Can Arms Races Lead to the Outbreak of War?

MICHAEL D. INTRILIGATOR
Center for International and Strategic Affairs
and Departments of Economics and of Political Science
University of California, Los Angeles

DAGOBERT L. BRITO
Murphy Institute of Political Economy
and Department of Economics
Tulane University

The possible relationships of an arms race to the outbreak of war are treated in the framework of a dynamic model of a missile war that could be used by defense planners to simulate the outbreak of war between two nuclear nations. It is shown that, depending on the initial and final configuration of weapons on both sides, an arms race could lead not only to war but to peace. Conversely, a disarming race could lead not only to peace but to war. The analytic framework is also applied to a qualitative arms race to show that such a race can promote crisis instability. These results are applied both to questions of disarmament and arms control and to the U.S.-Soviet postwar arms race. A conclusion of this analysis is that the quantitative U.S.-Soviet arms race of the 1960s and 1970s not only reduced the chances of war outbreak but also provided insurance against qualitative improvements in weapons.

The relationships between an arms race and the outbreak of war have been debated over many years. Indeed, the question of war outbreak is the most basic and fundamental one in any study of arms races. This question obviously carries much more urgency when it becomes that of whether a *nuclear* arms race might lead to a *nuclear* war. That is the focus of the present article: the relationships between a

AUTHORS' NOTE: This research was completed under collaborative grants on "Behavioral and Economic Foundations of Arms Races" from the National Science Foundation, whose support is gratefully acknowledged. A preliminary version of this article was presented at the International Political Science Association World Congress in Rio de Janeiro, Brazil, August 1982.

JOURNAL OF CONFLICT RESOLUTION, Vol. 28 No. 1, March 1984 63-84

nuclear arms race and the outbreak of nuclear war in a bipolar world of two nuclear powers. The article also treats a qualitative arms race and presents a specific application to the U.S.-Soviet Union arms race of the postwar period.

In his classic study of arms races, Richardson identified an unstable arms race with war. He concluded that an unending upward spiral of armaments in two nations must inevitably culminate in war.[1] Neither Richardson's model nor his analysis, however, warrant this conclusion. His model lacks any explicit treatment of war, so treating upward instability as a situation of "war" appears no more than a convenient label to apply to a particular case of instability. His analysis includes no formalization of the basis for war outbreak, so there is no justification for treating the upward unstable case as one of war. By contrast, the model to be developed here provides an explicit formalization of the initiation of war, facilitating analysis of the possible relationships between an arms race and the outbreak of war.

More recent work on the relationships between arms races and war outbreak has been conducted by Lambelet, Smith, and Wallace.[2] This work, however, both lacks the formal structure of an analytic model of war outbreak as presented below and does not explicitly treat the possibility that an arms race could avoid a war. Lambelet (1975) discusses two extreme hypotheses: the Richardson view that an unstable arms race can only end in war and the hypothesis that arms races and war outbreak are largely independent. He does not explicitly consider the hypothesis suggested below that, in certain circumstances, arms races can prevent war outbreak (or that a disarming race can result in war outbreak), but he does present some historical cases illustrating these possibilities, such as these examples:

> At a more fundamental level, the Korean War may have broken out partly *because* the United States had disarmed unilaterally. . . . There are also on record several arms races which were *not* followed by an open conflict. Thus the larger South American countries were, at various times in the century, involved in a clear-cut naval armament race . . . yet, none of the capital ships involved were ever sunk as a

1. See Richardson (1960, 1951). See also Rapoport (1957), Intriligator and Brito (1976), and Intriligator (1975) for analyses and interpretations of the Richardson model. Richardson (1960) identifies an unstable arms race with war outbreak, but Richardson (1951) provides a qualification to allow for "submissiveness."

2. See Lambelet (1975), Smith (1980), and Wallace (1982). See also Huntington (1958), Saaty (1968, 1964), Pruitt and Snyder (1969), Gray (1976, 1975, 1971), Lambelet (1976, 1973, 1971), Holsti (1972), Russett (1972), Wallace (1979, 1972), and Singer (1981).

> result of hostile action. . . . The pre-1914 Anglo-German naval race suggests a somewhat similar conclusion" [Lambelet, 1975: 124-125; emphasis in original].

Lambelet goes on to provide the general framework of a theory of the relationship between an arms race and the outbreak of war. The theory presented below can be considered a realization of the Lambelet framework, in which arms races affect and are affected by potential war outcomes.

Smith (1980) rejects both of the hypotheses advanced by Lambelet in their more extreme variants, namely, that arms races always end in war and that arms races and war are totally unrelated. She then attempts to identify those arms races that are particularly war prone using statistical methods. She estimates the Richardson model and infers stability or instability from the estimated coefficients of this model. Thus she, in effect, uses the Richardson approach of identifying an unstable arms race with the outbreak of war. She presents no formal analysis of the outbreak of war.

Wallace (1982) compares the "preparedness" hypothesis ("if you seek peace, prepare for war") and the "arms race" hypothesis (that arms races lead to war), using data on 99 serious great power disputes since 1816. He concludes that "International conflicts and disputes which are accompanied by arms races are much more likely to result in war than those in which an arms race does not occur" (1982: 44) on the basis of comparing arms races with a war outcome (23 of 28 cases) and no arms race with a no-war outcome (68 of 71 cases). It should be noted, however, that of the 19 cases he cites of serious disputes in the nuclear era (since 1945), none have ended in a war outcome, suggesting an opposite conclusion, namely that serious disputes accompanied by arms races in the nuclear era do *not* lead to war. The remainder of this article provides a formal analysis of the relationships between arms races and war outbreak, with the purpose of identifying those situations in which an arms race could act to reduce the chances of war outbreak. The following five sections treat a quantitative arms race. Then one section treats a qualitative arms race, and conclusions are presented in the final section.

ARMS RACES

An "arms race" refers to the interactive acquisition of weapons by two or more nations. Assuming only a single weapon, called a "missile" (more precisely, a warhead of a missile), and only two nations in a bipolar world, an arms race can be represented by movements in the weapons plane. Letting $M_A(t)$ and $M_B(t)$ be the numbers of missiles in country A and B respectively at time t, the weapons plane (M_A, M_B) is shown in Figure 1. An arms race is described by the level of weapons on both sides over a time interval (e.g., over the interval from t_0 to t_1), so that the arms race is summarized by

$$\{M_A(t), M_B(t); t_0 \leq t \leq t_1\} \qquad [1]$$

Geometrically the arms race is described by a weapons trajectory in the weapons plane; various such trajectories are shown in Figure 1. A weapons trajectory can involve increases in weapons on both sides, decreases in weapons on both sides, increases on one side but decreases on the other side, and other complicated changes, such as cycles or loops. The parametric representation in equation 1 in which numbers of missiles M_A and M_B are functions of time over a relevant time interval allows for all possible trajectories of weapons in both countries.

