# Alliances in Anarchic International Systems

EMERSON M. S. NIOU

*Duke University*

AND

PETER C. ORDESHOOK

*California Institute of Technology*

Alliances play a central role in international relations theory. However, aside from applications of traditional cooperative game theory which ignore the issue of enforcement in anarchic systems, or interpretations of the repeated Prisoners' Dilemma in the attempt to understand the source of cooperation in such systems, we have little theory on which to base predictions about alliance formation. This article, then, builds on an *n*-country, noncooperative, game-theoretic model of conflict in anarchic systems in order to furnish a theoretical basis for such predictions. Defining an alliance as a collection of countries that jointly abide by "collective security strategies" with respect to each other but not with respect to members outside of the alliance, we establish the necessary and sufficient conditions for an alliance system to be stable. In addition, we show that not all winning or minimal winning coalitions can form alliances, that alliances among smaller states can be stable, that bipolar alliance structures do not exhaust the set of stable structures, and that only specific countries can play the role of balancer.

There is little disagreement over the proposition that the concept of *alliance* is central to international relations theory. In the realist view, "the historically most important manifestation of the balance of power . . . is to be found . . . in the relations between one nation or alliance and another alliance" (Morgenthau, 1959:169) because "alliances and regional coalitions among the weak to defend themselves from the strong have been the typical method for preserving . . . balance" (Wright, 1965:773). Hence, "it is impossible to speak of international relations without referring to alliances" (Liska, 1962:3). And although neoliberalism offers a formula for stability based on notions of economic interdependence rather than on off-setting military capabilities, alliances *qua* regimes play an important role there as well, to the extent that they facilitate the realization of mutually beneficial economic gains.

Our understanding of alliances is aided by the fact that definitions come within striking distance of acceptability by even rigorous standards. For example,

Walt's (1987:12) definition that "an alliance is a formal or informal arrangement for security cooperation between two or more sovereign states" and Snyder's (1990:104) view that "alliances . . . are formal associations of states for the use (or non-use) of military force, intended for either the security or the aggrandizement of their members, against specific other states . . ." matches the game-theorist's idea of a coalition in which people coordinate strategies to realize some outcome that cannot be realized through uncoordinated action. However, with respect to the development of a general theory, this conceptual agreement does not vitiate Liska's (1962:3) observation that "it has always been difficult to say much that is peculiar to alliances on the plane of general analysis" or Ward's (1981:26) assertion that "little work has probed the black boxes of decision making within . . . alliances . . . [or] has sought to examine, understand, or predict which alliance groupings were likely to form."

The difficulty with accounting for alliances is that because international politics is anarchic—because there are no exogenous mechanisms for enforcing agreements—any hypotheses about alliances must arise out of notions of individual self-interest, where that self-interest is mediated by expectations that agreements will be enforced. Hence, the traditional tools for theorizing about alliance formation—classical cooperative game theory in general and, for a particular example, Riker's (1962) size principle—can at best reveal only half the story. Such analyses can offer hypotheses about alliance formation if agreements are assumed *a priori* to be enforceable, but they cannot tell how those agreements are enforced. And they can generate wholly misleading inferences if alliance formation is itself influenced by any variability in expectations about the likelihood of meaningful enforcement.[1]

The central question we address here, then, is: If alliances arise and are sustained by self-interest alone in an anarchic world characterized by competition for scarce resources, what types of alliance structures are stable and what types are unstable? Our approach to this question is to first conceptualize alliances as limited collective security agreements in which members of an alliance pledge (but are not otherwise committed) to defend the interests of others in ·the alliance. Second, we formulate a two-stage game which proceeds thusly: In Stage 1, countries partition themselves into disjoint alliances. This partition identifies the strategies that countries pledge to implement in Stage 2. Stage 2 consists of a recursive game in which countries attempt to secure resources from each other by making and implementing threats. Thus, the stability of an alliance partition depends on the equilibrium consequences in Stage 2 of that partition as compared to the consequences of alternative partitions. An alliance is stable because its members find it in their self-interest to abide by their pledges or because it cannot dislodge members from other alliances in order to form a more beneficial partition. Conversely, alliance partitions are unstable if some set of countries, by defecting from their pledges, can coalesce to produce a better outcome in Stage 2 for each member of this set.

At first glance, our treatment of alliances looks like the application of the concept of the core from cooperative games (Ordeshook, 1986). But rather than base our analysis on classifications of coalitions as either winning or losing, a coalition's value here and the incentives to join and maintain it are determined by expectations of what follows in the Stage 2 noncooperative threat-counter-threat game. The conclusions we reach, moreover, differ from those offered by

---

[1] Most existing research that explores the sources of cooperation and enforcement in anarchic environments relies on a particular game, the repeated Prisoners' Dilemma, which cannot generally characterize international affairs whenever those affairs become purely conflictual (cf. Taylor, 1976, 1987; Axelrod, 1984; Bendor and Mookherjee, 1987).

more traditional analyses. First, we establish that a great many alliance systems possess core-like stability. Systems can look either like the competitive arena portrayed by realists or like the more cooperative one described by neoliberals. Second, anarchic international systems need not exhibit the instabilities that are commonly deduced from classical game-theoretic approaches. And third, we address a number of questions that have as of now gone unanswered, or which we believe have been answered incorrectly. These questions include:

- Are profitable alliances restricted to winning or minimal winning coalitions? Does an explicit model of endogenous enforcement cause us to modify Riker's minimum winning coalitions hypothesis?
- Can the largest (most militarily powerful) states form an alliance at the expense of smaller states, or will alliance structures necessarily divide the most powerful states into opposing camps? Equivalently, is there an inherent tendency toward bipolarity in anarchic international systems?
- Must a collective security equilibrium encompass all states, or can something less than unanimous agreement enforce a system devoid of threats against sovereignty? An answer to this question allows us to determine whether the League of Nations failed because the United States refused to participate or whether we should look to other explanations.
- Can alliances be purely defensive—are offensive alliances more attractive than defensive ones?
- Must alliances form at all? Is bargaining over alliance membership fundamental to international politics?

Section 1 begins by reviewing a model of anarchic systems developed previously to formalize the notions of balance of power and collective security and to establish conditions under which anarchic systems are stable in the sense that all countries can ensure their sovereignty (Niou and Ordeshook, 1990, 1991). But here we turn our attention from the survivability of individual states to the issue of the stability of alliance systems. Defining alliances as limited collective security arrangements, our purpose is to see what alliances might form in anticipation of the necessity for conducting international politics in an otherwise anarchic world. In Section 2 we define the notion of an advantaged coalition, which we use in Section 3 to reexamine the stability of two extreme cases—one in which no alliance forms and the other in which an all-encompassing collective security alliance forms. In Section 4 we define a stable alliance structure, and in Section 5 we provide a necessary and sufficient condition for stable alliance systems along with some examples and subsidiary results that allow us to interpret those conditions. We examine the role of balancing powers in Section 6, and in Section 7 we provide some concluding remarks. Appendix A offers some results under an alternative assumption about the actions of indifferent countries, and Appendix B contains the proofs of all results.

## 1. A Model

Without wanting to enter the fray of the debate over whether countries are best modeled as relative or absolute resource maximizers (see our discussion in Niou and Ordeshook, 1994), we assume that countries pursue a single transferable resource in constant supply and that this resource also measures their relative power, their weight in any alliance. Although the assumption might seem to bias our conclusions in favor of a realist view, we do not want to secure cooperation simply by making the gains from it too great. Moreover, sustaining

cooperation in the context of constant sum competition reveals better the role that institutions play in ameliorating conflict.

In regard to the decisions and choices that our model allows countries, we assume that processes unfold in two stages: an alliance formation stage (Stage 1), followed by a threat-counterthreat stage (Stage 2). Because some of the same notation is employed to describe both of these stages, we summarize it briefly here. First, $\mathbf{r}^o = (r_1^o, r_2^o, \ldots, r_n^o)$ denotes the initial distribution of resources across the set $S = \{1, 2, \ldots, n\}$ of $n$ countries, where $r_1^o \geq r_2^o \geq \ldots \geq r_n^o$; $r(C)$ denotes the total resources controlled by the subset of countries $C$; and $R = r(S)$ denotes the total resources in the system. Hence, $C$ is winning if $r(C) > R/2$, it is losing if $r(C) < R/2$, and $C$ is minimal winning if subtracting any one country from it renders it losing. Countries that are in at least one minimal winning coalition are *essential;* otherwise, they are *inessential.* Finally, if $r_i^o > R/2$, country $i$ is *predominant*—it is winning against all other countries and it can absorb their resources at will—whereas if $r_i^o = R/2$, then $i$ is *near-predominant.*

**Stage 1:** Prior to formulating and making threats in Stage 2, countries negotiate alliances, which entails making promises about the strategies they will pursue in Stage 2. Letting $P = (P_1, P_2, \ldots, P_k)$ be a disjoint and exhaustive partition of $S$, we want to label $P$'s elements "alliances." First, though, we must define what it means to ally. Recall, then, that the Warsaw Pact was held together not only by fear of invasion from the West, but also by military force. Defectors— Hungary, Czechoslovakia, and Romania—were punished by the alliance's other members. And although nothing as dramatic as tanks rolling through Paris, Tokyo, or Bonn cemented the Western alliance, the source of its durability was both the threat from Moscow and its economic profitability. The collective security arrangement among the United States, Japan, and Western Europe required the administration of no severe military punishment, but there existed the threat of economic reprisal and a withdrawal of military support in the event of any defection: "A hegemon may help to create shared interests by providing rewards for cooperation and punishments for defection, but where no hegemon exists, similar rewards and punishments can be provided if conditions are favorable" (Keohane, 1984:78). Finally, we have Lalman and Newman's (1991) empirically based conclusion that "so called *Realpolitik* considerations of security are crucial to alliance formation decisions" (see also McGowan and Rood, 1975). Thus, the following definition seems appropriate:

> **Definition:** An *alliance* is a collective security arrangement among states in which all members of the alliance agree to not threaten each other, to punish defectors from this agreement whenever possible, and to threaten countries outside of the alliance whenever it is in their individual interest to do so.[2]

To formalize this definition, let $D = (D_1, D_2, \ldots, D_k)$ begin as a vector of $k$ empty sets—one for each element of $P$. Next, suppose countries $i$ and $j$ are members of the same alliance in $P$, say, $P_k$. Then, the partition $P$ implies that $i$ promises never to threaten $j$ or to participate in such a threat in Stage 2, and $j$ reciprocates (which is not to say that $i$ and $j$ abide by these promises since they need not be in equilibrium). However, if $t$ is also in $P_k$, but if $t$ ever threatens $i$

---

[2]Our formulation of Stage 1 focuses on the security component of alliance processes without an explicit measure of the "cost of lost autonomy" (Altfeld, 1984; Morrow, 1987). However, notice that the decreased flexibility in the threats that a country is permitted in Stage 2 owing to its membership in an alliance can be viewed as a loss of autonomy that is enforced by the threat of punishment by other alliance members.

or $j$ or participates in a threat against $i$ or $j$ in Stage 2, then $t$ is added to $D_k$ to signify that it is a defector from $P_k$. Membership in $P_k$ presumes that, whenever possible (whenever they have the opportunity and interest), $i$ and $j$ (as well as other nondefecting members of $P_k$) punish $t$ by threatening it or by joining in threats against it. If a member of $P_k$ fails to punish a defector, then it becomes a defector and is added to $D_k$.

