BALANCE OF POWER VERSUS COLLECTIVE SECURITY: 
A GAME-THEORETIC ANALYSIS

EMERSON M. S. NIOU  
Duke University

and

PETER C. ORDESHOOK  
California Institute of Technology

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.

1. Introduction

On August 10, 1990, the Los Angeles Times offered the following news report: "Lack of Fear Over WWIII Leads to Rare Global Cooperation ... The extraordinary worldwide response to the call for sanctions ... against Iraq represents the fulfillment of an elusive dream of leaders past, Woodrow Wilson's recipe for achieving global security through collective action of all nations." For forty-five years international politics has been overshadowed by a balance of terror in which organizations such as the United Nations were no more successful at ensuring peace than was the abortive League of Nations. However, the world took a more hopeful view following the Iraq war, the overthrow of communist regimes in Eastern Europe and the USSR, and the increasing acceptance of democratic principles worldwide. Nevertheless, before we can proclaim any new world order, we must contend with the fact that Western Europe seems determined to maintain its time-honored tradition of exhibiting an inability to police its own neighborhood (e.g., Yugoslavia), with the fact that ethnic tensions in the

successor states of the USSR threaten further political instability, and with the possibility that the emerging economic competition between Japan and China will take a military route. Thus, we cannot easily dismiss the argument that states will continue to be concerned with relative military status, with the argument that they will continue to seek economic advantage over even ostensibly allies, and with the realpolitik view that a balance of power continues to provide a surer basis for theorizing about international political processes that does the concept of collective security.

The debate over these views -- the debate between realists who foresee stability arising only from a balance of power and neoliberal institutionalists who foresee the possibility of greater cooperation in the form of collective security arrangements -- focuses on two issues: (1) delineating the goals of countries that best account for their actions, especially patterns of cooperation and conflict; and (2) evaluating the possibility that institutions of various descriptions can ameliorate conflict in an otherwise anarchic environment. Realists argue that although states may be concerned in the long run with absolute welfare, the impossibility of eliminating the threat of conflict forces them when evaluating strategies and outcomes, to be predominantly concerned with relative position as measured by military capability, economic productivity, and the like. This concern with relative position, in turn, attenuates the opportunities for cooperation and the role of institutions as facilitators of cooperation and forces states to abide by strategies consistent with balance-of-power politics. In contrast, neorealists, drawing on the lessons of scenarios such as the repeated Prisoners' Dilemma and the myriad instances of actual cooperation in international politics, see less reason for supposing that states are concerned necessarily with relative gain, see greater opportunities for mutually beneficial cooperation, and see an expanded role for institutions as facilitators of that cooperation. Equivalently, they are more willing than realists to extend the notion of collective security beyond domain of economic cooperation so that it includes strategic military issues as well.

This essay seeks to contribute to the resolution of this debate by offering a more precise formulation of the distinction between balance of power and collective security. Reviewing some results proved elsewhere about a game-theoretic model of anarchic systems, we consider whether balance of power and collective security can each be supported by equilibrium strategies. Arguments about the functions served by international organizations such as the United Nations and GATT, the viability of international military cooperation, and the viability of the Commonwealth of Independent States are then recast in terms of the different properties of these equilibria. Section 2 summarizes our theoretical perspective, which consists of developing a model of international threat and counter-threat processes. Section 3 offers some formal notation and the game we use to model anarchic systems. Briefly, in that model "power" is the sole determinant of winning and losing coalitions, threats and counter-threats are the mechanisms whereby countries secure resources and ensure their sovereignty, alliances are formed or maintained because it is in the interest of individual countries to do so, and no exogenous constraints ameliorate conflict. Section 4 provides our main result about balance of power equilibria, which establishes the existence of such equilibria but which also establishes that the two largest countries vie for control of the system and that the sovereignty of the smallest countries is not guaranteed. Section 5 provides our main result about collective security equilibria, which is that such equilibria exist and that they ensure the sovereignty of all countries in the system. Finally, Section 6 argues that much of the discussion between realists and neorealists about balance of power versus collective security is, in fact, an argument over the ease with which states can coordinate to one equilibrium rather than the other and the signals they can offer to indicate their commitment to a particular equilibrium.