In order to relate arms races, represented algebraically by equation 1 and geometrically by a trajectory in the weapons plane, to the outbreak of war, it is necessary to specify the nature of war initiation. The next four sections present a model of a missile war, apply this model to the problem of war initiation, identify regions of deterrence and of war initiation, and use these regions to relate arms races to war intiation.

A MODEL OF A MISSILE WAR

The model of a missile war presented here analyzes the causes of changes over time in the number of missiles and casualties in both countries during a possible war.[3] The model consists of a system of four differential equations with initial boundary conditions. The model determines the evolution over time of the war as a result of initial numbers of weapons on both sides, strategic decisions made by both countries, and the effectiveness of weapons against both counterforce targets (enemy weapons) and countervalue targets (enemy cities).

3. See Intriligator (1975, 1968, 1967), Saaty (1968), Brito and Intriligator (1974, 1973), and Intriligator and Brito (1977, 1976).

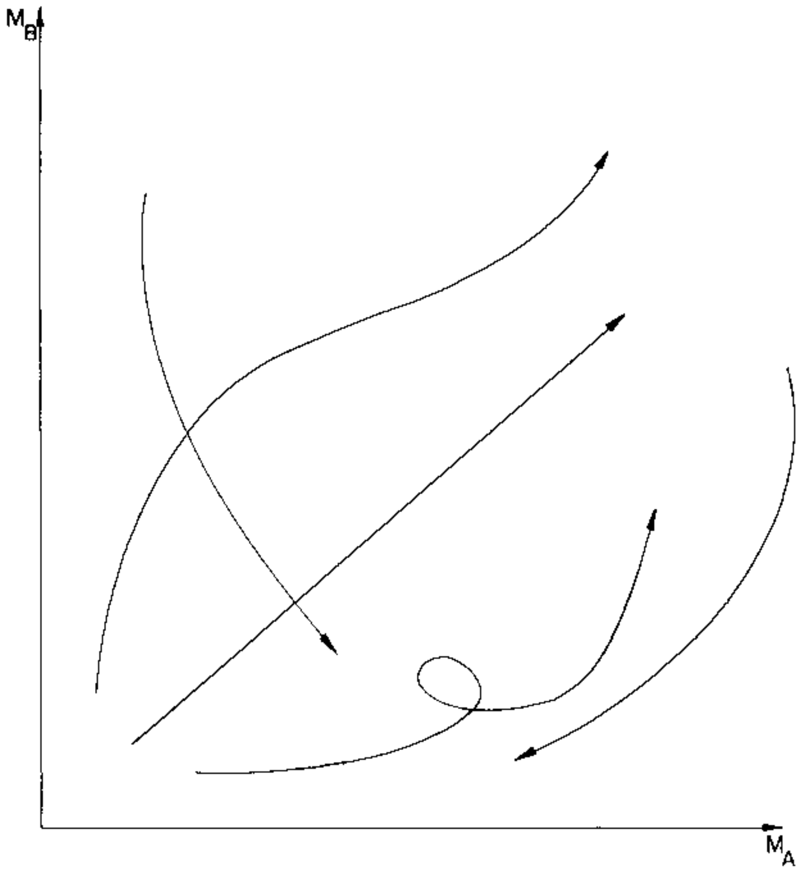

Figure 1: Weapons Plane and Various Weapons Trajectories

The variables of the model are $M_A(t)$ and $M_B(t)$, which, as before, represent the missiles in country A and B at time t. The model consists of the following differential equations for changes in each of these four variables and accompanying boundary conditions:

$$\dot{M}_A = -\alpha M_A - \beta'\beta M_B f_B \qquad M_A(0) = M_A^0 \qquad [2]$$

$$\dot{M}_B = -\beta M_B - \alpha'\alpha M_A f_A \qquad M_B(0) = M_B^0 \qquad [3]$$

$$\dot{C}_A = (1-\beta')\beta M_B v_B \qquad C_A(0) = 0 \qquad [4]$$

$$\dot{C}_B = (1-\alpha')\alpha M_A v_A \qquad C_B(0) = 0 \qquad [5]$$

The war starts at time t = 0, at which point, as shown in the boundary conditions, country A has M_A^0 missiles, country B has M_B^0 missiles, and there are no casualties on either side. Country A launches its missiles at rate $\alpha(t)$, so $-\alpha M_A$ in equation 2 represents the reduction in A missiles due to decisions by A to launch missiles. Similarly, $-\beta M_B$ in equation 3 represents the reduction in B missiles due to decisions by B to launch its missiles at the rate $\beta(t)$. (The dependence of M_A, M_B, α, β, and other variables on time is omitted in equations 2-5 for convenience.)

Missiles can be targeted counterforce at enemy missiles or countervalue at enemy cities. If A uses the counterforce proportion $\alpha'(t)$ at time t, then, of the αM_A missiles launched at this time, $\alpha'\alpha M_A$ are launched at B missiles while $(1-\alpha')\alpha M_A$ are launched at B cities. If $f_A(t)$ is the counterforce effectiveness of A missiles at time t (i.e., the number of B missiles destroyed per A counterforce missile), then $\alpha'\alpha M_A f_A$ represents the B missiles destroyed by A counterforce missiles, as is shown in equation 3. Similarly, the term $\beta'\beta M_B f_B$ in equation 2 represents the A missiles destroyed by B counterforce missiles, where $f_B(t)$ is the counterforce effectiveness of B missiles. If $v_A(t)$ is the countervalue effectivess of A missiles (i.e., the number of B casualties inflicted per A countervalue missile), then $(1 - \alpha')\alpha M_A v_A$ represents the B casualties inflicted by A countervalue missiles, as shown in equation 5. Similarly, the term $(1-\beta')\beta M_B v_B$ in equation 4 represents the A casualties inflicted by B countervalue missiles, where $v_B(t)$ is the countervalue effectiveness of B missiles.

