DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

REALISM VERSUS NEOLIBERALISM: A FORMULATION

Emerson M.S. Niou and Peter C. Ordeshook

SOCIAL SCIENCE WORKING PAPER 709

December 1989

Abstract

Although the debate between realism and neoliberalism offers deep insights and raises fundamental questions into the nature of international systems, it also offers the confusion that accompanies imprecisely formulated concepts and an imperfect application of subsidiary ideas. Using a noncooperative extensive-form game to model anarchic international systems, this essay seeks to resolve that debate by restating it in a more explicit and deductive context. Arguing that collective security corresponds to the system envisioned by neoliberals, we begin by differentiating between balance of power and collective security in terms of the strategies that characterize the foreign policies of countries. Next, we establish that both balance of power and collective security can correspond to equilibria in our game. Arguments about goals and institutions are then recast in terms of the different properties of these equilibria. In particular, a balance of power equilibrium does not guarantee every country's security, so in it countries must be vigilant about their relative share of resources. A collective security equilibrium, on the other hand, ensures everyone's sovereignty, and thereby allows absolute resource maximization. Unlike a balance of power equilibrium, however, a collective security equilibrium is not strong and it is not necessarily perfect, so the institutional structures facilitating the realization of mutual gains from the variety of cooperative "subgames" characterizing the world economy play a critical role in establishing the stability of that equilibrium.

Realism versus Neoliberalism: A Formulation

For those who are not a party to it, the debate between realists and neoliberals seems a curious circus. While realists struggle with the specification of state goals and with alternative conceptualizations of balance of power, neoliberals offer vague admonitions that goals depend on context. Realists see cooperation as secondary to the conflictual processes of politics even though stability requires some minimal level of cooperation in order to maintain alliances, whereas neoliberals, aside from references to examples in game theory that do not necessarily model any specific international process, fail to define precisely the necessary and sufficient conditions for cooperation. Realists have yet to contend successfully with the task of defining and measuring power unambiguously. Neoliberals, on the other hand, offer the idea of regime as a pivotal concept but provide no theoretically meaningful definition in terms of the rationalist paradigm they profess to adopt. Neoliberalism argues that institutions matter because they somehow modify the actions of decision makers both directly by altering the costs and benefits of actions and indirectly by modifying goals, whereas realism has difficulty explaining the institutions and patterns of cooperation that characterize human affairs. Neoliberals seeks a synthesis, but the one they offer is little more than a collage of disconnected ideas in a framework that is neither rigorous nor deductive; realists see a synthesis as inherently dangerous to the extent that it detracts from the view that power is the ultimate guarantee of sovereignty.

This debate, while it raises fundamental questions, promises little hope of resolution as long as it is accompanied by imprecisely formulated concepts and an imperfect application of subsidiary ideas. For example, although neither side rejects the premise of goal directed, strategic behavior, important distinctions, such as between cooperation and coordination or direct and indirect utility, are ignored, and the understanding of game theory's implications, as well as of its limitations, is restricted to scenarios formed by special cases like the Prisoners' Dilemma. Application of the rationalist paradigm, nevertheless, offers one route out of this confusion. Formal theorists, of course, are more familiar with narrowly defined issues that lend themselves to precise formalization (e.g., voting on agendas, 2-candidate elections, majority-rule coalition processes, deterrence), whereas those seeking to make sense of macro-politics must grapple with the full complexity of human affairs. In an anarchic context in which rules themselves are endogenous, the application of such tools as game theory, which requires a highly structured specification of subject matter, seems to inhibit understanding. Nevertheless, if we want to avoid the tendency to inconclusive polemical argument, formalism is essential.

Our starting point is a topic that normally arises within the realist's domain -- the distinction between collective security and balance of power as ways to organize international affairs and as alternative premises for foreign policy. Morgenthau and Claude distinguish between these two systems, but based on experiences with the abortive League of Nations and with the unsuccessful attempt to resurrect the League in the form of the United Nations, they see collective security as an idea whose time is passed -- one that military technology renders infeasible as a guarantor of peace and international security. Thus, Claude (1962) asserts: "Ideological commitment to the doctrine of collective security is overshadowed by the conviction ... [that it] is not an appropriate response to the problem of managing international power relations in the present era" (p. 192) because it was "conceived with reference to a kind of war which must now be designated old-fashioned" (pp. 192-4). With respect to balance of power, on the other hand, "Twentieth-century efforts to replace the system have at most introduced modifications of its operative mechanisms; today, the balance of power system operates by default" (p. 93).

Despite its asserted demise, the idea of collective security is germane to the realist-neoliberal debate. First, a collective security system, characterized by the suppression of all threats against any state's existence, seems much like the alternative to power politics envisioned by liberals and neoliberals. If every country's sovereignty is assured, then collective security allows the maximization of absolute welfare gains that balance of power considerations obstruct. Second, collective security's failures are ostensibly the failures of institutions -- of institutions that ensure timely and appropriate responses to threats by one state against another. Thus, when evaluating the argument that stability arises as much from the complex nexus of interdependencies characterizing the contemporary world economy and the institutions that service this complexity as it does from the operation of an "invisible hand" formed by the pursuit of national power, we should ask why institutions are seen as facilitating stability in one context, but are presumed to be inadequate in the classical conceptualization of collective security.

With respect to the contribution that we might make to answering such questions, our objective is to specify more precisely the distinction between balance of power and collective security so that we can learn why one idea seems to have survived better than the other. In this way we hope to assess whether the realist-neoliberal debate arises because of fundamentally different theoretical views, or whether it can be interpreted merely as differences in emphasis on key parameters. Specifically, we offer an extensive-form, game theoretic model of anarchic systems in order to learn whether balance of power and collective security can each correspond to an equilibrium. Arguments about goals and institutions are then recast in terms of the different properties of these equilibria.

1. Problems of Conceptualization and Definition

To be certain that our analysis does not miss the point of the realist-neoliberal debate we begin with a reexamination of the issues in it, with the understanding that realists and neoliberals agree on several matters. First, both sides of the debate see international politics as anarchic, and agree that the institutional structure and cooperation and coordination found in international affairs are endogenous -- are the product of self-enforcing action on the part of a system's constituent parts. Second, both sides, although appreciating the imperatives of domestic politics, see the necessity, as an abstraction, for conceptualizing nation-states as unitary actors and as the primary constituent parts of international systems. Third, both see the necessity for abiding by the assumption that these unitary nation-states are strategic, goal directed decision-makers. Beyond this, however, several problems impede the development of an appropriate theory and the determination of how the ideas of collective security versus balance of power pertain to the realist-neoliberal debate.

Goals and Equilibria: Perhaps nowhere is disagreement greater than it is over the goals we should impute to nation-states. Neoliberals argue "that states seek to maximize their individual *absolute* gains and are indifferent to the gains achieved by others ... [whereas] realists find that states are *positional* ... and worry that their partners might gain more from cooperation than they do" (Grieco 1988:487). Considerable confusion, however, accompanies this disagreement. Gilpin (1982:7), for example, asserts that "International relations continue to be a recurring struggle for wealth and power among independent actors in a state of anarchy." Aside from measurement problems, power is relative and in constant supply, so if nations maximize it, then international affairs is constant sum. But, although wealth is a component of power, wealth can be defined in absolute terms, and its maximization does not preclude all members of a system from gaining simultaneously. Waltz (1959:198) tries to avoid ambiguity by arguing that "in a condition of anarchy, relative gain is more important than absolute gain," but confounding matters is the issue of sovereignty and its role in the goal of power maximization. If increasing one's power also increases one's vulnerability (as when any such increase, while decreasing a second state's power, also increases a third state's power to a proportionally greater degree), then what actions should a decision maker take? Thus, Waltz (1979:126) adds the qualification that "the first concern of states in not to maximize power but to maintain their position in the system." Neoliberalism, on the other hand, offers a vaguer assumption: "Under different systemic conditions states will define their self-interest differently ... where survival is at stake efforts to maintain autonomy may take precedence over all other activities, but where the environment is relatively benign energies will also be directed to fulfilling other goals." (Keohane 1989:62. emphasis added).

To resolve this debate over goals we must first distinguish between ultimate ends and intermediate objectives as they manifest themselves in particular circumstances, since even if someone is motivated primarily by measures of their personal well being, it is not unreasonable to postulate a strategic concern with relative variables. This part of the realist-neoliberal debate, then, can be reformulated in terms of the viability of alternative equilibria, because it is equilibria that define the context of action. Realists perceive a single compelling equilibrium -- a balance of power -- in which states must be vigilant about relative position. Neoliberals see the complex interdependencies of contemporary affairs as occasioning a different kind of equilibrium in which states can pursue pure welfare maximization because threats to sovereignty are somehow not part of the character of the strategies that sustain the equilibrium: "Among republics, at any rate, military threats may be insignificant, expanding the potential area for cooperation and reducing both the role of force and the emphasis states place on their relative power positions in the international system" (Keohane and Nye 1989:247).

Resolution of the realist-neoliberal debate over goals, then, requires that we ascertain whether or under what conditions equilibria of both types exist. Our approach is to suppose that the primary goal of states is to maintain their sovereignty, and, simplifying matters, to assume that total resources are constant, so that maximizing relative and absolute gains are equivalent. Then, in partial support of the realist view, we establish the existence of a balance of power equilibrium in which relative resource maximization is the sole sustainable subsidiary goal because sovereignty is not assured for all players. On the other hand, we accommodate the possibility that "states concerned with self-preservation do not seek to maximize their power when they are not in danger" (Keohane 1989:47) by establishing the existence of a second equilibrium -- a collective security equilibrium -- in which everyone's sovereignty is assured, regardless of how resources are distributed. We then use this fact to infer, in accordance with the neoliberal view, that collective security equilibria allow the pursuit of auxiliary policies that maximize absolute welfare.

Thus, by finding two equilibria that establish different conditions for the preservation of sovereignty, we give theoretical meaning to Keohane's previously cited assertion that "Under different systemic conditions states will define their self-interest differently ..." With the alternative goals of relative versus absolute resource gains thus rendered endogenous and dependent on whether a balance of power or a collective security equilibrium prevails, the realist-neoliberal debate can then focus on the ease with which one equilibrium as against the other can be sustained.