Stage 1, then, is an open-ended negotiation in which states, in anticipation of playing Stage 2, partition themselves into alliances, and thereby into agreements about who is a legitimate target of threats and who, by threatening, should be labeled a defector and targeted for punishment. Of course, in accordance with the assumption that international systems are anarchic and that there are no exogenous mechanisms for enforcing agreements, the Stage 2 strategies implied by a partition need not be adhered to. Whether they will be followed depends on whether those strategies, taken as a package, are an equilibrium. Thus, each country's evaluation of a particular partition depends on what it can expect to get if everyone abides by their agreements or, in the event that the agreements are not an equilibrium, on the security value each country ascribes to its strategy. A partition is stable, in turn, if no subset of countries has an incentive to form an alliance other than one contained in it.

Our analysis of alliances, then, has some of the flavor of cooperative game theory—in particular, of application of the concept of the core (Ordeshook, 1986). However, it differs from the usual treatment of that theory in that the value of an alliance is not predicated on the assumption that any agreement is necessarily and somehow enforceable. Instead, an alliance's value is determined by what it implies about the play of a noncooperative game in which states abide by their promises only if it is in their individual interest to do so. It is the structure of that noncooperative game to which we now turn.

**Stage 2:** In accordance with Boulding's (1968:105) view that "threat systems are the basis of politics," we assume that threats and counterthreats are the mechanisms whereby countries secure resources from each other. Hence, avoiding mathematical niceties, Stage 2 of our model proceeds as follows:

1. A randomly chosen country, $i$, is given the opportunity to offer an initial threat or to "pass." An initial threat is a new resource distribution $\mathbf{r}$ and an implied threatening coalition $C$ that corresponds to the countries who do not lose resources by moving from $\mathbf{r}^o$ to $\mathbf{r}$. Of course, $r(C) > r(S\text{-}C)$.
2. If $i$ passes, we return to step 1.
3. If $i$ threatens, its partners in $C$ decide whether or not to accept participation in the threat. Only if all partners accept does $i$'s threat call for a response by the threatened countries. If one or more members of $C$ reject, we return to step 1.
4. Threatened countries are given the chance, in sequence, to respond to the threat. Responses are of two types: (1) a counterthreat that is a new threat; and (2) a proposal to transfer resources from one or more threatened countries to one or more members of the threatening coalition. If a counterthreat is proposed and accepted unanimously by the newly proposed coalition, it becomes the new current threat, and requires a response by the newly threatened countries. If a transfer is proposed and accepted by everyone involved, it determines a new status quo and we return to step 1.
5. Any threat that is not successfully countered is implemented—the threatened resource distribution becomes the new status quo, and the game proceeds as before by returning to step 1.

This model ignores a great many things, including the costs of conflict, uncertainty, and exogenous resource growth. Nevertheless, whatever cooperation is supported in it does not arise merely because we have made cooperation sufficiently profitable via some assumption about the value of a public good. Instead, it supposes only that countries join and maintain coalitions because it is in their individual interest to do so.[3]

## 2. Advantaged Coalitions

The feature of Stage 2 that makes its analysis "interesting" is that it allows for threats and counters that continue in sequence forever without any change in the status quo. So to check whether a combination of strategies is an equilibrium we must pretend that the game is finite and that we know the consequences of all branches in its extensive form (see, e.g., Baron and Ferejohn, 1989). After postulating these consequences, an equilibrium is characterized by (1) strategies in which no one has an incentive to defect unilaterally to some other strategy, and (2) the postulated consequences are consistent—the choices they imply yield those consequences.

The consequences that concern us are of the threats and counters that coalitions make. To evaluate a threat, a country must determine whether its partners will accept it, which depends on the counterthreat that might follow, and so on—all of which depends on each country's conjectures about the strategies of other countries. To analyze this situation we first form a two-way classification of coalitions. Coalitions in the first class can make threats such that, if they are accepted, the largest threatening country becomes near-predominant and no one in the coalition loses resources. Threats by all other coalitions, in contrast, can be countered so that no country in the originally threatening coalition is assured of gaining resources—indeed, some lose. We call the particular class of coalitions that can make profitable threats—the first class—advantaged, and they,: are defined thus:

> **Definition:** $C^*$ is the set of *advantaged* coalitions if, in addition to containing only winning coalitions,
>
> i   for every $C \in C^*$, $S$-$C$ has sufficient resources to render the largest member of $C$ near-predominant;
>
> ii  for every $C \in C^*$ there is no other winning coalition $C'$ such that the intersection of $C'$ and $C$ is the largest country in $C'$ but not in $C$;
>
> iii for no $C \in C^*$ is there a winning coalition $C'$ such that the intersection of $C$ and $C'$ equals $\{1\}$ and $r(C') > r(C)$;
>
> iv  $C^*$ is maximal—no additional coalitions satisfying conditions i and ii can be added without violating iii.

A *primary threat*, in turn, is a vector $\mathbf{r}$ proposed by a coalition $C$ that satisfies these four conditions such that no member of $C$ loses resources in $\mathbf{r}$ ($r_i \geq r_i^o$ for all $i \in C$), the members of $S$-$C$ are threatened with elimination ($r_i = 0$ for all $i \in S$-$C$), and the threat promises to make the largest member of $C$ near-predominant ($r_i = R/2$ for $i = \max[C]$). Reviewing these four conditions,

---

[3]Throughout our analysis we employ the concept of an equilibrium most appropriate to our model—subgame perfect equilibria. Thus, cooperation is enforced by individual self-interest, where that self-interest is defined by the strategies of other states.

- i ensures that the option of transferring resources so as to render someone near-predominant is available to the threatened countries—S-C cannot be too small or too large.
- ii ensures that there cannot be a counterthreat that makes the promise of near-predominance to a smaller country in C unless other members of C are included in the counter. Because these other members cannot simultaneously receive the same promise (it is always cheaper for S-C to transfer resources to a single country if a transfer is the only way to disrupt a threat), these members will reject the counter in favor of the current threat.
- iii and ii together ensure that a threat can be made such that there is no counter that can render the largest threatening country indifferent between acceptance and rejection.
- iv ensures that we do not overlook any coalitions that might make profitable threats.

This classification allows us to characterize the different equilibria that can prevail in Stage 2, but before we consider those equilibria, we want to illustrate $C^*$ here.

> **Example 1:** All 3-country systems in which everyone is essential are equivalent in the sense that $C^* = \{\{1,3\}, \{2,3\}\}$ and primary threats take the form (150,0,150) and (0,150,150). The winning coalition $\{1,2\}$ is never advantaged, because, in violation of condition ii, for $C' = \{2,3\}$, $C' \cap C = \{2\} = \max[C'] \neq [C]$. So in 3-country systems, the two largest countries can never coalesce for mutual gain—country 1 will lose resources from a successful counter by 3.

> **Example 2:** If $r^o = (110,80,60,50)$, $C^* = \{\{1,2,3\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\}\}$.[4] $\{1,2\}$ is not advantaged since, in violation of condition ii, $\{2,3,4\}$ is winning; nor is $\{1,3\}$ advantaged since, in violation of condition iii, $\{1,2,4\}$ is winning and $r(\{1,2,4\}) > r(\{1,3\})$. This example approximately matches the weights discounted for geography that Bueno de Mesquita (1981:105) assigns to Germany-Austria, Britain, Russia, and France in 1900.[5] Thus, neither {Germany,Britain} nor {Germany,Russia} is advantaged, and if we exclude as infeasible any Franco-German alliance, the advantaged coalitions reduce to {Germany,Britain,Russia} and {Britain,Russia,France}. Finally, even if we discount Britain's resources further owing to its geographical position so that Russia ranks above it, the same two alliances remain advantaged.

> **Example 3:** If $r^o = (70,65,60,55,50)$, then $C^* = \{\{1,4,5\}, \{1,3,4\}, \{1,3,5\}, \{2,3,4\}, \{2,3,5\}, \{2,4,5\}, \{3,4,5\}\}$. Notice that a 4-country coalition is not advantaged because condition i is not then satisfied.

> **Example 4:** If $r^o = (100,80,60,40,20)$, then $C^* = \{\{1,2,3\}, \{1,2,4,5\}, \{1,3,4\}, \{1,3,5\}, \{1,3,4,5\}, \{2,3,4\}, \{2,3,5\}, \{2,3,4,5\}\}$. $C^*$

---

[4]Notice that the threat (150,0,80,70) by {1,3,4} is, like (150,0,75,75), also a primary threat. But if two threats by the same coalition satisfy our requirements, then those threats are strategically equivalent.

[5]The actual weights that Bueno de Mesquita's equation assigns are Germany (.13) + Austria (.04) = .17, Britain = .16, Russia = .08, France = .07.

does not include {1,4,5} or {1,3} because $r(\{1,2,3\}) > r(\{1,4,5\})$ and because $r(\{1,2,4,5\}) > r(\{1,3\})$.