The evolution of the war over time, as summarized by the dynamic model equations 2-5, thus depends on the initial levels of missiles M_A^0, M_B^0; strategic decisions on both sides regarding rates of fire, $\alpha(t)$, $\beta(t)$; strategic decisions on both sides regarding targeting $\alpha'(t)$, $\beta'(t)$; the effectiveness of missiles of both sides against enemy missiles, $f_A(t)$ $f_B(t)$; and the effectiveness of missiles of both sides against enemy cities, $v_A(t)$, $v_B(t)$.

From the viewpoint of either one of the countries, the problem of "grand strategy" is that of choosing both a rate of fire and a targeting strategy. For country A, the rate of fire α can range between zero and some maximum rate $\bar{\alpha}$, determined on the basis of technical characteristics of the weapons. Similarly, the counterforce proportion α' can range between 0 and 1, where $\alpha' = 1$ is pure counterforce targeting (no cities), and $\alpha' = 0$ is pure countervalue targeting (only cities). Omitting intermediate values, which, in fact, would never be used, the two extreme values for each of the two strategic variables for country A yield four alternatives for grand strategy for A:

(1) Maximum rate/counterforce, where A fires its missiles at the maximum rate ($\alpha = \bar{\alpha}$ and targets only B missles ($\alpha' = 1$), destroying as many enemy missiles as possible—a "first-strike strategy."
(2) Maximum rate/countervalue, where A fires its missiles at the maximum rate ($\alpha = \bar{\alpha}$) and targets only B cities ($\alpha' = 0$), inflicting as many enemy casualties as possible— a "massive retaliation strategy."
(3) Zero rate/counterforce, where A holds its missiles in reserve ($\alpha = 0$) and targets only B missiles ($\alpha' = 1$), threatening to strike enemy missiles—a "limited strategic war strategy."
(4) Zero rate/countervalue, where A hold its missiles in reserve ($\alpha = 0$) and targets only B cities ($\alpha' = 0$), threatening to strike enemy cities—a "war of nerves strategy."

Assuming that country A treats the country B strategy as fixed, instead of trying to influence it, and that A has the goal of maximizing a payoff function

$$P_A = P_A[M_A(T), M_B(T), C_A(T), C_B(T)], \qquad [6]$$

$$\frac{\partial P_A}{\partial M_A(T)} > 0, \ \frac{\partial P_A}{\partial M_B(T)} < 0, \ \frac{\partial P_A}{\partial C_A(T)} < 0, \ \frac{\partial P_A}{\partial C_B(T)} \gtreqless 0,$$

which depends on missiles and casualties in both countries at the end of the war, time T, it has been shown that country A will select at any one time in the war t (where $0 < t < T$) one of these four grand strategies.[4] In particular, while A could choose some rate of fire intermediate between 0 and $\bar{\alpha}$, it is optimal for it to choose only one of these extreme values, firing missiles at either the zero or the maximum rate. Similarly, while A could choose some counterforce proportion intermediate between 0 and 1, it is optimal for it to choose only one of these extreme values, firing missiles at only enemy cities or enemy missiles. Thus, A is confined to the four possibilities listed above for grand strategy. Furthermore, given the payoff function in equation 6 and the differential equations and boundary conditions describing the evolution of the war in equations 2-5, it is optimal for A to make a single switch in its rate of fire, a switch

4. See Intriligator (1967) for proofs of this result and of the subsequent results reported in this paragraph.

from the maximum rate $\alpha = \bar{\alpha}$ to the zero rate $\alpha = 0$. Similarly, it is optimal for A to make a single switch in its targets, as summarized by the counterforce proportion, a switch from pure counterforce targeting $\alpha' = 1$ to pure countervalue targeting $\alpha' = 0$.

If country A uses its optimal strategy, the war proceeds in three stages. Country A starts the war with a first-stirke strategy ($\alpha = \bar{\alpha}$ and $\alpha' = 1$), a situation in which it has not yet switched either strategic choice variable, and it ends the war with a war of nerves strategy ($\alpha = 0$ and $\alpha' = o$), a situation in which it has switched both of the strategic choice variables. The middle stage of the war can be either one of a massive retaliation strategy ($\alpha = \bar{\alpha}$ and $\alpha' = 0$), if the switching time for targets precedes that for the rate, or it can be one of a limited strategic war strategy ($\alpha = 0$ and $\alpha' = 1$), if the switching time for the rate precedes that for targets. Country A thus starts the war using only counterforce targeting at the maximum rate of fire, so that it can destroy as many of the enemy missiles as possible; it ends the war by holding enemy cities as hostages, placing itself in the best possible position for extracting concessions or a desired settlement of the war and, depending upon which switching time occurs first, either inflicts massive casualties in country B during the middle stage of the war or hold its missiles in reserve, threatening enemy missiles.

A MODEL OF WAR INITIATION

A model of a missile war having been developed in the previous section, this model now will be specialized to treat the case of war initiation. The model is in fact, reinterpreted in an essential way. It is treated not as a model of an actual war but rather as a model of a hypothetical or potential war, a war that might break out at any time. Thus the model is treated not as a representation of actual missiles, casualties, strategic choices, and so forth, but as a representation of plans for a war, including expected missiles, casualties, and strategic choices that are foreseen as possible situations by defense planners. The model is thus a representation of the war simulations or war scenarios used by defense analysts, such as those in the Pentagon or the Kremlin.

From the vantage point of defense analysts in country A, the model can be used to simulate various possible scenarios. Of particular importance are two cases. The first case is that in which B attacks A, and

A simulates the effect of the B attack and its own retaliation. In this case, defense analysts in A would seek to have enough missiles to deter B by threatening it with unacceptable levels of casualties in its retaliatory strike. The second case is that in which A contemplates a first-strike attack on B, and its defense analysts would seek to destroy enough B missiles to disarm it and render a retaliatory strike ineffectual. These two cases are those of country A as a "deterrer" and country A as an "attacker," respectively.

