DOCUMENT RESUME

ED 266 993 SO 016 962

AUTHOR	Brams, Steven J.; Kilgour, D. Marc
TITLE	Rational Deescalation.
SPONS AGENCY	Ford Foundation, New York, N.Y.; National Science Foundation, Washington, D.C.
PUB DATE	Aug 85
GRANT	845-0354; SES84-08505
NOTE	36p.; Paper presented at the Annual Meeting of the American Political Science Association (81st, New Orleans, LA, August 29-September 1, 1985).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Conflict; Conflict Resolution; Decision Making; *Disarmament; *Educational Games; Games; Interdisciplinary Approach; *Mathematical Models; *Political Science; Postsecondary Education; Probability; Problem Solving; Simulation
IDENTIFIERS	*Prisoners Dilemma Game; Rules and Regulations

ABSTRACT

Based on the game "Prisoners' Dilemma," a game-theoretical model of the arms race suitable for postsecondary level mathematics and/or political science students is developed in which two players can initially choose any level of arms development. The purpose of the game is to show under what conditions deescalation rather than escalation is a rational response to the burdens that an unrestricted arms race imposes on both sides. The strategic problem that the players face is to choose both an initial level of action (with an associated escalation probability) and a subsequent level of response (with an associated retaliation probability). The higher the level of arming, the greater the probability that the choice will be viewed as escalatory. A matrix representation and the rules of the game are provided in the text which also explains the payoffs, strategic choices, and their interpretations. Quantitative, sequential choices define the game, which contains an Escalation Equilibrium analogous to the non-cooperative outcome in "Prisoners' Dilemma," The game also contains a Deescalation Equilibrium, which is analogous to the cooperative outcome in "Prisoners' Dilemma," except that it is stable. Separate sections provide an introduction, a description of "Prisoners' Dilemma" in relation to the superpower arms race, a discussion of the Deescalation Game and suggestions for rational play, and conclusions. An appendix presenting the details of the game's analysis and calculations of the players maximum strategies and values concludes the paper. (LH)

RATIONAL DEESCALATION

Steven J. Brams*
Department of Politics
New York University
New York, NY 10003

D. Marc Kilgour**
Department of Systems Design Engineering
University of Waterloo
Waterloo, Ontario
Canada N2L 3G1

*Steven J. Brams gratefully acknowledges the financial support of the Ford Foundation under Grant No. 845-0354 and the National Science Foundation under Grant No. SES84-08505.

**D. Marc Kilgour gratefully acknowledges the financial support of the Natural Sciences and Engineering Research Council of Canada under Grant No. A8974.

Prepared for delivery at the 1985 Annual Meeting of The American Political Science Association, The New Orleans Hilton, August 29 - September 1, 1985.

ABSTRACT

A game-theoretic model of arms races, based on Prisoners' Dilemma, is developed in which two players can initially choose any level of arming. The higher the level, the greater the probability that this choice will be viewed as escalatory by the other player, who can retaliate subsequently if his own initial choice was not considered escalatory.

The quantitative, sequential choices define a Deescalation Game, which contains an Escalation Equilibrium analogous to the noncooperative outcome in Prisoners' Dilemma. More auspiciously, this game also contains a Deescalation Equilibrium, which is analogous to the cooperative outcome in Prisoners' Dilemma, except that it is stable (i.e., a Nash equilibrium).

The latter equilibrium is better for both players than the Escalation Equilibrium. Moreover, unlike Prisoners' Dilemma, either player can initiate a move from the Pareto-inferior Escalation Equilibrium to the Pareto-superior Deescalation Equilibrium. The initial step is costless and induces subsequent rational moves that benefit both players, eventually leading to the Deescalation Equilibrium. The relevance of this analysis to the superpower arms race is discussed.

1. Introduction

The prevention of nuclear war is surely the most daunting problem facing the world today. The road to such a war, should one ever occur, will probably not be a "bolt from the blue" -- say, a massive nuclear strike by one superpower against the other and its allies. Rather, it is likely to erupt in a period of extreme crisis occasioned by a conventional conflict in which one side, facing imminent defeat, decides it has no recourse except to use nuclear weapons, or threaten their use. The conflict need not even involve a nuclear power directly but only as an ally that feels compelled to come to the aid of a threatened partner.

An arms race may trigger such a conflict. As tensions mount in such a race, verbal threats and provocative military maneuvers may precipitate war, which may then escalate as allies become involved. Then, if one side's position or very existence is jeopardized, there is a possibility that it would introduce or threaten to introduce nuclear weapons to try to avert disaster.

In a previous paper, we showed what kinds of probabilistic threats appeared to be optimal to prevent confrontation situations that could be modeled by the game of Chicken from exploding and wreaking destruction on both sides.[1] In this paper we shift the focus back to the progenitor of many crises that produce such perilous showdowns -- namely, arms races. Our aim is to show under what conditions deescalation rather than escalation is a rational response to the staggering burdens that an unrestrained arms race imposes on both sides.

For this purpose, we start from a model of an arms race based on the infamous game of Prisoners' Dilemma, but we make major emendations in the simple 2 x 2 version of this game to permit the players

(1) initially to choose any level of provocation along a disarm-arm dimension; and

(2) subsequently to retaliate at any level to a provocation if it is viewed as escalatory, or noncooperative, provided their initial choice was considered cooperative.

We interpret these initial and subsequent actions in terms of probabilities of escalation, and retaliation for escalation, which we assume each player chooses at the beginning of play from an infinite strategy space (specifically $[0,1] \times [0,1]$).

After calculating maximin strategies in this continuous game, we demonstrate that it contains two Nash equilibria, or stable outcomes. The one we call the "Escalation Equilibrium" corresponds to the unique Nash equilibrium in the classical 2 × 2 version of Prisoners' Dilemma (to be described in section 2). The other, which we call the "Deescalation Equilibrium," involves each side's cooperating initially with certainty but retaliating with a specified probability to noncooperation by the other side. Although the Deescalation Equilibrium is a promising addition to the finite version of Prisoners' Dilemma, it does not answer the nagging question of how one extricates oneself from the Escalation Equilibrium of the Deescalation Game, which by definition neither player has an incentive to depart from unilaterally.