The Role of Institutions: Neoliberals might believe that imposing the constant-sum condition in our model either ends debate or biases it in favor of realism. Cooperation, though, is possible in nearly all contexts, so the relevance of institutions should not depend on whether international politics is constant or nonconstant sum. Indeed, as Keohane (1989:109) observes, it is not the case that "the whole process leading to the formation of a new international regime will yield overall welfare benefits. Outsiders may suffer ... [and some regimes such as alliances and cartels] are specifically designed to impose costs on them." So the only general limitation placed on the domain of his perspective by the argument that "sufficient complementarily or common interests exist so that agreements benefiting all essential regime members can be made" is that international politics concerns three or more players. Neoliberals who believe that it is necessary to limit their domain to nonconstant sum situations, then, are giving away too much. Conversely, realists who see cooperation as impossible are drawing an incorrect inference from the presumed anarchic, constant-sum system they believe is most relevant. Constant sum games preclude the possibility of mutual gains only if there are two players, whereas n-person constant sum games can promise important gains to cooperation and coordination -- the sole difference being whether the gains necessarily come from some other player or can come from nature as well.

An additional source of confusion involves the distinction between cooperation and coordination, which is illustrated by Keohane's (1984:51) statement that "we can evaluate the impact of cooperation by measuring the difference between the actual outcome and the situation that would have obtained in the absence of coordination..." Two or more people cooperate whenever they choose strategies to take mutual advantage of nature or someone else, and cooperation is self-enforcing whenever such strategies correspond to a [Nash, subgame perfect, perfect, sequential, etc.] equilibrium in the noncooperative representation of the situation being modeled. Coordination, on the other hand, refers to the ability of players to achieve a particular equilibrium. Most games -- especially the repeated games that most appropriately model international affairs -- have a great many equilibria, and, barring exogenous considerations, there need not be any guarantee that a particular equilibrium or that any equilibrium whatsoever will be realized. So, we must be concerned not only with ensuring that cooperation is an equilibrium, we must also be concerned that such an equilibrium prevails in the uncoordinated action of international systems.

Turning specifically to the role of institutions, we must contend with the fact that neoliberals are ambiguous about what it is that institutions do, or even for that matter what institutions are. For Keohane (1989:162) "'institution' may refer to a general pattern or categorization of activity or to a particular human-constructed arrangement, formally of informally organized." We can infer, then, that institutions include formal organizations such as the United Nations, formally stated agreements such as GATT, and alliances themselves. Insofar as function is concerned, we are told that institutions can be merely agents for facilitating cooperation ("International institutions have the potential for facilitating cooperation" (Keohane 1989:174)) or that they are an essential component of cooperation ("International institutions make it possible for states to take action that would otherwise be inconceivable" (Keohane, 1989:5)). In contrast, realism argues that "institutions are unable to mitigate anarchy's constraining effects on inter-state cooperation" (Grieco 1988:485), so, in this view, if institutions have any role, it is merely that of facilitating the formation and maintenance of alliances that countries use to pursue power or to maintain sovereignty.

However, we can discern explicit roles for institutions if both a balance of power and a collective security equilibrium exist. First, institutions must facilitate the coordination of action so that some equilibrium is achieved. It is perhaps this function that Keohane and Nye (1989:247) have in mind when they argue that interdependence and institutions give rise to "groups of states which [develop] reliable expectations of peaceful relations and thereby [overcome] the security dilemma that realists see as characterizing international politics." Second, institutions are required to "service" alliances in a balance of power, and to realize the welfare gains that states might pursue under collective security. Whether institutions are more critical for maintaining one equilibrium as against another, however, depends on the properties of those equilibria, which, along with existence, is the issue to which we now turn.

2. Anarchy, Balance of Power, and Collective Security

The preceding discussion defines the questions that our analysis must confront: Even if self-enforcing cooperative equilibria exist, are coordination mechanisms necessary to ensure that a particular equilibrium prevails? Is a balance-of-power equilibrium more stable than the equilibria envisioned by neoliberals? Is there a special role for institutions in achieving and maintaining equilibria other than the type envisioned by realists and do these alternative equilibria ensure stability in such a way as to allow the pursuit of subsidiary welfare maximizing programs? Indeed, given the emphasis that neoliberals place on modeling Prisoners' Dilemmas and the like, we also ask whether the successful pursuit of subsidiary welfare maximizing policies are essential to "servicing" such alternative equilibria?

We approach these questions by reexamining a game theoretic model of international conflict (Niou and Ordeshook 1989) that assumes that, conditional on maintaining their sovereignty, countries pursue a single transferable resource in constant supply. We impose this assumption because, first, we do not want to secure endogenously enforced cooperation simply by making the gains from cooperating great. Sustaining cooperation in the context of constant sum competition reveals more clearly the role that institutions can play in ameliorating international conflict. Second, we do not want to resolve the realist-neoliberal debate merely by supposing that sovereignty is not an issue; instead, we want to learn whether the type of cooperation that neoliberals envision in which states pursue auxiliary welfare-maximizing policies can arise in an otherwise constant-sum context.

Next, we differentiate between two forms of stability so as to accommodate the fact that the issue of sovereignty is qualitatively different from that of how nations contend with the ebb and flow of economic and military capabilities. Briefly, system-stability implies that all countries can ensure their sovereignty -- that no country will have its resources reduced to zero -- whereas resource-stability implies that no reallocation of resources occurs.[1] System-stability, then, is the sort of stability envisioned as the consequence of balance of power, whereas resource-stability, which implies system-stability, corresponds to the presumed consequence of an all-encompassing collective security equilibrium.

Our model of anarchy, now, supposes that the resources controlled by each country are the sole determinant of winning and losing, that threats and counter-threats are the mechanisms whereby countries secure resources from each other, that countries join coalitions because it is in their individual interest to do so, and that no exogenous constraints ameliorate conflict. Informally, the game we use to model anarchic systems proceeds as follows: A country, say i, randomly chosen by nature, either offers an initial threat or it "passes," where the threat is a new resource distribution and a proposed threatening coalition C. If i passes, nature selects another country. If i threatens, its partners in C must decided whether or not to participate in the threat. Only if all such partners choose to participate does i's threat call for a response by the threatened countries. Responses are of two types. First, each threatened country, taken in sequence, can offer a counter-threat, which is a new threat, and which, if unanimously accepted by the newly proposed coalition, cancels the original threat and becomes the new current threat. The second type of response is a proposal by one or more threatened countries to surrender resources to one or more members of the originally threatening coalition. If a transfer is accepted by the countries involved, it determines a new status quo, and the game proceeds as before.

Of course, this model, which we formalize in the next section, greatly abstracts from reality and describes anarchy in a highly structured way. Nevertheless, our game does incorporate the assumptions that "power" is the ultimate determinant of who can win and who can lose, that threats and counters are the mechanisms whereby countries secure resources from others, and that countries join coalitions because it is in their individual interest to do so and not because exogeneously imposed constraints induce cooperation.

To analyze such a game, however, we must constrain the strategies that we admit as sources of equilibria, because with threats and counters allowed to sequence forever, there are infinitely many strategies. Moreover, folk theorems tell us that nearly any "reasonable" outcome can be sustained as an equilibrium, given an appropriate specification of strategies. However, because of the complexity of many of these strategies, the presumption that all strategies are feasible strains credulity. Fortunately, the notions of balance of power and collective security point us in the direction of the most intuitively plausible possibilities.

To model balance of power, we suppose that strategies are stationary -- that each country makes the same choices whenever it encounters the same threat, and that it thereby ignores who made a threat or who agreed to participate in a threat.[2] In a balance of power system, then, "all states are potentially fit alliance partners; none is seen as much more evil than any other" (Jervis 1986:60). To model collective security, on the other hand, we look at simple punishment strategies, where punishment is directed against those who try to upset the status quo either by making a threat or by agreeing to participate in it. The strategies forming a collective security equilibrium, then, are not stationary because they posit the formation of specific alliances, depending on who defects from the status quo.

Equilibria supported by stationary strategies, then, are of special interest because they correspond to the equilibria that realists perceive as the primary model of international stability. So, if we find such equilibria and learn also that only countries controlling some critical relative level of resources can ensure their sovereignty in it, then in accordance with realism's arguments about goals, countries must be vigilant about relative gains and losses. If, however, there also exists an equilibrium supported by punishment strategies in which no country offers an initial threat, then, in accordance with the neoliberal argument, realization of this equilibrium renders the issue of sovereignty and relative position less salient. And if the benefits that accrue through free trade and the like require a non-conflictual world, and if these benefits disappear when agreements to achieve them are disrupted by competition over relative position, then the issue confronting us, and which bears directly on the realist-neoliberal debate, is whether such an equilibrium can be sustained endogenously or whether its enforcement requires attending to the nurturing of these benefits. If this collective security equilibrium is as attractive as a balance of power equilibrium, then the sole critical role of institutions is to facilitate coordination so as to ensure that one equilibria in particular is achieved. But if such equilibria require "nurturing," then institutions must, in addition to coordinating action, ensure that this nurturing occurs.

3. A Formal Model and a Theorem About Balance of Power

To present our model formally requires some notation. Briefly, we let $(S,\mathbf{r}^o)$ denote a system, where $S = \{1,2,...,n\}$ is the set of countries and $\mathbf{r}^o = (r^o_1, r^o_2,...,r^o_n)$ is the distribution of resources across S. For convenience, we let $r^o_1 > r^o_2 > ... > r^o_n > 0$. If r(C) is the sum of resources controlled by the members of the coalition C, and if R denotes the total resources in the system, then the coalition C is winning if $r(C) > R/2$, it is losing if $r(C) < R/2$, and it is minimal winning if, for all i in C, C-{i} is losing. Countries in S who are part of at least one minimal winning coalition are <u>essential</u>; otherwise they are <u>inessential</u>. If $r^o_i > R/2$, i is <u>predominant</u> -- it is winning against all other countries and it can incorporate their resources at will -- so every country has an incentive to avoid the possibility that some other country becomes predominant. If $r^o_i = R/2$, i is <u>near-predominant</u>. With respect to the status quo distribution $\mathbf{r}^o$, max[C] denotes the country in C with the greatest total of resources. Finally, the game Γ that we use to model conflict proceeds as follows:

(1) Nature randomly selects i in S;

(2) i offers a threat $(\mathbf{r},C)$, i in C, or i passes. If i passes, we return to (1). If i threatens and if $r_j > r^o_j$, j is an active member of C, whereas $r_j = r^o_j$ for the passive members of C.

(3) The members of C-{i} simultaneously choose between approving or rejecting (r,C). If no j in C-{i} rejects, then (r,C) becomes the current threat; otherwise, return to (1).

(4) With (r,C) the current threat, nature randomly orders the members of S-C.

(5) With m in S-C offering the first counter, a counter takes one of two forms: a new threat, (r',C'), m in C'; or a resource transfer from S-C to one or more members of C. We denote those party to the transfer by C'.

(6) The members of C'-{m} simultaneously choose between approving or rejecting (r',C'). If a counter which is itself a threat is approved unanimously, it becomes the new current threat, and we return to (4). If one or more members of C'-{m} reject the counter, we select the next threatened country in the order chosen by nature and we return to (5). For counters that are resource-transfers, unanimous acceptance renders the transfer the new status quo and we return to (1).