2. Theoretical Perspective

The essential character of international systems that renders them anarchic is that if states cooperate, they must do so out of simple self-interest: Cooperation must characterize the equilibrium strategies of the non-cooperative game that models the system. In turn, balance of power politics as well as collective security arrangements, if they exist at all, must correspond to alternative equilibria in this model. Thus, assuming that we possess such a model, the questions we must confront include: Can balance of power and a stable collective security system be supported by non-cooperative equilibria? Does the existence of such equilibria require any special configuration of power relations among countries? If both equilibria exist, how can states coordinate to ensure that a particular equilibrium prevails? Is a balance-of-power equilibrium more stable than a collective security equilibrium? Is there a special role for institutions in achieving and maintaining equilibria and do either of these equilibria allow states to cooperate so as to realize mutual economic gains?

Rather than rely on metaphorical appeals to scenarios such as the repeated Prisoners' Dilemma, Chicken, or the Battle of the Sexes, we offer in this essay a game-theoretic analysis that attempts to model the essential features of anarchic international systems -- of systems without exogenous mechanisms of enforcement and which allows not only for cooperation but also for non-cooperation in the form of threats to eliminate players (countries) altogether. Briefly, our model assumes that, conditional on maintaining their sovereignty, countries pursue a single transferable resource in constant supply, which we can think of as either "power" or economic wealth. We assume constant supply because we do not want to secure cooperation simply by making the gains from it too great, and because sustaining cooperation in the context of constant sum competition reveals more clearly the role that institutions can play in ameliorating conflict. Next, so as to avoid confusion between the notion of equilibrium and that of stability, we say that a system is stable if and only if all countries can ensure their sovereignty in the sense that if all countries choose their equilibrium strategies, no country will have its resources reduced to zero.1

Our model of anarchic, now, supposes that each country's resources determines who can threaten whom, where threats and counter-threats are the mechanisms whereby countries secure resources from each other. 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 it passes, nature selects another country. If it threatens, its partners in C must decide 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.

With respect now to the types of equilibria that we consider for this portrayal of an anarchic international system, we know that nearly any "reasonable" outcome can be sustained as an equilibrium, given an appropriate specification of strategies.2 But the complexity of many of these strategies 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. In a balance of power system, then, "all states are potentially fit alliance partners; none is seen as much more evil than any other


other\(^3\). 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.\(^4\) The strategies forming a collective security equilibrium, then, are not stationary because they posit the formation of specific alliances, depending on who defects from a specified pattern of action.

Notice now that if we find equilibria supported by stationary strategies, and if only countries controlling some critical relative level of resources can guarantee their existence in it, then countries must be vigilant about the relative gains and losses that economic cooperation generates, which makes such cooperation difficult if not impossible. If, however, there is a collective security equilibrium supported by punishment strategies in which no country offers an initial threat, then 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-confrontational world, and if these benefits disappear when agreements to achieve them are disrupted by competition over relative position, then we should pay special attention to the circumstances under which such an equilibrium can be achieved and a balance of power equilibrium avoided.

3. The Model

With respect to notation, we let \((S, r)\) denote a system, where \(S = \{1, 2, \ldots, n\}\) is the set of countries and \(r = (r_1, r_2, \ldots, r_n)\) is the distribution of resources across \(S\). Without loss of generality, we let \(r_1 > r_2 > \ldots > r_i > 0\). If \(r(S)\) is the sum of resources controlled by the members of the coalition \(C \subseteq S\), and if \(R = r(S)\), then \(C\) is winning if \(r(C) > R/2\), it is losing if \(r(C) < R/2\), and it is minimal winning if, for all \(i \in C\), \(C\setminus\{i\}\) is not winning. Countries in \(S\) who are at least one minimal winning coalition are essential; otherwise they are inessential. If \(r_i > R/2\), \(i\) is predominant -- it is winning against all other countries and can incorporate their resources at will -- so every country has an incentive to avoid the possibility that some other country becomes predominant. If \(r_i = R/2\), \(i\) is near-predominant. With respect to the status quo distribution \(r^\ast\), \(\max(C)\) is the country in \(C\) with the greatest total of resources. Finally, the game \(\Gamma\) that we use to model conflict proceeds as follows:

