Do Opposites Attract? Military Complementarity and Alliance Formation*

Jessica Stoll†
Journeys in World Politics
October 9-11, 2008

Abstract

I hypothesize that a state's decision to form a military alliance is dependent on the benefit a potential ally will add, which will depend on the military strengths of the state and its potential allies. An alliance will be more valuable when the two states are relatively strong in different military branches, so that their military strengths are complementary. I devise three different measures of state military specialization and use these to create three measures of dyadic military complementarity. I argue that military specifics must be taken into account when making foreign policy decisions – it is not enough to refer to general ideas of power. I test the theory on 52 states from 1950-1988 using a bivariate probit model.

*Special thanks to Kevin Clarke, Hein Goemans, Jeremy Kedziora, Paulina Marek, Lukas Pfaff and Jun Xiang for comments and suggestions, and to Thomas Cusack for providing his coding of *The Military Balance* data.

†Department of Political Science, University of Rochester, Harkness Hall 304, Rochester, NY 14627. (js019m@mail.rochester.edu)

1 Introduction

Alliances offer a level of analysis between the state and the international system. They allows us to examine the dynamics of cooperation and competition in a formalized setting. Moreover, as Leeds et al. (2000) have shown, formal commitments made to allies are useful as indicators of state intentions and predictors of state actions – the high rate of commitment fulfillment demonstrates that alliance treaties are more than just paper-thin promises. Military alliances in particular allow us to ask about the balance of interests that go into making decisions about alliance formation: what is the influence of the military on military alliances? To what extent are alliance decisions dictated by military advantage, and to what extent by political concerns? If military commitments to allies are meaningful, then the impact of allies on military effectiveness must be seriously considered by foreign policy decision makers.

This paper addresses this impact of military considerations on alliance decisions. I argue that: 1) alliances with characteristics that increase military effectiveness are more likely to be formed than alliances with characteristics that decrease military effectiveness, and 2) military complementarity between two states increases the military effectiveness of an alliance between the two.

The paper proceeds as follows. In the next section I formulate my theory of military complementarity and alliance formation, building on these two intuitions, then review the relevant literature. I then explain my operationalization of military complementarity and methods of analysis and report the results of that analysis.

2 Theory

In order to develop a theory of why military alliances are formed, we start with the question of what purposes they serve. Though alliances may fulfill purposes unrelated to military action, such as constraining allies, bolstering domestic support or tying the hands of future leaders, explicit military obligations imply that the use of the military is a real possibility. An alliance with military obligations may be intended for actual military use, or it may be intended as a signal of willingness to use the military if provoked. To be a good tool for military use, the alliance must be militarily effective. To act as a signal, the alliance must assure both that if called upon, the state will fight (credible commitment) and that if the state fights, it will fight well (military effectiveness).

I assume that these two goals, fighting and signaling, are the motivations for military alliances at least some of the time. Now the question is, if these goals are present, how are they best pursued? The likelihood of formation of military alliances will still depend on what other possible tools could be used to accomplish the same goals. As an alternative to forming an alliance, a state can increase its own military strength. While this strategy is ultimately limited by the state's resources, it may provide domestic benefits that an alliance does not, such as a boost to the economy or as an outlet for the otherwise unemployed. Furthermore, an alliance may have domestic costs, such as a perceived loss of autonomy or reputation. Then because there are meaningful costs to alliances, in order for an alliance to be formed it must contribute efficiently to the foreign policy goals. To determine when a state is likely to choose alliance formation and when it is likely to increase its own military strength, we need to consider how well each option meets the requirements of credibility and military effectiveness. I argue that military complementarity increases the effectiveness and credibility of alliances and thus increases the likelihood of allying rather than arming.

Military complementarity occurs, unsurprisingly, when two states have complementary military strengths, such as if one is a land power and the other a sea power. Laying aside for a moment the issues of force integration and commitment credibility, an alliance with a complementary military power has obvious advantages. Not all states are specialized in one military area, and this paper does not speculate on what leads to military specialization. But whether production advantages lead to specialization or specialization, occurring for some other reason, leads to production advantages, it is safe to assume that if a state is specialized, it is cheaper to produce a fighting unit in its specialized area than in any other area. If the state is a sea power, it should be cheaper for the state to produce one new sailor than an equivalently trained and equipped soldier. For that sea power, then, allying with a land power would be more valuable than allying with another sea power.

Problems with force integration, as Bensahel (2007) persuasively argues, is one reason why alliances may not aggregate capabilities well. But a complementary alliance may ameliorate some of these problems. For instance, Bensahel cites the problem of differences in training, skill, and technology which can reduce the alliance to the lowest common denominator. But if two states are differently specialized, each can take leadership positions in its own area of expertise and minimize the need for intrabranch integration. The sea power will take the lead in naval operations, while the land power will fill the same role for army ones. The choice of whose long-standing operating procedures will predominate can be made similarly. The level of force integration achieved by even close alliances is unlikely to ever come close to the cohesiveness of a single national military, but if the problem can at least be lessened by complementarity, the resource advantages of an alliance may outweigh this cost.

Commitment credibility is another area in which complementarity can increase the utility of an alliance. Because complementarity provides military power in an area in which the state is weak, refusing to fulfill alliance obligations forfeits something of high value. The state may potentially replace the alliance forces with forces of its own, or go without, but the cost of doing so is higher for complementary alliances than noncomplementary ones. A higher cost of defection will make defection less likely, and anticipation of credibility will make the formation of complementary alliances relatively more likely than noncomplementary ones.

Therefore, I predict that we will be more likely to see states choosing to ally when the available alliance will exhibit military complementarity than if the available alliance will not. Though I have so far framed the choice as one between allying and arming, it is not a mutually exclusive choice – states can choose one, both or neither of these options. These possibilities are taken into account by the bivariate probit model.

Another thing to consider is what kind of threat the state is facing. The theory rests on the assumption that a military force spread across the three branches more or less equally is more valuable in a conflict than a military force specialized in only one or two areas. But this should depend on the type of enemy the state or alliance is facing. If the state or alliance is threatened by a specialized state or alliance, then specialization, not complementarity, will be the best choice of alliance; if your enemy is an army power and has a relatively weak navy, your primary concern should be matching their army strength, not adding to your navy in addition to the army. I predict that complementarity under these circumstances will make alliance formation less likely.

