total GDP, indicating that larger members contribute a larger share of their total income to defense. Olson and Zeckhauser (1966) also show that the correlation coefficient between a NATO member's defense budget as a percentage of GDP and its GDP is positive.[5] They conclude that "there is a significant positive correlation indicating that the larger nations in NATO bear a disproportionate share of the burden of the common defense" (Olson and Zeckhauser 1966: 275). Interestingly, even some critics of the free-riding hypothesis have used bivariate correlation analyses to cast doubt on the hypothesis. Russett (1970) was one of the first to show that the explanatory power of the free-riding hypothesis declines over time. Reinforcing this argument, Sandler and Forbes (1980) demonstrate that the correlation between defense expenditures and GNP began to diminish in the mid-1960s and loses statistical significance after 1967. While Olson and Zeckhauser reported a rank correlation between the military spending to GNP ratio and GNP of .490, Sandler and Forbes show that this correlation is only .319 in 1960 and declines to 0.099 in 1975. Sandler and Forbes attribute the declining correlation to changes in NATO's strategy. Again, however: such correlation does not corroborate the free-riding hypothesis as larger allies, and the USA in particular, also have costly military interests beyond the NATO area.

Only the third employed technique, multivariate regression analysis, controls for confounding factors that also influence a country's willingness to invest in defense. Here as well, studies have found a significant effect of country size on the share of military spending to GDP (e.g.,

[5] They look at defense spending relative to GNP rather than GDP.

Oneal and Diehl 1994). Multivariate regression analysis is more reliable since it can take some of the confounding factors into account. Yet, early models were notably parsimonious. Sandler and Murdoch (1990) include income and allied spending as well as Soviet military spending as explanatory variables. Oneal (1990) as well as Oneal and Diehl (1994) include economic size, the fraction of NATO's annual expenditures accounted for by contiguous allies (zero else) – a variable meant to control for the geographical distribution of NATO's capabilities – and the number of militarized disputes the NATO member countries are engaged in over the 5 years prior to the estimation year. Apparently, this latter variable marks a first attempt to account for heterogeneous interests. However, the operationalization of this variable violates the general idea that the vast majority of military expenditures aims at deterrence and not at actually engaging in militarized conflicts. This variable may thus account for the differences between Portugal and, say, Denmark, but not necessarily for the costly military interests the US entertains all around the world in order to deter actual and potential enemies. Oneal and Diehl (1994) also include military expenditures of the Soviet Union and a variable for the tensions between the US and the Soviet Union. They analyze the period from 1950 to 1986 and find a positive coefficient for country size. They also separately analyze the years before 1968 and after 1967 and find that the coefficient of country size becomes significantly smaller and drops from 0.21 to 0.05, but does not lose statistical significance. Finally, they show that NATO countries of smaller economic size respond less to Soviet military spending than larger NATO countries. Unless the broader geostrategic interests of larger alliance members are adequately accounted for, country size may still simply catch the effect of expensive broader military and geostrategic interests of larger allies like the USA, France, or the United Kingdom which have little effect on the safety of, say, Denmark and Norway who do not entertain such interests.

Palmer (1990a, 1990b) also applies multivariate regression. While he does not control for heterogeneous geostrategic interests, he goes some way in the direction of our proposed new method by estimating the reaction of two groups of the non-US NATO allies, namely the smaller ones and the bigger ones, to US defense spending per GDP. He finds that both groups of smaller allies respond positively to higher US defense spending per GDP but with a coefficient of less than one, thus supporting the bargaining-type model of free-riding rather than the Cournot-type model. He also finds that the group of bigger of the non-US NATO allies respond more strongly than the group of the smaller of the non-US NATO allies. However, as we will argue further below, one also has to take into account the reaction of allies to Soviet military spending and by estimating in levels of spending per GDP rather than growth in military spending, Palmer's test is again contaminated by the impact of heterogeneous geostrategic interests on the estimations.

3. An Alternative Test of the Free-Riding Hypothesis

The existing empirical evidence for the theoretically plausible albeit not uncontested theory of free-riding in alliances rests on weak testing strategies. They all suffer from an important identification problem. Since larger countries typically have broader geostrategic interests, the effect of country size on military spending could be caused by these different interests rather than by free-riding. In other words: Even in the absence of NATO, the US with its global military interests would still devote a larger share of her income to defense than, say, Denmark and Italy. More generally, differences across allies, both observed and unobserved, in the form of their history, political system and so on also cause differences in levels of military spending that have nothing to do with free-riding.

Clearly, the first-best strategy would be if one could sufficiently well measure and therefore control for these other reasons why countries' defense spending differ that have nothing to do

with free-riding in military alliances. Unfortunately, this is not possible. A persuasive test of the free-riding hypothesis therefore requires a different identification strategy, which we develop in the remainder of this section.

Importantly, in our proposed new method, we switch from estimating in levels of military spending to growth in military spending over time. By taking out level effects, we exploit the fact that heterogeneous geostrategic interests as well as other reasons which cause countries' levels of military spending to strongly differ are arguably much less likely to have a strong effect on changes in their military spending over time. For example, it takes a much larger military force to pursue expensive regional or global geostrategic interests, but it does not require a country to outgrow others in pursuing changes to their military spending year after year.

Equally importantly, we employ a quasi-spatial approach to testing augmented predictions of the free-riding in alliances theory. We argue that free-riding is best studied by looking at the responsiveness of smaller NATO members to changes in the military balance between the Soviet Union and the USA and that therefore any test of free-riding must take the responsiveness of smaller allies to growth in the military spending by both superpowers into account. The incentives to free-ride for smaller NATO members result not simply from the total defense burden that the USA musters, but from changes in the defense burden of the USA and the Soviet Union over time. Our alternative interpretation is based on the premise that incentives to free-ride are a function of the safety level of NATO members. Changes to this safety level are triggered by growth in US spending on the one hand and growth in Soviet spending if in excess of US spending on the other hand.[6]

[6] We use growth in Soviet Union military spending in our main estimations, but growth in Warsaw Pact military spending in robustness tests.