1. Nature randomly selects \(i \in S\).
2. Offers a threat \((r_i, C), i \in C\), or no passes. If no passes, we return to (1). Country \(j\) is assumed to be a member of \(C\) if \(r_j \geq r_i\).
3. The members of \(C\setminus\{i\}\) simultaneously choose between approving or rejecting \((r_i, C)\). If \(j \in C\setminus\{i\}\) rejects, \((r_i, C)\) becomes the current threat; otherwise, return to (1).
4. With \((r_i, C)\) the current threat, nature randomly orders the members of \(S\setminus C\).
5. With \(m \in S\setminus C\) offering the first counter, a counter takes one of two forms: a new threat, \((r_i', C')\), \(m < C\); or a resource transfer from \(S\setminus C\) to one or more members of \(C\). We denote those party to the transfer by \(C'\).
6. The members of \(C\setminus\{m\}\) simultaneously choose between approving or rejecting \((r_i', C')\). 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 \(C\setminus\{m\}\) 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).

Notice that \(\Gamma\) is a recursive game, because it allows threats and counters, as well as resource reallocations, to continue in sequence forever. This game's correct treatment, then, is to pretend that it 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 with these strategies -- the subgame perfect choices they imply must yield those consequences.\(^5\)

We are, of course, less interested in the mechanics of how this game ought to be analyzed than we are in the types of equilibria that can prevail in it. In fact, we are primarily interested in ascertaining whether two broad types of strategies can define alternative equilibria -- strategies that correspond to balance of power (competitive) politics and strategies that correspond to a collective security (universally cooperative) arrangement. Specifically, suppose for the moment that \(n = 3\) and suppose further that a threat or counter-threat can only take one of three forms -- \((150,150,0)\), \((150,0,150)\), and \((0,150,150)\). That is, suppose threats and counters (excluding alternative resource transfers) all have two countries eliminating the third, where those two countries agree to share the system's resources equally. Then a "balance of power equilibrium" has countries adopting the following types of strategies:

BP: If, for instance, country 1 is given the first move, then it threatens (150,0,150); 2 passes (0,150,150); 3 passes or threatens (150,0,150) or (0,150,150). If the initial threat is (150,150,0), then "reject"; otherwise "accept." If threatened, states 1 or 2 transfer to the largest threatening country. If state 3 is threatened, then 3 offers (150,0,150) or (0,150,150) as a counter-threat.

That is, in a balance of power equilibrium, countries threaten each other and others agree to participate in threats owing to the expectation that if they fail to agree, others will subsequently threaten them. On the other hand, in a collective security arrangement, strategies will look as follows:

CS: No state makes an initial threat, but if one is offered, the proposed partner "rejects." If the initial threat is rejected, then the "defecting" state is punished by being threatened in the next stage (and this threat is accepted). If two players defect by making and accepting an initial threat or by failing to punish, then play the game as described in BP.

The question, then, is whether each of these descriptions of strategies (suitably extended to \(n\)-country scenarios) can correspond to equilibria in our model of anarchic systems. And in the event that they are equilibria, we should ask whether one or the other is especially stable, whether existence depends on the actual initial distribution of resources, and whether the number of countries in the system, \(n\), plays any special role.