Military complementarity and specialization are topics that have long been intuitively understood and assumed to be meaningful in history, but not ones that have been pursued empirically. We think of Britain as a sea power, but for the years in this study there have always been at least twice, and sometimes as many as three times, as many British soldiers as sailors. Similarly, we know that states do consider the optimal allocation of their forces when making strategic decisions – consider Russia's perpetual quest for a warm water port in order to allow it to be a sea as well as land power. We consider the strategic value of the combination of forces in alliance decisions as well, such as the idea that Austria was Britain's "natural" ally on the continent.[1] By seriously considering the role of military specialization and complementarity in foreign policy decisions, this paper contributes to the literature that tries to "open up the black box of war" by acknowledging and studying how specific military decisions affect broader outcomes.

3 Previous Work

Using alliances to increase power falls under the classical Realist category of external balancing. Morgenthau (1985), for example, argues that alliances are chosen when internal balancing (arms races) is believed to be insufficient. Walt (1997) further argues that one reason alliances may endure long after the threat that motivated their formation has disappeared is if their type of institutionalization makes them adaptable to new or changing threats. This is exactly what Wallander (2000) credits NATO's persistence with – its structure and level of military integration, including division of labor, allow it to adapt to new threats. Bray and Moodie (1977) even advocate more military division of labor within NATO, though they are pessimistic about the possibility, citing, in part, concerns about the perceived effect on autonomy:

> Since no individual West European country is in a position to embark on a program to meet the full range of European requirements, this implies specialization and a much more explicit division of labor among the European members of NATO...While specialization thus offers a rational and attractive road, the practical obstacles are formidable.

[1] I am grateful to Jeremy Kedziora for this example.

> The biggest obstacle stems from abiding sovereignty preoccupations in Western Europe...Even if the principle of specialization were to be accepted by all European members of NATO, national pride would likely distort a rational implementation. (Bray and Moodie, 19–20)

Capability aggregation in alliances is neither perfect nor costless, something which Bensahel (2007) forcefully argues by proposing several ways in which allying may actually reduce military effectiveness. Alliances can hamper political and strategic integration and responsiveness. Differences in training, skill and technology can reduce joint operations to the lowest common denominator. Problems with gathering, analyzing, and sharing intelligence are significant, even within close and long-standing alliances, such as between the US and UK. Bensahel concludes that despite reduced military effectiveness, alliances may be acceptable because of the political benefits they provide, legitimacy in particular: "The decision to operate as part of an alliance is ultimately a political judgment, not a military judgment" (202).

Any alliance is also subject to the perpetual problem of free-riding, though Conybeare (1994) argues that the problem of free-riding is outweighed by the efficiency gains from alliance membership, while Koerner (unpublished) theorizes that, while free-riding may be present, alliance failure due to free-riding is unlikely because alliances insufficient to deter will not be formed at all. The possibility of unequal burden sharing may be either an incentive or disincentive to alliance formation versus arming, depending on whether the state anticipates benefiting or suffering from the inequality.

There are, of course, many other possible functions for alliances other than capability aggregation. Walt (1997) says that the otherwise irrational endurance of alliances may be explained by their impact on state reputation, their benefits for important political elites or groups, or ideological ties. Vincent, Strauss and Biondi (2001) use capability theory to argue that states join alliances to augment their power, but alliance rules may allow weak states to constrain the actions of strong ones (and thus high capability states will favor the relaxation of NATO's consensus requirements while low capability states will oppose it). Schroeder (1976) also believes alliances can let states exert control over other alliance members (and that this may be the reason for their formation, rather than being targeted at an outside threat), but that the control is likely to go in the other direction. For him, the impetus is political, not military: alliances have often been formed in order for strong states to exert control over weaker ones, though this is not always successful.

4 Operationalization and Methods

The hypotheses are tested on 52 states[2] from 1950-1988, for a total of 98,680 directed dyad-year observations. I use directed dyad-years as the unit of analysis in order to capture the idea of states choosing between potential alliances (each pairing). I chose directed rather than non-directed dyads in order to keep the bivariate model simple (limit the number of outcomes to consider). The dependent variable (outcome) relies on two determinations: form alliance or not and increase arms or not. A dyad-year is coded as having formed an alliance if an offensive or defensive alliance between the two states, as indicated by the Alliance Treaty Obligations and Provisions Project, began that year. A dyad-year is coded as having increased arms if the percent increase in the total number of members of the military (army, navy and air force, excluding reserves, paramilitary, etc.) increased by one standard deviation above the mean for all cases. Militaries grew an average of 4.7% across the sample, but the standard deviation was quite large – 37.9%. Therefore I coded the state as having increased arms if the size of its military personnel grew by at least 42.6%. Table 1 shows the frequency of outcomes across the 98,680 cases.

Table 1: Frequency of Outcomes

Form Alliance?	Increase Arms?	Frequency
No	No	88307
Yes	No	194
No	Yes	10139
Yes	Yes	40
	Total:	98680

[2]Argentina, Australia, Belgium, Brazil, Bulgaria, Canada, Chile, China, Cuba, Czechoslovakia, Denmark, Egypt, France, German Federal Republic, German Democratic Republic, Greece, Hungary, India, Indonesia, Iran, Iraq, Israel, Italy, Japan, Libya, Malaysia, Mexico, Netherlands, Nigeria, North Korea, Norway, Pakistan, Philippines, Poland, Romania, Russia, Saudi Arabia, South Africa, South Korea, Spain, Sweden, Switzerland, Syria, Taiwan, Thailand, Turkey, UK, USA, Venezuela, Yemen Arab Republic (North Yemen), Yemen People's Republic (South Yemen), and Yugoslavia. The states and years chosen were based on the availability of military data from *The Military Balance*. Many thanks to Thomas R. Cusack for his coding of the bulk of this data.

The measures of complementarity are the independent variables of most interest, but I do use several other control variables. For obvious reasons, I control for whether there was an ongoing dispute between the states in the dyad, using Maoz's indicator, with the assumption that an ongoing dispute between the two made an alliance between them less likely. I also include a dummy for an ongoing alliance between the two with the opposite assumption, and the difference in Composite Index of National Capability (CINC) scores,[3] since the military benefits of allying should decrease as the difference in power increases (the potential ally can add relatively less).

