R-3765-AF

First-Strike Stability

A Methodology for Evaluating Strategic Forces

Glenn A. Kent, David E. Thaler

RAND | PROJECT AIR FORCE

The research reported here was sponsored by the United States Air Force under Contract F49620-86-C-0008. Further information may be obtained from the Long Range Planning and Doctrine Division, Directorate of Plans, Hq USAF.

ISBN: 0-8330-1004-2

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

R-3765-AF

First-Strike Stability

A Methodology for Evaluating Strategic Forces

Glenn A. Kent, David E. Thaler

August 1989

A Project AIR FORCE report
prepared for the
United States Air Force

RAND

PREFACE

This report offers a logical and transparent methodology for evaluating strategic offensive forces on the basis of first-strike stability, which the authors define as a condition that exists when neither superpower perceives the other as motivated by the posture of strategic forces to launch the first nuclear strike in a crisis. The methodology may be used to examine alternative structures of strategic offensive forces relating to the formulation of arms control proposals and the avoidance of strategic force postures that may lead to instability in a crisis.

The report, written under the National Security Strategies Program of Project AIR FORCE, is addressed to policymakers in the Administration, Congress, and the military services. The suggested methodology expands and refines an earlier one presented in Glenn A. Kent, Randall J. DeValk, and David E. Thaler, *A Calculus of First-Strike Stability (A Criterion for Evaluating Strategic Forces)*, The RAND Corporation, N-2526-AF, June 1988. In a forthcoming study, the authors apply the methodology of first-strike stability to strategic nationwide ballistic missile defenses.

SUMMARY

Strategic nuclear forces should be evaluated on their ability to underwrite key U.S. national security objectives, namely to:

- Minimize Soviet influence and coercion short of war
- Enhance arms race stability
- Eliminate any temptations for a Soviet leader to launch a nuclear first strike against the United States (i.e., provide central deterrence)
- Reduce the pressures facing either superpower leader in a crisis to launch a nuclear first strike against the homeland of the other (i.e., increase first-strike stability)
- Deter aggression against friendly and allied nations and attack on U.S. forces abroad (i.e., provide extended deterrence)
- Should deterrence fail, limit damage and deny Soviet war aims to the extent possible and encourage termination of the conflict through negotiation while preserving U.S. security and sovereignty.

The objective of first-strike stability has recently assumed increasing prominence in the public debate over strategic forces. Such prominence demands a more concise and comprehensive understanding of first-strike stability than has been evident thus far in the national dialogue. This study offers a logical and transparent methodology for evaluating postures of strategic forces—i.e., their state of generation and dispersal—that enhance first-strike stability or create first-strike instability.

First-strike instability, stemming solely from the posture of U.S. and Soviet strategic forces, differs from crisis instability, arising from each leader's perception—influenced by emotional and psychological stress, miscalculation, and disinformation—of desperate alternatives in a crisis. Crisis instability includes first-strike instability, which along with many other factors might feed crisis instability. The methodology advanced in this report deals only with first-strike instability.

First-strike stability may be considered robust when, after considering the vulnerability of strategic forces on both sides, neither leader perceives the other as pressured by the posture of forces to strike first in a crisis. Neither leader sees an advantage in striking first to avoid the potentially worse outcome of

incurring a first strike if he waits. Conversely, first-strike instability arises when either leader perceives an advantage in launching first to avoid the potentially much worse consequence of waiting and incurring a first strike.

In developing this methodology, we first demonstrate the concept of first-strike options by illustrating the opportunities available to each side to limit damage to itself by conducting a counterforce first strike, i.e., a strike against the nuclear retaliatory forces of the opposing side. These options change as the structures and postures of U.S. and Soviet strategic forces change.

Second, we define the cost of a strategic nuclear exchange. Cost is measured by two terms: (1) the damage that a country incurs and (2) the damage that a country fails to inflict on the adversary in attempting to deny him his war aims; with regard to the second term, the cost increases as the damage to the opponent decreases. **Third, given specific first-strike options, we determine the cost to each side of (1) striking first and (2) incurring a first strike.**

Finally, we calculate four costs (the cost to each side of going first and second) for each of three postures within alternative structures of U.S. and Soviet forces and reduce these costs to a single numerical index. The index permits comparisons of alternative force structures, and postures of forces within these structures, on the basis of first-strike stability.