4. Balance of Power

Because our discussion of collective security builds on an analysis of balance of power, we focus initially on stationary strategies of type BP. Appreciating the fact, though, that readers may not want to wade


\(^4\) Friedman, James. 1986. Game Theory with Applications to Economics. New York, Oxford University Press.

through the notation required to summarize our conclusions rigorously, we note here that what follows is intended to establish some assumptions and a classification of threats that allows us to support the following conclusion:

If all countries are essential -- if all are members of at least one minimal winning coalition -- then there is a balance of power equilibrium in which no country is eliminated but in which one of the "larger" countries secures half the resources in the system at the expense of some losing coalition. But if there are essential countries, then those countries cannot assure their continued existence.

To provide a more precise statement of this result, let $\Gamma_i$ denote the game that follows the threat of $(r, C)$ and its acceptance, let $\varphi_i(\Gamma_i)$ denote the value that $i$ in $S$ associates with playing that sub-game $\Gamma_i$. Thus, $\varphi_i(\Gamma_i) = \varphi_i(\Gamma_i') \ldots \varphi_i(\Gamma_i')$ is the continuation value of $\Gamma_i$ and specifies what the countries believe follows from the approval of $(r, C)$. So $\varphi_i(\Gamma_i)$, when compared against whatever follows if $(r, C)$ is rejected, determines $i$'s preference for acceptance or rejection of $(r, C)$ 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 $\Pi$ 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. 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.

To posit continuation values we isolate one class of threats, where the members of one class are associated with one type of continuation value and all remaining threats are associated with a second type. We first identify Type 1 threats, which satisfy four conditions:

**Type 1 Threat:** $(r, C)$ is a Type 1 threat -- $(r, C) \in T^* --$ if:
1. $r^\text{max}(C) + r(S-C) \geq R/2$,
2. $r_i = 0$ for all $i \in S-C$,
3. $r^\text{max}(C) = R/2$,
4. There exists a $C \in W$ such that $C \cap C = \{k\} = \{\text{max}(C)\} + \{\text{max}(C)\}$.

Next, we refine $T^*$ further to define the set of primary threats, $T_\text{P} \subset T^*$. Letting $L$ be those countries in $S$ who can be the largest member of a minimal winning coalition, $T_\text{P}$ satisfies two conditions: First for no $(r, C)$ and $(r', C')$ in $T^*$ 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 that no winning coalition can offer a Type 1 threat that can be added to this set so as not to violate condition 1. Formally, letting $T_\text{S}$ be the power set of $T^*$, then:

**Primary Threats:** $T_\text{P} \subset T^*$ satisfies:
1. For no $(r, C) \in T^*$ there is an $(r', C') \in T^*$ such that $C \cap C' = \{\text{max}(C)\} + \{\text{max}(C)\} = \{k\}$ with both $C \setminus \{k\}$ and $C' \setminus \{k\}$ subsets of $S-L$;
2. There does not exist a $(r, C) \in T^*$ that can be included in $T_\text{P}$ without violating condition 1.

This definition renders $T_\text{P}$ unique. We now define two types of continuation values:

$C1$: $(r, C)$ in $T$ satisfies $C1$ if:
- $r_j \geq r'$ if $j \in C$ and $j \neq \text{max}(C)$
- $r_j = \text{R}/2$ if $j = \text{max}(C)$
- $r_j \leq r'_j$ if $j \notin S-C$

$C2$: $(r, C)$ in $T$ satisfies $C2$ if:
- $r_j \geq r'$ if $j \in C \cap S-L$
- $r_j \leq \text{R}/2$ if $j \in S-L$
- $r_j \leq r'_j$ if $j \notin S-L$ and if $r_j < r'_j < \text{max}(C)$

Before we can assign these continuation values, however, we must impose three assumptions.

**A1:** If $r_i = \text{R}/2$, then $r_i > \text{R}/2$ if any threat is made and implemented.

So systems with a near-predominant country are "frozen" -- no threat is made since everyone knows that the implementation of such a threat will allow the near-predominant country to become predominant, in which case it can eventually secure all of the system's resources.

**A2:** If $i$ can become near-predominant either from a transfer or from implementing a threat, $i$ prefers the transfer.