All other things equal, an increase in US spending raises the safety level of NATO members as it shifts the security balance between the alliance and its adversaries in the alliance's favor. Free-riding on the USA seemingly occurs if her allies increase their spending in response to higher US spending less than proportionally: they free-ride on the US effort to increase NATO's safety level if they match an increase of US military spending of one percent with an increase in their own military spending by less than one percent and the more so the further away from one percent.

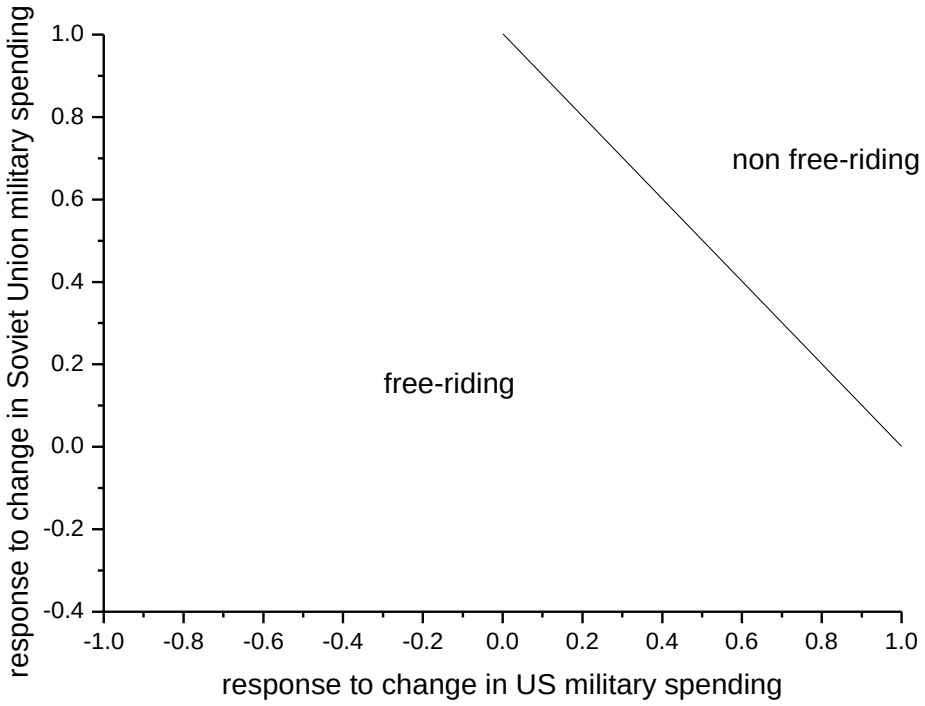
Yet, all other things are not equal since a stronger increase of Soviet above US spending meanwhile lowers the safety level of NATO members. For a full judgment on free-riding, one must therefore also take into account the responsiveness of the smaller NATO allies to situations of stronger Soviet relative to US spending growth. By increasing in certain years its spending less than the Soviet Union, the USA lets the security balance slip against the alliance. The smaller allies can thus make contributions to the public good of alliance security not merely by matching US military spending growth, but also by matching Soviet spending growth in years in which the Soviet Union increases military spending more than the US does. If, hypothetically, the smaller allies were totally unresponsive to higher US spending as such, but fully matched a stronger increase in Soviet relative to US spending with an equally strong increase in their own spending, then in total they have fully contributed toward preventing the security balance from moving in the adversary's favor.

The upshot is that in order to assess free-riding we need to examine a combination of the two types of responsiveness. Allies do not free ride on the US if they either fully match all US spending increases or they match all Soviet spending increases above US spending increases or they combine the two types of responsiveness with each other such that the combined sum fulfils these requirements. This gives us two polar cases, which can then be linearly combined with each other to create a free-riding threshold. To illustrate graphically, the point at the co-

ordinates (1.0; 0.0) in figure 1 below is the first point through which all plausible free-riding thresholds should go through. That is, a small NATO ally that fully responds to growth in US military spending with equal growth and simply ignores growth in Soviet Union military spending cannot be called a free-rider since it perfectly matches US spending changes. Second, the co-ordinate point (0.0, 1.0) should also be accepted as second plausible anchor point of the free-riding threshold. In words, a country that fully responds to any deterioration of the security balance in the adversary's favor – a consequence of Soviet spending increase in excess of US spending increase – by fully matching Soviet spending increases cannot be called a free-rider since it fully contributes toward stemming a shift in the security balance against the alliance, even if it simply ignores growth in US military spending in all other years.

The free-riding threshold is then simply the linear combination of the two co-ordinate points in figure 1: Free-riding occurs for points to the interior of the threshold line in the sense of lying between the co-ordinate axes and the threshold line, whilst absence of free-riding is represented by points that lie to the exterior of the threshold line. Expressed numerically, we therefore define free-riding as follows: If, over our estimation period 1956 to 1988, the estimated responsiveness of a NATO ally to US spending growth *plus* the responsiveness to growth in Soviet spending (if in excess of US spending) is smaller than 1.0, then an ally can be called a free-rider and the smaller this sum of the two estimated degrees of responsiveness the larger the extent of free-riding. Conversely, a sum of estimated degrees of responsiveness equal to or above one signals the absence of free-riding.

Figure 1: The Free-Riding Threshold



We do not claim that our definition is 'correct'. Definitions are not correct, they serve a purpose, they make a distinction, and they ought to be plausible. We have argued above that the two anchor points (1.0; 0.0) and (0.0, 1.0) are plausible. Let us further illustrate our definition. For years in which the US and the Soviet Union both increased their defense spending at the same rate of, say, 3 percent, a NATO member free rides if the growth rate of its defense spending is, on average in those years, less than 3 percent and it free rides the more the further below 3 percent its own growth rate is – controlling for relevant covariates. For years in which the US growth rate of military spending exceeds the Soviet Union's growth rate, our definition would similarly identify a smaller NATO ally as free-riding if its growth rate is, on average in those years, below that of the US. In years in which the Soviet Union increased its military spending by more than the USA does, a smaller NATO ally would be free-riding if, on average in those years, it increased its own military spending by less than the Soviet Union does.