I also control for each state's threat environment. Threat environment is computed for each state in the dyad using a modified version of Bolk's (1999) method, as follows:

$$\text{Threat Environment} = BH_{t-5} + BH_{t-4} + BH_{t-3} + BH_{t-2} + BH_{t-1} + MP_{t-5} + MP_{t-4} + MP_{t-3} + MP_{t-2} + MP_{t-1} + RO_{t-5} + RO_{t-4} + RO_{t-3} + RO_{t-2} + RO_{t-1}$$

where BH_{t-x} is the baseline hostility level (the number of MIDs the state was involved in x years before), MP_{t-x} is the number of major power opponents the state faced in those MIDs, and RO_{t-x} is the number of regional opponents the state faced in those MIDs.

For the initial analysis, these are the only controls I use: ongoing dispute, ongoing alliance, difference in CINC scores, and the threat environments for both states (along with the independent variable of interest, complementarity). After those initial regressions, I run additional ones with other suggested control variables, including the change in the threat environment of each state, the weighted global S-Score, the Polity scores for each state, the GDP[4] and GDP per capita for each state, distance between the state capitals, the number of borders of each state, and dummies for whether both states are NATO members or both states are Warsaw Pact members.

As the British example shows, raw numbers do not tell the whole story when it comes to

[3]CINC scores are a commonly used measure of state power, computed by summing data on energy consumption, iron and steel production, military expenditure, military personnel, total population and urban population.

[4]As an indicator of state power, I also considered CINC scores, major power status (as coded by the Correlates of War), possession of nuclear weapons, and military expenditures. Because these four were highly correlated with each other and with GDP, I chose to use just one. I chose GDP because military expenditures can be misleading (as just one example how, pay rates for service members will depend on whether military service is mandatory) and they are included in CINC scores, and CINC scores are used to compute the measure of specialization. Nuclear weapons possession is almost identical to the indicator of major power status, and both apply to a very limited number of cases in the data. My measure of complementarity also excludes several major power observations (and thus nuclear weapons observations) for reasons discussed below.

specialization. Because labor intensiveness varies across branches of the military, the optimal size of a branch varies, too, so comparisons across branches within a state are not the best way to gauge specialization. Instead, this measure of specialization compares across states within the same branch.[5] A measure of specialization should also be relative to the size of the state, otherwise we will simply get an indicator of major power status. Suppose a state is *not* an army power. Then we would expect that the state's share of the world army matches the state's share of world power, so:

$$\frac{\text{size of state 1 army}}{\text{sum of all world armies}} = \frac{\text{state 1 CINC score}}{\text{total world CINC score}}$$

If a state is an army power, then we expect its army to be larger than its power would indicate, so

$$\frac{\text{size of state 1 army}}{\text{sum of all world armies}} > \frac{\text{state 1 CINC score}}{\text{total world CINC score}}$$

I use this calculation as an indication of specialization.

To get a dummy variable for specialization, I compute specialization measures for each of the three branches, so

$$\text{army specialization} = \frac{\text{size of state 1 army}}{\text{sum of all world armies}} - \frac{\text{state 1 CINC score}}{\text{total world CINC score}}$$

and similarly for the navy and air force (all personnel numbers are in thousands). Then I (admittedly arbitrarily) count as specialized any state-year in which the measure is at least one standard deviation above the mean for all state-years. For the army, the mean is -0.000007 and the standard deviation 0.042, with a sum of 0.042. For the navy, the mean is -0.000000002 and the standard deviation 0.041, with a sum of 0.041. For the air force, the mean is 0.0000000534 and the standard deviation 0.043, with a sum of 0.043. For example, Iraq in 1986 had an army of 800,000 and a CINC score of 0.009667, while the total world army forces were 15,092,500 and the total world CINC score (for the states in the sample) was .9061626. Then Iraq's army specialization is $\frac{800}{15092.5} - \frac{0.009667}{0.9061626} = 0.042$ which is high enough for Iraq to be coded as an army power. I exclude

[5] Appropriately, this is how Britain traditionally gauged its own level of naval power – by comparing the size of its navy to its rivals'.

any dyad-year in which one or more state was double or triple specialized,[6] which eliminates 2,112 cases.[7] A dyad is considered complementary if both states are specialized, and in separate areas.

Because specialization has previously been a non-specific concept which could mean any of several things, I create two alternative measures to capture other possible definitions. The first alternative (labeled Complementary (2) in the regression tables) takes the ratio of the number of members in the largest branch of the military to the number of members in the second largest branch and codes the state as specialized if this ratio is at least one standard deviation (6.244) above the mean (6.459), or at least 12.704. For example, in 1957 Japan had an army of 150,000, a navy of 10,700 and an air force of 10,300. Then the ratio of its largest (army) to second largest (navy) branch is $\frac{150000}{10700} = 14.02$, large enough for Japan to be coded as specialized in the army. Because this computes one ratio and then asks which branch is the largest, there is no problem of double or triple specialization, so no cases are excluded.

The second alternative (labeled Complementary (3) in the regression tables) is designed to capture priorities in the state's capital investment. I compute three ratios of equipment to personnel numbers: for the navy, the number of principal surface combatants[8] to the number of naval personnel; for the army, the number of tanks to army personnel; and for the air force, the number of combat-ready aircraft to air force personnel. The mean ratio for the navy is 0.754 and the standard deviation 0.632, with a sum of 1.387. The mean ratio for the army is 9.603 and the standard deviation 9.723, with a sum of 19.326. The mean ratio for the air force is 12.655 and the standard deviation 11.547, with a sum of 24.203. I count as specialized any state-year in which the ratio is at least one standard deviation above the mean. For example, in 1970 Israel had an air force of 9,000 and 330 combat ready aircraft, which leads to a ratio of $\frac{330}{9} = 36.67$, easily large enough to be coded as an air power. I exclude any dyad-year in which one or more state was double or

[6]Triple specialized: Russia 55-60, 62; USA 51-54, 61. Navy and army specialized: Russia 55-62; USA 51-54, 61. Navy and air force specialized: Russia 55-60, 62, 65-69, 71-75; UK 50; USA 50-72. Army and air force specialized: France 50; Russia 55-60, 62; USA 51-54, 61. Interestingly, in contrast to the excluded cases for the third complementarity measure (below), all of these occur in major power states.

[7]There were 98,017 observations in which none of the six specialization measures were missing.