The methodology fosters a perceptive and disciplined discourse on matters surrounding first-strike stability as follows:

- **It requires an explicit definition of the costs of a nuclear exchange** and the attitudes and beliefs as to the trade-offs between limiting damage and denying the enemy any prospect of achieving his war aims.
- **It directs attention to the vulnerability or invulnerability of all force elements deployed by both sides rather than to the vulnerability of one particular force element.**
- **It promotes insights as to the relevant questions decisionmakers would ask in a deep crisis.** Rather than dwell on target lists, a leader is more likely to ask such questions as What incentives does the enemy have for launching a first strike? What portion of our forces could the enemy believe that he could destroy if he preempted? What is the difference to this nation between striking first and running the risk of incurring a first strike if we wait?
- **It is effective for evaluating alternative force structures and postures and for examining the potential**

consequences of various arms control proposals, including breakout. Planners can readily determine how modernization programs in the presence or absence of treaty constraints could enhance or erode first-strike stability.

The consideration of first-strike stability provides the basis for the following observations:

- **The index of first-strike stability developed in this study applies equally to the United States and the Soviet Union.**

There is no U.S. index separate and distinct from the Soviet index. If the index suggests relative stability or instability, that stability or instability applies equally to each side. The United States and the Soviet Union share the national security objective of first-strike stability, and maintenance of it requires joint efforts.

- **The attacker's incentive to catch the adversary in an unprepared position when the attacker's weapons arrive drives the requirement for each side to maintain a buffer of nontargetable forces on a day-to-day basis.**

Whether or not an attack is a surprise is irrelevant. What matters is that decisionmakers might not react to impending signs of attack (so-called strategic warning) by generating forces. Thus, forces that depend on such reactions for survivability may not contribute to first-strike stability. An attacker might decide to preempt earlier rather than later to ensure that his weapons catch the adversary in a vulnerable posture. Accordingly, each side should sustain a day-to-day posture to guarantee that, although the side incurring the first strike may be surprised, it will also be prepared.

- **First-strike stability under current conditions is relatively robust.**

Even if each leader believed that he could catch the forces of the adversary at current peacetime rates of generation, these rates are currently high enough that neither leader feels pressure to strike first to avoid the potentially worse consequence of waiting, i.e., neither country could limit damage to itself significantly by initiating an attack. At the same time, each can deny the other's war aims in a first strike or second strike by inflicting devastating damage.

- **Increasing the size of U.S. and Soviet strategic nuclear inventories does not necessarily erode first-strike stability.**

Although the opportunity for massive counterforce attacks by both sides might arise with larger forces, such attacks would yield little in terms of reducing the cost to each side by striking first if each side maintained an adequate buffer of nontargetable forces. An unconstrained arms buildup, however, would undercut the objective of arms race stability.

- **Reducing the size of U.S. and Soviet strategic nuclear inventories does not necessarily enhance first-strike stability.**

Without proper attention to the posture of strategic nuclear forces—i.e., the portion of these forces that an attacker believes that he can effectively attack in a counterforce first strike—reductions in inventories could erode first-strike stability. As inventories are reduced, forces usually would require higher day-to-day generation rates to maintain the current level of first-strike stability. Any arms reduction regime, therefore, must account as much for postures of forces as for total numbers of weapons and should encourage each side to make a large portion of its forces more survivable under peacetime conditions.

- **The United States and the Soviet Union should realize the importance to both sides of generating forces early in a crisis.**

The most favorable situation obtains when no generation from peacetime rates is required to sustain an adequate number of weapons in a survivable posture. If additional generation is required, however, such generation could be viewed as a confidence-building measure in the context of first-strike stability rather than as a confrontation.

- **If U.S. policymakers believe in enhancing first-strike stability, the United States should spend money to alleviate the vulnerabilities of U.S. strategic nuclear forces rather than to increase the vulnerabilities of Soviet strategic nuclear forces.**

The vulnerabilities of Soviet forces are not necessarily in the U.S. interest, nor are U.S. attempts to create further vulnerabilities in these forces. Survivable forces on both sides remain the key to the shared objective of enhancing first-strike stability.