At first sight, this last example appears to be very exacting of the smaller NATO allies, seemingly demanding them to grow their military spending by more than the USA does in response to Soviet spending if the latter's growth is in excess of US spending growth. This would appear placing higher demand on the smaller allies than on the USA. This is not the case, however. Keep in mind that our definition does not demand that smaller NATO allies grow their military spending both proportionally to US spending increases and proportionally to Soviet spending increases (if in excess of US spending increases) at the same time. Instead, all our definition requires is that the sum of responsiveness to the two growth rates is 1.0 or above. This would allow NATO members to have a growth rate of military spending that falls behind the US growth rate in years the USA grows its military spending faster than the Soviet Union as long as the smaller ally makes this up by larger than US spending increases in years in which the Soviet Union grows its military spending by more than the USA does. In other words, being a little responsive to US spending growth can be compensated by being strongly responsive to Soviet spending growth when Soviet spending growth exceeds that of the US, and vice versa.

We understand that by focusing on growth in military spending we change the perspective on free-riding. We do so because our alternative strategy reduces the identification problem that we have discussed in section 2. Our modeling strategy is not without drawbacks. Perhaps most importantly, our approach would not identify free-riding if the initial distribution of defense spending is very unequal with the smaller clearly exploiting the largest alliance member, but – starting from there – the smaller members of an alliance implementing similar growth rates over time as the largest ally or even higher growth rates than the hegemon. As we will see further below, our findings suggest free-riding by all but one of the smaller allies. Hence, it is not the case that we find results which could spuriously suggest the absence of free-riding whereas in truth there is free-riding. However, the estimated degrees of free-riding

and the ordering of countries according to their estimated degrees of free-riding may be misleading if a country that contributed relatively little to the alliance public good in the beginning increases its military spending relatively more over time than a country that contributed relatively more to the alliance public good in the beginning of our study.[7] As we said before, no method is perfect and this represents a drawback of our proposed new method.

We also accept that our new method is not entirely free of some of the issues and problems we raised against the old methods. Even if arguably much less affected, heterogeneous geostrategic interests can still impact growth rates in military spending as well. Bias occurs if changes in geostrategic interests are correlated with levels in geostrategic interests, a possibility that we cannot rule out entirely. However, the bias from this possible correlation will be smaller than the omitted variable bias of neglecting the variance in geostrategic interests altogether. We will also control for some of these confounding factors in our empirical research design and in a robustness test we additionally allow for time-invariant country heterogeneity and country-specific linear time trends. The latter should account for varying geo-strategic interests over time and their effect on military spending growth rates as long as they approximately follow a linear trend, such as for example fewer and fewer regional and global security interests over the period of study. More importantly, however, we contend that our proposed method represents a step in the right direction. Like all scientific progress, it cannot address all existing problems at once and still leaves many questions unanswered.

[7] We are grateful to an anonymous referee for drawing our attention to this caveat.

4. Empirical Research Design

We test the augmented predictions derived from the theory of free-riding in alliances in a quasi-spatial model. Accordingly, we regress the growth rate in military spending of NATO members other than the USA on the US growth rate in military spending and the Soviet Union growth rate in military spending if in excess of US spending increases (this variable is therefore set to zero in years in which US spending increases exceed Soviet spending increases). We call these estimation models quasi-spatial because growth in military spending by the smaller NATO allies is modeled as a function of growth in military spending by other countries (here: the USA and the Soviet Union). Different from spatial models, we assume no feedback from growth in spending by the smaller NATO allies on either US or Soviet spending, assuming the latter in effect to be exogenous. We have argued above that, as a first approximation, the spending of each single one of the smaller allies is irrelevant to the US and, by implication, to the Soviet Union, which justifies the assumption of treating US and Soviet spending decisions as exogenous. Any single one of the allies did not contribute more than at most 7.4 percent (the United Kingdom) to NATO defense spending during the period of our study.

Recall that the theory predicts that smaller countries free-ride more strongly than others. The standard approach toward testing this hypothesis would be to interact the US and Soviet growth rate variables with a variable measuring the country size of the smaller NATO allies. This would clearly test the hypothesis that free-riding is a function of country size. However, it would impose the assumption of a fixed and linear influence of country size on the degree of free-riding. Instead, we opt for a superior alternative and let the data tell us the degree to which each of the NATO allies free ride, if at all. We do so by estimating separate response functions for each of the countries in our sample. Doing so still allows us to investigate whether the country-specific degrees of free-riding are correlated with country size.

Our dependent variable is the growth rate in absolute military spending in real US$ rather than growth in military spending per GDP. Firstly, governments directly control military spending, not spending per GDP and, secondly, security is determined by military spending, not by the ratio of spending to GDP.[8] Country-specific response functions to the US growth rate in military spending and the Soviet growth rate (if in excess of the US growth rate) are our central explanatory variables. It would be theoretically superior to estimate the reaction of the small NATO allies to changes in the military threat they are exposed to, for which military capabilities and hence military spending are only a proxy variable.[9] However, military spending is observable, whereas threat and the true intentions of the enemy are not. Military expenditure estimates for Warsaw Pact countries are notoriously uncertain and we use only Soviet spending in our main estimations, but our results are robust to using spending by all Warsaw Pact nations instead, as our robustness section shows.

A number of control variables account for other factors determining growth in military expenditures. The growth rate of real GDP accounts for the fact that economic growth tends to lead to higher tax revenues which makes it easier to increase military budgets. We also include a measure of the intensity of armed conflicts in which a NATO country was involved in during a year, the lagged dependent variable to account for temporal dynamics and a linear year variable to account for any potential residual global trend. We include the level of military spending (relative to GDP) in the initial period (average of first three years) to account for the possibility that countries with a higher level of military spending at the outset grew their expenditures more slowly over time than countries with a lower initial level of military

[8] For example, if the Soviet economy shrank during the 1980s, as it probably did, this would increase its "defense burden", but would not increase the military threat if absolute expenditures did not increase.