(7) If the counter of the last threatened country is rejected, the resource distribution of the current threat becomes the status quo and we return to (1).

Requiring in (3) and (6) that members of C-{i} and C'-{m} act simultaneously rather than sequentially is arbitrary, so we show later that this assumption is important to the relative stability of collective security equilibria, but not to balance of power equilibria. However, the feature of Γ that we want to emphasize is that it is a recursive game, because it allows for the possibility that threats and counters, as well as resource reallocations, continue in sequence forever. Thus, the analysis of such games cannot proceed in the same way that we treat, say, finite voting agendas or infinitely repeated Prisoners' Dilemmas. With a finite agenda, we can "work backwards" from the game's terminal nodes to deduce equilibrium choices at earlier nodes, but here some branches of the extensive form do not lead to terminal nodes. Similarly, the infinitely repeated Prisoners' Dilemma is effectively rendered finite by supposing that the stream of payoffs accruing to the players are subject to a time discount. But in our game there is no "stream" of payoffs, and, in fact, all action may simply take place in the minds of participants as they contemplate possibilities, in which case there is no natural discount that can be applied to evaluate alternative resource distributions.

Wagner (1986:551) outlines our game's correct treatment: "the basic question that concerns us is whether states will act so as to eliminate other states. If one state is eliminated from a four-actor game, for example, the result is to precipitate a three-actor game. If a value can be assigned to such a subgame for each player, it is possible to determine whether any players have an incentive to eliminate other players." We proceed, then, by pretending that Γ is finite by supposing that we know the consequences of all branches in its extensive form. After postulating these consequences, an equilibrium is characterized by strategies in which no one has an incentive to defect unilaterally to any choice not dictated by that player's strategy, and the postulated consequences are consistent in that they are "self-fulfilling prophesies" -- the subgame perfect choices they imply must yield those consequences.

Because our discussion of collective security builds on an analysis of balance of power, we focus initially on stationary strategies. Supposing, then, that all countries respond identically to ($\mathbf{r}$,C) regardless of who makes this threat and regardless of the histories associated with each player, we let Γ_r denote the game that follows the threat of ($\mathbf{r}$,C) and its acceptance. Suppose country i associates the value $v_i(\Gamma_r)$ with playing that sub-game, then the vector $\mathbf{v}(\Gamma_r) = (v_1(\Gamma_r),...,v_n(\Gamma_r))$ -- the continuation value of Γ_r -- specifies what the countries believe follows from the approval of ($\mathbf{r}$,C). Thus, $v_i(\Gamma_r)$, when compared against whatever follows if ($\mathbf{r}$,C) is rejected, determines i's preference for acceptance or rejection of ($\mathbf{r}$,C) or for making this threat in the first place. Once values for all threats are specified we can assume that the acceptance of a threat or counter is a terminal node with its continuation value as the "final outcome." We then analyze Γ like a finite extensive-form game of complete information and we deduce sub-game perfect equilibrium strategies by working backwards from the terminal nodes in the same way we treat finite agendas in majority voting games -- we deduce what each country ought to do any time it must choose a threat, a counter, or accepting or rejecting a threat or counter.

The particular difficulty here is that continuation values are endogenous to the situation -- what follows depends on strategies, and strategies depend on what follows. An equilibrium, then, is a set of continuation values -- one for each threat -- and a set of strategies for each country such that these values and strategies are consistent. Thus, in equilibrium, the choices that the continuation values imply -- the strategies that are a subgame perfect equilibrium given the continuation values -- must, in turn, imply those continuation values.

The way in which we form continuation values and establish consistency is to first offer a 2-way classification of threats, where all the members of one class are associated with one type of continuation value and all remaining threats are associated with a second type of continuation value. We proceed by first identifying Type 1 threats, which satisfies four conditions: (i) the largest member of the threatening coalition C (max[C]) can in principle at least coalesce with the threatened countries S-C to form a new winning coalition; (ii) the threatened countries are threatened with elimination; (iii) the threat proposes to render the largest country in C near-predominant; and (iv) there is no other winning coalition C' such that C' and C share only one member where that member is the largest country in C' but not in C.[3] Next, we define primary threats, which are particular Type 1 threats. Letting L be those countries in S who can be the largest member of a minimal winning coalition, the set of primary threats satisfy two conditions: First for no two primary threats ($\mathbf{r}$,C) and ($\mathbf{r}$',C') is it the case that C and C' have a unique common member corresponding to the largest member of C and C' such that the remaining members of C and C' are in S-L. This requirement, after continuation values are assigned, ensures that no primary threat is an effective counter against another such threat. Second, the set of primary threats is maximal in the sense that there not exist a winning coalition with a Type 1 threat that can be added to this set so as not to violate its first characteristic, which assures us that nothing is precluded from this set that cannot be effectively countered.[4]

Assuming that countries abide by stationary strategies, the continuation values that can be shown to be consistent with subgame perfection imply the following: if (**r**,C) is not a primary threat, then no country is assured of becoming near-predominant, members of S-L do not gain resources, and no member of $C \cap L$ gains resources if it is smaller than the largest threatened country. If (**r**,C) is a primary threat, then the largest threatening country becomes near-predominant at the expense of threatened countries, and the remaining members of the threatening coalition retain the resources they currently control.[5]

The additional assumptions we require to establish this consistency are, **first**, that a near-predominant country can take advantage of conflicts among other countries so as to become predominant. Hence, systems with a near-predominant country are "frozen." **Second**, to accommodate war costs, we assume that if i can become near-predominant by implementing a threat or by having countries transfer enough resources to render i near-predominant, i prefers the transfer. So if max[C] denotes the largest member of C with respect to the current, status quo, resource distribution, then the system is frozen if S-C offers to render max[C] near-predominant. Clearly, if S-C prefers freezing the system, it should transfer to max[C], since this choice minimizes the resources that S-C must surrender. And max[C] accepts the offer: Because attempts to secure more than R/2 will be blocked, securing R/2 by transfer is max[C]'s most preferred feasible outcome. **Finally**, because elimination is different from merely having a small amount of resources, we suppose that countries are risk averse with respect to the possibility of elimination -- a country attributes a value of $-\infty$ to any outcome in which it is eliminated. These assumptions and the preceding assignment of continuation values allow us to establish the following important result:

If all countries are essential, and if all countries abide by stationary strategies, then $(S,\mathbf{r}^{o})$ *is necessarily system-stable* (Niou and Ordeshook 1989).

Furthermore, if we allow countries to make sequential threats (i.e., i proposes that C threatens $j \in$ S-C, then $k \in$ S-C, etc.), then inessential countries as well as smaller essential ones will be unable to assure their sovereignty in any n-country system.

These facts establish the existence of a balance of power equilibrium ensuring the sovereignty of "larger" states, but not of smaller ones. Hence, countries must be vigilant about their relative share of resources. Thus, aside from the institutional mechanisms that may be required to coordinate strategies and to service alliances, we do not find in this equilibrium much room for those international institutions that allow countries to pursue welfare maximizing policies without regard for the gains that others might realize. Moreover, in addition to being perfect, and thus self-enforcing, this equilibrium is also strong in the following sense: The "largest" countries (those in L), if given the opportunity to make a threat, prefer to do so because they gain and thereby avoid the possibility of loss, whereas smaller countries, although unable to gain by participating in a threat, avoid the possibility of losses by doing so. So if i believes that all others in S will choose their equilibrium strategies, then i has a positive incentive to make or to agree to primary threats that include it in the threatening coalition, since not doing so diminishes i's utility. A balance of power equilibrium, then, is attractive.

4. The **Example of 4-Country Systems**

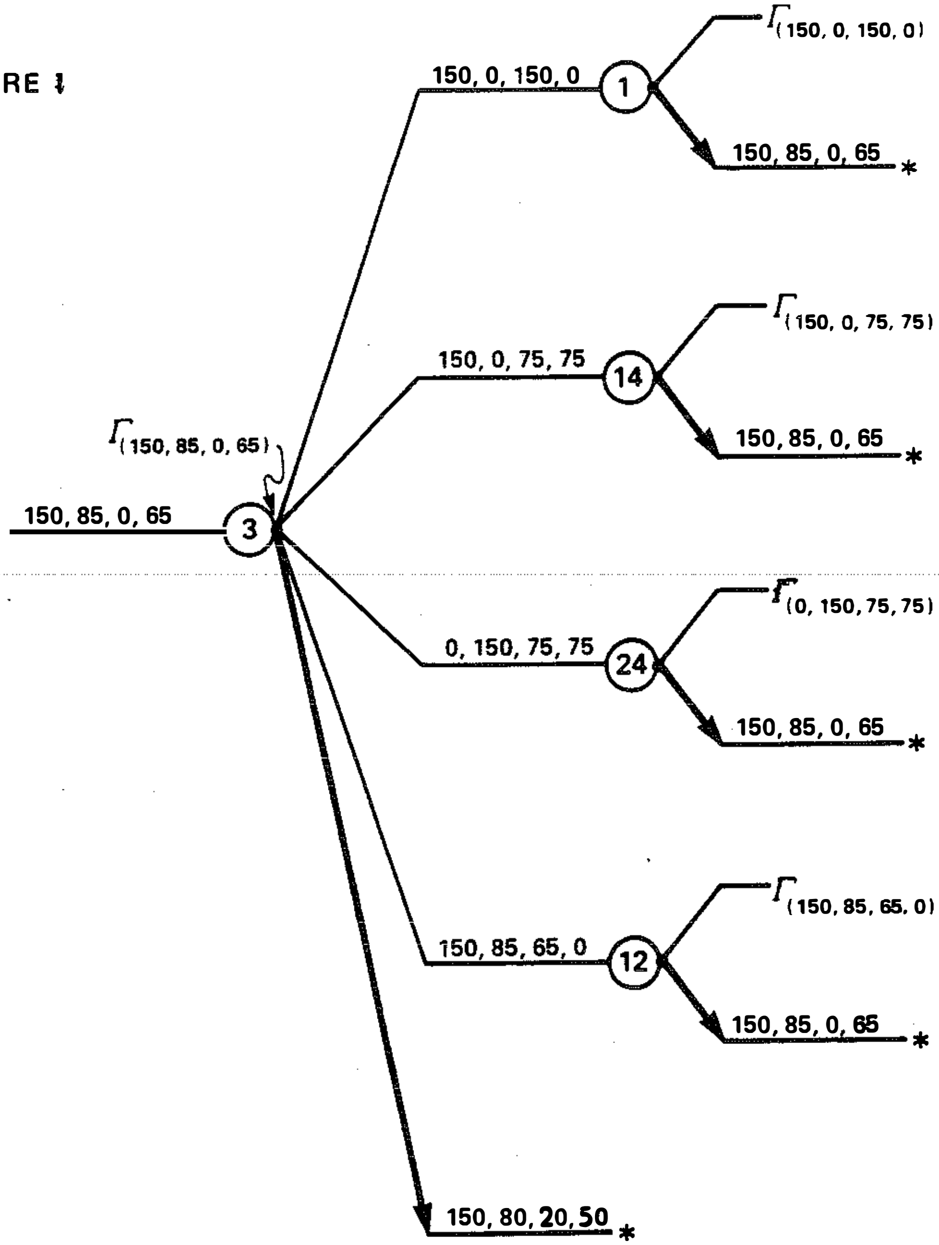
Before we proceed to an analysis of collective security equilibria, let us look at the nature of a balance of power equilibrium in two alternative 4-country examples -- one for which $\mathbf{r}^o = (110,80,60,50)$ and a second in which $\mathbf{r}^o = (100,95,75,30)$ -- where the critical difference between these examples is that country 4 is essential in the first but not in the second example. What we propose to show is that in a balance of power equilibrium, the first distribution occasions system-stability whereas country 4 is eliminated if the second distribution applies.