The superpowers seem stuck at this noncooperative equilibrium today. Happily for the players in the Deescalation Game, however,

there is a trajectory or path by which they can travel from the Escalation Equilibrium to the Deescalation Equilibrium. Surprisingly, either player can initiate such a sequence with impunity, triggering subsequent rational moves by the players that redound to the benefit of both, eventually reaching the Deescalation Equilibrium. We briefly compare this resolution of the trying dilemma posed by arms races -- particularly that between the superpowers -- to other game-theoretic approaches, arguing that our model offers a more realistic representation of the superpower arms race than others, some of which, nonetheless, suggest a similar resolution to our own.

2. Prisoners' Dilemma and the Superpower Arms Race

The 2 x 2 game of Prisoners' Dilemma, in which two players (Row and Column) each have two strategies and can rank the resulting four outcomes from best (4) to worst (1), is illustrated in Figure 1. The first number in the ordered pair that specifies each outcome is assumed

FIGURE 1

OUTCOME MATRIX OF PRISONERS' DILEMMA

		Column	
		Cooperate (C)	Do not cooperate ($\bar{C}$)
Row	Cooperate (C)	(3,3) Compromise	(1,4) Column wins
	Do not cooperate ($\bar{C}$)	(4,1) Row wins	(2,2) Trap

Key: (x,y) = (rank of Row, rank of Column)

4 = best; 3 = next best; 2 = next worst; 1 = worst

Circled outcome is Nash equilibrium

to be the ranking of Row, and the second number the ranking of Column.

Thus, the outcome (3,3) is next-best for both players, but no presumption is made about whether this outcome is closer to each player's best (4) or next-worst (2) outcome. (Later we assume that players can assign numerical values, or cardinal utilities, to the outcomes.) Because the two players do not rank any two outcomes the same -- that is, there are no ties between ranks -- this is a strictly ordinal game.

The short-hand verbal descriptions given in Figure 1 for each outcome are intended to convey the qualitative nature of the outcomes, based on the players' rankings. Because this game is symmetrical (i.e., the players rank the two outcomes along the main diagonal the same, and the ranks of the off-diagonal outcomes are mirror images of each other), the two players face the same problems of strategic choice.

Each player is assumed to be able to choose between the strategies of cooperation (C) and noncooperation ($\bar{C}$). Each obtains his next-best outcome of 3 ("compromise") by choosing C -- if the other player also does -- but both have an incentive to defect from this outcome to obtain their best outcomes of 4 by choosing $\bar{C}$ when the other player chooses C. Yet, if both choose $\bar{C}$, they bring upon themselves their next-worst outcome ("trap"). On the other hand, should one player choose $\bar{C}$ when the other chooses C, the $\bar{C}$-player "wins" by obtaining his best outcome (4) at the same time that the C-player suffers his worst (1) outcome.

The dilemma in Prisoners' Dilemma is that both players have a dominant strategy of choosing $\bar{C}$: whatever the other player chooses (C or $\bar{C}$), $\bar{C}$ is better. But the choice of $\bar{C}$ by both leads to (2,2), which is Pareto-inferior since it is worse for both players than (3,3). In addition, (2,2) is a Nash equilibrium -- that is, neither player has an incentive to deviate unilaterally from this outcome because he would do worse, or at least not better, if he did -- whereas (3,3) is not stable in this sense.[2]

Presumably, rational players would choose their dominant, or unconditionally best, strategies of $\bar{C}$, leading to the Pareto-inferior (2,2) Nash equilibrium. Because of its stability, neither player would be motivated to depart from (2,2), even though (3,3) is a better outcome for both than (2,2). In fact, (3,3) is Pareto-superior since any other outcome which is better for one player is worse for the other. Should (3,3) somehow manage to be chosen, however, both players would be tempted to depart from it to try to do still better, rendering it unstable.

Other concepts of equilibrium distinguish (3,3) as a stable outcome, but the rules of play they assume require that players act non-myopically or farsightedly; moreover, they do not rule out (2,2) as stable, too.[3] If threats are possible in repeated play of Prisoners' Dilemma under still different rules, however, the stability of (3,3) is reinforced.[4] Preplay negotiations can also lead to the (3,3) outcome.[5]

We shall introduce shortly the notion of a probabilistic threat as well as a probabilistic initial strategy choice. But before doing

that, it is worth pointing out that Prisoners' Dilemma is not a constant-sum game, in which what one player wins the other player loses. Rather, it is a variable-sum game because the sum of the players' payoffs at each outcome (if measured cardinally by utilities rather than ordinally by ranks) may vary.

A variable-sum game is also a game of partial conflict, as opposed to a (constant-sum) game of total conflict in which one player cannot benefit except at the expense of another. Prisoners' Dilemma is not a game of total conflict, for both players do worse at (2,2) than at (3,3), which perhaps belies the name "partial conflict" since (2,2) is, unfortunately for the players, both the product of dominant strategies and the unique Nash equilibrium. It is hard to see how the players can avoid it without risking their worst outcomes.

As a model of the superpower arms race, this recalcitrant game supports the logic of both sides' arming (noncooperation), even though this outcome is Pareto-inferior to their disarming or, less ambitiously, pursuing more limited policies of arms control (cooperation). Cooperation is problematic because, as Garthoff put it, "they [the Soviets] would like to have an edge over us [at (1,4) if they are Column], just as we would like to have an edge over them [at (4,1) if we are Row]."[6]

Prisoners' Dilemma elegantly captures this temptation to defect from the cooperative outcome that, it seems, has inexorably led the superpowers into a very costly arms race. Nevertheless, at the same time that it offers a striking explanation of the fundamental intractability of this continuing conflict -- based only on the rational behavior of the players -- it drastically simplifies the realities of the

superpower arms race.

Prisoners' Dilemma omits two salient features of the superpower arms race that we believe need to be incorporated into a more realistic model, the focus of our attention in the remainder of this paper. First, a player does not make a dichotomous choice between cooperation (disarming) or noncooperation (arming) but rather chooses a kind or level of action, or arms expenditures, that may be interpreted as being escalatory or deescalatory. Second, in response to an initial choice viewed as escalatory by his opponent, a player who was not viewed as escalatory at the start may subsequently choose a new level of expenditures that itself may be seen as escalatory or not.