In the first case, of country A as a deterrer, its defense analysts should anticipate, as in the last section, that B will strike first using a first-strike strategy ($\beta = \bar{\beta}$, $\beta' = 1$) over the time interval from 0 to θ_B, before A can retaliate. They would then retaliate using a massive retaliation strategy ($\alpha = \bar{\alpha}$, $\alpha' = 0$), inflicting casualties in B over the retaliation time interval from θ_B to $\theta_B + \psi_A$.[5] If defense planners in A simulate the result of this scenario, they can solve the differential equations 2-5 using these strategic values if, in this simulated war outbreak, f_B can be assumed to be constant over the retaliation interval ψ_A. These are reasonable assumptions if both intervals are relatively short (e.g., measured in minutes). The solution for the casualties in country B at the end of the retaliatory interval is then

$$C_B(\theta_B + \psi_A) = v_A[M_A^0 - f_B(1 - \exp(-\bar{\beta}\theta_B))M_B^0][1 - \exp(-\bar{\alpha}\psi_A)] \quad [7]$$

showing explicitly the dependence of simulated casualties on the initial number of missiles M_A^0, M_B^0; rates of fire $\bar{\alpha}$, $\bar{\beta}$; missile effectiveness ratios f_B, v_A; and the time intervals θ_B, ψ_A. If this expected number of casualties is sufficiently large, then A would deter B. In particular, if the minimum unacceptable number of casualties in B is $\bar{C}_B$ then solving

$$C_B(\theta_B + \psi_A) \geq \bar{C}_B \quad [8]$$

for M_A^0, the number of A missiles required to deter B (by threatening an unacceptable level of casualties) is given as

$$M_A \geq f_B\,(1 - \exp(-\bar{\beta}\theta_B))M_B + \frac{\bar{C}_B}{v_A(1 - \exp(-\bar{\alpha}\psi_A))} \quad [9]$$

5. This retaliation strategy is consistent with the switching strategy outlined in the last section; in this case, the switch in targets occurs at time θ_B, so after θ_B the A grand strategy is one of massive retaliation, with $\alpha = \bar{\alpha}$, $\alpha' = 0$.

(Here M_A^0 and M_B^0 are replaced by M_A and M_B, respectively, since the model is one of a simulated war that can start any time.) This inequality shows the number of missiles required for A to deter B as an explicit function of the number of missiles (and also the technical parameters $\bar{\alpha}$, $\bar{\beta}$, f_B, v_A; timing parameters θ_B, ψ_A; and the minimum unacceptable number of B casualties C_B). Geometrically, this inequality is the area to the right of the line marked "A deters" in the weapons plane in Figure 2, with intercept $C_B/[v_A(1-\exp(-\bar{\alpha}\psi_A))]$ on the M_A axis and with slope $1/[f_B(1-\exp(-\bar{\beta}\theta_B))]$.

In the second case, of country A as an attacker, defense analysts in A anticipate attacking B using its first-strike strategy ($\alpha = \bar{\alpha}$, $\alpha' = 1$) over the time interval from 0 to θ_A before B can retaliate. They then have to consider the effects of B retaliating using a massive retaliation strategy ($\beta = \bar{\beta}$, $\beta' = 0$) over the retaliatory time interval from θ_A to $\theta_A + \psi_B$. When they solve the differential equations, once again assuming that the time intervals are sufficiently short that f_A can be treated as constant over the first strike interval and f_B can be treated as constant over the retaliation time interval, the casualties that defense planners in A anticipate in this simulated war outbreak are $C_A(\theta_A + \psi_B)$, which is similar to equation 7 if the roles of A and B are interchanged. If defense planners in A regard $\hat{C}_A$ as a maximum acceptable level of casualties, then, solving

$$C_A(\theta_A + \psi_B) \leq \hat{C}_A \qquad [10]$$

for M_A ($= M_A^0$) yields

$$M_A \geq \left[\frac{\bar{1}_B}{f_A(1 - \exp(-\bar{\alpha}\theta_A))}\right] M_B - \frac{\hat{C}_A}{f_A(1 - \exp(-\bar{\alpha}\theta_A))v_B(1 - \exp(-\bar{\beta}\psi_B))} \qquad [11]$$

This inequality shows the number of A missiles required for A to attack B as an explicit function of the number of B missiles (and also the technical parameters $\bar{\alpha}$, $\bar{\beta}$, f_A, v_B; timing parameters θ_A, ψ_B; and the maximum acceptable number of A casualties $\hat{C}_A$). Geometrically, this inequality is the area above the line marked "A attacks" in the weapons plane in Figure 2, with intercept $\hat{C}_A/[v_B(1-\exp(-\overline{\beta}\psi_B))]$ on the M_B axis and with slope $f_A(1-\exp(\overline{\alpha}\theta_A))$.

Figure 2 also shows "B deters" and "B attacks" lines, indicating regions in which defense analysts in country B, on the basis of their simulated war outbreak using the same model of a missile war, have enough missiles, respectively, to deter A from attacking (by threatening unacceptable casualties) and to attack A (suffering an acceptable level of casualties). If the technical and timing parameters are estimated to be the same in both countries, then the "B deters" line is parallel to the "A attacks" line and the "B attacks" line is parallel to the "A deters" line. For example, the common slope of the "B deters" and "A attacks" line is $f_A(1-\exp(-\overline{\alpha}\theta_A))$. If defense planners in the two countries have different assessments of these parameters (f_A, f_B, $\overline{\alpha}$, $\overline{\beta}$, θ_A, θ_B), then the lines need not be parallel.

REGIONS OF DETERRENCE AND OF WAR INITIATION

The model of war initiation leads to regions of the weapons plane in Figure 2 that can be identified as ones of deterrence and of war initiation.[6]

The upper shaded cone bounded by the "A deters" and "B deters" lines is one of mutual deterrence, in which each country deters the other. This region is one of stability against war outbreak since in it each country has the ability to inflict unacceptable damage on the other in a retaliatory attack (Szilard, 1964). The region is one of "mutual assured destruction" or "mutual deterrence." In it, each country has enough missiles to deter the other, and neither country has enough missiles to attack the other.