[9] We are grateful to an anonymous referee for pointing this out.

spending. Lastly, since the military expenditures of contiguous allies is likely to be spatially clustered beyond what we control for in our model, we introduce a spatial error term into the estimations based on the predicted errors from the baseline model and a weighting matrix that uses contiguity (sharing a land border or being separated by sea by less than 150 miles) as connectivity variable. All data are taken from Whitten and Williams (2010), complemented with data from the Correlates of War project (http://www.correlatesofwar.org/) and from Gleditsch et al. (2002). Table 1 provides summary descriptive statistics. We cluster standard errors on countries.

Table 1. Summary descriptive statistics (N=395).

	Mean	Std. Dev.	Min	Max
Military spending growth rate	0.033	0.128	-0.454	0.519
US military spending growth rate	0.017	0.061	-0.097	0.237
Soviet Union spending growth rate (if > US rate)	0.031	0.045	-0.015	0.155
GDP growth	0.031	0.031	-0.083	0.177
Intensity of armed conflict involvement	0.261	0.849	0	6
Initial level of military spending to GDP	2.713	0.786	1.289	3.979
Spatial error term	0.000	0.105	-0.397	0.445

Iceland, Luxembourg and Spain are the only three NATO members not in the sample. Iceland has no independent army, we have no data for Luxembourg and Spain joined NATO only in 1982, becoming progressively integrated over time, which is too short given we restrict the analyses to the period 1956 to 1988. The years prior to 1956 were heavily affected by the Korean war and its aftermath, while the fall of the Berlin wall and the end of communism in Eastern Europe fundamentally changed the East-West antagonism for which NATO was originally created.

5. Results

Our empirical specification allows for country-specific responses to the US growth rate and Soviet growth rate (if in excess of the US growth rate) in military spending, for which table 2 reports results. Before we turn to discussing the country-specific response functions, let us briefly report results for the control variables. The lagged growth rate has a positive coefficient that is far from being statistically significant, however. This suggests that there is no temporal persistence in military spending growth rates controlling for the other explanatory variables in the model. We find the expected positive effect for the GDP growth rate on military budgets. As countries grow faster, they also increase their military budgets more than if they grow more slowly.[10] A one percentage point increase in GDP is associated, on average, by an 0.72 percentage point increase in military spending. We find no significant effect of conflict involvement on the growth rate of military budgets. Probably the reason for this is that for most countries defense budgets are so large even in peace years that fighting a limited armed conflict does not put much extra stress on countries' military expenditures. With the exceptions of the colonial wars of France and Portugal, the smaller NATO allies did

[10] This could dampen any degree of free-riding: If the resources not contributed to defense in period 1 were spent in a way that fosters economic growth, then in period 2 military spending increases because of a higher economic growth rate. However, the effect is too small to be relevant. Assume that all resources not spent on defense are used for investment. The coefficient with which investment translates into economic growth usually varies between 0.3 and 0.4. Thus, if one percent of GDP not spent on defense were invested, it would increase GDP growth by 0.4 percent, which according to our estimates would translate into an increase of merely slightly less than 0.3 percent in the defense budget in period 2. If we relax the unrealistic assumption that all non-military spending is invested and assume that only half of it is invested the effect would also halve. A countervailing force would be, however, if smaller countries had a higher income elasticity of military spending, for which Palmer (1991) finds some evidence.

not fight any major prolonged wars during our period of study and entertained militaries large enough to fight the very few and limited wars such as the one in the Falklands with only a small increase in military expenditures during the conflict period. We find the expected positive coefficient for the spatial error term, which captures remaining spatial clustering in military spending growth. Lastly, the initial level of military spending has no statistically significant impact on growth rates.

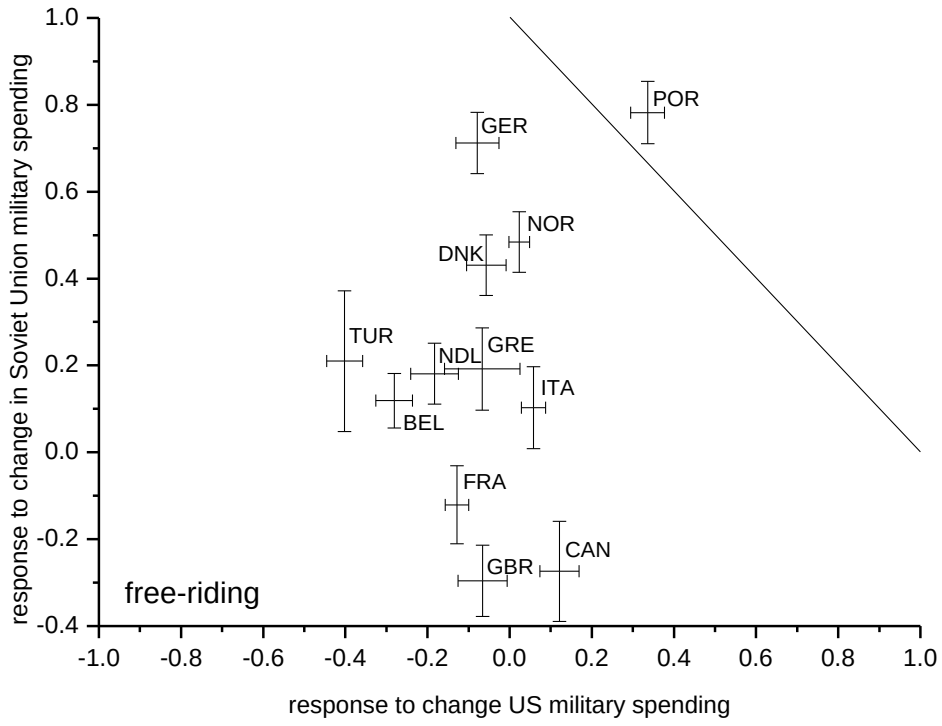
Table 2. Estimation results for entire period 1956 to 1988.