In effect, players in the Deescalation Game that we shall describe in the next section can choose initially to provoke or not provoke an opponent at any level; if provoked, they can retaliate or not retaliate at any level. Thereby we incorporate into our model not only quantitative choices of any level of cooperativeness/noncooperativeness but also sequential choices that permit players to respond if provoked. The additional structure of quantitative and sequential choices in Prisoners' Dilemma not only better mirrors, in our view, real-world choices in the superpower arms race, but it also will enable us to derive conditions under which it is rational for the players to be cooperative in the Deescalation Game and thereby escape the (2,2) trap.

3. The Deescalation Game

The Deescalation Game is defined by the following rules:

(1) The final outcome will be one of the four outcomes of Prisoners' Dilemma. The payoffs are the same as those

of Prisoners' Dilemma, except that cardinal utilities replace ordinal rankings. Thus r_4 and c_4 signify the highest payoffs for Row and Column, respectively, r_1 and c_1 the lowest, etc.

(2) The players do not choose initially between C and $\bar{C}$, as in Prisoners' Dilemma, but instead choose (unspecified) actions that have associated a nonescalation probability (s for Row and t for Column) and a complementary escalation probability (1-s for Row and 1-t for Column). With these probabilities, their actions will be interpreted as cooperative (C) and noncooperative ($\bar{C}$) strategy choices, respectively.

(3) If <u>both</u> players' initial choices are perceived as the same, the game ends at that position (i.e., CC or $\overline{CC}$). If one players' choice is perceived as C and the other's as $\bar{C}$, the former player then chooses <u>subsequent</u> actions with an associated nonretaliation probability (p for Column and q for Row) and a complementary retaliation probability (1-p for Column and 1-q for Row). With the retaliation probability, the conflict is escalated further to the final outcome $\overline{CC}$; otherwise it remains as before (at $C\bar{C}$ or $\bar{C}C$).

(4) The players choose their escalation probabilities and retaliation probabilities before play of the game. Play commences when each player simultaneously chooses initial actions that may be interpreted as either C or $\bar{C}$, with associated escalation probabilities. One player may then choose subsequent actions, according to rule 3, with the associated retaliation probability specified at the beginning of play.

The Deescalation Game is represented in Figure 2. Note that

FIGURE 2

MATRIX REPRESENTATION OF DEESCALATION GAME

		Column	
		t	$1-t$
Row	s	(r_3,c_3)	$q(r_1,c_4)+(1-q)(r_2,c_2)$ $= ((1-q)r_2, q+(1-q)c_2)$
	$1-s$	$p(r_4,c_1)+(1-p)(r_2,c_2)$ $= (p+(1-p)r_2, (1-p)c_2)$	(r_2,c_2)

<u>Key</u>: (r_i,c_j) = (payoff to Row, payoff to Column)

r_4,c_4 = best; r_3,c_3 = next best; r_2,c_2 = next worst; r_1,c_1 = worst

s,t = probabilities of nonescalation; p,q = probabilities of nonresponse

Normalization: $0 = r_1 < r_2 < r_3 < r_4 = 1$; $0 = c_1 < c_2 < c_3 < c_4 = 1$

besides the fact that the initial strategy choices of the two players are probabilities (with assumed underlying actions), rather than actions (C and $\bar{C}$) themselves, this payoff matrix differs from the Figure 1 outcome matrix in having expected payoffs rather than (certain) payoffs as its off-diagonal entries. This is because we assume that if one player is perceived to escalate, the other player's (probabilistic) retaliation will be virtually instanteous, so it is proper to include in the off-diagonal entries a combination of payoffs -- reflecting both possible retaliation and possible nonretaliation -- by means of an expected value.

We assume, of course, that $0 \leq s, t, p, q \leq 1$ because they represent probabilities. To simplify subsequent calculations, we normalize the payoffs of the players so that the best and worst payoffs are 1 and 0, respectively. Hence,

$$0 = r_1 < r_2 < r_3 < r_4 = 1;$$

$$0 = c_1 < c_2 < c_3 < c_4 = 1.$$

Because we assume the escalation and retaliation probabilities are chosen independently by the players, the expected payoffs for Row and Column are simply the sums of the four payoffs (expected payoffs) in the Figure 2 matrix, each multiplied by the probability of its occurrence:

$$E_R(s,q;t,p) = str_3+(1-s)t[p+(1-p)r_2]+s(1-t)(1-q)r_2+(1-s)(1-t)r_2; \quad (1)$$

$$E_C(t,p;s,q) = stc_3+(1-s)t(1-p)c_2+s(1-t)[q+(1-q)c_2] + (1-s)(1-t)c_2. \quad (2)$$

The introduction of escalation and retaliation probabilities into the expected-payoff calculations requires some explanation and interpretation. Essentially we assume that every initial action that a player may take carries with it a probability of being interpreted as escalatory by his opponent and, if it is, possibly drawing a response. This response, like the initial action that may escalate the conflict, is probabilistic in that it is not certain to constitute retaliation. Rather, both initial actions and subsequent responses have probabilities associated with their being viewed as escalatory and retaliatory, respectively, thereby leading to different outcomes in the game.

Thus, for example, the probability that Row will provoke Column

by his choice is some escalation probability 1-s. If Column is provoked, and providing that he did not also provoke Row initially (with escalation probability 1-t), Column will respond with a subsequent action that (further) escalates the conflict to mutual noncooperation with retaliation probability 1-p.

If neither player provoked the other [with joint probability st] or each provoked the other [with joint probability (1-s)(1-t)], then the retaliation probabilities never come into play, for we assume there is (i) no need to retaliate for the choice of CC and (ii) no possibility of retaliating for the choice of $\overline{CC}$. Hence, the first and last terms of E_R and E_C given by (1) and (2) do not include retaliation probabilities.

The strategic problem that the players face is to choose both an initial level of action (with an associated escalation probability) and a subsequent level of response (with an associated retaliation probability). We assume, in interpreting probabilities in the Deescalation Game, that the higher the level of (initial) escalation or (subsequent) retaliation, the greater the probability that these actions will be perceived as escalatory/retaliatory. Formally, then, we assume a linkage between the degree of escalation/retaliation and the probability that it will be interpreted as such by one's opponent.