Although our theoretical results apply even when we allow all possible threats, we limit discussion here for purposes of an example to specific threats and counters. Specifically, consider the following threats and conjectured continuation values (where $a < 150$, $b < 150$, $c < 60$, and $d < 50$, and where the values of a, b, c, and d depend on the threat in question), which consists of a sample of non-primary threats and all primary threats for this game:

non-primary threats	primary threats
$v(\Gamma_{(150,150,0,0)}) = (a,b,60,50)$	$v(\Gamma_{(150,85,65,0)}) = (150,80,60,10)$
$v(\Gamma_{(150,0,150,0)}) = (a,b,c,50)$	$v(\Gamma_{(150,85,0,65)}) = (150,80,20,50)$
$v(\Gamma_{(150,0,0,150)}) = (a,b,60,d)$	$v(\Gamma_{(150,0,75,75)}) = (150,40,60,50)$
	$v(\Gamma_{(0,150,75,75)}) = (40,150,60,50)$

Referring to Figure 1, suppose country 1 proposes the primary threats, (150,85,0,65), and that countries 2 and 4 accept. Country 3 has five alternative actions, but: (1) if 3 counters with (150,0,150,0), then 1 rejects since, by assumption, $v_1(\Gamma_{(150,0,150,0)}) = a < 150$ and rejection implements a threat giving 150 to 1; (2) if 3 proposes (150,0,75,75), then 4 rejects since $v_4(\Gamma_{(150,0,75,75)}) = 50 < 65$; (3) if 3 proposes (0,150,75,75), then 4 rejects for the same reason; and (4) if 3 proposes (150,85,65,0), then 2 rejects since $v_2(\Gamma_{(150,85,65,0)}) = 80 < 85$. Hence, the only alternative available to 3 is to transfer resources to 1, in which case, as conjectured, $v(\Gamma_{(150,85,65,0)}) = (150,80,20,50)$. An equivalent analysis holds for the threats

FIGURE 1



Γ(150, 0, 150, 0)
150, 0, 150, 0
1
150, 85, 0, 65 *
Γ(150, 0, 75, 75)
150, 0, 75, 75
14
150, 85, 0, 65 *
Γ(150, 85, 0, 65)
150, 85, 0, 65
3
Γ(0, 150, 75, 75)
0, 150, 75, 75
24
150, 85, 0, 65 *
Γ(150, 85, 65, 0)
150, 85, 65, 0
12
150, 85, 0, 65 *
150, 80, 20, 50 *

corresponding to the distributions (150,85,65,0), (150,0,75,75), and (0,150,75,75).

The situation is more complicated if two countries are threatened, as when {1,2} threatens {3,4} with (150,150,0,0). Figure 2 shows the part of the extensive form that pertains after such a threat is accepted, and after nature selects 3 to offer the first counter (the situation is symmetric if 4 counters first). As before, 3 has four counter-threats, but their rejection, rather than leading to the implementation of the threat, gives 4 an opportunity to offer a counter. Since 4's options are independent of 3's choice, Figure 2 portrays only one instance of 4's decision. Working backwards on the extensive form so as to identify subgame-perfect equilibrium strategies, and looking at 4's decision, we see that: (1) if 4 offers (150,0,0,150), 1 is certain to reject since $v_1(\Gamma_{(150,0,0,150)}) = a < 150$; (2) if 4 offers (150,85,0,65), then 2 rejects since $v_2(\Gamma_{(150,85,0,65)}) = 80 < 150$; and (3) if 4 offers (150,0,75,75) or (0,150,75,75), then {1,3} and {2,3}, respectively, accept. Since 4 is indifferent between offering (150,0,75,75) and (0,150,75,75), we suppose that 4 chooses one or the other with equal probability. In any event, there is no reason for 4 to consider a transfer to 1.

Looking now at 3's decision in Figure 2, because rejection of any counter by 3 yields a lottery between $v(\Gamma_{(150,0,75,75)}) = (150,40,60,50)$ and $v(\Gamma_{(0,150,75,75)}) = (40,150,60,50)$, if 1 prefers $v_1(150,0,150,0)$ to a lottery between 150 and 40, 3 does not counter with (150,0,150,0) -- $v_3(\Gamma_{(150,0,150,0)}) = c < r^o_3$ and, as we see shortly, 3 has better choices. But if 1 prefers the lottery, then the counter (150,0,150,0) is equivalent to the counter (150,85,65,0), since 2 rejects in favor of the lottery. Finally, the counters (150,0,75,75) and (0,150,75,75) -- both of which yield 3 a payoff of 60 -- are accepted by {1,4} and {2,4}, respectively. Now, though, we can introduce an assumption that does not affect our conclusions here, but which simplifies the generalization of our analysis; namely, suppose $i \in S\text{-}C$, to counter $(\mathbf{r},C)$, chooses $(\mathbf{r}',C')$ such that $S\text{-}C \subset C'$ whenever it is otherwise indifferent. The rationale for this assumption is that the threat against S-C makes the formation of S-C less costly (also, if i is indifferent, such a choice can characterize an equilibrium strategy since i has no positive incentive to choose differently). Presently, this assumption implies that 3 counters with (150,0,75,75) or (0,150,75,75). Since an identical argument holds if 4 counters first, $v(\Gamma_{(150,150,0,0)})$ is a lottery between (150,40,60,50) and (40,150,60,50) as originally asserted.

We can repeat this analysis for the other listed threats. So, establishing in this way that the posited continuation values are consistent with subgame perfection, we must now identify a symmetric stationary equilibrium for the full game. Consider the following two characterizations of equilibria with respect to the game's first move:

Case (1): no player makes or accepts an initial threat unless such action promises a gain;

Case (2): players make or accept initial threats if doing so promises no loss.

For case (1), 1 and 2 prefer to coalesce with 3 and 4, but, under the assumption of the presumed equilibrium, 3 and 4 do not accept initial offers since no offer yields them a gain.

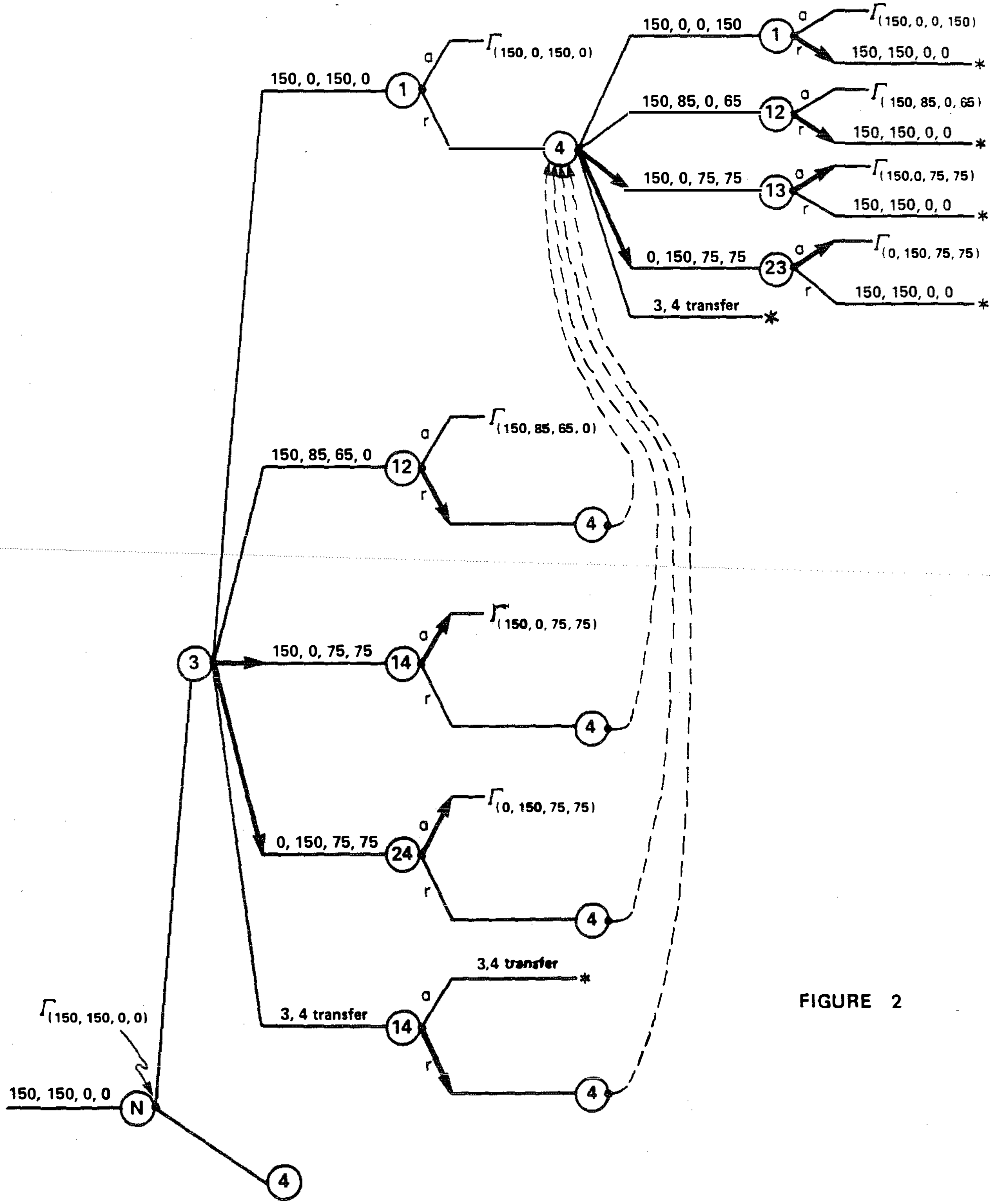

150, 150, 0, 0
N
Γ(150, 150, 0, 0)
4
3
150, 0, 150, 0
1
a
Γ(150, 0, 150, 0)
r
4
150, 0, 0, 150
1
Γ(150, 0, 0, 150)
150, 150, 0, 0
150, 85, 0, 65
12
Γ(150, 85, 0, 65)
150, 150, 0, 0
150, 0, 75, 75
13
Γ(150,0, 75, 75)
150, 150, 0, 0
0, 150, 75, 75
23
Γ(0, 150, 75, 75)
150, 150, 0, 0
3, 4 transfer
150, 85, 65, 0
12
Γ(150, 85, 65, 0)
4
150, 0, 75, 75
14
Γ(150, 0, 75, 75)
4
0, 150, 75, 75
24
Γ(0, 150, 75, 75)
4
3, 4 transfer
14
3,4 transfer
4

FIGURE 2

Moreover, neither 3 nor 4 gains by defecting unilaterally from its strategy. Hence, a situation in which no threats are made is an equilibrium. However, case (1) cannot be to a perfect equilibrium:[6] If there is a chance that 4 will accept a threat in which it does not lose, then 3 should not forego participating in threats that freeze the system. This argument bears on case (2). If everyone accepts threats in which they do not lose, then no one has an incentive to switch to a strategy of accepting or making threats only if it gains. Thus, given the limitations on threats we impose, our 4-country example, assuming stationary strategies, is necessarily system-stable, but only resource-instability corresponds to a perfect equilibrium.