The lower shaded sawtooth-shaped area is one of forced initiation. In it, neither country deters the other, and one or both can attack the other. In the middle portion of this area, the cone between the origin and the point I, A can attack B, and B can attack A, while neither can deter the other. Thus this portion is one of virtually forced preemption in which it is greatly advantageous to initiate rather than to retaliate. The "reciprocal fear of surprise attack" based on the tremendous advantage in striking first forces both sides to initiate, each trying to preempt the

6. For related diagrams, see Kybal (1960) and Beaufre (1965).

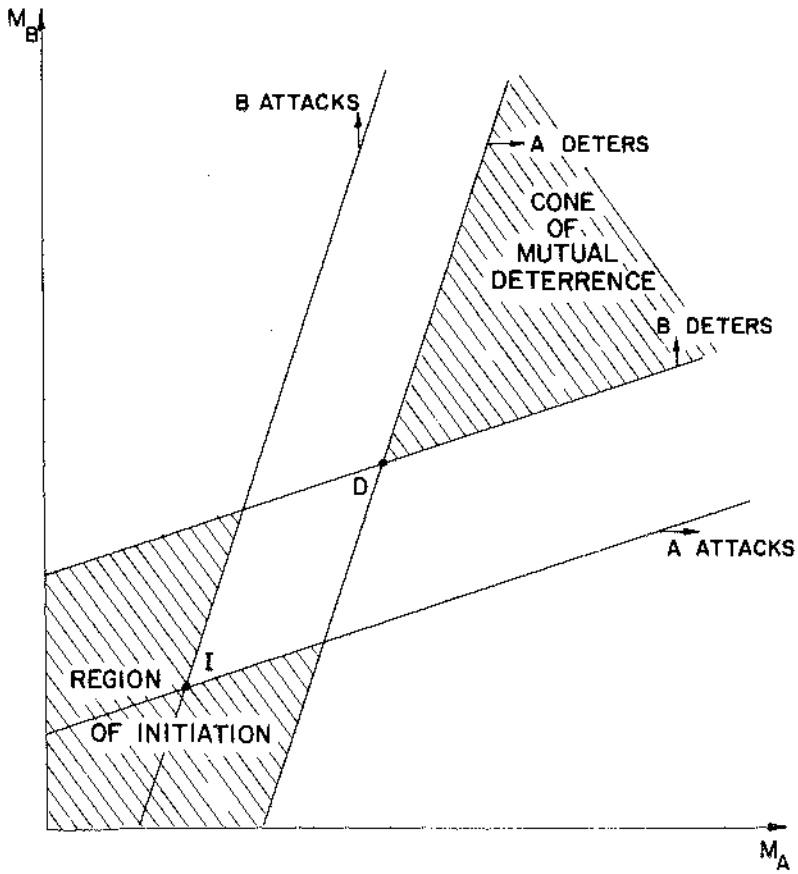

Figure 2: Deterring and Initiating Regions in the Weapons Plane

attack of the other (Wohlstetter, 1959). The other portions of this area are ones of asymmetry; in one (on the lower right), A can attack B but B cannot attack A, while in the other (upper left), B can attack A but A cannot attack B. Thus in these two asymmetric regions, stability relies on the intentions of one side. The entire sawtooth-shaped area, however, is one of initiation. This region exhibits instability against war outbreak: neither side has an adequate retaliatory capability to ensure deterrence, and one or both countries has sufficient capability, relative to the missiles held by the opponent, to attack the other.

The deterring and attacking lines for both countries that define the upper cone of mutual deterrence and the lower sawtooth-shaped area of initiation are themselves defined in equations 9 and 11 in terms of:

f_A, f_B: the effectiveness of missiles in destroying enemy missiles (which depend on missile accuracy and yield and enemy missile dispersion and hardness against attack).

v_A, v_B: the effectiveness of missiles in inflicting enemy casualties (which depend on missile accuracy and yield and enemy civil defense and dispersion).

$\overline{\alpha}, \overline{\beta}$: the (maximum) rates of fire of missiles (which depend on missile characteristics and the command and control structure).

θ_A, θ_B: the time intervals during which one country can attack the other without a response from the other (which depend on detection capabilities and command and control).

ψ_A, ψ_B: the time intervals for a retaliatory strike (which depend on missile characteristics and command and control).

C_A, C_B: the minimum unacceptable level of casualties which would deter each country, as estimated by the other country.

$\hat{C}_A, \hat{C}_B$: the maximum acceptable level of casualties that each country could accept in a retaliatory strike, as estimated by itself.

To give a numerical example, if it takes two missiles to destroy an enemy missile ($f_A = f_B = 0.5$), each missile can inflict 250,000 casualties ($v_A = v_B = 250{,}000$), the maximum rate of fire is 10% per minute ($\overline{\alpha} = \overline{\beta} = 0.10$), the first strike interval is 15 minutes ($\theta_A = \theta_B = 15$), the retaliatory strike interval is 10 minutes ($\psi_A = \psi_B = 10$), the minimum unacceptable level of casualties is 40 million ($C_A = C_B = 40$ million), and the maximum acceptable level of casualties is 5 million ($\hat{C}_A = \hat{C}_B = 5$ million), then point D in Figure 2 shows that the minimum number of missiles required for mutual deterrence is 414 missiles on each side, while point I in Figure 2 shows that the maximum number of missiles in the area of forced preemption (where each can attack the other) is 52 missiles. Thus, in this symmetrical example, if each side has more than 414 missiles, there is the stability of mutual deterrence, while if each side had fewer than 52 missiles, there would be the instability of forced preemption.

Obviously, all the numbers used in these calculations are subject to uncertainly, particularly the estimates of the minimum unacceptable and maximum acceptable levels of casualties. Thus it would be more appropriate to refer not to a cone of mutual deterrence and a region of initiation but rather to equal probability contours of war outbreak, showing relatively high probability of war outbreak in the region of initiation and relatively low probabilities of war outbreak in the cone of mutual deterrence. Alternatively, the lines of deterrence and attack in Figure 1 can be replaced by bands, reflecting the uncertainty in the values of the parameters defining deterring and attacking regions. For purposes of exposition, however, the lines, the cone of mutual deterrence, and the region of initiation will be used in the next two sections to analyze arms races and war intitiation.

ARMS RACES AND WAR INITIATION

The analysis of arms races and war initiation combines the weapons trajectories describing an arms race in the "Arms Races" section, as shown geometrically in Figure 1, with the regions of deterrence and of war initiation in the "Regions of Deterrence and of War Initiation" section, as shown geometrically in Figure 2. The result is Figure 3, which shows several trajectories and the regions of deterrence and of war initiation.

Trajectory 1 starts from the totally disarmed state of the origin and, as a result of an arms race, both countries move to a highly unstable situation of forced initiation. This is an example of the case in which an arms race leads to war.

Trajectory 2 involves an arms race that, as a result of a weapons buildup in both countries, moves the situation from one in which A could attack B to one in which each side deters the other. This is a situation in which an unstable situation is stabilized as a result of an arms race, in which "if you seek peace, prepare for war."

Trajectory 3 is one in which country A reduces its arms from an extremely high level while country B simultaneously increases its arms from a relatively low level. The result is a reduced chance of war outbreak, moving from a situation in which A can attack B to one in which each deters the other.