[8]This includes ships such as aircraft carriers, battleships, cruisers, destroyers, frigates and corvettes and excludes fast attack craft, patrol craft, sub chasers, and any kind of non-combat (support) vessels such as minelayers, minesweepers, troop or equipment carriers.

triple specialized,[9] which excludes 5,016 cases.[10] As before, the dyad is considered complementary if both states are specialized, and in separate areas. Tables 2 and 3 show summary statistics for continuous variables and counts for dummy variables.

Table 2: Counts, Dummy Variables

Variable	Obs.	1	0
Complementary	94568	144	96424
Complementary (2)	98680	1164	97516
Complementary (3)	93664	1522	92142
NATO Member	98680	23243	75437
Warsaw Pact Member	98680	11646	87034
Both NATO	98680	5130	93550
Both Warsaw Pact	98680	2796	95884
Major Power Status	98680	9705	88975
Nuclear Weapons	98680	9242	89438
Ongoing Dispute	98680	634	98046
Ongoing Alliance	98680	4212	94468
Civil War	96389	13520	82869

Table 3: Summary Statistics, Continuous and Categorical Variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
CINC	98680	.0184	.0384	.00012	.3195
Difference in CINC	98680	0	.0548	-.3194	.3194
Threat Environment	98680	21.2069	30.4432	0	271
Change in Threat Environment	96315	.4969	7.0410	-72	102
S-Score	98680	.4922	.3383	-.5741	1
Distance	98680	4281.408	2797.852	108	12071
Polity	95568	.5141	7.9378	-10	10
GDP	93435	266119.1	54792.4	915	5512845
GDP per Capita	93435	5.9194	4.4594	.439	22.499
GDP Growth	92722	.0283	.0482	-.2667	.3129
Military Expenditure	94842	6974075	2.75e+07	3920	1.18e+08
Number of Borders	96739	3.7031	2.6669	0	12

To address the issue of what type of threat the state or alliance is facing, I compute alliance measures of specialization for NATO and the Warsaw Pact. I do this by summing the force and

[9]Triple specialized: Libya 82, 85, 88. Navy and army specialized: East Germany 86-88; Libya 82, 85, 88. Navy and air force specialized: South Africa 68-70, 82; Sweden 68-69; Libya 82, 85, 88. Army and air force specialized: Czechoslovakia 70; Israel 65, 68-69, 72, 75-79, 85; Libya 76-88; Sweden 78, 81-82, 85-88; Switzerland 70-76; Yemen Arab Republic 80-81. Unlike the original measure of specialization, none of these are major powers.

[10]There were 39,558 observations in which none of the six specialization measures (three for each state in the dyad) were missing.

equipment totals and CINC scores for all the states in the two alliances and then computing the three indicators of specialization as I did for individual states. Only the original measure of specialization yields interesting results; the second measure codes neither alliance as specialized at any time, and the third only codes the Warsaw Pact as an army power from 1971-1988. The original measure is more generous, coding NATO as a naval power for the entire period, an army power from 1951-1954, and an air power from 1950-1973, 1982-1984, and 1986-1988. It also codes the Warsaw Pact as an army power in 1950 and 1955-1962, a naval power in 1955-1963, and an air power from 1950-1960, 1962, and 1965-1980. My intuition here is that if the Warsaw Pact is double or triple specialized or not specialized at all, a member of NATO will prefer complementary alliances, while if the Warsaw Pact is specialized in only one area, a member of NATO will prefer non-complementary alliances.

To test this, I split the time period into four shorter periods: 1950-1959, 1960-1969, 1970-1979 and 1980-1988 and add dummies for the state's membership in either of the two major alliances and interaction terms between NATO or Warsaw Pact membership and complementarity. To be clear, my predictions are determined by the number of areas in which the alliance is specialized per decade.[11] I expect that the interaction between NATO membership and complementarity will make an alliance more likely in 1950-1959 and 1980-1988 and less likely in 1960-1969 and 1970-1979, while the interaction between Warsaw Pact membership and complementarity will make an alliance more likely in 1950-1959, 1960-1969 and 1980-1988 but less likely in 1970-1979. For the second measure of complementarity, the predictions are no different from the original predictions, as neither alliance is considered specialized. For the third measure of complementarity, there is only a slight difference in predictions: from 1970-1979 and 1980-1988, if a state is a member of NATO and the potential alliance partner is complementary, an alliance should be less likely.

I use a bivariate probit model with robust standard errors. In this two equation model,

$$y_1* = x_1\beta_1 + \epsilon_1 \text{ with } y_1 = 1 \text{ if } y_1* \geq 0 \text{ and } y_1 = 0 \text{ if } y_1* < 0$$

$$y_2* = x_2\beta_2 + \epsilon_2 \text{ with } y_2 = 1 \text{ if } y_2* \geq 0 \text{ and } y_2 = 0 \text{ if } y_2* < 0$$

[11]As noted above, the two alternative measures of specialization are coded positively so sparingly that there is not much of a problem determining this number. For the original measure, however, the predictions are based on which amount of specialization was present during a majority of the decade. These are the numbers reported in Table 4. Table 5 shows the frequency of the interaction term across time periods.

Table 4: Number of Areas Specialized in, by Decade

	1950-1959	1960-1969	1970-1979	1980-1988
NATO (1)	2	3	1	2
Warsaw (1)	3	1	1	0
NATO (2)	0	0	0	0
Warsaw (2)	0	0	0	0
NATO (3)	0	0	0	0
Warsaw (3)	0	0	1	1

Table 5: Frequency of Interaction by Decade

	1950-1959	1960-1969	1970-1979	1980-1988
NATO*Complementary (1)	30	4	14	10
Warsaw*Complementary (1)	3	4	10	2
NATO*Complementary (2)	264	24	0	0
Warsaw*Complementary (2)	86	14	0	0
NATO*Complementary (3)	0	57	139	112
Warsaw*Complementary (3)	0	24	166	126

In this case, the outcome is arm without allying if $y_1* \geq 0$ and $y_2* < 0$, ally without arming if $y_1* < 0$ and $y_2* \geq 0$, arm and ally if $y_1* \geq 0$ and $y_2* \geq 0$, and do neither if $y_1* < 0$ and $y_2* < 0$.