When making their choices of initial and subsequent levels of action (and hence probabilities) before play of the game, we assume that the players know that their opponents will judge the level of these actions exactly as they do themselves. Consequently, each player's probability assessment of each level of action will coincide with his opponent's. Thus, the players can assume that the four escalation

and retaliation probabilities in the two expected-payoff equations are identical.

These probabilities become common knowledge once the levels of action (with which they are in correspondence) are selected in the Deescalation Game. This information that is introduced into the play of the game does not mitigate the problem of choosing the probabilities -- in ignorance of one's opponent's choices -- before play commences.

With respect to the retaliation probabilities, it should be noted that they are not assumed to be a function of the escalation probabilities. To be sure, the higher one player's escalation probability, the more likely his opponent's retaliation probability will come into play, and hence the more likely retaliation will occur. But since the retaliation as well as the escalation probabilities are chosen before the start of the game, the former (for one player) are necessarily independent of the latter (for the other).[7]

It is fair to ask why retaliation is ever a problem in Prisoners' Dilemma; it would seem, on the contrary, always to be a rational response by a player once he perceives his opponent has escalated the conflict by choosing $\overline{C}$. In the case of Row, for example, if Column has escalated to (4,1), he (Row) does immediately better by moving the game to (2,2), from which neither player would have an incentive to depart, as we showed earlier.

This logic does not hold in Chicken, which reverses the two worst outcomes of the players in Prisoners' Dilemma. Thus, the $C\overline{C}$ outcome is (4,2), and $\overline{CC}$ is (1,1), in Chicken. Now Row, at (4,2), would appear

irrational in threatening to retaliate by moving to (1,1), which is the principal problem we analyze in our quantitative, sequential analysis of Chicken as a model of deterrence.[8]

In theory, players solve this problem by precommitting themselves to carry out threats, despite the irrationality of doing so for the threatener. In practice, one of us has argued, this takes form in terms of the operational procedures the superpowers have set in place to respond, if attacked, to a nuclear first strike.[9]

We assume that the same kind of precommitments to retaliate can be made in the case of the Deescalation Game. In this game, however, it is the combination of escalation and retaliation probabilities that may make initial escalation for, say, Row, from (r_3,c_3) -- rather than subsequent retaliation by Column from (r_4,c_1) -- irrational.

In the absence of an adequate precommitment to retaliate on the part of Column, Row may think that he can impose a small probability of escalation without serious repercussions, although this subjects Column to his worst outcome. But in our model Column's retaliation probability assures Row that "too high" an escalation probability would be irrational for Row, because it would carry the game from (r_3,c_3) to (r_2,c_2).

Put another way, a precommitment to retaliate with a probability above a particular level -- to be specified later -- renders initial escalation unprofitable. This is a precommitment that seems unproblematic, unlike in Chicken. More relevant to the problem of commitment in Prisoners' Dilemma is a player's ability to precommit himself never

to escalate, which we show has a surprising and salutary consequence in the Deescalation Game under certain conditions. In either event, we assume that players can precommit themselves to strategies -- escalatory, probabilistic, or certain -- so that there is never any doubt on the part of an opponent that they will be implemented.

The quantitative questions we next address in our game-theoretic analysis are what combinations of escalation and retaliation probabilities (i) maximize the payoff a player can guarantee himself of, whatever his opponent does, (ii) lead to Nash equilibria, and (iii) induce cooperative choices that allow players to escape the trap of mutual noncooperation. We in fact show that there are escape routes, which is why in the title of this paper we call deescalation "rational" and refer to our extension and refinement of Prisoners' Dilemma as the Deescalation Game.

4. Rational Play in the Deescalation Game

Consider the Deescalation Game from Row's vantage point. In Prisoners' Dilemma, by choosing his dominant strategy $\overline{C}$, he can guarantee himself a payoff of at least r_2, whatever Column chooses. This guaranteed minimum is Row's security level. By comparison, because Row chooses probabilities of certain actions and reactions, rather than strategies themselves, in the Deescalation Game, it is by no means obvious what he can guarantee himself of, independent of Column's (probabilistic) choices.

In the Appendix we show that in fact Row can guarantee himself the same value he can in Prisoners' Dilemma, namely r_2. We do this by calculating, first, the value of Row's expected payoff, E_R, when Column,

by his choice of t and p, makes it as small as possible. We then assume that Row, by his choice of s and q, seeks to maximize this minimum value of E_R. The resulting maximin of E_R is Row's security level, for it is the value that Row can assure himself of even if Column seeks to minimize E_R.

There are two ways that Row can guarantee himself at least his maximin value: by choosing any of his strategies with (i) $s = 0$ and q arbitrary, or (ii) $q = 0$ and s arbitrary. In the former case, Row escalates with certainty; if Column also escalates or retaliates with certainty, Row obtains r_2, otherwise a higher expected payoff (because it includes r_4 with some positive probability when Column does not retaliate). In the latter case, Row never escalates but always retaliates; if Column escalates with certainty, Row ensures himself of r_2; otherwise his expected payoff is greater when Column does not (because it includes r_3 with some positive probability).

Only when Column always escalates ($t = 0$) does Row suffer his security level of r_2 when he chooses any of his maximin strategies. When $t > 0$, by contrast, Row always can do better than r_2. In this case, however, which maximin strategy serves him best depends on Column's choice of p, as shown in the Appendix. Column's maximin strategies and security level are analogous, because of the symmetry of the Deescalation Game.

Maximin strategies, especially in variable-sum games like the Deescalation Game, are conservative in the extreme, for they presume that one's opponent desires to minimize one's payoff, even if it hurts him to do so. By contrast, in constant-sum games maximin strategies

(which are also minimax strategies -- minimize an opponent's maximum payoff) are more defensible because hurting an opponent always helps oneself.

If perhaps overly conservative, however, each player's maximin strategy of escalating with certainty,

$$s = 0,\ q \text{ arbitrary};\ t = 0,\ p \text{ arbitrary}, \tag{3}$$

results in a Nash equilibrium, which we call the Escalation Equilibrium. This equilibrium, of course, corresponds to the unique Nash equilibrium at (r_2, c_2) in Prisoners' Dilemma. Since a player who escalates forgoes any opportunity to retaliate in the Deescalation Game, the Escalation Equilibrium is independent of whatever retaliation probabilities the players choose in this game.