ACKNOWLEDGMENTS

The authors thank their RAND colleagues Timothy J. Webb and Jonathan K. Cave for reviewing this study and David A. Ochmanek, Robert A. Levine, and Dean A. Wilkening for suggesting improvements and additions to earlier drafts.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS ix

FIGURES xiii

TABLES xv

GLOSSARY xvii

Section

I. INTRODUCTION 1
   First-Strike Stability and Deterrence 1
   First-Strike Stability as a Criterion for Evaluating Strategic Forces 4
   Methodology for Examining First-Strike Stability 6

II. ASSESSING FIRST-STRIKE STABILITY 7
   Draw-Down Curves in Weapons Domain 7
   Lines of Constant Cost in Damage Domain 11
   Damage Curves Relating "Potential Damage to Value" to "Number of Weapons Available to Attack that Value" 17
   Lines of Constant Cost in Weapons Domain 20
   Coupling Lines of Constant Cost with Draw-Down Curves 22
   Index of First-Strike Stability 24

III. APPLYING THE METHODOLOGY 30
   Force Structures and Postures to be Evaluated 30
   Results 32
   Measures of National Security Objectives 43

IV. CONCLUDING REMARKS 47

Appendix

A. U.S. AND SOVIET STRATEGIC OFFENSIVE FORCE STRUCTURES 51
B. U.S. AND SOVIET STRATEGIC OFFENSIVE FORCE POSTURES 59

C. ALTERNATIVE COST FUNCTIONS 61
D. ALTERNATIVE DAMAGE CURVES 71

FIGURES

1. Generic weapons domain 9
2. Notional U.S. and Soviet first-strike attack options 10
3. Generic damage domain, given nuclear exchange 12
4. U.S. costs at damage domain corners, given nuclear exchange 14
5. U.S. lines of constant cost in damage domain, given nuclear exchange 15
6. U.S. and Soviet lines of constant cost in damage domain, given nuclear exchange 17
7. U.S. and Soviet damage curves 18
8. U.S. and Soviet lines of constant cost in weapons domain . . 21
9. Notional U.S. and Soviet first-strike attack options with lines of constant cost 23
10. U.S. decision tree 25
11. Generic probability domain 27
12. Probability domain with notional ratios of U.S. and Soviet costs 28
13. First-strike stability index of force postures in alternative U.S. and Soviet strategic offensive force structures—perception of "ride-out" 33
14. First-strike stability index of force postures in alternative U.S. and Soviet strategic offensive force structures—perception of prompt retaliatory launch 33
15. U.S. and Soviet first-strike attack options for Case I forces operating in posture A (without cost lines) 34
16. U.S. and Soviet first-strike attack options for Case I forces operating in posture A 36
17. U.S. and Soviet first-strike attack options for Case II forces operating in posture A 38
18. U.S. and Soviet first-strike attack options for Case III forces operating in posture A 39
19. U.S. and Soviet first-strike attack options for Case III forces operating in posture B 41
20. U.S. and Soviet first-strike attack options for Case IV forces operating in posture A 42
21. Measures of national security objectives 44

C.1. Soviet lines of constant cost in damage domain, given nuclear exchange and warlike Soviet leader 62

C.2. U.S. and Soviet lines of constant cost in weapons domain, given warlike Soviet leader 64
C.3. U.S. and Soviet lines of constant cost in damage domain, given nuclear exchange—damage limitation as only operational objective 65
C.4. U.S. and Soviet lines of constant cost in weapons domain—damage limitation as only operational objective . . . 66
C.5. Function based on ratios in weapons domain 68
C.6. Function based on ratios in damage domain 69
C.7. Function based on constant difference in damage domain . . 69
C.8. Function based on discounted difference in damage domain . 70
D.1. Alternative U.S. and Soviet damage curves 71
D.2. U.S. and Soviet lines of constant cost in weapons domain, given steeper damage curves 72

TABLES

A.1. Case I—Current U.S. strategic offensive force structure 52
A.2. Case I—Current Soviet strategic offensive force structure 53
A.3. Case II—Postulated U.S. strategic offensive force structure: no treaty constraints, modernized forces, and improved basing modes 54
A.4. Case II—Postulated Soviet strategic offensive force structure: no treaty constraints, modernized forces, and improved basing modes 55
A.5. Case III—Postulated U.S. strategic offensive force structure: START constraints, modernized forces, and improved basing modes 56
A.6. Case III—Postulated Soviet strategic offensive force structure: START constraints, modernized forces, and improved basing modes 57
A.7. Case IV—Alternative postulated U.S. and Soviet strategic offensive force structures: START constraints, modernized forces, and improved basing modes 58
B.1. Percentage of U.S. and Soviet forces considered nontargetable for three postulated offensive force postures 60

GLOSSARY

ALCM—Air-launched cruise missile.