5 Analysis

Table 6 shows the results of the first bivariate probit regression. Complementarity has the expected sign, so that when a dyad exhibits military complementarity, the pair is more likely to begin an alliance. However, the marginal effects in Table 7 show that the impact of complementarity on its own is minimal to non-existent, especially on the ally/not ally decision.

Next I expanded the set of control variables in the regression to include: major power status, the CINC score of each state, the weighted global S-Score, the distance between the capitals, the Polity score of each state, the GDP and GDP per capita of each state, the number of borders of each state, whether there was an ongoing civil war in either state, and indicators of whether both states were members of NATO or of the Warsaw Pact. Table 8 shows the results of this regression.[12] Complementarity still has the predicted sign (complementarity increases the likelihood of alliance

[12] Distance between capitals has a coefficient and standard error of 0 for both equations and is omitted from Table 8 for space reasons.

Table 6: Bivariate Probit

Variable	**Coefficient**	**(SE)**
Equation 1: Ally		
Complementarity	0.062	(0.319)
Difference in CINC	-2.111**	(0.543)
Ongoing Dispute	-5.078**	(0.200)
Ongoing Alliance	0.960**	(0.056)
Threat Environment (1)	0.002**	(0.001)
Threat Environment (2)	0.000	(0.001)
Intercept	-3.074**	(0.031)
Equation 2: Arm		
Complementarity	-0.054	(0.203)
Difference in CINC	-2.543**	(0.156)
Ongoing Dispute	0.012	(0.095)
Ongoing Alliance	-0.020	(0.032)
Threat Environment (1)	-0.007**	(0.000)
Threat Environment (2)	-0.003**	(0.000)
Intercept	-1.262**	(0.009)
Equation 3: ρ		
Intercept	0.193**	(0.037)

Table 7: Marginal Effects

	Neither		Ally Only		Arm Only		Both	
Measure	$\frac{\partial y}{\partial x}$	(SE)	$\frac{\partial y}{\partial x}$	(SE)	$\frac{\partial y}{\partial x}$	(SE)	$\frac{\partial y}{\partial x}$	(SE)
Complementary	0.007	(0.026)	0.000	(0.001)	-0.007	(0.026)	0.00	(0.000)
Difference in CINC	0.358	(0.021)	-0.006	(0.002)	-0.349	(0.021)	-0.003	(0.001)
Ongoing Dispute	0.000	(0.013)	-0.001	(0.000)	0.002	(0.013)	0.000	(0.000)
Ongoing Alliance	-0.012	(0.005)	0.014	(0.002)	-0.005	(0.004)	0.003	(0.000)
Threat Environment (1)	0.001	(0.001)	0.000	(0.000)	-0.001	(0.000)	0.000	(0.000)
Threat Environment (2)	0.004	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)

formation), but neither the significance nor the marginal effects (Table 9) have increased.

Next I compare my main measure of complementarity to the two alternatives.[13] Both these measures affect the propensity to ally in the opposite of the predicted way – here, complementarity makes an alliance less likely, except for the third measure when extra control variables are included. The marginal effects, reported in Tables 11 and 13, show that these alternative measures of complementarity also have little to no impact on the alliance decision.

The difference in results between the three measures indicates that they are getting at different state characteristics. The third measure addresses priorities in capital investment, but it may be misleading. The measure does not distinguish between old and new equipment, nor does it take into account domestically manufactured versus purchased equipment, which may be subsidized by an actual or potential alliance partner. The second measure, though it does not exactly compare raw numbers, does rely on them in a way that does not take into account the differences between the personnel requirements of different services.

To compare the fit of the three measures, I use the Akaike Information and Bayesian Information Criteria. Since an accurate comparison relies on using the same set of observations, I re-run the regressions on the observations used by all three, so that any observation which is coded as double or triple specialized by any of the measures was excluded. This leaves 91,583 observations. Table 14 shows a comparison of the frequency of complementarity in the new and old data.

Table 15 shows the log likelihood, AIC and BIC for each of the models using the limited data. The measures are so close that they are nearly indistinguishable, though the third measure of complementarity, which takes into account the ratio of the largest to second largest military branch, appears to fit the data slightly better. This closeness is striking given that there is virtually no overlap in the measures of complementarity. There is only one dyad-year which both the second and third measures code as complementary, and there is no overlap at all between the first and second or first and third measures.

Next I check the second set of predictions, which take into account the specialization of the main enemy. We know from Table 5 that in many periods the interaction term did not occur at all.

[13]GDP of each state is included in the regression for Table 10, but has coefficients and standard errors of 0 in every instance, and so is omitted from the table for space reasons.

Table 8: Bivariate Probit, Additional Controls

Variable	**Coefficient**	**(SE)**
Equation 1: Ally		
Complementary	0.224	(0.311)
Ongoing Dispute	-4.964**	(0.095)
Ongoing Alliance	0.693**	(0.070)
Threat Environment (1)	0.003**	(0.001)
Threat Environment (2)	0.002**	(0.001)
Δ Threat Environment (1)	-0.002	(0.003)
Δ Threat Environment (2)	0.000	(0.002)
S-Score	0.384**	(0.112)
Polity (1)	0.001	(0.002)
Polity (2)	0.002	(0.003)
GDP (1)	0.000	(0.000)
GDP (2)	0.000**	(0.000)
GDP per Capita (1)	-0.045**	(0.009)
GDP per Capita (2)	-0.033**	(0.008)
Number of Borders (1)	-0.042**	(0.010)
Number of Borders (2)	-0.028**	(0.009)
Civil War (1)	-0.272*	(0.097)
Civil War (2)	-0.294**	(0.100)
NATO Dyad	-0.031**	(0.127)
Warsaw Pact Dyad	-3.947**	(0.100)
Intercept	-2.319**	(0.116)
Equation 2: Arm		
Complementary	-4.607**	(0.064)
Ongoing Dispute	-0.559**	(0.191)
Ongoing Alliance	-0.304**	(0.067)
Threat Environment (1)	0.007**	(0.001)
Threat Environment (2)	0.001**	(0.000)
Δ Threat Environment (1)	-0.008**	(0.002)
Δ Threat Environment (2)	-0.001**	(0.000)
S-Score	0.240**	(0.039)
Polity (1)	0.050**	(0.002)
Polity (2)	0.005**	(0.001)
GDP (1)	0.000**	(0.000)
GDP (2)	0.000	(0.000)
GDP per Capita (1)	-0.027**	(0.003)
GDP per Capita (2)	-0.039**	(0.004)
Number of Borders (1)	-0.057**	(0.006)
Number of Borders (2)	-0.028**	(0.005)
Civil War (1)	-0.014	(0.024)
Civil War (2)	-0.283**	(0.039)
NATO Dyad	-1.171**	(0.125)
Warsaw Pact Dyad	0.698**	(0.058)
Intercept	-0.594**	(0.049)
Equation 3: ρ		
Intercept	0.159**	(0.058)