This last example approximately matches Bueno de Mesquita's capability numbers if we divide Germany and Austria and let $S = \{$Britain, Germany, Russia, France, Austria$\}$. But suppose we allow ourselves the luxury of speculating that $S = \{$Iraq, Syria, Iran, Saudi Arabia, Kuwait$\}$. Ignoring geography and the potential intervention of other powers, {Iraq,Syria,Iran} is advantaged as are {Syria,Iran,SA,Kuwait} and {Syria,Iraq,SA,Kuwait}. However, notice that {Iraq,Iran,Kuwait} and {Iraq,Iran,SA} are also advantaged even though we can find no historical parallel to these seemingly improbable coalitions. Although we would not want to defend this assignment of capability numbers, the example suggests that predicting alliances requires something more than the mere identification of advantaged coalitions.

Note also that the set of advantaged coalitions does not exhaust the set of minimal winning coalitions. For example, in the 3-country case the minimal winning coalition {1,2} is not advantaged. Letting $L$ be the countries that can be the largest member of a minimum winning coalition ($L = \{1,2\}$ in Examples 1, 2, 3, and 5, and {1,2,3} in Example 4), the following remark summarizes what we know about advantaged coalitions:

> **Remark 1:** *(1) Not all winning or minimal winning coalitions are advantaged; (2) Every advantaged coalition must contain at least one member of $L$ and one member of $S-L$; (3) A coalition consisting of any one country in $L$ and all of $S-L$ is advantaged; (4) Every essential country is a member of some advantaged coalition; (5) If there are 4 or more essential countries, then only coalitions with 3 or more members are advantaged.*

Each of these facts warrants greater empirical evaluation than we can give it here. Fact 1 tells us that by concerning ourselves with the issue of endogenous enforcement, significant asymmetries in the values of coalitions arise that the usual applications of cooperative game theory fail to accommodate. Fact 2 implies that offensive coalitions cannot be limited to an alliance of only the "great powers" in $L$, whereas fact 3 shows that a single, aggressive "great power" can profitably threaten other large states with an alliance of small states. Both of these facts, taken together, then, provide smaller states with their measure of protection in a balance of power environment. Fact 4 reenforces this proposition because it tells us that every essential state has at least one primary counterthreat to every nonprimary threat. Thus, threats against any state must be crafted carefully to be profitable. Finally, fact 5 tells us that in systems with four or more relevant states, the formulation of primary threats will necessarily involve bargaining among three or more states. Only in 3-country systems is bilateral bargaining sufficient to generate offensive or defensive coalitions.

To this list we can add one more fact. Notice that in our 3-country example, country 3 cannot be the target of a primary threat. On the other hand, every country is the target of such a threat in our other examples. Lemma 1 generalizes this fact:

> **Lemma 1:** *If there are four or more essential countries, then every country is excluded from at least one advantaged coalition.*

Thus, if $n > 3$, every country must be concerned about the danger of having to cede resources in the threat-counterthreat process of Stage 2.

## 3. Two Extreme Cases

The preceding discussion does not mean that countries will form advantaged coalitions in Stage 1, and to this end it is useful to review what happens in Stage 2 for two special cases. The first case corresponds to the partition $P = (\{1\},\{2\}, \ldots,\{n\})$, which means that no country allies initially with any other country. The other extreme sets $P = (\{1,2, \ldots,n\})$, which corresponds to an all-encompassing collective security system in which everyone agrees in Stage 1 to not threaten anyone else and to punish any defector from this agreement— George Bush's "new world order."

If we find equilibria in Stage 2 supported by the strategies implied by $P = (\{1\},\{2\}, \ldots,\{n\})$, and if we learn also that only countries controlling some critical relative resource level can ensure their sovereignty, then countries must be vigilant about relative gains and losses. On the other hand, if there is an equilibrium supported by the strategies implied by $P = (\{1,2, \ldots,n\})$ in which no country offers an initial threat, then realization of this equilibrium renders the issue of sovereignty and relative position less salient and allows for greater flexibility in the design of cooperative arrangements. Moreover, if the benefits that accrue through free trade and the like require a nonconflictual world, and if these benefits disappear when agreements to achieve them are disrupted by competition over relative position, then the issue that bears directly on the realist-neoliberal debate is whether such an equilibrium is more or less attractive than the one supported by stationary strategies.

Skirting formalism, we need only review the essential conclusions we prove elsewhere about these two cases.[6] First, in order to make the analysis tractable, we assume that countries only employ *stationary [punishment] strategies.*[7] Specifically, notice that if we ignore the history of play with the exception of who has defected from an alliance agreement, then any point of the Stage 2 game can be described by the 3-tuple $(\mathbf{r}^o,D,\mathbf{r})$ if there is a current threat or $(\mathbf{r}^o,D,\emptyset)$ if there is no current threat. Stationarity then requires that a country choose the same

---

[6]Our results require four assumptions. *First,* outcomes with a near-predominant country are terminal. This assumption can be derived from the supposition that third parties can take advantage of conflicts among others, in which case outcomes with a near-predominant country are terminal since no one makes a new threat for fear that it will become predominant. Notice that this assumption is consistent with the requirement that resources be destroyed among warring parties. *Second,* if $i$ can become near-predominant by implementing a threat or by a resource transfer, $i$ prefers the transfer. This assumption, then, can be justified by the supposition that countries prefer to avoid costly conflict whenever possible. The relevant implication here is that a terminal node is reached if $S$-$C$ offers to render any $i$ near-predominant; and if $S$-$C$ prefers to end the game, it should transfer to max[$C$] since this choice minimizes the resources that it must surrender. *Third,* as a further characterization of strategies, we suppose that when countering a threat and when it is possible to do so, $i$ chooses a counter that includes all jointly threatened countries in the newly proposed coalition. The rationale here is that the threat against $S$-$C$ makes the coordination of their actions less costly. This assumption, though, merely facilitates some of our original proofs and is not essential to the analysis. *Finally,* for terminal nodes (when some country is near-predominant), $i$'s payoff, $u_i(\mathbf{r})$, equals $r_i$. But for nonterminal nodes we assume risk aversion in this sense: if $\mathbf{R}(\mathbf{r})$ is the set of terminal and nonterminal resource distributions that might be reached from $\mathbf{r}$, given each country's strategy, then $u_i(\mathbf{r}) = \min[R_i(\mathbf{r})]$. This second part of our assumption about utility is designed to avoid complex expected utility calculations that arise owing to random choices by nature later in the game. This assumption precludes leaders who would place any share of their countries' resources and even themselves at risk for the promise of gain and presumes, in effect, that they are minmax regret decision makers. (Parenthetically, we note that $\min[R_i(\mathbf{r})] > 0$ even though some branches of the game's extensive form yield $i$'s elimination since the calculation of utility is contingent on prespecified strategies; also, this assumption is stronger than what is actually required, although risk aversion is essential to the analysis.)

[7]Of course, in a game such as ours, there are infinitely many pure strategies, including some that might be deemed reasonable. We know that allowing certain strategies, such as requiring that countries target threats and counters at those who have threatened them previously, labeling as a defector only those countries that propose an initial threat, or allowing a country to be labeled a defector for only a few periods, leave our analysis unaffected. But this does not prove that *all* strategies are equivalent to the ones we consider, and so ascertaining the sensitivity of our analysis to the allowable strategy domain remains a potentially valuable extension of our analysis.

action at any two decision points described by the same 3-tuple. Of course, if $P = (\{1\},\{2\}, \ldots,\{3\})$, then $D$ is always empty, in which case the following result holds:

> **Result 1:** *If all countries are essential, if $P = (\{1\},\{2\}, \ldots,\{n\})$, and if countries participate in threats whenever doing so does not lead to a reduction in their resources, then the implied strategies yield a strong equilibrium in Stage 2 in which no country is eliminated. But if we allow sequential threats (i.e., $i$ proposes that $C$ threatens $j$, then $k$, etc.), then inessential countries and perhaps even "small" essential ones cannot assure their sovereignty* (Niou and Ordeshook, 1990).[8]

The logic of this result follows from the values we associate with the threats made by advantaged coalitions. With a primary threat, the largest threatening country becomes near-predominant from a resource transfer and its partners do not lose resources; a nonprimary threat allows for primary counterthreats and thereby cannot ensure that the threatening countries retain even their current resource endowments. Since Result 1 states, in effect, that these values are consistent with equilibrium strategies, if we again let $L$ be the set of countries that can be the largest member of a minimum winning coalition, we have the following corollary:

> **Corollary 1:** *For the equilibrium described by Result 1, only members of $L$ can gain resources, and if $n > 3$, all members of $S$ can lose resources.*[9]

Thus, the "largest" countries, if allowed to threaten, do so because they gain and avoid the possibility of loss, whereas smaller countries, although unable to gain from a threat, avoid the possibility of loss. So if a country believes others will abide by their equilibrium strategies, it has an incentive to make or agree to threats that include it in the threatening coalition, since not doing so diminishes its expected utility. To illustrate this argument, let us return now to one of our earlier numerical examples.

> **Example 4 (continued):** With $\mathbf{r}^o = (100,80,60,40,20)$, every country is essential and no country can be eliminated. For example, if $\{1,2,3,4\}$ threatens 5, 5 can counter with a primary threat against, say, 2, since $\{1,3,4,5\}$ is advantaged but $\{1,2,3,4\}$ is not. Country 1 accepts since the counter forces 2 to render it near-predominant, and 3 and 4 accept since freezing the system guarantees that they cannot lose resources (which is not a guarantee that a threat against 5 alone provides since 5's elimination merely results in a 4-country game in which, from Corollary 1, all countries can lose resources). Similarly, if $\{1,2\}$ threaten $\{3,4,5\}$, 3, 4, and 5 can counter with $\{2,3,4,5\}$ or $\{1,3,4,5\}$. But if $\{2,3,4\} \in C^*$ proposes $(0,150,75,75,0)$, 1 and 5 cannot offer a primary threat that causes a member of $\{2,3,4\}$ to defect—2 rejects such an offer since doing so implements the threat and

---

[8]Although the alternative "countries make or participate in threats only if doing so promises them a gain" yields an equilibrium, this equilibrium is unstable in that countries have an incentive to defect if there is any chance that others will defect. That is, the equilibrium is not perfect.