Auspiciously, the Escalation Equilibrium is not unique in the Deescalation Game. As shown in the Appendix, there is a second Nash Equilibrium,

$$s = 1,\ q \leq \frac{c_3 - c_2}{1 - c_2}\ ;\ t = 1,\ p \leq \frac{r_3 - r_2}{1 - r_2}\ , \tag{4}$$

which we call the Deescalation Equilibrium. It says that a player (say, Column) will never escalate ($t = 1$); but in response to escalation by Row, sometimes Column will not retaliate (with nonretaliation probability $p \leq \frac{r_3 - r_2}{1 - r_2}$) and other times he will (with retaliation probability $1-p > \frac{r_3 - r_2}{1 - r_2}$). More accurately, Column will choose actions in response to any prior (escalatory) actions by Row with a retaliation probability

greater than the threshold value, $\frac{r_3-r_2}{1-r_2}$.

Why this threshold value? As shown in the Appendix, this is the value that makes Row's expected payoff, E_R, independent of his choice of s. If Column's retaliation probability exceeds this threshold, however, Row would (irrationally) decrease E_R should he deviate from $s = 1$ (i.e., by choosing $s < 1$). Hence, given Column's retaliation probability is above the threshold value, Row maximizes E_R by choosing $s = 1$ and will not have an incentive to deviate. For analogous reasons, Column will not deviate from the Deescalation Equilibrium, rendering the resulting outcome stable. This outcome, of course, corresponds to the (r_3,c_3) compromise in Prisoners' Dilemma.

Perhaps the most significant feature of the Deescalation Game is that it makes the compromise outcome stable, even though this outcome is highly unstable in the underlying Prisoners' Dilemma game. This stability is due to the fact that the values of the two off-diagonal outcomes of Prisoners' Dilemma, which give Row 4 at one outcome (lower left in Figure 1) and Column 4 at the other (upper right in Figure 1), are diminished to expected values less than r_3 and c_3 by the Deescalation Equilibrium strategies. The high probability of retaliation substantially dilutes the value of a win, r_4 or c_4, with the value of the much less desirable trap outcome, r_2 or c_2. Meanwhile, the payoffs at compromise, r_3 and c_3, are unaffected in the passage from Prisoners' Dilemma to the Deescalation Game, making them, in relative terms, the most attractive when retaliation is likely. When both sides are prepared to retaliate, nonescalation is each player's best strategy, and compromise the mutually best outcome.

We demonstrate in the Appendix, using an exhaustive search for Nash equilibria, that there are none other than the Escalation Equilibrium and the Deescalation Equilibrium in the Deescalation Game. One effect, then, of high retaliation probabilities in this game is to transform the cooperative outcome from a next-best nonequilibrium (in the underlying Prisoners' Dilemma game) to a best equilibrium (in the Deescalation Game) -- without changing the payoffs to the players. The Deescalation Equilibrium, however, is not the product of dominant strategies in the Deescalation Game, for one player's Deescalation Equilibrium strategy is best if the other player chooses his, but definitely not best if the other player chooses certain other strategies.

At the same time that (r_3,c_3) is stabilized in the Deescalation Game, the stability of (r_2,c_2) is called into question -- even though it corresponds to a Nash equilibrium. To see why, assume that the players begin at the outcome defined by

$$s = 0,\ q = q_0;\ t = 0,\ p = p_0;\ \text{or}\ (0,q_0;0,p_0)\ , \tag{5}$$

where p_o and q_o are arbitrary. Since the players escalate with certainty, they receive payoffs (r_2,c_2) at this Escalation Equilibrium.

Now let Column change his strategy to $t = t_0$, $p = 0$, so the strategies become

$$(0,q_0;t_0,0)\ , \tag{6}$$

where Column escalates with arbitrary probability $t_0 > 0$ and always retaliates $(p = 0)$. The players still receive (r_2,c_2), but Column has changed his Nash-equilibrium strategy (i.e., probabilities) without cost

to himself.

If Row next changes his Nash-equilibrium strategy to never escalate but always retaliate, giving

$$(1,0;t_0,0), \tag{7}$$

his expected payoff will be

$$E_R(1,0;t_0,0) = t_0 r_3 + (1 - t_0) r_2 \;.$$

This is clearly better for him (since $t_0 > 0$) than r_2 that he receives at the Deescalation Equilibrium and at (6), so he would be motivated to switch from (6) to (7). In fact, switching from $s = 0$, $q = q_0$ to $s = 1$, $q = 0$ maximizes Row's expected payoff as long as Column plays $t = t_0, p = 0$.

But now, if $t_0 < 1$, Column can respond to the situation at (7) by changing his strategy to never escalate but always retaliate, too, giving

$$(1,0;1,0). \tag{8}$$

This raises E_C for him from

$$E_C(t_0,0;1,0) = t_0 c_3 + (0 - t_0) c_2$$

at (7) to

$$E_C(1,0;1,0) = c_3$$

at (8), which is a Deescalation Equilibrium with payoffs (r_3, c_3) for both players. Again, Column's move from (7) to (8) maximizes Column's return, assuming Row's strategy is fixed.

Thereby the players can move progressively along the path defined by

$$(5) \xrightarrow[\text{to Column}]{\text{Costless}} (6) \xrightarrow[\text{to Row}]{\text{Beneficial}} (7) \xrightarrow[\text{to Column}]{\text{Beneficial}} (8),$$

with only the first step that triggers the process not positively beneficial to the player (Column) who makes the initial move from the Escalation Equilibrium. But it is a costless change for Column,[10] so presumably he will make it if he anticipates that it will trigger the subsequent (beneficial) moves by Row and Column, respectively.