Table 9: Marginal Effects of Complementarity, Additional Controls

	Neither		Ally Only		Arm Only		Both	
Measure	$\frac{\partial y}{\partial x}$	(SE)	$\frac{\partial y}{\partial x}$	(SE)	$\frac{\partial y}{\partial x}$	(SE)	$\frac{\partial y}{\partial x}$	(SE)
Complementary	0.000	(0.001)	0.001	(0.001)	-0.001	(.000)	0.000	(.000)

Table 10: Bivariate Probit

Variable	**Coefficient**	**(SE)**	**Coefficient**	**(SE)**
Equation 1: Ally				
Complementary (2)	-4.667**	(0.085)		
Complementary (3)			-0.189	(0.211)
Difference in CINC	-0.056	(0.558)	-0.057	(0.577)
Ongoing Dispute	-4.920**	(0.070)	-4.928**	(0.068)
Ongoing Alliance	0.919**	(0.053)	0.877**	(0.056)
Threat Environment (1)	0.002**	(0.001)	0.002**	(0.001)
Threat Environment (2)	0.002**	(0.001)	0.002**	(0.001)
Intercept	-3.046**	(0.029)	-3.033**	(0.029)
Equation 2: Arm				
Complementary (2)	0.388**	(0.043)		
Complementary (3)			-0.602**	(0.075)
Difference in CINC	-2.060**	(0.140)	-1.894**	(0.164)
Ongoing Dispute	0.252**	(0.068)	0.028	(0.092)
Ongoing Alliance	-0.073*	(0.029)	-0.065*	(0.033)
Threat Environment (1)	-0.003**	(0.000)	-0.007**	(0.000)
Threat Environment (2)	-0.002**	(0.000)	-0.003**	(0.000)
Intercept	-1.173**	(0.008)	-1.237**	(0.009)
Equation 3: ρ				
Intercept	0.145**	(0.035)	0.169**	(0.037)

Table 11: Marginal Effects of Complementarity 2 & 3

	Neither		Ally Only		Arm Only		Both	
Measure	$\frac{\partial y}{\partial x}$	(SE)	$\frac{\partial y}{\partial x}$	(SE)	$\frac{\partial y}{\partial x}$	(SE)	$\frac{\partial y}{\partial x}$	(SE)
Complementary (2)	-0.084	(0.011)	-0.001	(0.000)	0.086	(0.011)	0.000	(0.000)
Complementary (3)	0.055	(0.004)	-0.001	(0.001)	-0.055	(0.004)	0.000	(0.000)

Table 12: Bivariate Probit, Additional Controls

Variable	**Coefficient**	**(SE)**	**Coefficient**	**(SE)**
		Equation 1: Ally		
Complementary (2)	-5.242**	(0.138)		
Complementary (3)			0.037	(0.224)
Ongoing Dispute	-4.936**	(0.068)	-4.941**	(0.061)
Ongoing Alliance	0.789**	(0.071)	0.636**	(0.068)
Threat Environment (1)	0.002*	(0.001)	0.002**	(0.001)
Threat Environment (2)	0.002*	(0.001)	0.002**	(0.001)
Δ Threat Environment (1)	$0.004^{\dagger}$	(0.002)	0.001	(0.002)
Δ Threat Environment (2)	$0.004^{\dagger}$	(0.002)	0.001	(0.002)
S-Score	0.369**	(0.108)	0.411**	(0.102)
Distance	0.000**	(0.000)	0.000**	(0.000)
Polity (1)	$0.006^{\dagger}$	(0.004)	0.002	(0.002)
Polity (2)	0.008*	(0.004)	0.003	(0.003)
GDP per Capita (1)	-0.045**	(0.007)	-0.044**	(0.008)
GDP per Capita (2)	-0.045**	(0.007)	-0.042**	(0.008)
Number of Borders (1)	-0.039**	(0.010)	-0.030**	(0.009)
Number of Borders (2)	-0.039**	(0.011)	-0.032**	(0.009)
Civil War (1)	-0.278**	(0.099)	-0.318**	(0.097)
Civil War (2)	-0.296**	(0.100)	-0.317**	(0.097)
NATO Dyad	-0.050	(0.121)	-0.036	(0.119)
Warsaw Pact Dyad	-4.099**	(0.108)	-3.936**	(0.095)
Intercept	2.240**	(0.111)	-2.320**	(0.108)
		Equation 2: Arm		
Complementary (2)	0.078	(0.065)		
Complementary (3)			-0.517**	(0.143)
Ongoing Dispute	0.120	(0.088)	-0.556*	(0.192)
Ongoing Alliance	-0.159**	(0.051)	-0.305**	(0.067)
Threat Environment (1)	-0.002**	(0.000)	-0.006**	(0.001)
Threat Environment (2)	0.001*	(0.000)	0.001**	(0.000)
Δ Threat Environment (1)	0.019**	(0.001)	-0.009**	(0.002)
Δ Threat Environment (2)	-0.001	(0.001)	0.001	(0.002)
S-Score	0.081**	(0.028)	0.231**	(0.040)
Distance	0.000**	(0.000)	0.000**	(0.000)
Polity (1)	0.025**	(0.001)	0.049**	(0.002)
Polity (2)	0.017**	(0.001)	0.005**	(0.001)
GDP per Capita (1)	-0.037**	(0.002)	-0.024**	(0.003)
GDP per Capita (2)	-0.047**	(0.003)	-0.036**	(0.004)
Number of Borders (1)	-0.009*	(0.004)	-0.052**	(0.006)
Number of Borders (2)	-0.016**	(0.003)	-0.027**	(0.005)
Civil War (1)	0.270**	(0.021)	0.037**	(0.025)
Civil War (2)	-0.268**	(0.026)	-0.286**	(0.040)
NATO Dyad	-0.571**	(0.057)	-1.174**	(0.125)
Warsaw Pact Dyad	0.341**	(0.045)	0.379**	(0.058)
Intercept	-0.841**	(0.034)	-0.641**	(0.050)
		Equation 3: ρ		
Intercept	0.128	(0.056)	0.161**	(0.058)