[9]The case of $n = 3$ is special in that, in equilibrium, the smallest country cannot lose resources, and one of the two largest countries gains resources at the expense of the other.

renders it near-predominant; and 4 and 5 reject since doing so
yields each of them a secure 75. Thus, 1 and 5 must cede
resources.

Turning now to an all-encompassing collective security arrangement, so that
we can use Result 1, we assume that if the sum of the resources of the countries
in the $D_k$'s equals or exceeds $R/2$, then everyone plays as if $P = (\{1\},\{2\}, . . .,\{n\})$.
That is, if states with a majority of resources in the system defect from the
alliance agreement reached in Stage 1, then all states act as if the system were
"a wholly anarchic free for all." In this event, we have the following result:

> **Result 2:** *If there are four or more essential countries, and if $P =$
> $(\{1,2, . . .,n\})$, then the implied strategies yield a strong equilibrium in
> Stage 2 such that no country makes an initial threat and the status quo
> is preserved; but if there are only three such countries, or if countries
> must sequentially choose between accepting and rejecting participation
> in threats, then the collective security equilibrium is not strong* (Niou
> and Ordeshook, 1991).

To see how such an equilibrium is supported it is sufficient to examine a
single example:

> **Example 4 (continued):** With $\mathbf{r}^o = (100,80,60,40,20)$, suppose
> country 1 defects from $P = (\{1,2,3,4,5\})$ and proposes a primary
> threat by $\{1,3,4\}$. If 3 and 4 must choose to accept or reject
> simultaneously, then each must fear that if it alone accepts (in
> which case the threat is rejected), it will be the target of a
> punishment by a primary threat in the next round—4 can pun-
> ish with $\{1,2,4,5\}$ and 3 can punish with $\{2,3,5\}$. And since, from
> Corollary 1, agreeing to a threat will only leave 3 and 4's re-
> sources unchanged, there is nothing lost by rejecting 1's initial
> offer. On the other hand, if threats are sequentially considered
> by 3 and 4, then an initial acceptance by one of them renders
> the other indifferent between accepting and rejecting (since
> neither gains resources from a primary threat regardless of
> whether it is an initial one or a punishment) and thereby renders
> the equilibrium weak.

In this way, then, punishment strategies support equilibria in which no one
makes a threat. But such equilibria are vulnerable if there are only three essential
countries or if countries sequentially reveal their willingness to participate in
threats. The particular problem occasioned by that weakness is that it is difficult
to judge adherence to a punishment strategy (since commitment is revealed
only after the fact) and countries cannot be certain that they are in such an
equilibrium. If everyone presumes that all others have some chance of defecting
from administering punishments if they are indifferent about doing so, then
the collective security equilibrium can break down. Thus, collective security
requires "nurturing" by mechanisms that facilitate the realization of those mu-
tual benefits that disappear when countries compete for relative position rather
than pursue the pure objective of absolute resource maximization.

## 4. Stable Alliance Structures: A Definition

Thus far we have merely identified the coalitions that assure a beneficial resource transfer, but we cannot determine what agreements can be reached in Stage 1 in anticipation of Stage 2. Suppose, then, that prior to Stage 2, countries partition themselves into exhaustive and disjoint alliances. We appreciate, of course, that such bargaining lies at the heart of the most interesting processes in international affairs. Indeed, we argue later that the period 1871–1914 consisted of just such an "out of equilibrium" process and that therefore "balance of power politics" does not correspond to the attainment of an equilibrium but rather to the process whereby particular equilibria are achieved.

Because any model of bargaining is necessarily *ad hoc,* we approach the analysis of Stage 1 using a classical cooperative game-theoretic approach to identify the potential "sticking points" of bargains. In accordance with our view that systems are anarchic, though, we continue to assume that countries participate in alliances because it is in their individual interest to do so, not because of any exogenous enforcement. Thus, a partition is *a stable alliance structure* if no set of countries can gain by coordinating their actions so as to reform the partition into a different alliance structure, either by defecting from their current alliances or by joining two or more alliances to form a single alliance.

To remove any ambiguity from this definition, let $P$ be a partition of $S$ and let $s_i^P$ be country $i$'s *security value with respect to P,* where,

> **Definition:** Country $i$'s security value with respect to the partition $P$, $s_i^P$, corresponds to $i$'s minimum payoff if everyone subsequently abides by subgame perfect strategies, with the assumption that alliance partners play punishment strategies with respect to each other unless, in playing our threat-counterthreat game, they prefer to defect unilaterally from doing so.

This definition is incomplete because we have some flexibility in the specification of certain actions when countries are indifferent. But before we examine alternatives, let us consider a situation that entails no ambiguity. If no alliances form in our 3-country example—if $P = (\{1\},\{2\},\{3\})$—then $s_1^P = s_2^P = 70$, because neither 1 nor 2 can preclude the possibility that one or the other will join with 3. But $s_3^P = 80$, because there is no primary threat against 3 and it can always offer a primary threat as a counter to any threat against it. Indeed, since country 3 cannot be threatened by a primary threat and since it can never gain resources, it never has an incentive to admit or to expel a partner from an alliance. Countries 1 and 2, on the other hand, have a considerable incentive to ally with 3.

To generalize our discussion to larger systems, we say that the partition $P$ is stable if there is no alternative partition $P'$ such that the security value of all members of $C'$ in $P'$ is greater than their security value in $P$. Formally,

> **Definition:** The partition $P$ is stable if there does not exist a $P'$ such that for any $C'$ in $P'$, $s_i^{P'} > s_i^P$ for all $i$ in $C'$.

A stable alliance system, then, looks like an element of the core of a cooperative game—an outcome in which no coalition has the ability and a unanimous incentive to upset. An alliance partition is stable if no collection of countries has a unanimous positive incentive to establish a different partition. But unlike the

usual applications of the core that presuppose the exogenous enforcement of agreements, alliances here are enforced by the mutual self-interest that arises from subsequently playing the threat-counterthreat game we use to model anarchic systems.

## 5. Stable Alliance Systems: Existence

To characterize stable and unstable alliance systems, we offer the following general theorem. Letting $G(P) = \{ i \mid s_i^P < r_i^o \}$ denote the set of countries with security values less than their current resource distribution, then,

> **Theorem 1:** *The alliance structure P is stable if and only if G(P) does not contain all countries in S.*

By itself, Theorem 1 is difficult to interpret, but at least one immediate implication follows from it and Results 1 and 2:

> **Corollary 2:** *The all-encompassing collective security system P = ({1,2, . . .,n}) is stable, whereas P = ({1},{2}, . . .,{n}) is stable only if n = 3.*

Our analysis, then, is not inconsistent with Snyder's (1991:124) argument that "a multipolar system structure in which capabilities are distributed evenly does not by itself imply any alignments. However, moderate differences in capability may generate some alignment expectations in a multipolar system. For example . . . two strong states with a weaker state between them will tend to be rivals and protective of the weaker state against each other." Snyder's argument about the propensity to form alliances, however, is valid in our model only if $n = 3$. Specifically,

> **Remark 2:** *If n > 3, then some alliance is certain to form in Stage 1.*

That is, only 3-country systems are impervious to the forces of alliance formation. This fact is a consequence of two subsidiary facts: If $n = 3$, then any advantaged alliance requires the participation of the smallest country and the smallest country cannot gain resources if it participates in a threatening advantaged alliance. Consequently, the smallest country has no positive incentive to join a threatening coalition. So, if it is common knowledge that country 3 participates in an initial threat only if it gains and 3 would assist in defending any threatened country (lest it become the next victim), then ({1},{2},{3}) is stable. However, once we move to larger systems, all countries can lose resources (Lemma 1) and the forces for alliance formation in some form become irresistible.

Corollary 2, though, does not imply that collective security enjoys any advantage over a balance of power type system, since we have not shown that alliance partitions with offensive alliances are unstable. To see that they are in fact stable requires that we clarify what it is that countries do when they are indifferent between abiding by and defection from an equilibrium strategy in Stage 2.

To see the problem, notice that if $\mathbf{r}^o = (70,65,60,55,50)$ and if $P = (\{1,3,4\},\{2,5\})$, then, because {1,3,4} is advantaged, $G(P) = \{2,5\}$—that is, countries 2 and 5 can be the targets of a primary threat by {1,3,4}. However, {2,4,5} and {2,3,5} are also advantaged and 3 and 4 are both indifferent as to which primary threat they participate in. Moreover, 3 and 4 are indifferent between

staying with {1,3,4} and defecting to form {2,4,5} or {2,3,5} since, if they defect, they cannot be punished—the only response to a threat by an advantaged coalition is a resource transfer, not another threat (Niou and Ordeshook, 1990).