Country-specific response of:	to US growth	to Soviet growth (if in excess of US growth)
Canada	0.121**	-0.274**
	(0.0477)	(0.115)
Great Britain	-0.0659	-0.296***
	(0.0601)	(0.0820)
Netherlands	-0.183***	0.181**
	(0.0581)	(0.0704)
Belgium	-0.281***	0.119*
	(0.0446)	(0.0626)
France	-0.128***	-0.121
	(0.0285)	(0.0897)
Portugal	0.336***	0.782***
	(0.0411)	(0.0718)
West Germany	-0.0786	0.712***
	(0.0524)	(0.0708)
Italy	0.0583*	0.103
	(0.0293)	(0.0942)
Greece	-0.0669	0.192*
	(0.0918)	(0.0946)
Norway	0.0233	0.484***
	(0.0250)	(0.0699)
Denmark	-0.0568	0.431***
	(0.0479)	(0.0697)
Turkey	-0.402***	0.210
	(0.0437)	(0.162)
Lagged dependent variable	0.0635	
	(0.0469)	
GDP growth	0.723*	
	(0.386)	
Intensity of armed conflict involvement	0.000650	
	(0.00613)	
Initial level of military spending to GDP	0.00963	
	(0.00651)	
Linear year trend	0.000263	
	(0.000348)	
Spatial error term	0.555***	
	(0.0981)	
Constant	-0.542	
	(0.686)	
Observations	395	
R-squared	0.305	

Note: Dependent variable is growth in military expenditures. Standard errors clustered on countries in parentheses. * statistically significant at 0.1, ** 0.05 level, *** 0.01 level.

Turning to the ally-specific response functions, recall that we have argued that the response to the US growth rate plus the response to the Soviet growth rate in case the latter exceeds the US growth rate indicates the degree of free-riding, if any – a sum of coefficients of 1.0 or above suggests no free-riding. This information is best represented graphically. Figure 2 plots the responsiveness to the Soviet growth rate (if in excess of the US growth rate) on the y-axis against the responsiveness to the US growth rate on the x-axis for each of the NATO countries in the sample, together with their respective 95 percent confidence intervals. It also displays the threshold for our definition of free-riding.

Figure 2: Response Functions for 12 NATO Members with 95 Percent Confidence Intervals.



The distance to this threshold marks the degree to which countries free ride. According to our definition, our results reject the free-riding hypothesis for only one country clearly, namely Portugal, which is above the free-riding threshold. However, Portugal is an outlier in many respects. Most importantly, for most of our estimation period, Portugal was governed by a

right-wing dictatorship first under Salazar and then Caetano.[11] If we only estimate the response functions during Portugal's democratic period, then this country would also be detected as a free-rider. Figure 2 thus supports the first part of the free-riding hypothesis in that with the exception of Portugal all of the smaller allies free-ride. It does not, however, support the second part of the hypothesis that the degree of free-riding is correlated with country size: larger countries do not appear to be systematically closer to the free-riding threshold than smaller countries. This is confirmed by bivariate correlation analysis between the estimated degrees of free-riding and the average size of GDP of the NATO allies over the estimation period, which suggests that the degree of free-riding is not correlated with country size ($r = 0.10$, p-value 0.76 if Portugal is included and $r = -0.12$, p-value 0.74 if Portugal is excluded from the sample). Instead, once Portugal is removed from the sample as an outlier, the degree of free-riding is negatively correlated with either sharing a land border with the Warsaw Pact or being contiguous to a Warsaw Pact country (i.e., sharing a land border or being separated by sea with a distance of less than 150 miles). The respective correlations are $r = -0.54$, p-value 0.09 and $r = -0.72$, p-value 0.01. Consistent with this picture of the degree of free-riding being determined by geographical location, we find that the ordering of countries by degree of free-riding is positively correlated with geographical distance of the ally's capital city to Moscow (Spearman's rho 0.61, p-value 0.05).

6. Robustness

All regression analyses suffer from uncertainty about whether the estimation model has been correctly specified. Table 3 therefore reports results from testing the robustness of our

[11] It is also the case that contrary to most other NATO allies, Portugal fought colonial wars during our estimation period. However, the effect of these wars on Portugal's military expenditure growth should be captured by our "intensity of involvement in armed conflict" control variable.

inferences to plausible changes in model specification. To facilitate interpretation of the results from the robustness tests, figure 3 displays the responsiveness parameters and their 95-percent confidence intervals for each of the small NATO member included in our analysis for the five robustness test models.

Not all military spending growth by the US and the Soviet Union is relevant to the safety of the smaller NATO allies. Both superpowers have fought major wars outside the North Atlantic area which affected their military spending growth. The most important of these are the Vietnam war and the invasion of Afghanistan by the Soviet Union, respectively. Model 2 reports the country-specific response rates for the period outside the main Vietnam war activity (1965-73), while model 3 does the same for the period outside the period of Soviet occupation of Afghanistan (1980-88).

In model 4, we take into account that the smaller allies' own geostrategic interests may affect not only their level of military spending, but also the growth in their military spending, and that this effect can have a heterogeneous impact over time. To do so we allow each country to have its own intercept plus its own linear year trend. In other words, we estimate a unit fixed effects model together with unit-specific linear year trends.

In model 5 we replace military spending growth rates of the Soviet Union by military spending growth rates of the Warsaw Pact nations. Lastly, in model 6 we include further control variables from Whitten and Williams (2011), namely various measures of government composition such as the number of government parties, whether the government is a minority government as well as the left-right position of the government and the presence of an election in any one year. For this model, we had to drop Portugal and Greece from the sample as they do not have available data for these political variables until late into the period of our estimations.

The Vietnam war period captured by model 2 allows us to explore the influence of abnormal growth rates of military expenditure on our inferences.[12] One might expect that abnormal growth rates of military spending by the USA during the Vietnam war era push the response functions of NATO members somewhat down. This would then spuriously lend additional support for the free-riding hypothesis even though the NATO members merely do not respond to higher military spending for wars they do not fight – which seems a quite reasonable strategy and does not amount to free-riding. However, we do not find that the response functions of the countries are systematically smaller outside the Vietnam war period than in the main estimations and on the whole there is little effect on the response functions of NATO allies of removing the Vietnam war years from the estimations. This exercise in turn demonstrates that our inferences of free-riding are not invalidated by years of abnormal growth rates of military spending in the USA.

[12] The Afghanistan war does not lead to abnormal growth rates in military spending in the Soviet Union.

Table 3. Robustness tests.