Table 13: Marginal Effects of Complementarity 2 & 3, Additional Controls

Measure	Neither $\frac{\partial y}{\partial x}$	Neither (SE)	Ally Only $\frac{\partial y}{\partial x}$	Ally Only (SE)	Arm Only $\frac{\partial y}{\partial x}$	Arm Only (SE)	Both $\frac{\partial y}{\partial x}$	Both (SE)
Complementary (2)	-0.005	(0.005)	0.000	(0.000)	0.006	(0.005)	0.000	(0.000)
Complementary (3)	0.000	(0.001)	0.000	(0.001)	0.000	(0.000)	0.000	(0.000)

Table 14: Frequency of Complementarity by Measure in Limited Data

Measure	Original Yes	Original No	Original Total	Limited Yes	Limited No	Limited Total
1	144	96424	94568	144	91439	91583
2	1164	97516	98680	1164	90419	91583
3	1522	92142	93664	1487	90096	91583

In addition, one model (1950-1959) would not converge, and some variables were dropped due to collinearity. The results are shown in Tables 16 and 17. Because there are no different predictions for the second measure of complementarity and there are no results (due to collinearity) for the different predictions for the third measure, it is only on the original one that we can look for support for this hypothesis. The results are in the expected direction in three out of the six cases: the interaction between NATO membership and complementarity does make an alliance less likely in 1960-1969, and the interaction between Warsaw Pact membership and complementarity does make an alliance less likely in 1970-1979 and more likely in 1980-1988.

6 Conclusion

Alliance theory has been a mainstay of international relations since Thucydides, but it is only recently that theory and empirics have begun to catch up with what we instinctively understand about the importance of the military. This paper follows the trend of such works by considering the concrete military benefits of an alliance. It looks at a lower level military decision – not just

Table 15: Measures of Fit, Limited Data

Measure	LL	AIC	BIC
1	-32129.4	64288.8	64430.18
2	-32092.04	64214.07	64355.45
3	-32105.8	64241.6	64382.97

Table 16: Bivariate Probit by Decade

	1950-1959	1960-1969	1970-1979	1980-1988
Equation 1: Ally				
Complementary		-4.587** (0.189)	-4.291** (0.183)	-4.084** (0.400)
NATO (1)		0.151 (0.162)	-0.321 (0.333)	-0.296 (0.265)
Warsaw (1)		0.502** (0.128)	0.371† (0.194)	-4.953** (0.404)
NATO*Complementary		-0.145 (0.262)	1.683** (0.502)	-0.145 (0.695)
Warsaw*Complementary		-0.386 (0.376)	-0.328 (0.218)	4.720** (0.251)
Equation 2: Arm				
Complementary		-5.542** (0.091)	-5.150** (0.100)	-5.140** (0.367)
NATO (1)		-1.245** (0.063)	-5.683** (0.036)	-6.477** (0.062)
Warsaw (1)		-0.279** (0.042)	-5.734** (0.068)	-6.987** (0.090)
NATO*Complementary		0.914** (0.136)	5.841** (0.146)	7.082** (0.395)
Warsaw*Complementary		0.778** (0.149)	5.467** (0.130)	7.037** (0.213)

Table 17: Bivariate Probit by Decade

	1950-1959	1960-1969	1970-1979	1980-1988
Equation 1: Ally				
Complementary (2)	-4.904** (0.141)	-4.284** (0.128)		
NATO (1)	0.165* (0.073)	0.132 (0.152)	-0.340 (0.331)	-0.295 (0.262)
Warsaw (1)	0.429** (0.082)	0.507** (0.116)	0.321† (0.175)	-5.175** (0.399)
NATO*Complementary (2)	0.183 (0.156)	0.594** (0.223)		
Warsaw*Complementary (2)	-0.105 (0.170)	0.226 (0.172)		
Equation 2: Arm				
Complementary (2)	-0.192** (0.061)	-1.274** (0.369)		
NATO (1)	-0.417** (0.025)	-1.048** (0.045)	-5.536** (0.029)	-6.644** (0.052)
Warsaw (1)	0.098** (0.029)	-0.194** (0.036)	-5.673** (0.059)	-6.857** (0.080)
NATO*Complementary (2)	0.319** (0.106)	-3.002** (0.373)		
Warsaw*Complementary (2)	0.411* (0.161)	1.097† (0.627)		
Equation 1: Ally				
Complementary (3)		-3.101** (0.135)	-4.392** (0.190)	-4.398** (0.216)
NATO (1)	0.166* (0.072)	0.131 (0.151)	-0.189 (0.310)	-0.061 (0.240)
Warsaw (1)	0.434** (0.082)	0.509** (0.116)	0.205 (0.195)	-4.919** (0.532)
NATO*Complementary (3)		-0.081 (0.175)	1.024* (0.486)	0.329** (0.106)
Warsaw*Complementary (3)		-0.471** (0.140)	5.455** (0.491)	4.347** (0.319)
Equation 2: Arm				
Complementary (3)		−0.295† (0.179)	0.142 (0.102)	-6.448** (0.054)
NATO (1)	-0.438** (0.027)	-1.191** (0.061)	-5.712** (0.038)	-6.890** (0.076)
Warsaw (1)	0.146** (0.030)	-0.259** (0.042)	-5.908** (0.071)	-7.278** (0.102)
NATO*Complementary (3)		-3.830** (0.189)	-0.276* (0.110)	5.953** (0.098)
Warsaw*Complementary (3)		-4.669** (0.187)	-0.003 (0.216)	7.091** (0.141)

power aggregation in general, but in precise ways.

The findings show that the way in which we operationalize specialization matters, and that at least one possible measure does make alliance formation more likely. The results using the other two measures of complementarity do not support the theory, which indicates that the three operationalizations are getting at different state characteristics which may be of concern to decision makers. The results from taking the type of threat faced into consideration are ambiguous, though here the analyses are working from a very limited number of data points.