Appendix A considers what occurs when states always abide by equilibrium strategies even if they are indifferent between doing so and defecting. But in matters of survival, decision makers are unlikely to be comforted by technical arguments about equilibria. They should plan for worst-case scenarios, which includes the defection of indifferent allies. So here we assume that in the case of indifference, countries have some probability of unilaterally defecting from an alliance's collective security arrangement. Thus, in our example, everyone's security value in $P = (\{1,3,4\},\{2,5\})$ is less than their current endowment, whereas in the partition $P' = (\{3,4,5\},\{1,2\})$, $s_3^{P'} \geq r_3^o$, $s_4^{P'} \geq r_4^o$, and $s_5^{P'} \geq r_5^o$ because a unilateral defection from {3,4,5} does not yield an alternative coalition with a primary threat—neither {1,2,3}, {1,2,4}, nor {1,2,5} is advantaged. Assuming, then, that indifferent countries "play with a shaky hand," Theorem 2 provides our first specific result about stable alliances:

> **Theorem 2:** *If "countries play with a shaky hand," and if there are more than three essential countries, then $P = (C,S\text{-}C)$ is stable if and only if:*
>
> 1. *$C \in C^*$ and for no $i \in C$ is $S\text{-}C + \{i\} \in C^*$; or*
> 2. *$C \in C^*$ and for no $i \in S\text{-}C$ is $C + \{i\} \in C^*$.*

Notice that by letting $S\text{-}C$ be empty, Theorem 2 subsumes Corollary 2's assertion that an all-encompassing collective security alliance is stable, since, from part 5 of Remark 1, no single country is an advantaged coalition and thus no country has an incentive to defect. But notice also that this part of Remark 1 asserts that if $n \geq 4$, advantaged coalitions must have three or more members. So the partition $(S\text{-}\{i\},\{i\})$ is stable—no member of $S\text{-}\{i\}$ has an incentive to defect and join $i$ since doing so does not generate a profitable threat. That is,

> **Corollary 3:** *If there are four or more essential countries, $P = (S\text{-}\{i\},\{i\})$ is stable; and if $\max[S\text{-}\{i\}] + \{i\} < R/2$ so $S\text{-}\{i\}$ is not advantaged, $P$ yields a collective security equilibrium.*

This corollary, then, answers our earlier question about the possibility of enforcing a collective security agreement with something less than unanimity and raises a new question, which we examine later, about how many countries can be excluded from such an agreement before it becomes unstable. But first, letting Table 1 summarize our discussion, we return to several of our earlier numerical examples in order to illustrate Theorems 1 and 2 (Table 1 also summarizes the results in Appendix A). In illustrating this result it is useful to distinguish between profitable and unprofitable alliances, because such a distinction removes some of the ambiguity between bipolar and multipolar systems and reveals that Theorem 2 applies to certain types of "multipolar" systems. If $\mathbf{r} = (70,65,60,55,50)$, then $P = (\{3,4,5\},\{1\},\{2\})$ and $P' = (\{3,4,5\},\{1,2\})$, in addition to being stable, are equivalent in the sense that it matters little whether 1 and 2 ally to transform $P$ into the bipolar system $P'$. Allied or not, both countries are certain to be targets subsequently of a primary threat by {3,4,5}. Hence, to sort through these equivalences, we offer the following definition:

> **Definition:** Suppose $P_k = \{k,i, \ldots,m\} \in P$. If we substitute $\{k\},\{i\}$, $\ldots,\{m\}$ for $P_k$ in $P$ to form $P'$, then $P$ and $P'$ are *equivalent* if

$s_j^P = s_j^{P'} < r_j^0$ for all $j \in P_k$, and if $s_j^P = s_j^{P'}$ for all other $j \in S$. In this event $P_k$ is *unprofitable*.

Thus, $P = (C, P_1, \ldots, P_k)$ is equivalent to $(C, S\text{-}C)$ if $S\text{-}C = P_1 \cup \ldots \cup P_k$ is unprofitable.

> **Example 2 (continued):** If $\mathbf{r}^o = (110,80,60,50)$ and $P = (\{1,2,3\},\{4\})$ then, since $\{1,2,3\}$ is advantaged, $s_4^P < r_4^o$. The only potential defectors from $\{1,2,3\}$ are 2 and 3, because neither gains by threatening 4, but neither has an incentive to defect unilaterally since both can be punished. Hence, $P$ is stable. Now let $P = (\{1,2\},\{3,4\})$, in which case 1 and 2 must each be concerned that its partner will defect since both $\{1,3,4\}$ and $\{2,3,4\}$ are advantaged. Thus, $s_j^P < r_j^o$, $j = 1,2$. And although neither 3 nor 4 has a positive incentive to defect, each is indifferent between maintaining the alliance and defecting, so if both play with a shaky hand, $s_j^P < r_j^o$, $j = 3,4$. Hence, $S = G(P)$, and, from Theorem 1, $(\{1,2\},\{3,4\})$ is unstable. For similar reasons, the partitions $(\{1,3\},\{2,4\})$ and $(\{1,4\},\{2,3\})$ are unstable.

> **Example 3 (continued):** With $\mathbf{r}^o = (70,65,60,55,50)$, let $P = (\{3,4,5\},\{1\},\{2\})$. Notice that $\{3,4,5\}$ is advantaged, but no unilateral defection yields an advantaged coalition. Thus, any defection can be punished and $P$ is stable. In contrast, let $P = (\{1,4,5\},\{2,3\})$. Although $\{1,4,5\}$ is advantaged, 4 and 5 can each unilaterally defect to $\{2,3\}$ to form an advantaged coalition, which is a possibility we cannot preclude if countries play with a shaky hand whenever indifferent. Indeed, this argument applies to all bipolar systems in which one country has a primary threat except $(\{3,4,5\},\{1\},\{2\})$. And to illustrate Corollary 3, let $P = (\{1,2,3,4\},\{5\})$. If $\{1,2,3,4\}$ proposes to eliminate 5, the result is a 4-country game in which no country is immune from being the target of a primary threat. But if 1 moves first in our threat-counterthreat game and proposes that $\{1,4,5\}$ threaten $\{2,3\}$, 5 accepts, but 4 is indifferent between accepting and rejecting and so there is some probability that 4 rejects and 1 and 5 are targets of a subsequent threat. Thus, no threats are made and no one has an incentive to defect from their alliance, so $P$ is stable.

> **Example 4 (continued):** If $\mathbf{r}^0 = (100,80,60,40,20)$, then 4-country coalitions against 1, 2, or 3 are advantaged, as is $\{1,2,3\}$. And, since no unilateral defection from any of these coalitions generates an advantaged alliance, the four alliance structures portrayed in Table 1 are stable. On the other hand, $(\{1,3,4\},\{2,5\})$ is not stable even though $\{1,3,4\}$ is advantaged, because if 3 defects, $\{2,3,5\}$ is advantaged.

Earlier we illustrated Example 4 by supposing that $S = \{$Iraq,Syria,Iran,Saudi Arabia,Kuwait$\}$ and by noting that although certain advantaged coalitions "made sense," others such as $\{$Iraq,Iran,Kuwait$\}$ did not because we could not find historical parallels. However, aside from an all-encompassing collective security alliance, only alliances against $\{$Iraq$\}$, $\{$Syria$\}$, $\{$Iran$\}$, or $\{$SA,Kuwait$\}$ are stable alliances (see Table 1). Of these, only the alliance against Syria has not been explicitly proposed. The alliance $\{$Syria,Iraq,SA,Kuwait$\}$ was thwarted

TABLE 1. Stable alliance structures with profitable alliances.

| $r^o$ | With "Shaky Hand" | Without "Shaky Hand" |
|---|---|---|
| (120,100,80) | ({1,3},{2})* <br> ({2,3},{1})* <br> ({1},{2},{3}) | same as with shaky hand <br> + <br> ({1,2,3}) |
| (110,80,60,50) | ({1,2,3,4}) <br> ({1,2,3},{4})* <br> ({1,2,4},{3})* <br> ({1,3,4},{2})* <br> ({2,3,4},{1})* | same as with shaky hand <br> + <br> ({1},{2},{3,4}) <br> ({1},{3},{2,4}) <br> ({1},{4},{2,3}) |
| (70,65,60,55,50) | ({1,2,3,4,5}) <br> ({3,4,5},{1},{2})* <br> ({1,2,3,4},{5}) <br> ({1,2,3,5},{4}) <br> ({1,2,4,5},{3}) <br> ({1,3,4,5},{2}) <br> ({2,3,4,5},{1}) | same as with shaky hand <br> + <br> ({1,4,5},{2},{3})* <br> ({2,4,5},{1},{3})* <br> ({1,3,4},{2},{5})* <br> ({1,3,5},{2},{4})* <br> ({2,3,5},{1},{4})* <br> ({1},{2},{3},{4,5}) |
| (100,80,60,40,20) | ({1,2,3,4,5}) <br> ({1,3,4,5},{2})* <br> ({2,3,4,5},{1})* <br> ({1,2,4,5},{3})* <br> ({1,2,3},{4},{5})* <br> ({1,2,3,4},{5}) <br> ({1,2,3,5},{4}) | same as with shaky hand <br> + <br> ({1,3,4},{2},{5})* <br> ({1,3,5},{2},{4})* <br> ({2,3,4},{1},{5})* <br> ({2,3,5},{1},{4})* <br> ({1},{2},{3,4,5})* <br> ({1},{4},{5},{2,3}) |

*denotes partition in which one alliance is advantaged.

by Hussein and Assad's conflicting ambitions, whereas {Iraq,Syria,Iran} was thwarted by the intervention of outside powers who succeeded in forming {Syria,Iran,SA,Kuwait}.

Note now that there is only one stable partition with a 3-country offensive alliance in Example 4 and only one advantaged alliance in Example 3 yields such a partition. The pattern here, then, is that smaller offensive alliances are more likely to be vulnerable to being upset in Stage 1 since a single defection tips the balance in favor of a competing advantaged alliance. These alliance structures become stable only if countries are unconcerned about the possibility of defections when their alliance partners are indifferent about defecting. So, in contrast once again to Riker's size principle, our analysis provides a reason for countries to construct winning offensive alliances that are bigger than minimum size.

One last example illustrates some of the possibilities that arise in larger systems with two "super powers" and many smaller states.

**Example 5:** If $r^o$ = (100,80,20,20,20,20,20,20), then notice first that any bipolar partition $(C,S\text{-}C)$ is stable if $C$ is advantaged, because no unilateral defection can generate a countercoalition that is advantaged. Thus, there are a great many stable alliance systems involving advantaged alliances, but all of them pit one of the two larger countries against the other. In this instance, if purely offensive alliances are somehow inevitable, then we

should anticipate considerable bargaining for the loyalties of the smaller states.

Example 5 is interesting also because it reveals the different alliances that can impose a collective security arrangement. From Corollary 3 we know that any alliance of the form $S$-$\{i\}$ containing 1 and 2 can enforce such an arrangement. But consider the partition $(\{1\},\{2\},\{3,4,5,6,7,8\})$. That this partition is stable follows from the fact that no unilateral defection from $C = \{3,4, \ldots,8\}$ yields an advantaged coalition—$\{1,2,j\}$ violates condition ii in the definition of such coalitions. And since they are never the recipients of a transfer, no member of $C$ has a preference for a different alliance system that allows for a primary threat in Stage 2. For the same reason the partition $(\{1\},\{2\},\{j\},C$-$\{j\})$ is also stable. Hence, a collective security system can be imposed by an alliance that controls as little as one-third of the system's resources. Nevertheless, if we make the small-state alliance too small, as in the partition $(\{1\},\{2\},\{3\},\{4\},\{5,6,7,8\})$, then the corresponding alliance structure is unstable.[10]

Collective security can be secured, then, if a sufficient number of small states coalesce. But this is not the only possibility. For instance, $(\{1,2, \ldots,7\},\{8\})$, and $(\{1,2, \ldots,6\},\{7,8\})$ are stable for the same reason that $(\{1,2,3,4\},\{5\})$ is stable in Example 4. Thus, if the two largest states deem it in their interest to forego competition, then they and some of the smaller states can impose a collective security system even if there are a few "renegade" states.