Trajectory 4 is one in which both sides disarm with the result that they move from a stable situation in the cone of mutual deterrence to an unstable one in which each can attack the other. This trajectory illustrates the potential dangers of bilateral disarmament that is carried too far.

Trajectory 5 is one in which, like the previous case, both sides reduce their level of armaments, but, unlike the previous case they stay within the cone of mutual deterrence. Such selective reductions in arms need

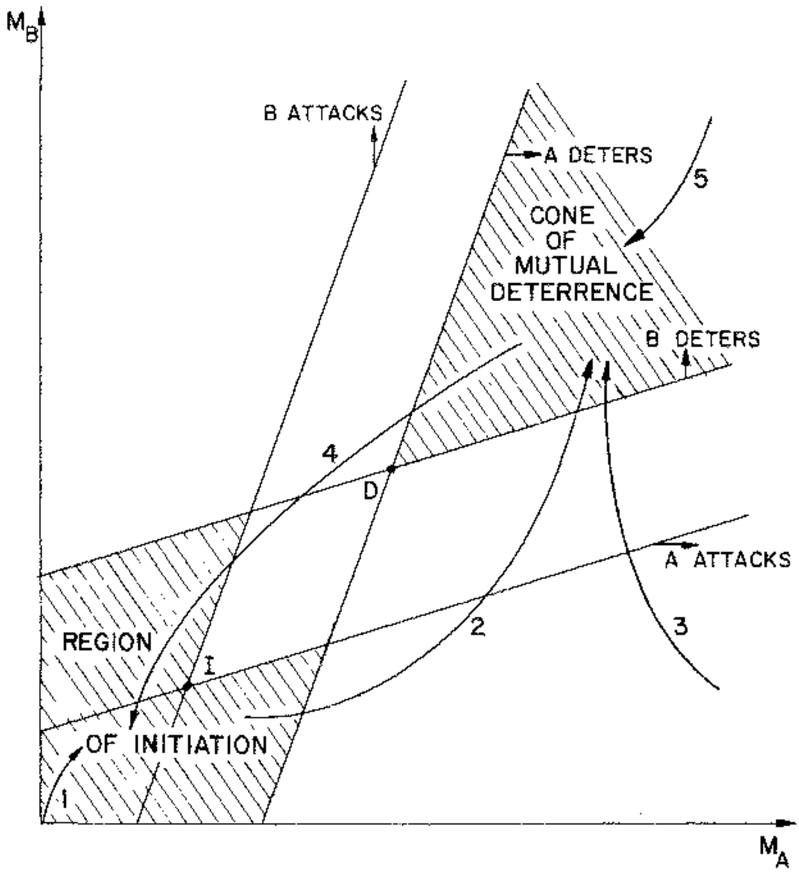

Figure 3: Weapons Trajectories and Regions of Deterrence and War Initiation

not cause instability as long as the situation remains in the cone of mutual deterrence.

Thus, the answer to the question, "Can arms races lead to the outbreak of war?" is not a simple one. The frequently expressed opinion that arms races lead to war is neither right nor wrong; it depends on the circumstances. In certain situations, an arms race can lead to war (as in trajctory 1), while, in other situations, an arms race can lead to peace (as in trajectory 2). This conclusion is a new one that has not appeared in the previous formal literature on arms races.[7] Conversely a disarming race can result in war (as in trajectory 4) or can preserve peace (as in trajectory 5), again a conclusion that has not appeared in the formal literature on disarmament. In fact, the contrast between trajectories 4 and 5 illustrates the difference between disarmament, which, despite its lofty goals, can be highly dangerous in leading to unstable situations, and arms control which, by selectively reducing weapons but retaining mutual deterrence, can promote greater stability (e.g., in eliminating obsolete or unstable weapons systems but retaining enough weapons for deterrence).

QUALITATIVE ARMS RACE AND CRISIS STABILITY

The discussion thus far has concentrated on a quantitative arms race, that is, increases (or decreases) in the number of missiles in each of two countries. The same framework can, however, be used to analyze a qualitative arms race, that is, changes in the characteristics of the missiles in both countries without a change in the number of these missiles. Huntington (1958) has suggested that a qualitative arms race can be more dangerous than a quantitative arms race. Furthermore, there have been various qualitative improvements in missiles in both the United States and the Soviet Union, particularly since the 1970s. Indeed, while the Soviets have been increasing both the quality and the quantity of their weapons, the main changes in the U.S. weapons have been qualitative ones, particularly in the strategic land-based missiles (Minuteman replaced by Minuteman III and possibly to be augmented the MX) and in submarine-based missiles (Polaris replaced by Poseidon and due to be replaced by Trident). There has been concern that these qualitative changes in weapons may have affected crisis stability and thereby increased the chances of war. In particular, the presence of multiple warheads in MIRV systems and the greater accuracy of newer missiles may have decreased crisis stability.

7. For an analytic overview of this literature, see Intriligator (1982). While the conclusion presented here is a new one as far as the formal arms race literature is concerned, related conclusions have appeared in the general literature on deterrence (e.g., in Schelling, 1966), where it is argued that major cuts in retaliatory capabilities can contribute to instability.

The framework introduced earlier can be used to study this qualitative arms race, in particular, the presence of MIRVs and greater accuracies. The primary impact of these qualitative changes has been to raise the counterforce effectiveness of missiles in destroying enemy missiles, f_A and f_B, which, as already noted, depends on missile accuracy and yield and enemy missile dispersion and hardness against attack.[8] For example, multiple warheads can increase the probability of destroying enemy missiles by independently targeting several enemy missiles, and greater accuracy is of critical importance in destroying hardened enemy missiles.

Qualitative changes in missiles such as MIRVs and improved accuracy increase the counterforce effectiveness ratios f_A and f_B, and thus the impacts of these changes can be determined by a comparative statics analysis of the system without and with these changes.