Hence, if $\text{max}(C)$ is the largest member of $C$ with respect to the status quo, then the system is frozen if $S-C$ offers to render $\text{max}(C)$ near-predominant. Clearly, if $S-C$ prefers freezing the system, it should transfer to $\text{max}(C)$, since this choice minimizes the resources that $S-C$ must surrender. And $\text{max}(C)$ accepts the offer: Because attempts to secure more than $\text{R}/2$ will be blocked, securing $\text{R}/2$ by transfer is $\text{max}(C)$'s most preferred feasible outcome.

**A3:** Letting $u_i$ denote country $i$'s payoff, $u_i(r) = r_i$ at terminal nodes. For non-terminal nodes, if $R(r)$ is the set of terminal and non-terminal resource distributions that might be reached from $r$, then $u_i(r) = \text{min}(R(r))$.

A3 "prunes" non-terminal transfers from the extensive form. If choosing between $r_i$ with certainty (as when another country transfers to a third party that ends the game) and playing a game with $r'$, resources, $r_i > r'$, then $i$ prefers the certainty of $r$ since $i$ may be threatened subsequently with the necessity for transferring resources itself to a level below $r_i$. But if $i$ is threatened and if in countering this threat $i$ must choose between retaining $r_i$ after it transfers to end the game and retaining $r_i', r_i' > r_i$ by transferring a smaller amount so as to attract someone away from the threatening coalition, then $i$ prefers the terminating threat.

Finally, we use the following assumptions to refine our description of stationary strategies:

**A4:** If $(r, C)$ is the current threat, then $i \in S-C$ chooses a counter, $(r', C')$, such that $C \cap C = \{i\}$ and $C' \cap C = \{i\}$.

To illustrate, let $(r, C)$ be the current threat and let $i \in S-C$ offer a counter-threat. Any counter, of course, must either entail a transfer, since $C$ is necessarily winning, or it must coopt one or more members of $C$ into a new coalition, $C'$. As supposes that if $i \in S-C$ can form a counter that coopts only one member of $C$ so that all of $i$'s other coalition partners in the counter are in $S-C$, then $i$ chooses that counter. More valuable alternatives are not ignored, but, whenever it is indifferent, $i$ takes advantage of the fact that $S-C$ is a coalition that, because of $C$'s threat, is "already nearly formed."
Three lemmas follow now from this construction.6

**Lemma 1**: For each $i \in I$, there is at least one $(r, c) \in P$ such that $i = \max C$ (in particular, if $L_i \neq L$, there is a threat in $P$); and for each $j \in I$, there is at least one $(r, c) \in P$ such that $j \in C$.

**Lemma 2**: If all threats not in $P$ satisfy C1, and if all threats in $P$ - $(r, c)$ satisfy C2, then for any stationary equilibrium, $(r, c) \in P$ satisfies C2.9

**Lemma 3**: If $T$ is in $P$ satisfy C2, if $(r, c) \in P$, and if all other threats satisfy C1, then for any stationary equilibrium, $(r, c)$ satisfies C1.

We turn now to our central result -- the characterization of a stationary equilibrium. What remains at issue is a specification of a country's choice whenever it is selected to make the initial threat, and the responses of its partners in a proposed initial threat. Postponing the question of the fate of inessential countries, assume that all S are essential. Limiting the discussion to symmetric strategies -- strategies in which all countries in $L_i$ and all those in $L_s$ abide by the same strategy, consider this statement as a characterization of equilibrium:

If $i \in S$ is chosen to make the initial threat, $i$ randomly chooses $(r, c) \in P$, $i \in C$ and all $j \in C - i$ accept; and if $i \in L$, then $i$ chooses $(r, c)$ such that $i = \max C$.