This analysis leads us to reject the hypothesis that an American defection from the League of Nations precluded an effective collective security arrangement. If the United States was most interested in returning to its isolationist position following World War I, then even the defection of a major European power might not have precluded such an arrangement. On the other hand, since the corresponding equilibrium in Stage 2 is weak without effective economic incentives, perceptions play an especially important role. That is, if everyone believes that collective security requires unanimity, then such an arrangement is unlikely to appear without it. Thus, we are led to speculate that the League of Nations proved ineffective not because of specific individual defections, but more because the world war had so effectively destroyed the beliefs that rendered such an equilibrium attainable.

If we let $\mathbf{r}^o = (75,65,60,55,50)$, the following conclusions summarize our examples:

1. alliances are certain to form—$(\{1\},\{2\},\{3\},\{4\},\{5\})$ is unstable;
2. not all winning or advantaged coalitions can establish a stable alliance structure—$(\{1,2,3\},\{4\},\{5\})$ is not stable;
3. the consequence of an all-encompassing collective security system is achieved by a collective security agreement that encompasses "nearly all" states—$(\{1,2,3,4\},\{5\})$ is stable and no country offers a threat in the subsequent play of our threat-counterthreat game;
4. stable alliance systems need not be bipolar—$(\{3,4,5\},\{1\},\{2\})$ is stable—but there always exists a bipolar system that is strategically equivalent to a non-bipolar one—$(\{3,4,5\},\{1,2\})$ in our particular example;

And, using $\mathbf{r}^0 = (100,80,20, \ldots,20)$ as our example,

---

[10]To see this, notice first that the defection of any $j$ from $\{5,6,7,8\}$ renders $\{1,2,3,4,j\}$ advantaged, so $s_j^P < r_j^o$, $j = 5,6,7,8$. Second, $\{1,5,6,7,8\}$ is advantaged, so $s_j^P < r_j^o, j = 2,3,4$. Finally, given their security values, $2,3, \ldots,8$ have an incentive to threaten 1 (that such a threat exists follows from Lemma 1), so $s_1^P < r_1^o$. The instability of $P$ follows from Theorem 1.

5. a stable alliance system need not contain any winning alliances—({1},{2}, {3, . . .,8}) is stable.

So our analysis identifies three types of alliances as parts of a stable alliance system:

A. *advantaged alliances* like {3,4,5} in ({1,2},{3,4,5}), which are offensive;
B. *cooperative alliances* such as ({1,2,3, . . .,7},{8}), which support a collective security outcome; and
C. *blocking alliances* like {3,4, . . .,8} in ({1},{2},{3,4, . . .,8}), which are not winning and which are defensive (because they can block any primary threat) and offensive (because they can join with others to form such a threat).

Type A alliances are likely to be short-lived, because they are designed to upset the status quo and to reallocate resources. Type B alliances model the abortive League of Nations, and, more recently, Bush's "new world order." Historically, such alliances are short-lived as well, and indeed, we already know from Result 2 that they require a special form of "nurturing" if they are to compete against alliances of the first type. Finally, type C alliances are perhaps most congruent with classical balance of power notions in that they play the role of balancer and can either prevent profitable threats or determine which threat is eventually made and accepted. We cannot say, however, whether such alliances have any advantage over other types in terms of durability.

## 6. Balancers

Type C alliances that can block the formation of primary threats are necessarily pivotal between coalitions that can make such a threat, which suggests that the concept of a balancer can be given precise meaning with our analysis. Specifically,

> **Definition:** A country or an alliance is a *balancer* in a given alliance structure if it is pivotal between any two coalitions of alliances with primary threats.

The examples of Britain in the 19th and China and the United States in the 20th century raise several questions about balancers. Owing ostensibly to its geographical isolation and its desire to preserve the status quo so as to maintain profitable trading relationships, Britain is credited with playing the role of balancer in the 19th century, and thus we can ask: Does our analysis predict Britain's role and does it rationalize the relevance of geography? China in this century sought a similar role for itself, although, unlike Britain, it sought to form an alliance of third world states to offset the American and Soviet-led blocs. And now with the Soviet bloc dissolved, the United States is seen as the balancer in East Asia, with China, Japan, the Koreas, and Taiwan seen at least as parts of some potential arrangement of competing alliances. Is there any reason to suppose that China would find a balancing alliance of third world states especially valuable; and why is the U.S. assumed to be uniquely positioned to play a balancing role in a region far removed from its territory?

Looking at the issue of whether any country can be a balancer, two remarks help us answer such a query. First,

**Remark 3:** *The two largest countries in S can never individually be balancers in any alliance structure.*

This remark appears to be contradicted by Britain's 19th century role and the one portrayed for the United States in Asia. Looking first at Britain, by most measures, Britain's military capability exceeded that of any continental power despite the fact that Germany was closing fast at the end of the century. However, it is here that Britain's geographical position with respect to the continent is relevant. Although her navy ensured far greater force projection than Germany with respect to Africa, India, the Far East, and even, perhaps, the Balkans, the events of World War I confirm that Britain was handicapped in any military engagement close to Germany. Britain's potential on the continent as compared with Russia or France is less clear, but, referring to previous numerical examples, let us consider two possibilities.

Suppose first that language and culture rendered Germany and Austria "natural" allies, that (ignoring Italy) $r^o = (110,80,60,50)$ as in Example 2, and that military capability on the continent was ordered G(ermany + Austria) > B(ritain) > R(ussia) > F(rance). If a Franco-German alliance is infeasible, there are two offensive alliance structures: ({R,B,F},{G}) and ({G,R,B},{F}), plus one "blocking" structure, ({G},{F},{R,B}). In all three instances, then, Britain and Russia pivot together. Although Russia sought the role of balancer, it attempted to play this role using one alliance that was not winning—{R,F}—and one that was winning but not advantaged—{G,R}. Also, its efforts were hampered by its ineptitude and perfidity—certainly, Britain's leaders played their role with considerably greater skill.

A second possibility is to allow Germany and Austria to be independent players and to discount Britain's resources on the continent still further so that Russian resources exceed those of Britain—a reasonable assumption if we want to talk about the "European balance" in Germany's neighborhood and if we want to understand German fears of a joint Franco-Russian mobilization. Suppose, in particular, that G(ermany) = 100, R(ussia) = 80, B(ritain) = 60, F(rance) = 40, and A(ustria) = 20, so that the situation corresponds to Example 4. Then the five 3-country advantaged alliances are {G,R,B}, {G,B,A}, {R,B,F}, {R,B,A}, and {G,B,F}. Thus, given our assignment of weights, only Britain is a balancer; namely, in the partitions ({G,A},{R,F},{B}) and ({R,A},{G,F},{B}). This second partition, however, was deemed infeasible, and, interestingly, the remaining partition is the one that dictated events prior to war.

Thus, regardless of how we discount its resources, Britain plays a special role. However, more problematical from the perspective of asserting that Britain's role was preordained is this result:

**Remark 4:** *If the number of essential countries equals three, then the smallest state is a balancer; but if the number of such countries exceeds three, then no individual country can be a balancer in any stable alliance structure.*

Reassessing Britain as balancer, recall that 1870–1914 was marked by considerable alliance instability. The League of Three Emperors, the Triple Alliance, Germany's courting of Britain, Russia's vacillation between alliance with Germany and alliance with France, and Italy's uncertain status contrast with 40+ years of stability exhibited by NATO and the Warsaw Pact (McGowan and Rood, 1975). Thus, although geography assisted Britain in its role, balancing also required a fluidity of alignments that disappeared when Germany threatened continental predominance (Kaplan, 1957).

A nearly equivalent circumstance appears to prevail today in Eastern Asia. China's relationship to Taiwan, despite the rhetoric, is hardly set in concrete, and her relationship with the two Koreas remains in flux. Japan and China, on the other hand, seem destined to be direct competitors in Asia, especially for the resources and markets of other states (including eastern Russia) in the region. Only the United States can offset the resources of either of these powers. But, although militarily predominant on paper, simple geography discounts America's power as well as the extent to which she threatens the two primary regional competitors.

Insofar as China's strategy of trying to form an alliance of "unaligned states" is concerned, Remark 4 suggests that in a 3-state system dominated by America and the USSR, China alone could play the role of balancer. But as other states became relevant actors, China's unique position vanished. China's alternative was to forge a balancing alliance of "unaligned" states, but the states that were available to it for this purpose were too weak economically and militarily to allow for any balancing role.[11]

## 7. Conclusions

We have ignored several important things in our analysis. We do not allow countries to invest resources so that relative resources (power) change over time. Aside from the assumption that countries prefer to have resources ceded to them over securing them by implementing threats, we do not fully accommodate the costs of war. And although we allow indifferent countries to choose probabilistically, a second and perhaps more important way in which uncertainty can affect matters is bypassed by our assumption that every state knows the point at which a state becomes predominant. Hence, there is no risk in allowing a state to become near-predominant, which is not an assumption that we can comfortably assert characterizes reality.