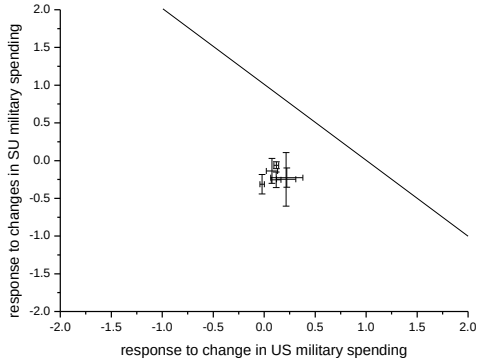
Robustness test:	Model 2 Outside Vietnam war period (1965-73)		Model 3 Outside Afghanistan invasion period (1980-88)		Model 4 Controlling for unit fixed effects and unit-specific year trend		Model 5 Replacing Soviet with Warsaw Pact spending growth		Model 6 Further political control variables	
Country-specific response of:	to US growth	to SU-US growth diff	to US growth	to SU-US growth diff	to US growth	to SU-US growth diff	to US growth	to SU-US growth diff	to US growth	to SU-US growth diff
Canada	0.222	-0.225	-0.0201	-0.313**	0.119***	-0.0618	0.117**	-0.254**	0.0763	-0.136
	(0.158)	(0.127)	(0.0211)	(0.129)	(0.0230)	(0.0456)	(0.0470)	(0.103)	(0.0562)	(0.165)
Great Britain	-0.0942	0.0474	0.0347	-0.343**	-0.0422	-0.161**	-0.0698	-0.250***	0.0468	0.0308
	(0.182)	(0.100)	(0.0378)	(0.118)	(0.0486)	(0.0623)	(0.0608)	(0.0756)	(0.111)	(0.173)
Netherlands	-0.436**	0.254***	-0.0830**	0.129	-0.271***	0.132**	-0.179**	0.153**	-0.0902	0.0290
	(0.164)	(0.0761)	(0.0370)	(0.0878)	(0.0485)	(0.0434)	(0.0581)	(0.0608)	(0.0776)	(0.116)
Belgium	-0.610***	0.0458	-0.179***	0.158**	-0.239***	0.0979***	-0.279***	0.214***	-0.331***	0.0770
	(0.126)	(0.0748)	(0.0300)	(0.0683)	(0.0358)	(0.0150)	(0.0435)	(0.0611)	(0.0882)	(0.0952)
France	-0.553***	-0.0202	-0.0442	-0.201	-0.103***	0.0111	-0.127***	-0.144*	-0.0740	-0.0457
	(0.0921)	(0.239)	(0.0259)	(0.124)	(0.0287)	(0.0805)	(0.0291)	(0.0757)	(0.0871)	(0.108)
Portugal	0.317***	1.113***	0.564***	0.681***	0.502***	1.073***	0.351***	0.827***		
	(0.0859)	(0.0823)	(0.0190)	(0.0620)	(0.0437)	(0.0572)	(0.0399)	(0.0667)		
West Germany	0.256*	1.085***	-0.00772	0.691***	-0.0865**	0.429***	-0.0660	0.753***	-0.00993	0.666***
	(0.142)	(0.0839)	(0.0342)	(0.0820)	(0.0377)	(0.0238)	(0.0510)	(0.0673)	(0.0665)	(0.119)
Italy	0.404***	0.0132	0.0725	0.0270	-0.0682***	-0.217**	0.0610*	0.164	-0.0209	-0.0232
	(0.129)	(0.126)	(0.0478)	(0.123)	(0.0212)	(0.0915)	(0.0291)	(0.0917)	(0.0572)	(0.127)
Greece	-0.502	0.341***	0.105*	0.123	-0.131	0.0625	-0.0646	0.238**		
	(0.292)	(0.0986)	(0.0547)	(0.0899)	(0.0970)	(0.0429)	(0.0892)	(0.0989)		
Norway	0.284***	0.402***	0.00348	0.473***	-0.0159	0.476***	0.0339	0.504***	0.0219	0.476***
	(0.0846)	(0.0775)	(0.0221)	(0.0836)	(0.0125)	(0.0310)	(0.0256)	(0.0653)	(0.0219)	(0.0751)
Denmark	-0.441***	0.422***	0.0815***	0.410***	-0.107***	0.452***	-0.0479	0.408***	-0.0583	0.267**
	(0.136)	(0.105)	(0.0219)	(0.0848)	(0.0184)	(0.0275)	(0.0465)	(0.0679)	(0.0448)	(0.0995)
Turkey	-0.917***	0.210	-0.304***	0.257	-0.337***	0.214**	-0.396***	0.133	-0.481***	0.192
	(0.172)	(0.163)	(0.0716)	(0.170)	(0.0352)	(0.0829)	(0.0422)	(0.147)	(0.0863)	(0.136)

Note: Control variables included, but not reported. Standard errors clustered on countries in parentheses. * statistically significant at 0.05, or ** 0.01 level.

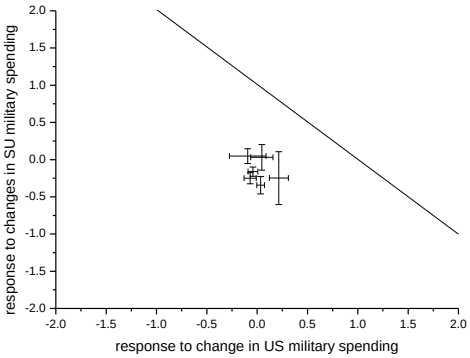
Figure 3 shows the estimated responsiveness parameters with their associated 95 percent confidence intervals for the main estimations and all robustness test models. For many of the countries, the estimated coefficients are not robust – if we define robustness as the absence of a significant difference in point estimates across all tests. This follows from the confidence intervals of some of the estimated responsiveness parameters not overlapping with the remaining ones. Thus defined, we would have to conclude that the influence of growth in military spending by the USA and the Soviet Union on growth of military spending of the other NATO members is not robust. However, that is not the question we are interested in. Instead, we want to know whether smaller countries free-ride and this is the causal inference subjected to the robustness tests. Our baseline model provided evidence that all NATO members free-ride bar Portugal. A robustness test should therefore investigate whether this inference is robust. And indeed, the robustness tests affect our inference on the free-riding behavior of NATO members only in one case: Given our definition of free-riding, we cannot be certain whether West Germany free-rides: one of her estimated responsiveness parameter is fully in the non free-riding space. The inference for all other countries, however, remains robust to changes in the model specification that we conducted in the robustness tests: all of their estimated responsiveness parameters together with their 95 percent confidence intervals lie to the interior of the free-riding threshold.