Now consider the second example in which $\mathbf{r}^o = (100,95,75,30)$, so country 4 cannot form a minimal winning coalition. It is straightforward to show, now, that threats by 1 and 2 against 4 in the form (110,105,85,0) eliminate 4; however, if, for example, 1 is selected to make the first move, it should threaten (150,0,150,0) or (150,0,120,30), in which case 2 and 4 or 2 alone transfer to 1; but in either case 4 is not eliminated. On the other hand, suppose country 3 moves first and threatens (110,105,85,0), i.e., 3 proposes to eliminate 4 and distribute its resources evenly among 1, 2, and 3. Referring to Figure 3, if 1 and 2 accept, 4 cannot respond with a threat such as $(150,0,150-\delta,\delta)$ or $(0,150,150-\delta,\delta)$ that attracts 3, because 3 knows that such threats merely lead to transfers between 1 and 2 whereas rejection increases 3's resources to 85. On the other hand, if 4 responds with a threat against 3, such as $(150,150-\delta,0,\delta)$, then 3 can counter either with a threat against 1 or a threat against 2. Thus, 1 or 2 reject 4's counter. Thus, 4 must transfer to one country or another. If 4 transfers all of its resources, it is, of course, eliminated; on the other hand, if it only offers 1, 2, or 3 something more than ten units of resources, then this action merely opens the door to the possibility that 4's resources will incrementaly approach zero. Thus, ({1,2,3,4},(100,95,75,30)) is not system-stable, because country 4 can be eliminated.

5. Collective Security

Turning now to the notion of collective security, modeled by punishment strategies designed to preserve the status quo, we first simplify the extensive game-forms with which we must contend, by modifying the assumption that nature chooses randomly from S after a threat is rejected -- the assumption that we return to step (1) if a threat is not unanimously accepted in step (3). Without altering our conclusions about the nature of balance of power equilibria, we let D denote the countries that are the potential targets of punishment as determined by the strategy specified below, and we assume that nature chooses from the set S-D. That is, nature chooses from those players who will administer a punishment (if a point in the game is ever reached in which D = S, then suppose that nature thereafter chooses from S with uniform probability).

Our next step is to isolate a particular type of punishment strategy that matches the simplicity of stationary strategies, because we do not want to confront the objection that

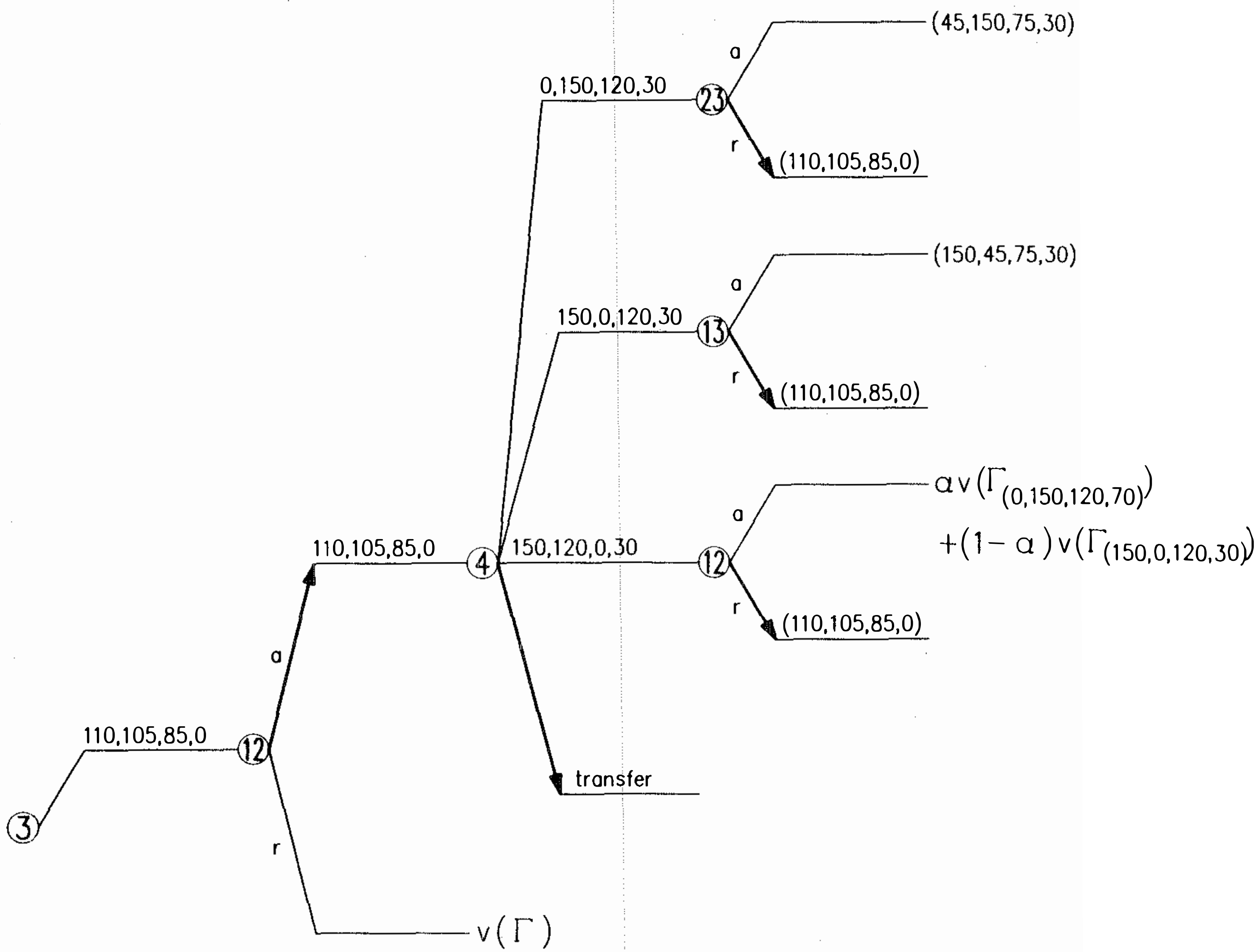

110,105,85,0
a
r
v(Γ)
110,105,85,0
transfer
0,150,120,30
150,0,120,30
150,120,0,30
a
r
(45,150,75,30)
(110,105,85,0)
a
r
(150,45,75,30)
(110,105,85,0)
a
r
αv(Γ(0,150,120,70))
+(1− α)v(Γ(150,0,120,30))
(110,105,85,0)

Figure 3

balance of power equilibria are somehow easier to compute and realize. Hence, we restrict our attention to the following characterization:

(a) No country proposes an initial threat;

(b) No country accepts an initial threat if one is offered;

(c) Threats are directed against one or more defectors;

(d) Countries accept threats that are punishments;

(e) Whenever any threat is accepted, all countries use stationary strategies thereafter.

Players defecting from (a)-(d) are added to D and are thereafter subject to punishment.

To proceed we must now modify our notation for continuation values. Assume that subgames begin after countries pass, or after all relevant countries choose between accepting and rejecting the last offered threat, we let $v^D(\Gamma_r) = (v^D_1(\Gamma_r),...,v^D_n(\Gamma_r))$ denote the continuation value of the subgame beginning after the threat (r,C) is made and accepted, where D is the current set of defectors. If there is no current threat (e.g., a country passes), then $v^D(\Gamma_o)$ denotes the corresponding continuation value. The general form we assume for $v^D_j(\Gamma_r)$ and $v^D_j(\Gamma_o)$ is as follows:

$$v^D_j(\Gamma_r) = v_j(\Gamma_r) \tag{1}$$

$$v^D_j(\Gamma_o) \begin{cases} = r^o_j \text{ if } j \notin D \text{ and } j \neq \max[S\text{-}D] \\ = a, \text{ where } r^o_j < a \leq R/2, \text{ if } j \notin D \text{ and } j = \max[S\text{-}D] \\ < r^o_j \text{ otherwise.} \end{cases} \tag{2}$$

Expression (1) states that, in accordance with (e), if there is a standing threat, then all countries play stationary strategies thereafter and continuation values are as specified in footnote 5. On the other hand, if there is no current threat, the specification of $v^D_j(\Gamma_o)$ in expression (2) states that if j is not in D -- if the presumed equilibrium does not target j for punishment -- and if j is not the largest member of S-D, then j merely retains its current resource allocation. If j is not a target for punishment but if it is the largest member of S-D, j receives a transfer that either renders j near-predominant or which eliminates S-C.[7] Finally, if $j \in D$, then j's expected payoff is less than its current resource holdings.

These definitions yield the following general result about collective security:[8]

If $|S| > 3$, and if there are four or more essential countries, the strategy described in (a)-(e) yields a strong subgame perfect equilibrium in which no country makes an initial threat and no country is eliminated, regardless of whether S contains inessential countries.[9]

We limit the domain of this result to systems with four or more essential countries, because otherwise, although a collective security equilibrium exists, it is neither strong nor subgame perfect (Niou and Ordeshook 1989). The weakness of such equilibria here arises because if there are only three essential countries, only 2-country coalitions can form primary threats. For example, if $r^o = (120,100,80)$ and if country 3 defects by proposing the primary threat (150,0,150), then 1 has a positive incentive to accept this threat -- in accordance with (e) and with the postulated continuation values for primary threats when stationary strategies are used, the eventual outcome is (150,70,80). Country 3, of course, is indifferent between threatening (150,0,150) or passing, which renders the equilibrium weak rather than strong, and 1's willingness to accept the threat rather than punish 3 precludes subgame perfection.