Indeed, the "trigger condition" can be relaxed to $t_0 > 0$, $p_0 < \frac{r_3-r_2}{1-r_2}$ at (6) in the sense that any such (t_0,p_0) chosen by Column would motivate Row to choose $s = 1$, $q = 0$ at (7). However, use of any p_0 satisfying $0 < p_0 < \frac{r_3-r_2}{1-r_2}$ would reduce (temporarily) Column's payoff to

$$E_C(t_0,p_0;0,q_0) = c_2 - t_0p_0c_2 .$$

As noted previously, $p_0 = 0$ is costless, so the (5)--->(6)--->(7)--->(8) path is the most persuasive -- no player would ever suffer any loss in departing from his Nash-equilibrium strategies, making the need for irrevocable precommitments less. Obviously, the roles of the players that are indicated above can be reversed to trace another path from the Escalation to the Deescalation Equilibrium.

It is interesting to note in the Deescalation Game that it is the Escalation Equilibrium which exhibits some instability, for a costless perturbation by one player induces an immediate shift away from the Escalation Equilibrium toward the Deescalation Equilibrium. The perturbation triggering the shortest path to deescalation is $t = t_0 = 1$, $p = 0$. It is also noteworthy that this particular perturbation strategy -- never escalate, but always retaliate -- bears a strong resemblance to the tit-for-tat strategy recommended by Axelrod for iterated Prisoners' Dilemma.[11]

We do not, however, assume repeated conditioned play of Prisoners' Dilemma but only an ability to retaliate for an initial untoward action of an opponent. Remarkably, this retaliatory ability turns out to be sufficient both to deter an opponent and induce him to shift to the same deterrent strategy, from which he also will benefit. Once both players have adopted -- and precommitted themselves to -- this posture, their payoffs at the Deescalation Equilibrium are not only better for both than at the Pareto-inferior Escalation Equilibrium but they are also highly stable: both players would do immediately worse by deviating from $s = t = 1$ (never escalate) because of possible retaliation.

However, they can afford to raise $p = q = 0$ (certain retaliation) up to the threshold values given earlier [see (4)] for the Deescalation Equilibrium -- thereby making retaliation less than certain -- and still maintain stability. In other words, each player's retaliatory threat need only be probabilistic -- or perhaps a certain equivalent (i.e., a lower-level retaliatory action) that signals more serious

retaliation is possible. This possibly may make it more credible, as we argued in the Deterrence Game based on Chicken,[12] except that each player in the Deescalation Game -- and Prisoners' Dilemma on which it is based -- has an evident incentive to carry out a threat because he immediately benefits, even if the resulting outcome is Pareto-inferior.

In fact, whether the underlying game is Chicken or Prisoners' Dilemma, the purpose of threatening retaliation is to deter an opponent from deviating from (r_3,c_3), whether or not it is costly to carry out a threat once he does. Thus, the logic underlying threats that stabilize (r_3,c_3) in both games is exactly the same. But beyond the use of retaliatory threats to render this outcome an equilibrium in the Deescalation Game, we believe even more hopeful is our finding that there is a costless, and in general beneficial, way for the players to escape the (r_2,c_2) trap and reach the (r_3,c_3) compromise outcome in this game.

6. Conclusions

Arms races are not only terribly costly but also may increase the probability of war between two states under certain conditions.[13] When these states are the superpowers, and the costs are in the hundreds of billions of dollars -- with nuclear holocaust a possible consequence of fighting that may erupt in an extreme crisis -- then there is good reason to ponder how to deescalate the superpower arms race.

The arms race has persisted, we believe, because both sides see it as a Prisoners' Dilemma, with little hope of escaping the (2,2) trap. To be sure, the superpowers have been able to reach some arms-control agreements. For the most part, however, they have been of a

very limited nature, and even some of these seem in trouble today because of mistrust and suspicions of cheating as well as new technological developments.[14]

The Deescalation Game, insofar as it reflects the quantitative choices about arms expenditures that each side makes -- and the possible responses to the other side's perceived expenditures -- gives some basis for being sanguine. First, by stabilizing the compromise outcome (r_3,c_3), and, second, by showing that there is a rational path from the trap (r_2,c_2) to the compromise (r_3,c_3), it suggests how the Deescalation Equilibrium might supplant the Escalation Equilibrium as the rational outcome of this game. Essentially each side must precommit itself to respond to, but not initiate, escalation. Retaliation, while rational in Prisoners' Dilemma (as opposed to Chicken) once one side has escalated, nevertheless hurts both players at the resulting Escalation Equilibrium, at least compared to the Deescalation Equilibrium.

The fact that the players can extricate themselves from the Escalation Equilibrium by a series of rational moves and responses in the Deescalation Game is what makes this game a much more pleasant one to play than Prisoners' Dilemma. If it is also a more realistic model of Prisoners' Dilemma-type conflicts such as the superpower arms race, then it suggests a solution, at least at a conceptual level, to the pathology of such conflicts when they have the quantitative, sequential character of the Deescalation Game.

We think that arms races, particularly the arms race between the superpowers, have this character. It requires a leader of imagination

to commit himself to deescalatory policies, though to be effective our model suggests that these need to be combined with the threat of possible retaliation if the other side does not follow suit. Given such a carrot-stick combination, there is no great daring that this posture demands, because, at least in theory, it is costless. In reality, this may not be entirely so -- for domestic political reasons, among others -- but we believe our model goes a long way toward justifying a more conciliatory posture if the threat of retaliation is also present and real.

APPENDIX

We present the details of our analysis of the Deescalation Game in this Appendix. We begin by calculating the players' maximin strategies and values, and then we determine all Nash equilibria by an exhaustive search.

The rules of the Deescalation Game are given in the text, where the payoffs and strategic choices, and their interpretations, are made explicit. The game is depicted in Figure 2. For convenience, the expected payoffs of Row (R) and Column (C) are repeated here:

$$\begin{aligned} E_R(s,q;t,p) &= str_3+(1-s)t[p+(1-p)r_2]+s(1-t)(1-q)r_2 \\ &\quad + (1-s)(1-t)r_2 \\ &= r_2+st(r_3-r_2)+(1-s)tp(1-r_2)-s(1-t)qr_2; \qquad (1) \\ E_C(t,p;s,q) &= stc_3+(1-s)t(1-p)c_2+s(1-t)[q+(1-q)c_2] \\ &\quad +(1-s)(1-t)c_2 \\ &= c_2+st(c_3-c_2)-(1-s)tpc_2+s(1-t)q(1-c_2). \qquad (2) \end{aligned}$$