So if $i \in L$ is chosen by nature to make the initial threat, then $i$ proposes a primary threat in which it is the largest member of the threatening coalition; if $i \in L$, then $i$ proposes a randomly selected primary threat in which it is a member of the threatening coalition, and all members of the proposed threatening coalition accept. Using this characterization of strategies, we learn that for the game $T$ as specified:7

**Result 1**: If all $i \in S$ are essential, there is a stationary equilibrium such that $(S, r)$ is stable such that some country in $I$ is rendered near-prominent.

Furthermore, if we allow 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. Thus, there is a balance of power equilibrium that ensures the sovereignty of "larger" states, but not of smaller ones, and with sets the two largest countries in the system in competition against each other for near-prominence. It is important to note that this equilibrium is also strong in this 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. So a balance of power equilibrium is attractive.

5. **Collective Security**

Turning now to collective security and punishment strategies designed to preserve the status quo, we simplify matters by modifying the assumption that nature chooses randomly from $S$ after a threat is rejected. 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).

Next, we formulate a punishment strategy that matches the simplicity of stationary strategies, because we do not want to confront the objection that balance of power equilibria are easier to compute: (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 (members of $D$); (d) Countries accept threats that are punishments; and (e) Whenever any threat is accepted, all countries use stationary strategies thereafter. Players defecting from $(S \cup D)$ 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 $v_i(T) = (v_i(T), v_j(T), \ldots, v_n(T))$ 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 $v_i(T)$ denotes the corresponding continuation value. The general form we assume for $v_i(T)$ and $v_i(T)$ is as follows:

$$v_i(T) = \begin{cases} T & \text{if } j \notin D \text{ and } j = \max(S - D) \\ r_j & \text{if } j \notin D \text{ and } j < \max(S - D) \\ a & \text{otherwise} \end{cases}$$

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 previously. But if there is no current threat, the specification of $v_i(T)$ 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 a target for punishment but if it is the largest member of $S - D$, then $j$ receives a transfer that either renders $j$ near-prominent or which eliminates $S - C$. Finally, if $j \in D$, then $j$'s expected payoff is less than its current resource holdings.

These definitions yield the following result about collective security:8

**Result 2**: 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.

We limit the domain of this result to systems with four or more essential countries, because although a collective security equilibrium exists otherwise, it is neither strong nor subgame perfect. The weakness of such equilibria if $|S| = 3$ arises because only 2-country coalitions have primary threats. For example, if $r = (120, 100, 80)$ and if country 3 defects by proposing the primary threat $(150, 0, 150)$, then it 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, more than one country must accept a proposed primary threat, and this fact renders a collective security equilibrium both strong and subgame perfect.

6. Implications

The principal lessons of our analysis are these:

1. A balance of power and a collective security equilibrium can exist simultaneously in an anarchic system that does not allow for universal gains from cooperation.
2. 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.
3. 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.
4. In a balance of power one of the two largest states becomes near-predominant.
5. Collective security equilibria are strong only if the number of countries exceeds three.
6. If defection from a collective security equilibrium implies not only a punishment administered by other states but also the inability to pursue gains from cooperation, then collective security equilibria become attractive. Thus, to the extent that international organizations facilitate trade and cooperation, collective security becomes a more secure alternative to balance of power.

Our analysis, moreover, reveals a critically important function served by international organizations. Because we have identified two substantively plausible equilibria, countries must explicitly coordinate to achieve an equilibrium. To illustrate this problem in its simplest form, suppose there are only three countries, that is, r* = (120,100,80), and that each country must choose between a balance of power foreign policy (BP) and a collective security foreign policy (CS). This simple characterization of foreign policy decision-making yields a normal form like the one in Table 1. 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. Thus, (BP,BP,CS) 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).

<table>
<thead>
<tr>
<th>BP</th>
<th>CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>110, 110, 80</td>
<td>150, 70, 80</td>
</tr>
<tr>
<td>70, 150, 80</td>
<td>110, 110, 50</td>
</tr>
</tbody>
</table>

Table 1

This coordination task is especially important in a collective security equilibrium, 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 1, notice that if country 3 abides by BP, then BP 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.