In larger systems, on the other hand, the lemma in footnote 8 establishes that more than one country must accept a proposed primary threat, and this fact renders a collective security equilibrium both strong and subgame perfect. To illustrate, suppose again that $r^o = (110,80,60,50)$, and referring to Figure 4, if 3 and 4 accept 2's offer, then, in accordance with (e), (40,150,60,50) prevails (the dashed lines denote information sets so that, for example, when 4 chooses between accept and reject, 4 does not know 3's choice). But if 3 alone accepts 2's threat, since $D = \{2,3\}$, nature chooses between 1 and 4 to make the next move. Suppose 1 is chosen. If, in accordance with (c), 1 proposes (150,85,0,65) to punish 3, if 2 and 4 accept, and if the preceding continuation values hold, then (150,80,20,50) prevails. Figure 4 shows, moreover, that 2 and 4 have no incentive to reject 1's proposal. Similarly, if 1 punishes 2 by proposing (150,0,75,75), then (150,40,60,50) prevails. Because 150 is 1's greatest feasible payoff, it has no positive incentive to do anything other than propose to punish 3 or 2 (1 will not propose to punish both simultaneously because (150,0,0,150) is not a primary threat and yields 1 an expected return of less than 150). A parallel argument shows that if 4 moves instead of 1, 4 has no incentive to offer different threats. Thus, assuming that 1 and 4 choose between their two punishment strategies with probability α and $1-\alpha$, then $v^{\{2,3\}}(\Gamma_o) = \alpha(150,80,20,50) + (1-\alpha)(150,40,60,50)$, which is consistent with the assumption in expression (2) that $v^{\{2,3\}}_{\max[S-D]=1}(\Gamma_o) = R/2$, $v^{\{2,3\}}_2(\Gamma_o) < r^o_2$, $v^{\{2,3\}}_3(\Gamma_o) < r^o_3$, and $v^{\{2,3\}}_4(\Gamma_o) = r^o_4$. A parallel argument holds if 4 defects by accepting 2's offer, so consider the final possibility -- that 3 and 4 reject 2's offer. At this point, $D = \{2\}$, so nature chooses 1, 3 or 4 for the next move. As Figure 4 reveals, however, none of these countries has a positive incentive to propose anything other than (150,0,75,75). Thus, neither 3 nor 4 has any unilateral incentive to defect from (b), in which case 2 has no incentive to defect from (a).

Similar analyses of the remaining initial threats establishes the overall consistency of the proposed continuation values, and in this way we see that the proposed punishment strategies, in conjunction with the postulated continuation values, supports a subgame perfect equilibrium in which no country makes an initial threat. It is also straightforward to see that this equilibrium is perfect.

Now let $r^o = (100,95,75,30)$ so that 4 is inessential, and suppose 3 defects from (a) by proposing the same threat against 4 that eliminates 4 in a balance of power system,

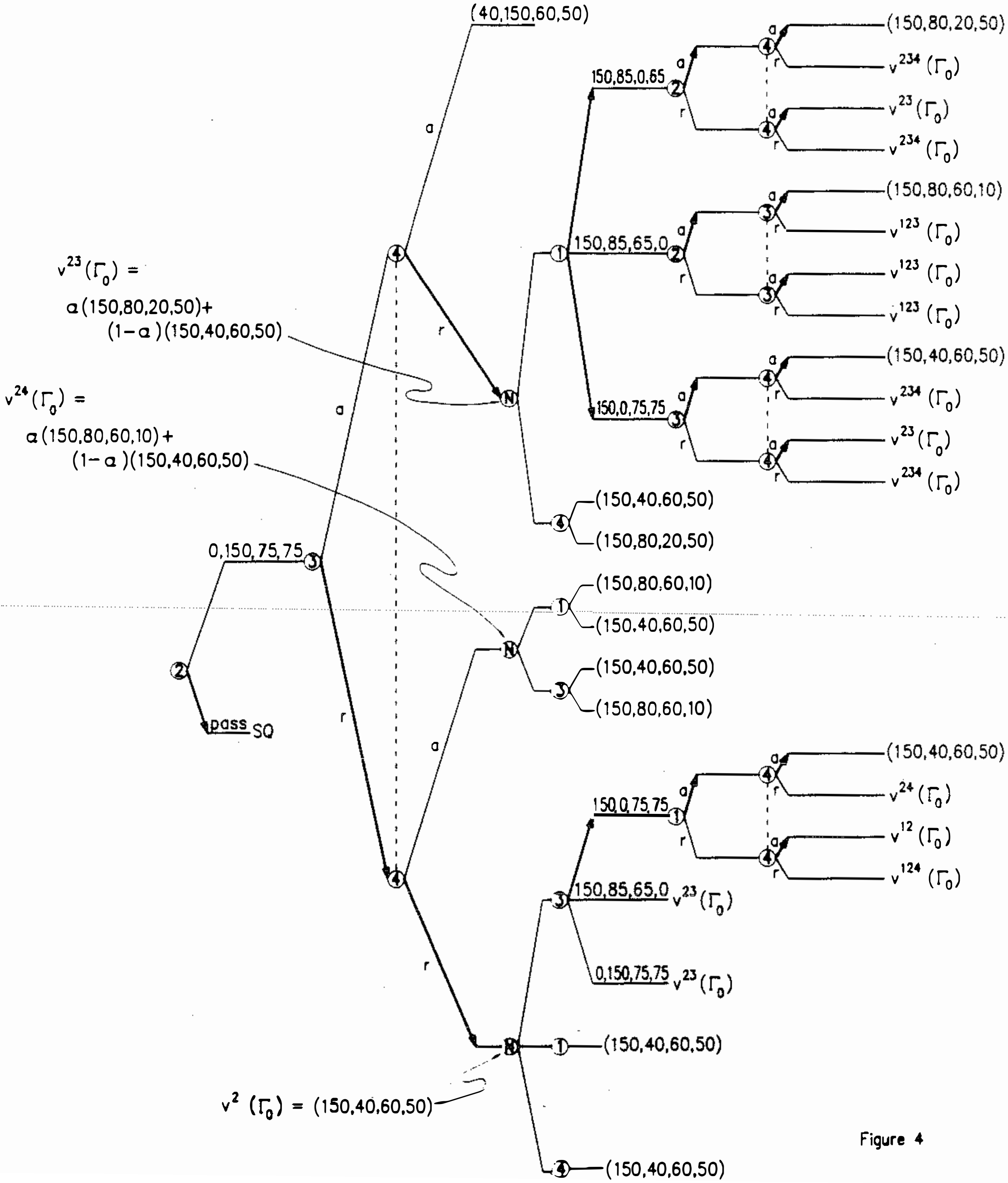

(40,150,60,50)
(150,80,20,50)
$v^{234}(\Gamma_0)$
$v^{23}(\Gamma_0)$
$v^{234}(\Gamma_0)$
150,85,0,65
(150,80,60,10)
$v^{123}(\Gamma_0)$
$v^{123}(\Gamma_0)$
$v^{123}(\Gamma_0)$
150,85,65,0
$v^{23}(\Gamma_0)$ =
a(150,80,20,50)+
(1−a)(150,40,60,50)
(150,40,60,50)
$v^{234}(\Gamma_0)$
$v^{23}(\Gamma_0)$
$v^{234}(\Gamma_0)$
150,0,75,75
$v^{24}(\Gamma_0)$ =
a(150,80,60,10)+
(1−a)(150,40,60,50)
(150,40,60,50)
(150,80,20,50)
0,150,75,75
(150,80,60,10)
(150,40,60,50)
(150,40,60,50)
(150,80,60,10)
pass SQ
(150,40,60,50)
$v^{24}(\Gamma_0)$
150,0,75,75
$v^{12}(\Gamma_0)$
$v^{124}(\Gamma_0)$
150,85,65,0 $v^{23}(\Gamma_0)$
0,150,75,75 $v^{23}(\Gamma_0)$
(150,40,60,50)
$v^{2}(\Gamma_0)$ = (150,40,60,50)
(150,40,60,50)