First assume that only country A (e.g., the United States) improves its missiles by increasing f_A and that the other country B (e.g., the Soviet Union) keeps its missiles unchanged (e.g., not MIRVed) and that f_B is thus unchanged. None of the terms in equation 9 are changed, so the number of A missiles required to deter B shown geometrically in Figures 2 and 3 as the line marked "A deters," is unchanged. The corresponding equation for B to deter A, namely,

$$M_B \gtrsim f_A\,(1 - \exp(-\bar{\alpha}\theta_A))\;M_A + \frac{\bar{C}_A}{v_B(1 - \exp(-\bar{\beta}\psi_B))} \qquad [9']$$

shown as the line "B deters" in Figures 2 and 3, is changed, however, because of the increase in f_A. The slope of this line increases as f_A increases, but the intercept remains unchanged. Rotating the "B deters" line about its intercept on the M_B axis implies that configurations of missiles (M_A, M_B) that were in the cone of mutual deterrence would no longer be in the cone. Configurations near the old "B deters" line are no

8. Qualitative improvements in missiles have some other impacts, such as increasing the countervalue effectiveness of missiles in inflicting enemy casualties, v_A and v_B; increasing the maximum rates of fire of missiles, $\bar{\alpha}$ and $\bar{\beta}$; decreasing the time interval of the initial attack, θ_A and θ_B; and changing the time interval of the retaliatory strike, ψ_A and ψ_B. Any or all of these could be treated, but all are less significant both in terms of actual changes in weapons characteristics and in terms of impacts than changes in the counterforce effectiveness of missiles f_A and f_B. Also note that in equations 9 and 11 and in the corresponding equations for B to deter and to attack, f_A enters only as $f_A(1 - \exp(-\bar{\alpha}\theta_A))$, so increasing f_A is equivalent to increasing $\bar{\alpha}$ or θ_A or can be offset by decreases in $\bar{\alpha}$ or θ_A.

longer stable in the sense that each side has enough missiles to deter the other. In particular, the point of minimum mutual deterrence, shown as D in Figures 2 and 3, moves as a result of an increase in f_A to a new configuration involving larger numbers of missiles and proportionately even larger numbers of B missiles. Consider the numerical example presented earlier in which point D, the minimum numbers of missiles required for mutual deterrence, was 414 on each side. If, because of qualitative improvements in A such as MIRVs and improved accuracy, f_A increases from 0.5 to 0.75 (so that it takes only one and a half A missiles to destroy a B missile), the new point D is (454, 517). Thus a 50% improvement in the counterforce effectiveness of A missiles, from f_A = 0.5 to f_A = 0.75, implies that the minimum number of missiles required to deter increases by 40 or 9.7% from 414 to 454, for A but by 103, or almost 25%, from 414 to 517, for B. In this asymmetrical situation, B needs proportionately more missiles because of the fact that more of its missiles would be destroyed in the counterforce A first strike, given the greater counterforce effectiveness of A missiles.

In addition to shrinking the cone of mutal deterrence, the qualitative improvement in A missiles also enlarges the sawtooth-shaped area of forced initiation in Figures 2 and 3. An increase in f_A, other parameters remaining unchanged, will not affect either the "A deters" or the "B attacks" lines, but it will affect both the "B deters" and the "A attacks" lines. The "A attacks" line is defined by equation 11, so increasing f_A affects the slope of this line, raising the slope (the line "A attacks" remaining parallel to the line "B deters") but keeping the intercept on the M_B axis unchanged. Using the same numerical example, changing f_A from 0.5 to 0.75 shifts the point I, the maximum number of missiles in the area of forced preemption from 52 on each side to (57, 65). A 50% improvement in the counterforce effectiveness of A missiles therefore implies that the maximum number of missiles for forced preemption increases by 5, from 52 to 57, for A, but by 13 from 52 to 65 for B, representing increases again of 9.6% for A and 25% for B. Thus improvements in A missiles could lead to forced preemption by both sides. If both sides have very few missiles (e.g., if both have 55 missiles), then each side could attack the other if A improves the counterforce effectiveness of its missiles. Thus there may be value in holding additional reserves of missiles as a hedge to insure against possible qualitative improvements in missiles.

Consider now what happens to points D and I as a result of qualitative improvements in missiles on both sides. If both A and B improve their missiles by MIRV, and the improved accuracy and so forth leads to increases in both f_A and f_B, the slopes of all four lines "A deters", "B deters." "A attacks", and "B attacks" in Figures 1 and 2 will change. The result will be that the cone of mutual deterrence will shrink, with points that previously were ones of mutual deterrence no longer having this property. Furthermore, the regions of initiation will expand, with points that previously were not ones of initiation becoming such points. Continuing the same numerical example, suppose that both sides increase their counterforce effectiveness by 50% from $f_A = f_B = 0.5$ to $f_A = f_B = 0.75$. The result is that the point D, the minimum point of mutual deterrence, increases from 414 missiles on each side to 606 missiles on each side, an increase of 192 missiles or about 46%. Similarly, the point I, the maximum point of forced preemption, increases from 52 missiles on each side to 76 missiles on each side, an increase of 24 missiles or, again, about 46%.

It is clear that qualitative improvements in missiles such as MIRV and an increased accuracy that has the effect of increasing the counterforce effectiveness of missiles lead both to a shrinking of the region of mutual deterrence and an expansion of the regions of initiation, lending support to the argument that such a qualitative arms race can be destabilizing or promote crisis instability. On the other hand, barring very substantial increases in counterforce effectiveness, there will continue to exist a cone of mutual deterrence and the region of initiation will be confined to the area around the origin.[9] Furthermore, there is a relationship between a qualitative and a quantitative arms race in that, assuming a cone of mutual deterrence exits, enough missiles on both sides can ensure that the situation is one of mutual deterrence. Indeed, one interpretation of the very substantial numbers of missiles on both sides is that each wants to hedge against a possible technological breakthrough on either or both sides that could substantially increase counterforce effectiveness, insuring that in any foreseeable situation, the

9. A cone of mutual deterrence exists as long as the stability condition:

$$f_A(1 - \exp(-\bar{\alpha}\theta_A))\, f_B(1 - \exp(-\bar{\beta}\theta_B)) < 1$$

is met. This condition will always be met if it takes more than one misile to destroy an enemy missile (i.e., $f_A, f_B < 1$). Even if f_A and/or f_B exceeds unity, however, the condition could still be met. For example, if $\bar{\alpha} = \bar{\beta} = 0.10$ and $\theta_A = \theta_B = 15$, as in the numerical examples in the sections on regions of deterrence and war initiation and on a qualitative arms race and stability, this stability condition is met as long as, in the symmetric case, both f_A and f_B are less than 1.28.

configuration of missiles will be in the cone of mutual deterrence or, at least, not in the region of initiation.