To identify Row's maximin strategy, suppose first that s and q are fixed and notice from (1) that

$$\frac{\partial E_R}{\partial t} = s(r_3-r_2)+(1-s)p(1-r_2)+sqr_2 \geq 0,$$

with equality if and only if s = 0 and p = 0. Thus, if Row chooses s > 0,

$$\min_{t,p} E_R(s,q;t,p) = \min_p E_R(s,q;0,p) = \min_p \{r_2-sqr_2\} = r_2(1-sq).$$

Also, since $t \geq 0$, $p \geq 0$, and $r_2 < 1$,

$$\min_{t,p} E_R(0,q;t,p) = \min_{t,p} \{r_2+tp(1-r_2)\} = r_2.$$

Therefore, for all permissible values of s and q,

$$\min_{t,p} E_R(s,q;t,p) = r_2(1-sq),$$

so that Row's maximin value is

$$\max_{s,q} \min_{t,p} E_R(s,q;t,p) = \max_{s,q} \{r_2(1-sq)\} = r_2.$$

Furthermore, Row can achieve his maximin value r_2 by choosing any of his strategies with $s = 0$ (and q arbitrary), or with $q = 0$ (and s arbitrary). It is interesting to note that any of these maximin strategies yields to Row exactly his maximin value when $t = 0$ [see (1)]; if $t > 0$, Row may receive more. Specifically, if $t > 0$ and $p > (r_3-r_2)/(1-r_2)$, a maximin strategy of the form ($s = 0$, q arbitrary) gives Row his best payoff, whereas if $t > 0$ and $p < (r_3-r_2)/(1-r_2)$ Row's preferred maximin strategy is $s = 1$, $q = 0$. If $p = (r_3-r_2)/(1-r_2)$, Row would be indifferent among his maximin strategies because

$$E_R = r_2 + t(r_3 - r_2)-s(1-t)qr_2,$$

which yields the same value, $r_2+t(r_3-r_2)$, in every case.

It follows from the symmetry of the Deescalation Game that Column's maximin value is c_2 and that any of Column's strategies with $t = 0$ or $p = 0$ are maximin strategies, guaranteeing him a payoff of

at least c_2. The properties of Column's maximin strategies are analogous to those of Row's, as discussed above.

We turn now to the search for Nash equilibria in the Deescalation Game. Our search is organized according to the values of s and t at the equilibrium.

Case 1: $t = 0$.

If $t = 0$, then (1) becomes

$$E_R(s,q;0,p) = r_2(1-sq),$$

so that R's best reply to $t = 0$ is either $s = 0$ (and q arbitrary) or $q = 0$ (and s arbitrary). By symmetry, $t = 0$ (and p arbitrary) is also a best reply for C against $s = 0$. It is easy to verify directly that all strategy combinations

$$s = 0,\ q \text{ arbitrary};\ t = 0,\ p \text{ arbitrary}, \tag{3}$$

are equilibria. We call (3) the Escalation Equilibrium, since it is characterized by both players' escalating with certainty. At the Escalation Equilibrium, the outcome of the Deescalation Game is always the trap outcome of the underlying Prisoners' Dilemma game, yielding the players (r_2,c_2).

We now show that (3) are the only equilibria consistent with Case 1 by considering C's response to R's strategy choice $s > 0$, $q = 0$. By (2),

$$E_C(t,p;s,0) = c_2+st(c_3-c_2)-(1-s)tpc_2.$$

If $0 < s < 1$, then

$$\max_{t,p} E_C(t,p;s,0) = \max_t E_C(t,0;s,0)$$

$$= \max_t\{c_2+st(c_3-c_2)\} = c_2+s(c_3-c_2),$$

which occurs at $t = 1$ and $p = 0$. If $s = 1$, this maximum is also $c_2+s(c_3-c_2)$, occurring at $t = 1$. Therefore, C's best reply to $s > 0$, $q = 0$ includes $t = 1$, so that no strategy combination including $s > 0$, $q = 0$ and $t = 0$ (as assumed in Case 1) is an equilibrium.

Case 2: $s = 0$.

By an argument analogous to that for Case 1, the only equilibrium consistent with $s = 0$ is the Escalation Equilibrium (3).

Case 3: $t = 1$.

If $t = 1$, then (1) becomes

$$E_R(s,q;1,p) = r_2+s(r_3-r_2)+(1-s)p(1-r_2)$$

$$= [r_2+p(1-r_2)]+s[r_3-r_2-p(1-r_2)]. \tag{9}$$

From the final expression of (9) it follows that $s = 1$ is R's best reply if

$$p \le \frac{r_3-r_2}{1-r_2} .$$

Symmetry places an analogous condition on q in order that $t = 1$ be C's best response to $s = 1$. It is easy to verify directly that

$$s = 1,\ q \le \frac{c_3-c_2}{1-c_2}\ ;\ t = 1,\ p \le \frac{r_3-r_2}{1-r_2}\ , \tag{4}$$

is an equilibrium, which we call the Deescalation Equilibrium. Observe that the Deescalation Equilibrium always results in the compromise out-

come of the underlying Prisoners' Dilemma game, yielding the players (r_3,c_3).

To show that there are no equilibria other than (4) consistent with Case 3, we note first that, if $p \geq (r_3-r_2)/(1-r_2)$, (9) implies that $s = 0$ would be R's best reply; but we have already proven (Case 2) that there are no equilibria with $s = 0$ and $t = 1$. The only remaining possibility is the combination $0 < s < 1$, $t = 1$, and $p = (r_3-r_2)/(1-r_2)$. But now (2) gives

$$\frac{\partial E_C}{\partial p} = -(1-s)tc_2 < 0$$

since $0 < s < 1$ and $t = 1$. Thus $p = 0$ at any equilibrium with $0 < s < 1$ and $t = 1$, contradicting the inference [from (9)] that $p = (r_3-r_2)/(1-r_2) > 0$.

Case 4: $s = 1$.

As in Case 3, the only equilibrium with $s = 1$ is the Deescalation Equilibrium (4).