Despite these limitations (which characterize other analyses of coalitions), we can provide conclusions that are more powerful than those offered by previous research. First, we see that hypotheses such as the size principle require modification—indeed, stable alliance structures need not even contain winning alliances. Blocking coalitions have value, at least for the smaller states that cannot gain resources from "great power" confrontations, and they have value also for larger states if they cannot otherwise seize the advantage. Second, the realization of a universal collective security arrangement does not require the acquiescence of all states. Something other than a coalition-of-the-whole can enforce such an equilibrium if states fear the uncertainty that prevails after excluded states are "eliminated." Of course, all forms of collective security share the weakness of being weak equilibria, at least in our analysis. But the varied alliance structures that can support such equilibria ought to give hope that these equilibria can be realized with appropriate nurturing. Indeed, because there are many stable alliance systems, there is no need to choose between the "state of nature of pure balance of power" and the seemingly unrealizable utopia of an all-encompassing collective security system. Third, although alliances can be both offensive and balancing, balancing alliances are most easily formed by collections of smaller

---

[11] Admittedly, though, the balancing role we have outlined for smaller states is not altogether supported by historical evidence. As Fox (1959:185) observes, "Attempts to add to the power of the small states by combining with other small and presumably disinterested small states regularly failed. . . ." On the other hand, Fox also reveals that a balancing role was not precluded as a possibility: "None of the small states . . . dared go so far in using the strength of one side to oppose another . . . [but] the possibility of such a move was frequently in the minds of the great-power leaders."

states that cannot aspire to near-predominance or by geographically distant states. Thus, our analysis reveals a fundamental divergence in the foreign policy objectives of "small" versus "large" states (Rothstein, 1968) or between "land" and "sea" powers.

Of course, we should prefer more specific predictions about alliances. The opportunity to make such predictions, however, is attenuated by our focus on the outcomes that end bargaining rather than the particularistic details of bargaining itself. This is not to say that we reject the view that the study of politics concerns process. In assessing events in the 19th century, for example, we must appreciate that Britain, as a trading state, would see the implementation of any threat as jeopardizing its position. Thus, unable to establish a collective security system owing to the apparent permanence of a Franco-German conflict, Britain would foster a system without a stable, threatening alliance. Thus, process rather than descriptions of particular equilibria becomes the focus of diplomacy and historical scholarship.

However, rather than lament our failure to model bargaining, we note that we have already made heroic assumptions, and a formalized model of bargaining would reduce generality even further. Thus, we cannot say whether systems are more likely to move toward competing, blocking, or collective security alliances, because all three can correspond to equilibria or to points on the path to a particular equilibrium. On the one hand, though, it is foolhardy to suppose that a purely game-theoretic approach can unambiguously uncover an explanation for the selection of one equilibrium over another. Game theory tells us that such selection involves a great many things, including the initial beliefs of players and the opportunities for revising a game's structure through invention as events unfold—inventions that cannot be portrayed beforehand in an extensive or strategic form representation of things. On the other hand, our analysis does suggest that the alternative conceptualizations of international affairs that realists and neoliberals offer are not, in fact, alternative universes. Rather, they are part of the same universe, and both views may be more or less relevant as a function of circumstances that lie wholly outside of this model or any tractable formal analysis of alliances and international conflict and cooperation.[12]

# Appendix A

If indifferent countries "play with a sure hand" and always make equilibrium choices, then the set of stable alliance structures necessarily expands, and ascertaining the extent of this expansion allows us to evaluate the effect of the form of uncertainty that our analysis admits. Our central result in this circumstance is this:

> **Theorem 3:** *If countries defect from alliances only if they gain from doing so, then any bipolar alliance system is stable.*

This result appears to admit too much, but recall the three types of stable systems identified previously. For example, if $\mathbf{r}^o = (70,65,60,55,50)$, then $P = (\{1,2,4\},\{3,5\})$, $P' = (\{1,2,3,4\},\{5\})$, and $P'' = (\{1,2,3\},\{4,5\})$ are stable. In $P$ the alliance $\{1,2,4\}$ has a primary threat, $P'$ is equivalent to an all-encompassing collective security arrangement, and $P''$ establishes $\{4,5\}$ as a potential blocking

---

[12]For an elaboration of the argument that realist and neoliberal views can coexist within the same model, so that cooperative versus competitive politics is more a process of equilibrium selection, see Niou and Ordeshook (1994). And for a discussion of the limits of formal modeling see Ordeshook (1993).

alliance. Thus, if we again ignore unprofitable alliances, $P$ illustrates Remark 5, $P'$ illustrates Remark 6, and $P''$ illustrates Remark 7:

> **Remark 5:** *If countries defect from alliances only if they gain from doing so, then any alliance system in which one alliance has a primary threat is stable.*

> **Remark 6:** *If countries defect from alliances only if they gain from doing so, and if alliances must be renegotiated after any reallocation of resources, then the bipolar alliance system $(C,S-C)$ is stable and is equivalent to an all-encompassing collective security arrangement if $r(S-C) + r_{max[C]} < R/2$.*

> **Remark 7:** *As long as the members of $S - A$, $r(S-A) > r(A)$ and $A \in P$, require some member of $A$ to form a primary threat, then $P$ is stable.*

These remarks do not imply that the only interesting nonbipolar stable alliance structures are those that entail coalitions of countries in $S-L$, or that only the smallest countries can form defensive alliances. Our next theorem allows us to establish circumstances under which other types of alliance systems are stable, including systems in which members of $L$ join a defensive alliance.

> **Theorem 4:** *If countries defect from alliances only if they gain from doing so, then $P$ is stable if there is a $C \in P$ such that there does not exist a $K \subseteq S-C$ in which $K \in C^*$ and for no $i \in C$ is $\{i\} + K \in C^*$ with $i = max[K,i]$.*

> **Example:** If $r^o = (70,65,60,50,30,7,6,6,6)$, then $S-L = \{4,5,6,7,8,9\}$, and from Remark 7, $(\{1\},\{2\},\{3\},\{4,5,6,7,8,9\})$ is stable. But let $P = (\{1,6,7\},\{2,8,9\},\{3,4,5\})$. Because no combination of 1, 2, 6, 7, 8, and 9 is advantaged, $\{3,4,5\}$ cannot be threatened and it has no incentive to accept another country into it. Hence, $\{3,4,5\}$ is a purely defensive alliance, and $P$ is stable.

> **Example:** If $r^o = (100,80,60,40,20)$, then $S-L = \{3,4,5\}$, and, from Remark 7, $(\{1\},\{2\},\{3,4,5\})$ is stable. But if $P = (\{1\},\{4\},\{2,3\},\{5\})$, then $\{2,3\}$ cannot be threatened by a primary threat—indeed, as long as $\{2,3\}$ maintains itself, only $\{2,3,5\}$ and $\{2,3,4\}$ are advantaged. However, $\{2,3\}$ is not purely defensive since it can join other countries to form an advantaged coalition.

Thus, if countries play with a sure hand, we can observe blocking alliances that include members of $L$—$\{2,3\}$ when $r^o = (100,80,60,40,20)$ and $\{3,4,5\}$ when $r^o = (70,65,60,50,30,7,6,6,6)$. But the stability of such alliances rests on a precarious assumption, and although "great powers" (those in $L$) might try to construct an alliance that blocks threats, they are likely to be thwarted by any uncertainty of commitment.

## Appendix B: Proofs

**Proof of Lemma 1:** The lemma is clearly true $\forall \, k \in L$ since $|L| \geq 2$ and since, for any $k \in L$, $\exists \, C \in C^*$ that excludes $k$, namely, $C = L_o + \{j\}$, $j \neq k$, $j \in L$ where $L_o = S-L$. For any $j \in L_o$, if $r_j + r_1 \geq R/2$, then clearly $S-\{i\} \in C^*$. So

suppose $r_j + r_1 < R/2$. Then construct the coalition $C'$ by adding members of $S$-$\{j\}$ to $j$, beginning with the largest (country 1), then the next largest, and so on to $m+1$, until $r(C') + r_m \geq R/2$. Then $S$-$C' \in C^*$—by construction, condition i is satisfied; condition ii cannot be violated since $r_1 + r_j < R/2$ implies that $r_m + r_j < R/2$; and condition iii cannot be violated since $1 \notin S$-$C'$.

We proceed now to establish three additional lemmas.

> **Lemma 2:** If $C \notin C^*$ because condition iii is violated, then both $S$-$C + \{\max[C]\}$ and $S$-$C + K$ are advantaged, where $K \subseteq C - \{\max[C]\}$.

**Proof:** That $S$-$C+\{\max[C]\} \in C^*$ follows from the assumption that condition iii is violated. Next, notice that $r(C$-$\{\max[C]\}) + r(S$-$C) > R/2$, otherwise $\max[C] > R/2$. So add members from $C$-$\{\max[C]\}$ to $S$-$C$, beginning with the smallest members of $C$-$\{\max[C]\}$, until the resulting coalition is winning. This coalition is advantaged—neither condition ii nor iii can be violated.

> **Lemma 3:** If $C \notin C^*$ because condition ii is violated, then $\exists\, j \in C, j \neq \max[C]$, such that $S$-$C+\{j\} \in C^*$, with $j = \max[S$-$C + \{j\}]$.

**Proof:** If condition ii is violated, then $\exists\, j \neq \max[C]$ in $C$ such that $S$-$C + \{j\} \in W$. Let $j$ be the smallest country in $C$ for which this is true. $S$-$C + \{j\} \notin C^*$ either because it fails to satisfy condition ii or iii. It cannot violate condition ii, though, since $\exists\, C' \in W$ that excludes $S$-$C$ has $j$ as its largest member, and is advantaged. Nor can condition iii be violated; otherwise, $C$-$\{\max[C],j\} + \{h\} \in W$, where $h \neq \max[S$-$C]$. However, $S$-$C + \{j\} \in W$ by construction, so $C$-$\{j\}$ and $C - \{\max[C],j\} \notin W$. And since $r^o_{\max[S\text{-}C]} < r^o_j$ by construction, $C - \{\max[C],j\} + \{h\} \notin W$.

> **Lemma 4:** If all $i \in S$ are essential and $r_{\max[C]} + r(S$-$C) < R/2$, then $|C| \geq 4$.

**Proof:** The lemma is clearly true if $|S| = 4$, otherwise country 4 is inessential. By the same token, if $|S| \geq 5$ and if $r_{\max[C]} + r(S$-$C) < R/2$, then $|C| \geq 4$; otherwise, members of $S$-$C$ are inessential.