The relevance of this example, then, is four-fold. First, both cooperative and non-cooperative equilibria -- world orders -- can exist within a scenario other than a repeated Prisoners' Dilemma in which all countries have an incentive to avoid some mutually destructive outcome. Second, since we have already incorporated the influence of power into the analysis by way of defining legitimate threats and counter-threats, we cannot now use power to predict which of these two equilibria prevail. Equilibrium selection -- coordination to a particular equilibrium -- must occur on some other basis. Third, neither of the equilibria we identify here is Pareto-dominated by the other. Thus, there is no reason to suppose a priori that states will gravitate to one rather than the other -- whether they thus gravitate will depend on things other than the relative efficiency of one equilibrium as compared to another. Finally, this analysis illustrates that learning how states coordinate to particular equilibria is an essential part of any explanation for final outcomes. If there are multiple equilibria in so simple a model as the one we offer, then we can be certain that this multiplicity characterizes an even more complex reality.

6. Conclusions

Having thus identified some of the equilibria that can exist within a single scenario, let us now consider the problems associated with coordinating to particular equilibria. Referring again to the game in Table 1, consider the prospect of the three countries in this example coordinating merely with some pre-play discussion. It is here, however, that we can discern the sources of the realist's disagreement with neoliberalism, because there are good reasons for supposing that mere pre-play discussion is ineffectual with respect to ensuring the collective security equilibrium (CS,CS,CS). First, this equilibrium calls for states to "do nothing" until there is a defection that warrants punishment. Hence, regardless of the verbal agreements they reach, each state, as the game unfolds, may question whether others are abiding by their collective security strategies or whether they are merely postponing making a threat until circumstances (not modeled here but presumably including exogenously induced changes in the distribution of resources) are favorable to that purpose.

Second, collective security requires that states punish defectors; but proposing a punishment (as opposed to some other threat) may be rational only if it is certain beforehand that the ostensible partners in the punishment will maintain their commitment to it. Because a collective security equilibrium is subgame perfect in our model, doing so is rational here. But we should not ignore the possibility, as a practical matter, that states might be concerned that a defection of one type increases the perceived likelihood of yet other defections, so that defection becomes a self-fulfilling prophecy. Our example, after all, assumes that all countries have perfect foresight, whereas if there is always something left to chance, then, barring a perfectly functioning coordination mechanism, the viability of pursuing a punishment strategy may be reduced.

Third, that collective security is an equilibrium means only that no state has an incentive to defect unilaterally from the agreement. This does not mean that states cannot gain if two or more of them defect simultaneously -- if there are coordinated defections. For example, if states 2 and 3 defect from (CS,CS,CS) to (CS,BP,BP), then 2 gains and 3 loses nothing. And, stepping outside the limits of our formal analysis for a moment, country 2 can presumably reward 3 somehow for its compliance. Indeed,
if we are willing to assume that states can coordinate to achieve one type of equilibrium, then, barring other considerations, we should be willing to assume that subsets of them can coordinate to achieve other ends -- if $n$ countries can coordinate, then it is reasonable to assume that $m < n$ can also coordinate.

Although the realist's objection to the neoliberal argument takes the form of a discussion of these issues, we can see its theoretical content better by referring to the game in Table 2, which, like the Battle of the Sexes and like the game in Table 1, has two non-equivalent, non-interchangeable equilibrium strategy pairs, $(a_b, b_1)$ and $(a_b, b_2)$. At first glance we might suppose that, barring any prior asymmetrical beliefs about strategies, neither equilibrium is more likely to prevail than the other. But suppose we consider the outcome that prevails if the two players start at arbitrary strategy pairs and if they adjust their strategies sequentially. For example, if they begin at $(a_1, b_2)$, they arrive at $(3,4)$ if column chooser moves first, whereas they arrive at $(4,3)$ if row chooser moves first, followed by column chooser (via the route $a_1$ to $a_2$, $b_1$ to $b_2$). Counting the number of ways each equilibrium can be reached from some other pair of strategies, there are four routes to $(3,4)$ and ten to $(4,3)$. Thus, we might suppose that in the absence of coordination, $(4,3)$ is more likely to prevail than $(3,4)$.