Case 5: $0 < s < 1$, $0 < t < 1$

In this case, it follows from (2) that

$$\frac{\partial E_C}{\partial p} = -(1-s)tc_2 < 0,$$

so that $p = 0$ is necessary at any equilibrium. Similarly, $q = 0$ is necessary also. But now (1) shows that

$$E_R(s,0;t,0) = r_2+st(r_3-r_2)$$

which, since $t > 0$, R can maximize only at $s = 1$. Hence there are no equilibria consistent with Case 5.

FOOTNOTES

1. Steven J. Brams and D. Marc Kilgour, "Optimal Deterrence," Social Philosophy and Policy (forthcoming 1985).

2. John Nash, "Non-cooperative Games," Annals of Mathematics 54 (1951), pp. 286-295.

3. Steven J. Brams and Donald Wittman, "Nonmyopic Equilibria in 2 × 2 Games," Conflict Management and Peace Science 6, no. 1 (Fall 1981), pp. 39-62; D. Marc Kilgour, "Equilibria for Far-sighted Players," Theory and Decision 16, no. 2 (March 1984), pp. 135-157; see also Frank C. Zagare, "Limited-Move Equilibria in 2 × 2 Games," Theory and Decision 16, no. 1 (January 1984), pp. 1-19.

4. Steven J. Brams and Marek P. Hessel, "Threat Power in Sequential Games," International Studies Quarterly 28, no. 1 (March 1984), pp. 15-36.

5. Ehud Kalai, "Preplay Negotiations and the Prisoners' Dilemma," Mathematical Social Sciences 1 (1981), pp. 375-379.

6. Raymond L. Garthoff, "The Role of Nuclear Weapons: Soviet Perceptions," in Nuclear Negotiations: Reassessing Arms Control Goals in U.S.-Soviet Relations, ed. Alan F. Neidle (Austin, TX: Lyndon B. Johnson School of Public Affairs, 1982), pp. 1-11. In a review of several different game-theoretic representations of the superpower arms race, Hardin concluded that Prisoners' Dilemma reflects "the preference ordering of virtually all articulate policy makers and policy analysts in the United States and presumably also in the Soviet Union." Russell Hardin, "Unilateral Versus Mutual Disarmament," Philosophy and Public Affairs 12, no. 3 (April 1983), p. 248.

7. One complication in our model would be to assume that players choose not retaliation probabilities but retaliation functions before play commences. Such a function might specify, for example, that the greater the level (probability) of escalation by an opponent, the greater the probability of retaliation. What functions might be optimal in maximizing players' expected payoffs -- presumably by deterring escalation, given precommitments to these functions can be made -- requires further investigation. Dacey suggests an interesting decision-theoretic approach to this question through the use of probabilistic bribes, threats, and tit-for-tat combinations. See Raymond Dacey, "Ambiguous Information and the Manipulation of Plays of the Arms Race Game and the Mutual Deterrence Game," in Interaction and Communication in Global Politics, ed. Claudio Cioffi-Revilla, Richard L. Merritt, and Dina A. Zinnes (Beverly Hills, CA: Sage, 1985).

8. Steven J. Brams and D. Marc Kilgour, "Optimal Deterrence."

9. Steven J. Brams, Superpower Games: Applying Game Theory to Superpower Conflict (New Haven, CT: Yale University Press, 1985), pp. 45-46.

10. This is not the case in the Deterrence Game, wherein the player who triggers rational moves from a Preemption Equilibrium to the Deterrence Equilibrium will incur a temporary cost. See Steven J. Brams and D. Marc Kilgour, "The Path to Stable Deterrence," (mimeographed, 1985).

11. Robert Axelrod, The Evolution of Cooperation (New York: Basic, 1984). Axelrod's analysis, however, is based on a very different game-theoretic model from ours. He found that when many computer programs giving strategies were matched against each other in computer tournament play of Prisoners' Dilemma, "tit-for-tat" did better than

any other program. That is, if one starts out by cooperating, but retaliates on the next round with noncooperation if the other player does not cooperate initially -- and imitates his previous-round behavior in all subsequent play -- then one does better on average than never cooperating or choosing most other strategies. (However, there is no strategy that is unequivocally best in tournament play -- it depends on one's opponents' strategies.) Generally speaking, it pays to be "nice" (begin by cooperating), "provocable" (be ready to retaliate quickly if provoked), and "forgiving" (by returning to cooperating, after retaliating, as soon as the other player does). These conclusions, it should be emphasized, follow from a model that assumes repeated play against different (randomly chosen) opponents, which is hardly descriptive of superpower conflict that involves one continuing opponent. Possibilities for cooperation in iterated (but nontournament play of) Prisoners' Dilemma, with possible discounting of future payoffs, are investigated in, among other places, Michael Taylor, Anarchy and Cooperation (London: Wiley Ltd., 1976); and Stephen J. Majeski, "Arms Races as Iterated Prisoners' Dilemmas," Mathematical Social Sciences 7, no. 3 (June 1984), pp. 253-266. In this work, as in Axelrod's, we find the dichotomous nature of choices (cooperation versus noncooperation) an unrealistic representation of arms-race decisions, which we believe are less qualitative choices over an indefinite series of trials and more quantitative choices in single-play-with-the-possibility-of-retaliation sequences. Finally, we think metagame theory, which posits qualitative choices contingent on the choices of others, is plagued by being both nonquantitative and making heroic demands on the predictive capabilities of players about the choices of others. See Nigel Howard, Paradoxes of Ratio-

nality: Theory of Metagames and Political Behavior (Cambridge, MA: MIT Press, 1971).

12. Steven J. Brams and D. Marc Kilgour, "Optimal Deterrence."

13. There is by no means unanimity on this point. For example, Intriligator and Brito argue, on the basis of a dynamic model of a missile war from which they derive conditions for stable deterrence, that the chances of the outbreak of war have, paradoxically, been reduced because of the recent U.S.-Soviet quantitative arms race. See Michael D. Intriligator and Dagobert L. Brito, "Can Arms Races Lead to the Outbreak of War?" Journal of Conflict Resolution 28, no. 1 (March 1984), pp. 63-84. For a review of the literature and a theoretical and empirical assessment of this question, see James D. Morrow, "A Twist of Truth: A Re-Examination of the Effects of Arms Races on the Occurrence of War" (Department of Political Science, Michigan State University, mimeographed, 1984).

14. Steven J. Brams, Superpower Games, ch. 3 and 4.