Perhaps a better way of understanding how military needs impact the choice of alliance partner is to take a step back and ask what affects the military allocations within a state. In other words, what leads to specialization in the first place? If this can be mainly attributed to resource availability or other non-strategic facts (the UK is an island, therefore it has a large navy; Russia has a very large land mass relative to its navigable coastline, therefore it has a large army) then the above is theoretically appealing – states make strategic decisions to compensate for natural disadvantages. However, if the state's allocation decision is heavily influenced by strategic factors, then alliance decisions can be directed at different considerations. In this case, Bensahel's conclusion, that the decision is ultimately political, not military, would seem to be correct.

References

[1] Altfield, Michael F., "The Decision to Ally: A Theory and Test," *The Western Political Quarterly,* **37** 4, Dec. 1984: 523-544.

[2] Bennett, D. Scott, "Testing Alternative Models of Alliance Duration, 1816-1984," *American Journal of Political Science,* **41** 3, July 1997: 846-878.

[3] Bennett, D. Scott and Timothy Nordstrom, "Foreign Policy Substitutability and Internal Economic Problems in Enduring Rivalries," *Journal of Conflict Resolution,* **44** 1, Feb. 2000: 33-61.

[4] Bensahel, Nora, "International Alliances and Military Effectiveness: Fighting Alongside Allies and Partners," in Brooks, Risa A. and Elizabeth A. Stanley (eds.), *Creating Military Power: The Sources of Military Effectiveness.* Stanford University Press, 2007.

[5] Bolks, Sean M., "Security policy choice: Foreign policy behavior as a function of threat, governmental structure, and state capability," Rice University Ph.D. Thesis, 1999.

[6] Bray, Frank T.J. and Michael Moodie, *Defense Technology and the Atlantic Alliance: Competition or Collaboration?* Institute for Foreign Policy Anaysis, Inc., 1977

[7] *CIA World Factbook.* https://www.cia.gov/library/publications/the-world-factbook/index.html. (2008).

[8] Cioffi-Revilla, Claudio and Harvey Starr, "Opportunity, Willingness and Political Uncertainty: Theoretical Foundations of Politics," *Journal of Theoretical Politics,* **7** 4, Oct. 1995: 447-476.

[9] Conybeare, John A.C., "The Portfolio Benefits of Free Riding in Military Alliances," *International Studies Quarterly,* **38** 3, Sept. 1994: 405-419.

[10] Koerner, Craig, "Free Riding and Deterrence Failure: The False Linkage," Unpublished manuscript.

[11] Leeds, Brett Ashley, Andrew G. Long and Sara McLaughlin Mitchell, "Re-Evaluating Alliance Reliability: Specific Threats, Specific Promises," *Journal of Conflict Resolution,* **44**, 2000: 686-699.

[12] Leeds, Brett Ashley, Jeffrey M. Ritter, Sara McLaughlin Mitchell, and Andrew G. Long, "Alliance Treaty Obligations and Provisions, 1815-1944," *International Interactions,* **28**, 2002: 237-260.

[13] Maoz, Zeev. Dyadic MID Dataset (version 2.0): http://psfaculty.ucdavis.edu/zmaoz/dyadmid.html. (2005)

[14] Marshall, Monty G., and Keith Jaggers, *Polity IV project: Political regime characteristics and transitions, 1800-20026* Dataset users manual. University of Maryland: Integrated Network for Societal Conflict Research (INSCR) Program and Center for International Development and Conflict Management (CIDCM), 2002.

[15] *The Military Balance.* International Institute for Strategic Studies, 1962-1988.

[16] Morgan, T. Clifton and Glenn Palmer, "A Model of Foreign Policy Substitutability: Selecting the Right Tools for the Job(s)," *Journal of Conflict Resolution,* **44** 1, Feb. 2000: 11-32.

[17] Morgenthau, Hans J. and Kenneth W. Thompson, *Politics Among Nations: The Struggle for Power and Peace.* (Sixth Edition) Alfred A. Knopf, Inc., 1985.

[18] Morrow, James D., "Arms versus allies: trade-offs in the search for security," *International Organization,* **47** 2, Spring 1993: 207-233.

[19] Most, Benjamin A. and Harvey Starr, "International Relations Theory, Foreign Policy Substitutability, and 'Nice' Laws," *World Politics,* **36** 3, Apr. 1984: 383-406.

[20] Palmer, Glenn and T. Clifton Morgan, *A Theory of Foreign Policy.* Princeton University Press, 2006.

[21] Regan, Patrick M., "Substituting Policies During U.S. Interventions in Internal Conflicts: A Little of This, a Little of That," *Journal of Conflict Resolution,* **44** 1, Feb. 2000: 90-106.

[22] Sandler, Todd, Jon Cauley and John F. Forbes, "In Defense of a Collective Goods Theory of Alliances," *Journal of Conflict Resolution,* **24** 3, Sept. 1980: 537-547.

[23] Sarkisian, Natasha, Class notes, SC 704: Topics in Multivariate Statistics, Boston College.

[24] Schroeder, Paul W., "Alliances, 1815-1945: Weapons of Power and Tools of Management," in Knorr, Klaus (ed.), *Historical Dimensions of National Security Problems.* National Security Education Program, 1976.

[25] Singer, J. David, Stuart Bremer, and John Stuckey, "Capability Distribution, Uncertainty, and Major Power War, 1820-1965," in Bruce Russett (ed.), *Peace, War, and Numbers.* Sage, 1972.

[26] Sorokin, Gerald L., "Arms, Alliances, and Security Tradeoffs in Enduring Rivalries," *International Studies Quarterly,* **38** 3, Sept. 1994: 421-446.

[27] Starr, Harvey, "Substitutability in Foreign Policy: Theoretically Central, Empirically Elusive," *Journal of Conflict Resolution,* **44** 1, Feb. 2000: 128-138.

[28] Vincent, Jack E., Ira L. Straus and Richard R. Biondi, "Capability Theory and the Future of NATO's Decisionmaking Rules," *Journal of Peace Research,* **38** 1, Jan. 2001: 67-86.

[29] Wallander, Celeste A., "Institutional Assets and Adaptability: NATO After the Cold War," *International Organization,* **54** 4, Autumn 2000: 705-735.

[30] Walt, Stephen M., *The Origins of Alliances.* Cornell University Press, 1987.

[31] Walt, Stephen M., "Why Alliances Endure or Collapse," *Survival,* **39** 1, Spring 1997: 156-79.