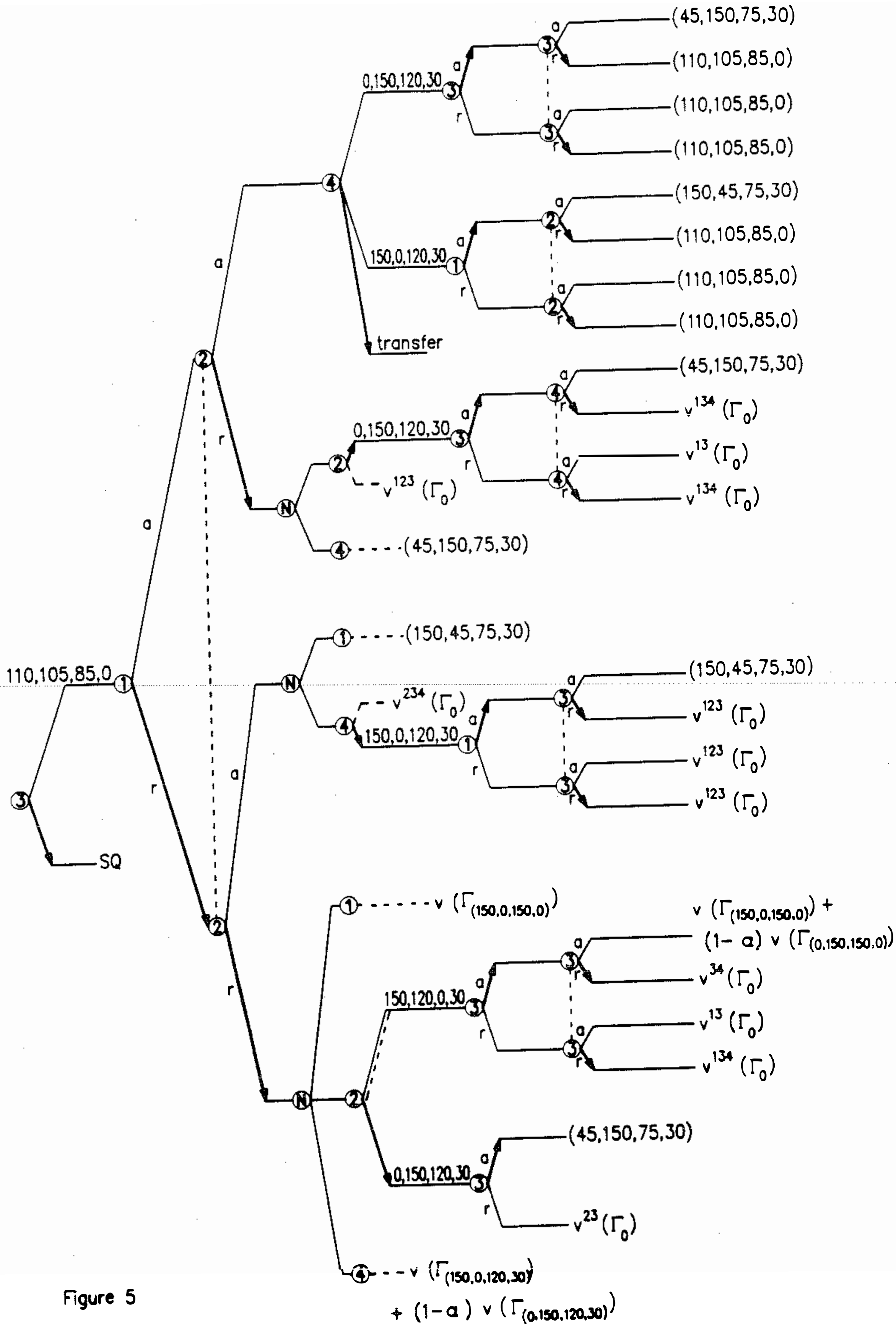
Figure 4

(110,105,85,0). Referring to Figure 5: (1) If 1 and 2 defect and accept 3's offer, then 4 cannot counter effectively and is eliminated; (2) If 2 rejects but 1 accepts, then 2 and 4 will punish 2 (they cannot punish 3) and (45,150,75,30) prevails; (3) If 1 rejects and 2 accepts, then 1 and 4 will punish 2 and (150,45,75,30) prevails. Finally, if 1 and 2 reject, the situation becomes especially complex because neither 1 nor 2 has an incentive to try to punish 3 alone. As Figure 5 illustrates in the event that nature picks 2 to punish 3, 2 prefers the threat (0,150,150,0) (or (45,150,75,30), which is equivalent from 2's point of view). That is, punishing 3 alone is not a primary threat, whereas if 2 defects and threatens 1 instead, 3 accepts and 1 or 1 and 4 must transfer to 2. Thus, if 1 and 2 reject, the game evolves into a lottery in which either 1 eventually transfers to 2 or 2 transfers to 1. Nevertheless, regardless of the probabilities over these outcomes, both 1 and 2 have a dominant choice in responding to 3's initial offer -- reject. Thus, unlike a balance of power equilibrium in which an inessential country 4 is eliminated, 3's initial threat is rejected and country 4 survives. Indeed, since 3 does not gain from threatening 4, 3 has no positive incentive to defect from the presumed equilibrium of not offering any initial threat (although, because there are only three essential countries in this example, 3's preference is not strong).

6. The Weakness of Collective Security Equilibria

The preceding analysis establishes that the postulated punishment strategies support an equilibrium in which no one makes a threat, and no one is eliminated. Thus, collective security ensures the sovereignty of everyone so that, unlike balance of power, countries can pursue absolute welfare maximization. However, collective security equilibria share a weakness that renders them vulnerable.

Returning to Figure 4, if 3 chooses before 4 and accepts 2's offer, then 4 is indifferent between rejecting and accepting, and it appears that sequential choice has no effect on collective security. However, 4's indifference rests on the supposition that if it rejects, then 1 will punish 2 or 3 even though 1 is indifferent between punishing and defecting to the primary threat (150,85,65,0). Thus, if there is even the slightest chance that 1 will defect -- if 1 "plays with a shaky hand" -- then 4 should follow 3's lead of accepting 2's initial offer. This argument, of course, does not eliminate collective security as an equilibrium since it merely renders 3 indifferent between accepting and rejecting 2's offer. And if 3 is indifferent and has no incentive to defect from its punishment strategy, then 2 should pass rather than threaten. The difficulty, however, concerns a country's ability to judge 3's commitment to a punishment. If, for reasons wholly exogenous to the analysis, 2 believes that 3 has a "sufficiently high" probability of defecting from prior agreements to punish and of agreeing to its offer, then the collective security arrangement breaks down. It is perhaps here then that we see the underlying intuition behind the notion of why regimes, consisting of "sets of implicit or explicit principles, norms, rules, and decision-making procedures around which



110,105,85,0
SQ
0,150,120,30
150,0,120,30
transfer
(45,150,75,30)
(110,105,85,0)
(110,105,85,0)
(110,105,85,0)
(150,45,75,30)
(110,105,85,0)
(110,105,85,0)
(110,105,85,0)
(45,150,75,30)
0,150,120,30
$v^{123}(\Gamma_0)$
$v^{134}(\Gamma_0)$
$v^{13}(\Gamma_0)$
$v^{134}(\Gamma_0)$
(45,150,75,30)
(150,45,75,30)
$v^{234}(\Gamma_0)$
150,0,120,30
(150,45,75,30)
$v^{123}(\Gamma_0)$
$v^{123}(\Gamma_0)$
$v^{123}(\Gamma_0)$
$v(\Gamma_{(150,0,150,0)})$
$v(\Gamma_{(150,0,150,0)})$ +
$(1-a)\ v(\Gamma_{(0,150,150,0)})$
150,120,0,30
$v^{34}(\Gamma_0)$
$v^{13}(\Gamma_0)$
$v^{134}(\Gamma_0)$
(45,150,75,30)
0,150,120,30
$v^{23}(\Gamma_0)$
$v(\Gamma_{(150,0,120,30)})$
$+\ (1-a)\ v(\Gamma_{(0,150,120,30)})$

Figure 5

actors' expectations converge in a given area of international relations (Krasner 1983:2), are an essential component of the neoliberal's portrayal of international systems.

Thus, while the stability of a collective security equilibrium is greatly enhanced by subgames that allow countries to realize mutual gains, they are damaged by subgames that allow subsets of countries to realize such gains at the expense of others. In our example, if exogenous factors assure 3 some benefit from coalescing with 2 or if 3 and 2 have some sort of ideological affinity so that 3 prefers the outcome (40,150,60,50) to (150,40,60,50), then an all-encompassing collective security cannot be sustained, and a balance of power equilibrium -- which is strong and perfect -- prevails. Thus, as Keohane and Nye (1989:248) argue, "the problem [for international stability] is how to generate and maintain a mutually beneficial pattern of cooperation in the face of competing efforts by governmental (and nongovernmental actors) to manipulate the system for their own benefit."

7. Implications

This essay's primary objective is to give some coherence to the realist-neoliberal debate, and, if possible, to resolve that debate, so Table 1 summarizes the lessons of the preceding discussion that are important for reevaluating the issues separating realism and neoliberalism.

Table 1

	Collective Security	Balance of Power
Does an equilibrium generally exist?	yes	yes
Is the equilibrium strong and perfect?	not necessarily	yes
Does the equilibrium ensure everyone's sovereignty?	yes	no
Is the equilibrium sensitive to sequential choices?	yes	no

The principal lesson of our analysis, then, is that anarchy occasions both a balance of power and a collective security equilibrium. This fact, in turn, allows us to see why the realist-neoliberal debate so easily degenerates into an inconclusive argument over goals. Goals are endogenous and depend on which equilibrium prevails. In a balance of power, nations must be concerned with relative resources, because a loss of sovereignty cannot be precluded if they become too weak; under collective security, nations can focus on absolute gains since no one makes threats against the sovereignty of any state, large or small. The debate over goals, then, is in fact a debate over the equilibrium that is thought to prevail at specific points in time, as well as a debate over the relative stability of these equilibria.

Perhaps the most disconcerting implication of our analysis to neoliberals is the fact that a balance of power equilibrium is attractive because it is both strong and perfect. If a country believes that all or nearly all other states abide by it -- if it believes that all or nearly all other states will coalesce freely and cannot be relied on to participate in punishments -- it will have a positive incentive to abide by it as well and to accept primary threats when they are offered and to make them when it is possible to do so. Collective security equilibria, on the other hand, are neither strong nor subgame perfect if $n = 3$ and they are not perfect if countries move sequentially in accepting or rejecting proposals. Countries, then, can "wander" from such equilibria and move the system to a balance of power equilibrium.

The neoliberal view, though, is resurrected if we take account of those welfare-maximizing subgames, which are not a part of our model, that collective security equilibria allow (e.g., multilateral investments and a focus on domestic economic investment as against military procurement) and that render a collective security equilibrium strong and perfect. If defection from such an equilibrium implies not only a punishment administered by other states but also the inability to pursue gains from cooperation, then defection is doubly costly. Collective security equilibria become attractive, then, when they are nurtured in such a way that all or nearly all states realize gains when they abide by them. Thus, to the extent that international organizations facilitate trade and cooperation in the ways envisioned by neoliberals, collective security becomes a more secure alternative to balance of power.

There is, however, another critically important function served by such organizations. Because there are at least the two equilibria that our analysis identifies, countries must solve a coordination problem similar to the one that arises in the Battle of the Sexes game. To illustrate this problem in its simplest form, suppose there are only three countries, that $\mathbf{r}^{o} = (120,100,80)$, and that each country must choose between a balance of power foreign policy (BOP) and a collective security foreign policy (CS). This simple characterization of foreign policy decision-making yields a normal form like the one in Table 2. That is, if either of the two larger countries defects from a collective security equilibrium, it is punished and must transfer resources to its largest opponent; but if only one such country abides by such a strategy, it alone is the target of threats.[10] Thus, (BOP,BOP,BOP) and (CS,CS,CS) are both equilibria, and international organizations must not only facilitate the realization of mutual welfare gains, they must also ensure that the countries can coordinate to (CS,CS,CS).

		Country 3: BOP		Country 3: CS	
		Country 2: BOP	Country 2: CS	Country 2: BOP	Country 2: CS
Country 1	BOP	110,110,80	150,70,80	110,110,80	70,150,80
	CS	70,150,80	110,110,80	150,70,80	120,100,80

Table 2: Coordination Problem with Three Countries

This coordination task seems especially important in a collective security equilibria, because even if no threats are observed initially, no country can be certain that others have not defected from or will otherwise fail to abide by appropriate punishment strategies. As presently formulated, our model offers no opportunity for signaling a commitment to such strategies, so if a sufficient number of states believe that others would participate in threats, the collective security arrangement is destroyed. With respect to the game in Table 2, notice that if country 3 abides by BOP, then BOP is a dominant choice for countries 1 and 2; but if 3 abides by CS, then CS is dominant for 1 and 2. Thus, the equilibrium that prevails depends critically on what 1 and 2 believe about 3. Correspondingly, the policies associated with collective security that maximize welfare without regard to relative gains are likely to be viewed as risky, with pessimists warning of dangers and questioning whether international organizations can perform their cooperative and coordinative functions. Indeed, we can even predict the permanency of realist-neoliberal debates.