CONCLUSIONS

The result of this analysis is that, when arms races are considered in the context of a model of a possible nuclear war, arms races can in certain circumstances lead to war but they can also lead to the avoidance of war. Conversely, disarmament can lead to the avoidance of war but it can also lead to war in certain circumstances. The result depends essentially on a comparison of the starting and ending points of the arms race relative to the regions of deterrence and war initiation. Points close to the totally disarmed state are highly unstable, so an arms race starting from the disarmed state would be highly dangerous (as illustrated in trajectory 1 in Figure 3). Conversely, a disarming race moving close to the disarmed state would also be highly dangerous (as illustrated in trajectory 4 in Figure 3). Arms control can be interpreted not only as selective decreases in arms that retain mutual deterrence but also as selective *increases* in arms leading to a situation of mutual deterrence (as illustrated respectively in trajectories 5 and 2 in Figure 3).

A qualitative arms race involving MIRV, increases in accuracy, and so forth can lead to crisis instability by both shrinking the region of deterrence and expanding the regions of war initiation. In this respect, a qualitative arms race can be more dangerous than a quantitative arms race. The potential instability of a qualitative arms race can, however, be offset by a quantitative arms race in that, if each country had enough missiles, a quantitative arms race could ensure that situations of war initiation via forced preemption could be avoided given any forseeable improvements in missiles.

To relate this analysis to the postwar situation for the United States and the Soviet Union, trajectory 2 perhaps comes closest to describing the history of superpower arms since 1945. Before the Soviet Union had a significant weapons capability, the situation was an asymmetric one in which the U.S. (country A in the figure) could attack the Soviet Union, but the Soviet Union could not attack the United States. In this asymmetrical situation, peace was preserved because of the U.S. commitment not to initiate a war although it should be noted that the Soviet Union had considerably more conventional forces than the United States during this period. The 1960s and 1970s witnessed a major buildup of weapons on both sides, leading to the present situation of mutual deterrence, as illustrated in trajectory 2, a situation of stability against war outbreak. In fact, the greatest instability was probably in the late 1950s and early 1960s, particularly during the 1962 Cuban missile crisis, in which neither side could deter the other. The 1970s and 1980s have witnessed substantial improvements in the quality of missiles, including MIRV, and improved accuracy of missiles. While these developments in themselves promoted crisis instability, they occurred after each side had already achieved levels of missiles that ensured stability despite the improvements in missile quality. Thus the quantitative arms race that had occurred earlier between the United States and the Soviet Union not only reduced the chances of war outbreak but provided insurance against later qualitative improvements in missiles.

REFERENCES

BEAUFRE, A. (1965) Deterrence and Strategy (R. H. Barry, trans.). New York: Praeger.

BRITO, D. L. and M. D. INTRILIGATOR (1974) "Uncertainty and the stability of the armaments race." Annals of Econ. and Social Measurement 3: 279-292.

——— (1973) "Some applications of the maximum principle to the problem of an armaments race." Modeling and Simulation 4: 140-144.

GRAY, C. S. (1976) The Soviet-American Arms Race. Lexington, MA: Lexington Books.

——— (1975) "Arms races and their influence upon international stability with specific reference to the Middle East," in G. Sheffer (ed.) Dynamics of a Conflict. Atlantic Highlands, NJ: Humanities Press.

——— (1971) "The arms race phenomenon," World Politics 24: 39-79.

HOLSTI, O. R. (1972) Crisis, Escalation, War. Montreal: McGill-Queen's Univ. Press.

HUNTINGTON, S. P. (1958) "Arms races: prerequisites and results," Public Policy 8: 41-86.

INTRILIGATOR, M. D. (1982) "Research on conflict theory: analytic approaches and areas of application." J. of Conflict Resolution 26: 307-327.

——— (1975) "Strategic considerations in the Richardson model of arms races." J. of Pol. Economy 83: 339-353.

——— (1968) "The debate over missile strategy: targets and rates of fire." Orbis 11: 1138-1159.

——— (1967) Strategy in a Missile War: Targets and Rates of Fire. Security Studies Paper No. 10. Los Angeles: Security Studies Project, University of California, Los Angeles.

——— and D. L. BRITO (1977) "Strategy, arms races, and arms control," in J. V. Gillespie and D. A. Zinnes (eds.) Mathematical Systems in International Relations Research. New York: Praeger Publishers.

——— (1976) "Formal models of arms races." J. of Peace Sci. 2: 77-88.

KYBAL, D. (1960) "Remarks," in Proceedings of the National Security Seminar, Asilomar, California, April 24-30.

LAMBELET, J. C. (1976) "Complementary analysis of the Anglo-German dreadnought race, 1905-1916." Papers of the Peace Research Society (International) 26: 49-66.

——— (1975) "Do arms races lead to war?" J. of Peace Research 12: 123-128.

——— (1973) "Toward a dynamic two-theater model of the East-West arms race," J. of Peace Sci. 1: 1-38.

——— (1971) "A dynamic model of the arms race in the Middle East, 1953-1965." General Systems 16: 145-167.

PRUITT, D. G. and R. C. SNYDER (1969) Theory and Research on the Causes of War. Englewood Cliffs, NJ: Prentice-Hall.

RAPOPORT, A. (1957) "Lewis Fry Richardson's mathematical theory of war." J. of Conflict Resolution 1: 249-304.

RICHARDSON, L. F. (1960) Arms and Insecurity. Pittsburgh: The Boxwood Press.

——— (1951) "Could an arms race end without fighting?" Nature, September 29: 567-568.

RUSSETT, B. M. [ed.] (1972) Peace, War, and Numbers. Beverly Hills, CA: Sage.

SAATY, T. L. (1968) Mathematical Models of Arms Control and Disarmament. New York: John Wiley.

———(1964) "A model for the control of arms. Operations Research 12.

SCHELLING, T. C. (1966) Arms and Influence. New Haven, CT: Yale Univ. Press.

SINGER, J. D. (1981) "Accounting for international war: the state of the discipline." J. of Peace Research 18: 1-18.

SMITH, T. C. (1980) "Arms race instability and war." J. of Conflict Resolution 24: 253-284.

SZILARD, L. (1964) "'Minimal deterrent' vs. saturation parity." Bull. of the Atomic Scientists 20: 6-12.

WALLACE, M. D. (1982) "Armaments and escalation: two competing hypotheses," Int. Studies Q. 26: 37-56.

——— (1979) "Arms races and escalation: some new evidence." J. of Conflict Resolution 23: 3-16.

——— (1972) War and Rank Among Nations. Lexington, MA: Lexington Books.

WOHLSTETTER, A. (1959) "The delicate balance of terror." Foreign Affairs 37: 211-234.