ATB—Advanced-technology bomber.

central deterrence—The objective of eliminating—by threatening retaliation (countervailing action) with U.S. strategic nuclear weapons—any temptations for a Soviet leader to launch a nuclear first strike against U.S. territory. He has an unambiguous choice between striking first and waiting; if he waits he knows that there will be no war.

cost—A measure of (1) the damage suffered by one's own country and (2) the damage that a country fails to inflict on the enemy in attempting to keep him from achieving his war aims if war occurs; cf. price to attack.

counterforce attack—An attack by Country A against Country B's strategic nuclear forces—forces that Country B can direct against Country A in a retaliatory attack—in an effort by Country A to limit damage to itself; applies only to a first strike.

countervalue attack—An attack to deny the adversary his war aims; applies to first or second strike. Although many analysts use this term to describe an attack against cities, we contend that an attack against theater projection forces, war-supporting facilities, and possibly leadership in an effort to deny the adversary his war aims is an attack against the assets that the adversary values.

crisis instability—The condition that exists when either leader feels pressure because of emotion, uncertainty, miscalculation, misperception, or the posture of forces to strike first to avoid the worse consequence of incurring a first strike. The national leader then has an ambiguous choice between striking first and waiting, when, if he waits, he does not know whether he will be avoiding war or incurring a first strike.

damage curve—A curve that depicts the relationship between potential damage to one's value and the number of weapons available to the enemy to attack that value (see value).

damage domain—A domain that portrays any combination of damage to U.S. value and Soviet value; used in this methodology in defining the cost to each side of a nuclear exchange.

diminishing marginal returns—A condition describing the diminishing utility of achieving each additional increment of a given good; i.e., the diminishing utility of destroying each additional increment of enemy value.

draw-down curve—A curve that depicts the potential first-order interaction of U.S. and Soviet strategic nuclear forces—specifically, counterforce attack options in the weapons domain against given force elements—and incorporates assumptions of force structure and force posture.

extended deterrence—The deterrence of military aggression against friendly and allied nations and attacks on U.S. forces abroad by extending to them the deterrence provided by U.S. strategic nuclear forces.

first-strike instability—The condition that exists when, owing solely to the posture of forces, either leader is perceived in a crisis to feel pressure to strike first to avoid the worse consequence of incurring a first strike. The enhancement of first-strike stability can eliminate force posture as a catalyst to crisis instability. The national leader has an ambiguous choice between striking first and waiting, when, if he waits, he does not know whether he will be avoiding war or incurring a first strike.

FO—follow-on

force posture—The state of generation and dispersal of strategic forces, reflecting the portion of such forces that are nontargetable, e.g., bombers on strip alert, SSBNs at sea, dispersed mobile ICBMs, or silo-based ICBMs that can (in the eyes of the attacker) be launched under attack.

force structure—The systems constituting an individual force element (missiles, bombers, and submarines); the capabilities of those systems (probability of kill, reliability, stealth, and hardening against nuclear attack); and the basing modes (mobility and redundancy) of those systems.

generation of forces—The placement of forces in a nontargetable, and therefore survivable, posture (see force posture).

ICBM—Intercontinental ballistic missile.

index of first-strike stability—A numerical index developed in this report to rank structures and postures of strategic nuclear forces on the basis of the relative stability inherent in them.

line of constant cost—A line along which various trade-offs in terms of damage suffered versus damage inflicted are considered to have the same utility (or disutility).

price to attack—The number of weapons required to attack effectively a given force; cf. cost.

PRL—See prompt retaliatory launch.

probability domain—A domain that portrays heuristically any combination of probabilities (and perceptions of probabilities) that a U.S. or Soviet leader would launch a first strike.

prompt retaliatory launch—The act of launching one's ICBMs after the opponent has launched his, but before the opponent's warheads have exploded on target (also described as launch under attack).

RV—Reentry vehicle.

SICBM—Small intercontinental ballistic missile.

SLBM—Submarine-launched ballistic missile.

SSBN—Nuclear-powered ballistic missile submarine.

START—Strategic arms reduction talks.

value—Theater projection forces, war-supporting facilities, and perhaps leadership; see countervalue attack.

weapons domain—A domain that portrays any combination of numbers of U.S. and Soviet weapons.

I. INTRODUCTION

Strategic nuclear forces should be evaluated on their ability to support key national security objectives. One of these objectives, first-strike stability, is gaining increasing prominence as a framework in which to discuss such issues as force structure, strategic arms control, and strategic defenses. This report offers a methodology for assessing strategic forces on the basis of first-strike stability.