<table>
<thead>
<tr>
<th></th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>3, 4</td>
<td>0, 0</td>
<td>0, 0</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0, 0</td>
<td>2, 2</td>
<td>1, 4</td>
</tr>
<tr>
<td>$a_3$</td>
<td>0, 0</td>
<td>4, 1</td>
<td>4, 3</td>
</tr>
</tbody>
</table>

Table 2

A similar calculation pertains to the situation portrayed in Table 1 if we suppose that country 3 gains some nominal amount from its coalition partner whenever it participates in an initial threat -- there are twice as many routes to $(E_1, E_1, E_1)$ as there are to $(E_2, E_2, E_2)$. The realist's objection to neoliberalism, then, can be restated thus: Although the equilibrium neoliberal institutionalists postulate require explicit coordination, the absence of effective coordination is more likely to yield realist's scenario than it is to any other outcome. Understanding this, states naturally prepare for a competitive and less than wholly cooperative or benign environment.

The game in Table 1 and its n-country counterparts, however, cannot resolve matters in favor of one side or the other until other matters are considered, including the costs of conflict, the positive externalities that accrue to all states from cooperative action, incomplete information, deception, and misperception. Neoliberalism continues to have available to it the response that not only is a cooperative equilibrium attractive because convergence it can be made mutually beneficial, and not only do events reveal that the requisite coordination is feasible, but the mechanisms of coordination can also expand the opportunities for convergence to that equilibrium. On the other hand, barring a compelling argument to the contrary, realists are justified in arguing that prudent states will be concerned that the promise of a wholly cooperative equilibrium is only that -- a promise -- and that those who fail to make appropriate preparations for a more conflictual system will be disadvantaged. These preparations, in turn, establish a set of beliefs that move outcomes away from those that neoliberals envision as equilibria. That is, the supposition of a conflictual environment may be more readily sustained as a self-fulfilling prophecy than a wholly cooperative one.

RUSSIA'S NUCLEAR STRATEGY: 90'S AND AFTER

A. PIONTKOWSKY, A. SKOROKHODOV
Strategic Studies Centre
Institute for System Analysis
117312 Moscow, Russia

In this paper we will try to review some characteristic features of the dramatic changes to the world structure, which made a number of traditional military-strategic conceptions obsolete and counter-productive.

1. Military-Political Stability Structures In The Cold War Period

Let us recall as a starting point those postulates which the military-strategic equilibrium in Europe and in the world were based on during the decades of the cold war. First and foremost the confrontation itself was generated by an ideologically-charged geopolitical conflict of two nuclear superpowers, their allies and satellites. Conventional forces of the antagonists faced each other on the European theatre and from time to time they got involved in local conflicts in different parts of the world either directly or more frequently through proxies.

The basic strategic confrontation between NATO and Warsaw Treaty countries has always been interpreted (whether explicitly or not) in the context of an escalation ladder (Fig.1). Both sides were modernizing and deploying weapons designed for potential use at each of the stages of escalation.

The 1972 agreements sealed the concept of strategic parity between the USSR and the US which has been interpreted as the impossibility of winning a full-scale nuclear war. MAD conditions (second-strike capability for both sides) ensured stability at the level 4 of the escalation ladder. But the problem of stability was not restricted only to the fulfillment of MAD conditions. In the author's view, the stability should be regarded as a multi-level conception. Such approach required the implementation of conditions which ensure stability at each level of the ladder of a potential conflict (Fig.1).

Let's suppose that these conditions are not observed for some of these levels, that is, some side can achieve a sensible advantage in military or political terms at some particular level of conflict. In this case, the other side would be tempted to consider an escalation, that is, transition to the next level so as to deprive the enemy of the real (or imaginary) advantage which it enjoys at the lower stage of the conflict.