**Proof of Theorem 1:** To prove sufficiency, notice first that if $G(P) = \varnothing$, then no $i \in L_o$ gains by defecting to some other alliance or by admitting someone to their alliance. Since threats in which members of $L$ gain require the participation of members of $L_o$, $P$ is stable. Second, suppose that $G(P) \neq \varnothing$ and that $S \neq G(P)$. Thus, $j \in G(P)$ can improve its security value if $G' \subseteq G(P)$ is advantaged or if $G'$ can form an advantaged coalition with $i \in S$-$G(P)$ such that $i = \max[G',i]$. But then $s^P_j < r^o_j$ for $j \in S - G(P) - \{i\}$, which is a contradiction. To prove necessity, we must show that if $S = G(P)$, then $P$ is not stable, which follows from the fact that in this instance, every country can improve its security value by reforming $P$ so that it is the member of an alliance with a primary threat.

**Proof of Theorem 2:** We already know that an all-encompassing collective security system is stable (Result 2). So suppose $S$-$C$ is not empty.

(Sufficiency, part 2): To see that no $i \in C$ has an incentive to defect in the threat-counterthreat game, suppose $i$ switches from $C$ to $S$-$C$ to form $P' = (C$-$\{i\},S$-$C+\{i\})$. If $r(C$-$\{i\}) \geq r(S$-$C+\{i\})$, then since $i$ may tremble back to $C$-$\{i\}$, $s^{P'}_i < r^o_j \,\forall\, j \in S$-$C$. And by Lemmas 2 and 3, $s^{P'}_j < r^o_j \,\forall\, j \in C$-$\{j\}$. So $\forall\, j \in S$-$\{i\}$, $s^{P'}_i < r^o_j$. By Lemma 1, $\exists\, C' \in C^*$ with $i \notin C'$, so $s^{P'}_i < r^o_i$. Thus, $i$ will not defect.

Alternatively, if $r(C-\{i\}) < r(S-C+\{i\})$, then given the conditions of the theorem, $S-C+\{i\} \notin C^*$, and by Lemmas 2 and 3, $s_i^{P'} < r_i^o \; \forall \; i \in S-C+\{i\}$. Thus, $i$ does not defect from $C$.

(Sufficiency, part 3): Suppose $C \notin C^*$. We have two cases. First, if $r(S-C) + r_{\max[C]} \geq R/2$, then $C \notin C^*$ because condition ii or iii is violated. Lemma 2 implies that all $j \in C$ are vulnerable to a primary threat. Lemma 3 implies that all members of $C-\{j\}, j \neq \max[C]$, are vulnerable to a primary threat—but then $j$ is vulnerable as well since $S-C + \{\max[C]\} \in C^*$. Hence, $s_j^P < r_j^o \; \forall \; j \in C$. Since, by the assumption of the theorem, $C + \{i\} \notin C^* \; \forall \; i \in S-C$, any defection from $S-C$ can be punished and no $i \in S-C$ has an incentive to defect. Hence, $s_i^o = r_i^o \; \forall \; i \in S-C$. By Theorem 1, $(C,S-C)$ is stable. Alternatively, let $r(S-C) + r_{\max[C]} < R/2$, so unilateral defection from $C$ can form an advantaged coalition with $S-C$. After the elimination of $S-C$, by Lemmas 4 and 1, everyone is the target of some primary threat if alliances are renegotiated. Thus, $P$ is stable (by Theorem 1).

(Necessity, part 2): If $C \in C^*$, then $s_j^P < r_j^o \; \forall \; j \in S-C$. However, if there is an $i \in C$ such that $S-C \cup \{i\} \in C^*$, then by the "shaky hand assumption," $s_j^P < r_j^o(\forall \; j \in C-\{i\})$, in which case all members of $C - \{i\}$ are willing to join a primary threat against $i$ in the play of our threat-counterthreat game. And since $i$ can be threatened by a primary threat (Lemma 1), $s_i^P < r_i^o$, so $P$ is not stable (by Theorem 1).

(Necessity, part 3): If $C \notin C^*$ but if $\exists \; i \in S-C$ such that $i$ can form an advantaged coalition with $C$, then if there is some probability that $i$ will join in a primary threat against $S-C-\{i\}$ whenever $i$ is indifferent, $s_j^P < r_j^o \; \forall \; j \in S-C-\{i\}$. So as before, $s_j^P < r_j^o \; \forall \; j \in S-\{i\}$, in which case $\exists \; C \in C^*$ such that $i \notin C$ (Lemma 1) and $P$ is unstable (by Theorem 1).

**Proof of Remark 3:** Country 1 cannot be a balancer, otherwise, condition iii is violated. But if we attempt to make 2 a balancer, then condition ii is violated.

**Proof of Remark 4:** Let $P = (A_1, A_2, \{j\})$, where $A_1 + \{j\}, A_2 + \{j\} \in C^*$. Then $s_j^P < r_j^o \; \forall \; j \in A_1$ and $A_2$, in which case, from Lemma 1, $s_j^P < r_j^o \; \forall \; j$, and from Theorem 1, $P$ is unstable.

**Proof of Remark 5:** If $C \in P$ and $C \in C^*$, then for no $i \in C$ and $K \subseteq S-C$ is $K + \{i\} \in C^*$ and $i = \max[K + \{i\}]$. So no $i \in C$ has a positive incentive to defect, and $s_i^P = r_i^o \; \forall \; i \in C$. By Theorem 1, $P$ is stable.

**Proof of Remark 6:** Members of $C$ have two choices: eliminate or not eliminate $S-C$. If $C$ eliminates $S-C$, then since alliances must be renegotiated, by Lemma 1, $s_j^P < r_j^o \; \forall \; j \in C$ in the new system provided that $|C| \geq 4$ (since from Remark 2, every $i \notin C'$ for some $C' \in C^*$), which is what Lemma 4 establishes. If $C$ does not eliminate $S-C$, then since no $i \in C$ can form an advantaged coalition with members of $S-C$, $s_i^P = r^o \; \forall \; i \in C$. By Theorem 1, $P$ is stable.

**Proof of Remark 7:** No subset of $S-A$ can coalesce to form an advantaged coalition, so the members of $A$ cannot be threatened with a primary threat. Clearly now, no $i \in A \cap L_o$. For $i \in A \cap L$, if $i$ is the largest country in the new advantaged coalition, then $r(A-\{i\}) + r_i > R/2$, which contradicts the assumption that $r(S-A) > r(A)$.

**Proof of Theorem 3:** Let $P = (C,S-C)$. If $C \in C^*$, then Remark 5 establishes that $P$ is stable. If $C \notin C^*$, but if $r(S-C) + r_{\max[C]} > R/2$, then Remark 7 establishes

that $P$ is stable. And if $r(S\text{-}C) + r_{\max[C]} > R/2$, then $P$'s stability follows from Remark 6.

**Proof of Theorem 4:** A direct corollary of Theorem 1.

# References

ALTFELD, M. F. (1984) The Decision to Ally: A Theory and Test. *Western Political Quarterly* **37**:523–544.

AXELROD, R. (1984) *The Evolution of Cooperation.* New York: Basic Books.

BARON, D., AND J. FEREJOHN (1989) Bargaining in Legislatures. *American Political Science Review* **83**(4):1181–1206.

BENDOR, J., AND D. MOOKHERJEE (1987) Institutional Structure and the Logic of Ongoing Collective Action. *American Political Science Review* **81**(1):129–154.

BOULDING, K. (1968) *Beyond Economics: Essays on Society, Religion and Ethics.* Ann Arbor, MI: University of Michigan Press.

BUENO DE MESQUITA, B. (1981) *The War Trap.* New Haven, CT: Yale University Press.

FOX, A. B. (1959) *The Power of Small States.* Chicago: University of Chicago Press.

KAPLAN, M. A. (1957) *Systems and Process in International Relations.* New York: John Wiley and Sons.

KEOHANE, R. O. (1984) *After Hegemony.* Princeton, NJ: Princeton University Press.

LALMAN, D., AND D. NEWMAN (1991) Alliance Formation and National Security. *International Interactions* **16**(4):239–253.

LISKA, G. (1962) *Nations in Alliance: The Limits of Interdependence.* Baltimore, MD: Johns Hopkins University Press.

McGOWAN, P. J., AND R. M. ROOD (1975) Alliance Systems in Balance of Power Systems: Applying a Poisson Model to Nineteenth-Century Europe. *American Political Science Review* **49**(3):859–870.

MORGENTHAU, H. J. (1959) *Politics among Nations.* New York: Alfred A. Knopf.

MORROW, J. (1987) Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances. *American Journal of Political Science* **35**(4):904–933.

NIOU, E. M. S., AND P. C. ORDESHOOK (1990) Stability in Anarchic International Systems. *American Political Science Review* **84**(4):1208–1234.

NIOU, E. M. S., AND P. C. ORDESHOOK (1991) Realism versus Neoliberalism: A Formulation. *American Journal of Political Science,* **35**(2):481–511.

NIOU, E. M. S., AND P. C. ORDESHOOK (1994) Less Filling, Tastes Great: The Realist-neoLiberal Debate. *World Politics,* **46**(2):209–234.

ORDESHOOK, P. C. (1986) *Game Theory and Political Theory.* Cambridge: Cambridge University Press.

ORDESHOOK, P. C. (1993) "The Development of Contemporary Political Theory." In *Political Economy,* edited by W. Barnett, M. Hinich, and N. Schofield. Cambridge: Cambridge University Press.

RIKER, W. H. (1962) *Theory of Political Coalitions.* New Haven, CT: Yale University Press.

ROTHSTEIN, R. L. (1968) *Alliances and Small Powers.* New York: Columbia University Press.

SNYDER, G. H. (1990) Alliance Theory: A Neorealist First Cut. *Journal of International Affairs* (Winter):103–123.

SNYDER, G. H. (1991) Alliances, Balance, and Stability. *International Organization* **45**(1):121–142.

TAYLOR, M. (1976) *Anarchy and Cooperation.* New York: John Wiley and Sons.

TAYLOR, M. (1987) *The Possibility of Cooperation.* Cambridge: Cambridge University Press.

WALT, S. M. (1987) *The Origins of Alliances.* Ithaca, NY: Cornell University Press.

WARD, M. D. (1981) Research Gaps in Alliance Dynamics. Monograph. Series in World Affairs, vol. 19, book 1. Graduate School of International Studies, University of Denver, Denver, CO.

WRIGHT, Q. (1965) *A Study of War.* Chicago: University of Chicago Press.