Figure 3. Response Functions with 95 Percent Confidence Intervals (robustness test models).

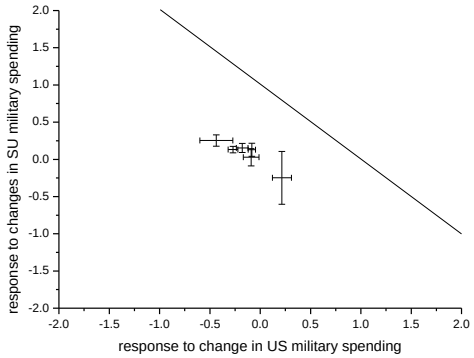
CAN



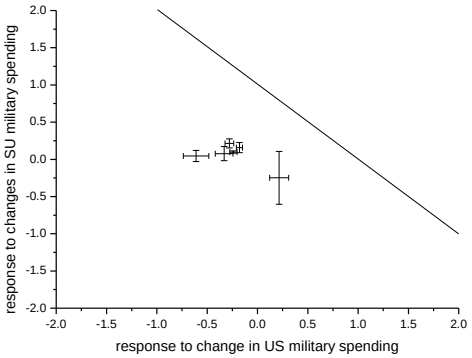
GBR



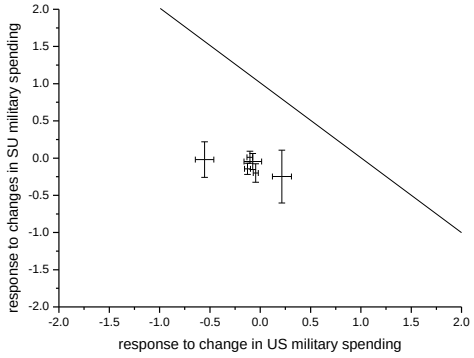
NDL



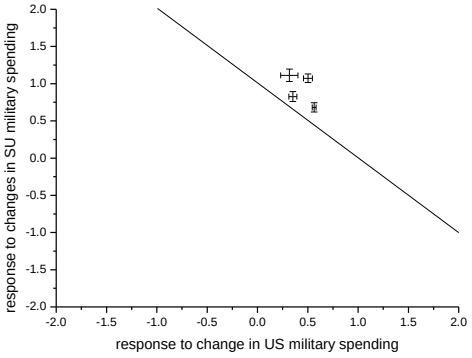
BEL



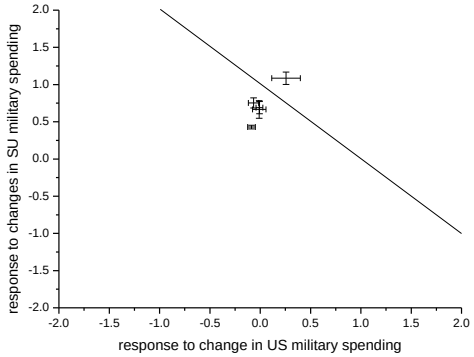
FRA



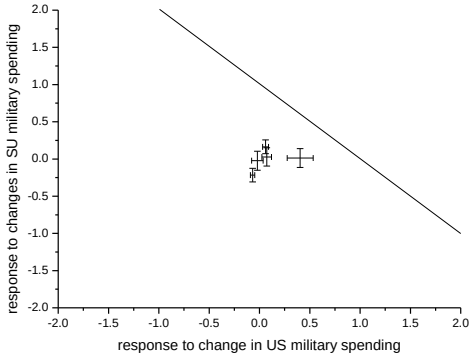
POR



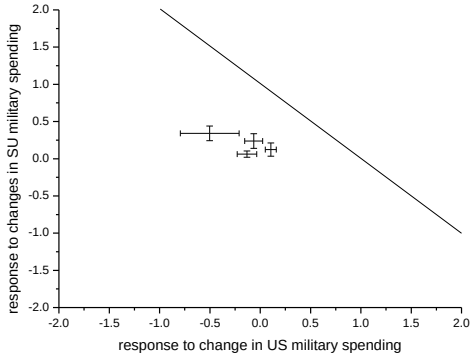
GER



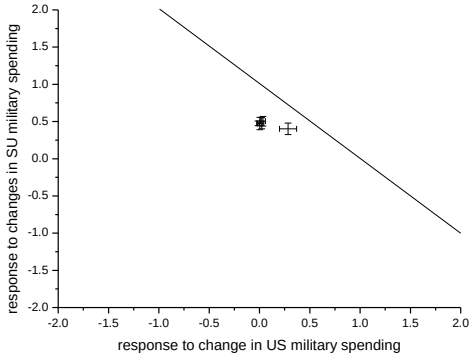
ITA



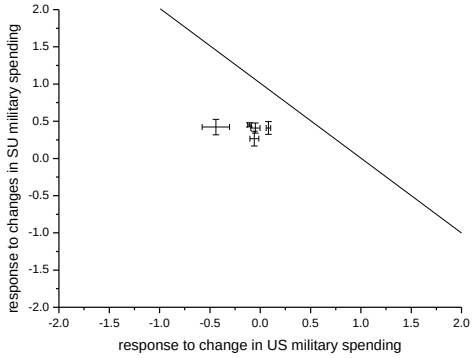
GRE



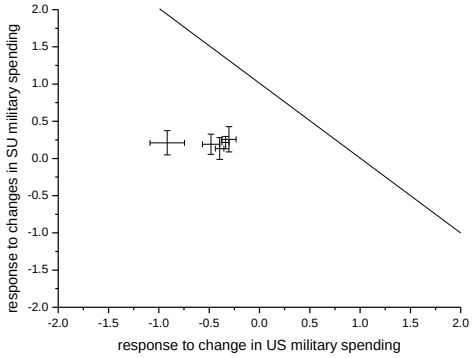
NOR



DNK



TUR



7. Conclusion

In this article, we have used a new method to test the old theory of free-riding in military alliances. The shift in methods is justified by the argument that differences in levels of

military spending, even when expressed relative to GDP, cannot be used to infer free-riding unless one could appropriately control for the difference between the global aspirations of a superpower like the USA and the more spatially limited military objectives of the smaller allies. Clearly, the relatively higher military spending in the USA mirrors her broader interests – and these broader interests make it impossible to attribute relatively lower military spending in Canada and European countries to free-riding.