Realist-neoliberal debates, then, are readily formalized, but we should not delude ourselves into believing that we have overcome the hardest problems. First, our model does not take account of investment and endogenous resources growth. Second, we ignore the costs of nuclear conflict that account importantly for the US-USSR balance of power that allows collective security to characterize both the US and Soviet alliances. Finally, our model ignores uncertainty and misperception. Until we confront this complexity (which is far easier said than done), we cannot comprehend fully the meaning of the signals that countries generate by their actions and words as they try, as part of their strategies, to reveal their commitment to particular actions or to deceive others about this commitment. Indeed, as we indicate earlier, the beliefs that are part of this complex signaling process not only give theoretical meaning to the concept of regime, but are also critically important for determining the types of equilibria that can prevail. Despite these limitations, however, we have accomplished one objective -- namely, we have embedded the alternative worlds perceived by realists and liberals in the same model, so that debate can now focus on a specific model's details and parameters.

Footnotes

1. For additional discussion of this distinction in the context of a cooperative game-theoretic analysis of balance of power, see Niou, Ordeshook, and Rose (1989).

2. For an example of the application of stationary strategies in a different substantive context see Baron and Ferejohn (1989)

3. Formally,

Type 1 Threat: $(\mathbf{r},C)$ is a Type 1 threat -- $(\mathbf{r},C) \in T^1$ -- if

i $r^o_{max[C]} + r^o(S-C) \geq R/2$,

ii $r_j = 0$ for all $j \in S-C$,

iii $r_{max[C]} = R/2$,

iv $\nexists\, C' \in W$ such that $C' \cap C = \{k\} = \{max[C']\} \neq \{max[C]\}$.

If we let T_C denote the threats that C can make, then for any C there is an infinity of Type 1 threats in T^C, which differ only in the distribution of $r^o(S-C) - [R/2 - r_{max[C]}]$ among C-{max[C]}. However, most such threats are strategically equivalent and have the same continuation values. So if we can associate a Type 1 threat with C, we focus on one such threat and ignore the others that C might make. Formally, then, let **C** denote all coalitions that have Type 1 threats, let $(\mathbf{r},C)$ be a particular Type 1 threat by C, and redefine the set of all threats T as $T - \cup_{C \in \mathbf{C}} [T_C - \{(r,C)\}]$. Reintroducing these excluded threats leaves our analysis unaffected.

4. Formally, letting $\mathbf{T}_o$ be the power set of T^1 (the set of all subsets of T^1), then,

Primary Threats: $T^p \in T^o$ is a set of primary threats if,

i for no $(\mathbf{r},C) \in T^p$ is there an $(\mathbf{r}',C') \in T^1$ such that $C \cap C' = \{max[C]\} = \{max[C']\} = \{k\}$ with both $C-\{k\}$ and $C'-\{k\}$ subsets of S-L;

ii $\nexists\, (\mathbf{r},C) \in T^1$ that can be included in T^p without violating condition i.

Notice that this definition necessarily renders T^p unique.

5. Formally, if $(\mathbf{r},C)$ is a primary threat, then

$$v_j(\Gamma_r) \begin{cases} = r^o_j \text{ if } j \in C \text{ and } j \neq max[C] \\ = R/2 \text{ if } j = max[C] \\ \leq r^o_j \text{ if } j \in S-C \end{cases}$$

On the other hand, if $(\mathbf{r},C)$ is not a primary threat, and if we let L be those countries in S who can be the largest member of a minimal winning coalition, then

$$v_j(\Gamma_r)\begin{cases}\leq r^o_j \text{ if } j\in C\cap S\text{-}L\\ < R/2 \text{ if } j\in C\cap L\\ \leq r^o_j \text{ if } j\in C\cap L \text{ and if } r^o_j < r^o_{\max[S-C]}.\end{cases}$$

6. An equilibrium is perfect if, for each $i\in S$, no arbitrarily small probability that others defect from their equilibrium strategies yields an incentive for i to defect.

7. That possibility that a is less than R/2 arises whenever a threat is not a primary threat in the sense that the threatened countries do not have sufficient resources to render their largest opponent near-predominant. In this instance, we suppose that countries use the credible threat of transferring all of their resources to the largest opponent. Certainly, such countries are indifferent as to the method of elimination (indeed, a transfer may be more peaceful and less costly than implementing a threat), whereas the recipient country have every incentive to accept the transfer. What renders such transfers credible threats, however, is that if it is believed, then it forestalls elimination -- countries will prefer to make other threats -- and if it must be implemented, then it yields an outcome that is no worse than what prevails if the threat is not implemented (the threat is subgame perfect).

8. Letting L denote the countries in S who can be the largest member of a minimal winning coalition, E the set of essential countries, and W the set of winning coalitions, then the proof of our result about collective security equilibria requires the following lemma:

If $|S|>3$ *and if* $|E|>3$*, then C has a primary threat only if* $|C|\geq 3$.

If $|S|>3$ and $|E|>3$, then $|S\text{-}L|\geq 2$, since, if $|S\text{-}L|=1$, then, by definition, $r(L\text{-}\{i\}) > R/2$, $i\in L$, and $r(\{i\}+S\text{-}L) > R/2$, which is impossible since otherwise $r(L\text{-}\{i\}+\{i\}+S\text{-}L) = r(S) > R$. From Lemma 0 in Niou and Ordeshook (1989), we know that for all $i\in L$, $C=\{i\}+S\text{-}L$ has a primary threat, but to see that C must have more than two members, suppose to the contrary that $C=\{i,j\}$. Clearly, at least one member of $\{i,j\}$ must be in L, since otherwise $\{i,j\}\notin W$. So supposing first that $i,j\in L$, since $\{i\}+S\text{-}L$ has a primary threat, condition (iv) in the definition of Type-1 threats is violated. Similarly, if only $i\in L$, let $\{i,j\}\in W$. Then $i = \max[S]$ and $\{i\}+S\text{-}L\text{-}\{j\}\in W$; otherwise $S\text{-}L\text{-}\{j\}$ are inessential. By condition (i) in the definition of primary threats, neither $\{i,j\}$ nor $\{i\}+S\text{-}L\text{-}\{j\}$ can form a primary threat. Q.E.D.

9. Let i be the initial defector who proposes the threat $(\mathbf{r},C)$. To eliminate consideration of cases for which we could easily extend the argument that follows, we assume that $(\mathbf{r},C)\in T^P$. The first thing we must show, now, is that no country defects to make an initial threat. Assuming

the continuation values in expression (1), we have three cases: Letting D' be the defectors in C-{i} who approve i's proposal,

1. D' = C-{i}, in which case, from (e), all countries use stationary strategies, and $v_j(\Gamma_r)$ equals r_j for all $j \in$ C-max[C], and equals R/2 for max[C].
2. D' $\subset$ C-{i}, in which case $v^D{}_j(\Gamma_o) < r^o{}_j$ if $j \in D$, and $v^D{}_j(\Gamma_o) \geq r^o{}_j$ otherwise.
3. D' = $\emptyset$, in which case $v^{\{i\}}{}_i(\Gamma_o) < r^o{}_i$ and $v^{\{i\}}{}_j(\Gamma_o) \geq r^o{}_j$ for all $j \neq i$.

Thus, the only country that might have a positive incentive to defect by offering an initial threat is max[C], so let i = max[C] $\in$ L. However, since, from the presumed continuation values, no $j \in$ C-{i} has an incentive to defect by approving (**r**,C), D' is empty, and i is punished. So no one defects initially. Next, we must establish the consistency of continuation values in expression (2). Again we look at our three cases, since now subgames begin after all relevant countries have chosen between acceptance and rejection of a proposed threat. For case 1, the results in Niou and Ordeshook (1989) establish consistency. For case 2, nature chooses a $k \in$ S-D to make the next (punishing) move. Clearly, if all other continuation values hold, k defects by not punishing one or more members of D only if it can form a primary threat (**r**',C') with members in D and if k = max[C'], since in this instance, from (e) and the continuation values assumed for stationary strategies, $v_k(\Gamma_{r'}) = R/2 > r^o{}_k$. In this instance, $v_j(\Gamma_{r'}) = r^o{}_j$ for all $j \in$ C'-{k} and $v_j(\Gamma_{r'}) \leq r^o{}_j$ for all $j \in$ S-C'. On the other hand, if nature chooses $j \in$ C-D, $j \neq k$, to make the next move, {j}+D cannot form a primary threat [if (**r**,C) $\in T^P$, then (**r**'C') $\notin T^P$ if C' $\subset$ C], so j has no incentive to defect from punishing D. Since no $j \in$ D' can gain by defecting, and since there is some chance that they will lose, the expected value from defecting is less than $r^o{}_j$ for all $j \in$ D'. Finally, for case 3, let nature pick $j \in$ S-{i} to make the punishing move. Since, from the lemma in footnote 8, {i,j} cannot form a primary threat, if j defects it will be punished; hence, it has no positive incentive to defect, which implies that $v^{\{i\}}{}_i(\Gamma_o) < r^o{}_j$. Q.E.D.

10. The smallest country cannot be punished because it alone cannot be the target of a primary threat. However, it is only 3-country systems that thus "protect" any country.

References

Baron, D. and J. Ferejohn. (1989). "Bargaining in Legislatures," **American Political Science Review**, forthcoming.

Claude, I. (1962). **Power and International Relations,** N.Y.:Random House.

Gilpin, R. (1981). **War and Change in World Politics**, Cambridge: Cambridge Univ. Press.

Grieco, J.M. (1988). "Anarchy and the Limits of Cooperation: A Realist Critique of the Newest Liberal Institutionalism," **International Organization**, 42(3), 485-507.

Jervis, R. (1976). "From Balance to Concert: A Study of International Security Cooperation," in K.A. Oye, ed., **Cooperation Under Anarchy**, Princeton: Princeton Univ. Press

Keohane, R.O. (1984). **After Hegemony**, Princeton: Princeton Univ. Press.

___________. (1989). **International Institutions and State Power**, Boulder (Col.): Westview.

Keohane, R.O. and J.S. Nye (1989). **Power and Interdependence**, 2nd ed., Boston: Little Brown.

Krasner, S.D. (1983). **International Regimes**, Ithaca: Cornell Univ. Press.

Niou, E.M.S., and P.C. Ordeshook. (1989). "Conflict and Stability in Anarchic International Systems," Calif. Inst. of Technology working paper #700.

Niou, E.M.S., P.C. Ordeshook, and G.F. Rose. (1989). **The Balance of Power**, Cambridge: Cambridge Univ. Press.

Waltz, K. (1987). **Theory of International Politics**, Reading (Mass.):Addison-Wesley.

Wagner, R.H. (1986). "The Theory of Games and the Balance of Power," **World Politics**, 38(4):546-76.