By contrast, we have developed a quasi-spatial method, which infers free-riding from the responsiveness of the smaller NATO members' growth in military expenditures to growth in US spending on the one hand and growth in Soviet spending when above US spending growth on the other hand. If, in summing up both responsiveness parameters, the smaller allies respond such that their total responsiveness is smaller than one, then they are detected as free-riding on the efforts to maintain the security balance between the alliance and its arch enemy.

Does our analysis support the theory of free-riding? Our results are mixed but still suggest free-riding by the vast majority of smaller allies – a result that is robust to a range of plausible changes to model specification. For only one county (West Germany) our results are inconclusive. Only for Portugal can we clearly reject the hypothesis of free-riding and we have argued above that Portugal is likely to be an outlier. Yet, our analysis does not support the hypothesis that the degree of free-riding is a function of country size. Smaller and larger NATO members do not significantly differ in their degree of free-riding. Instead, we find that the degree of free-riding is correlated with geographical position in relation to the Warsaw Pact. NATO countries closer to Moscow exhibit lower free-riding as do countries that share a land border or are contiguous with one of the Warsaw Pact countries.

Our results thus reject the second part of the free-riding in alliances theory that claims that the extent of free-riding is a function of country size, but lend support to the first and main part of the theory: the smaller allies free-ride on the superpower. However, we wish to stress that there is one condition under which this result and interpretation would be spurious, namely if for both the Soviet Union and the USA the military expenditures related to their global interests grew faster over time than the military expenditures related to the geographically more restricted NATO-Warsaw Pact constellation. This seems possible, but unlikely during our period of study and the best test for this possibility is to exclude the Vietnam and Afghanistan periods from the estimates – and our results are robust to these tests. Hence, whilst we are confident that our results correctly identify free-riding by the smaller NATO allies, we see a promising way forward in combining the traditional 'in levels' interpretation of free-riding with our dynamic perspective, namely in research that attempts to identify the dynamically changing share of US military expenditure devoted to the European and North-Atlantic area, and in case study research that seeks to identify the response of NATO members to 'shocks' in the security perception of a limited number of NATO members.

References

Diehl, P. F. 1994. "Substitutes or Complements – The Effects of Alliances on Military Spending in Major Power Rivalries." *International Interactions* 19(3): 159-176.

Gates, William R. and Katsuaki L. Terasawa. 2003. "Reconsidering Publicness in Alliance Defence Expenditures: NATO Expansion and Burden Sharing." *Defence and Peace Economics* 14 (5): 369-383.

Gleditsch, Nils Petter; Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg & Håvard Strand. 2002. "Armed Conflict 1946–2001: A new dataset." *Journal of Peace Research* 39 (5): 615–37.

International Herald Tribune. 2011. Gates Assails NATO over Willingness to Share Costs. 11-12 June 2011, p. 1.

Murdoch, James C. and Todd Sandler. 1984. "Complementarity, free riding, and the military expenditures of NATO allies." *Journal of Public Economics* 25 (1-2): 83-101.

Olson, Mancur and Richard Zeckhauser. 1966. "An Economic Theory of Alliances." *Review of Economic Statistics* 48:3: 266–79.

Olson, Mancur. 1965. The Logic of Collective Action. Cambridge, MA: Harvard U. Press.

Oneal, John R. and Paul F. Diehl. 1994. "The Theory of Collective Action and NATO Defense Burdens: New Empirical Tests." *Political Research Quarterly* 47 (2): 373-396.

Oneal, John R. 1990. "Testing the Theory of Collective Action: NATO Defense Burdens, 1950–1984." *Journal of Conflict Resolution* 34:3: 426–48.

Palmer, Glenn. 1990a. "Corralling the Free Rider: Deterrence and the Western Alliance." *International Studies Quarterly* 34 (2): 147-164.

Palmer, Glenn. 1990b. "Alliance Politics and Issue Areas: Determinants of Defense Spending." *American Journal of Political Science* 34 (1): 190-211.

Palmer, Glenn. 1991. “Deterrence, Defense Spending, and Elasticity: Alliance Contributions to the Public Good.” *International Interactions* 17 (2): 157-169.

Russett, Bruce M. 1970. *What Price Vigilance?* New Haven: Yale U. Press.

Sandler, Todd. 1977. “Impurity of Defense: An Application to the Economics of Alliances.” *Kyklos* 30 (3): 443-460.

Sandler, Todd. 1993. "The Economic Theory of Alliances – A Survey." *Journal of Conflict Resolution* 37(3): 446-483.

Sandler, Todd and James C. Murdoch. 1990. “Nash-Cournot or Lindahl Behavior?: An Empirical Test for the NATO Allies.” *Quarterly Journal of Economics* 105:4: 875–94.

Sandler, Todd and James C. Murdoch. 2000. “On Sharing NATO Defense Burdens in the 1990s and Beyond.” *Fiscal Studies* 21:3: 297–327.

Sandler, Todd and John F. Forbes. 1980. “Burden Sharing, Strategy, and the Design of NATO.” *Economic Inquiry* 18:3: 425–44.

Sandler, Todd and Keith Hartley. 1999. *The Political Economy of NATO: Past, Present, and into the 21st Century*. Cambridge: Cambridge U. Press.

Sandler, Todd and Keith Hartley. 2001. “Economics of Alliances: The Lessons for Collective Action.” *Journal of Economic Literature* 39: 869-896.

Whitten, Guy D. and Laron K. Williams. 2010. “Buttery Guns and Welfare Hawks: The Politics of Defense Spending in Advanced Industrial Democracies.” *American Journal of Political Science* 55 (1): 117-134.