FIRST-STRIKE STABILITY AND DETERRENCE

National security analysts have traditionally debated strategic forces and arms control in terms of deterrence. In such debate, they distinguish between *central* deterrence, defined as the objective of convincing the Soviet Union that the United States would use its strategic nuclear forces to redress in kind aggression with nuclear weapons against its territory, and *extended* deterrence, which U.S. strategic nuclear forces project over U.S. allies, principally those in Western Europe.

To provide central deterrence, the United States seeks to ensure that the Soviet Union would never conclude that the "gain would be worth the pain" should it initiate a nuclear exchange. By this one-sided calculus, the Soviet leader always should conclude that the *cost*—measured in terms of damage to *value* (i.e., to theater projection forces, war-supporting facilities, and perhaps leadership)—of striking first would exceed greatly the cost of maintaining the status quo (i.e., no nuclear exchange).[1] As long as the Soviet leader believes this, he will feel no temptation to strike the United States first.

The strategy underlying extended deterrence involves making the Soviets believe that large-scale military aggression against Western Europe could lead to escalation. By maintaining credible first-strike options (i.e., by reducing the U.S. cost of going first), the United States presumably erodes the Soviets' confidence that a deep crisis could be controlled.

Historically, U.S. policymakers have structured U.S. strategic forces to underwrite the objective of central deterrence by avoiding any posture that might tempt the Soviet Union to strike the United States.

[1]Thus, in this vernacular, the destruction of value corresponds to the denial of war aims, rather than to the devastation of cities, the more common use; see "cost" and "value" in the Glossary, above.

We believe, however, that *the criterion of central deterrence as described above is less relevant to minimizing the likelihood of strategic nuclear war than is the more demanding criterion of first-strike stability.*

If a leader believes (as the above construct of central deterrence implies) that his choice lies unambiguously between striking first and no war if he waits, the choice seems obvious—to wait. If, in contrast, a leader believes that in waiting he takes a substantial risk of incurring an enemy first strike, his choice is less certain and the basis for making this fateful decision is far more complex.

The criterion of first-strike stability involves the concept that the consequence of waiting could be an enemy first strike. A U.S. or Soviet leader might be perceived in a deep crisis as feeling pressure to preempt if the consequences of waiting and incurring an enemy first strike potentially were much worse. First-strike stability, therefore, captures some of the dynamics of this pressure by a two-sided calculus that involves the cost to each side of striking first compared with the cost to each side of waiting and potentially incurring a first strike.

Two factors, operating interactively, might create pressure in a crisis to strike first: (1) the perception by the leader of Country A of the likelihood that the leader of Country B will launch a first strike if he, the leader of Country A, waits and (2) the extent to which the expected cost to Country A of incurring a first strike exceeds the cost to Country A of initiating a first strike. The leader of Country B would simultaneously weigh these two factors.

The most relevant questions a leader might ask in a deepening crisis include: What is the likelihood of incurring an enemy first strike if my country waits? What fraction of my country's value could the enemy destroy if he struck first? What portion of my country's strategic nuclear forces does the enemy believe he could destroy or damage if he initiated? How much better off would my country be if it preempted that attack? The methodology presented in this report captures the sense of these questions.

The report distinguishes between crisis instability and first-strike instability. The latter arises solely from the strategic force structure and the force postures within that force structure. The much broader crisis instability arises from numerous factors that might induce instability in a crisis, including psychological stress, ambiguous or incorrect information, erroneous assessments of enemy intent, miscalculation, and misperception. Crisis instability, therefore, includes first-strike instability. The methodology suggested in this report addresses only first-strike stability (or instability) stemming from the posture of strategic forces.

In sum, first-strike stability involves a two-sided calculus of each side's cost of going first compared with its cost of striking second. The calculus accounts for each side's ability to (1) limit damage to itself in a first strike by striking the opponent's strategic retaliatory forces and (2) deny the opponent's war aims in a first or second strike by inflicting damage on the enemy's value.

The discussion of first-strike stability in this report relates to concepts first enunciated by Thomas Schelling in *The Strategy of Conflict.*[2] The following points, which he made 30 years ago, remain equally valid today:

- We live in an era in which a potent incentive on either side . . . to initiate total war . . . is the fear of being a poor second for not going first.
- The enemy's invulnerability to our own first strike could be to our advantage if it relieved him of a principal concern that might motivate him to strike first. If *he* has to worry about the exposure of *his* strategic forces to an attack by *us, we* have to worry about it too.[3]
- It might seem anomalous to insist to the Russians that they cover any nakedness of their strategic forces, or for them to suggest that we protect better some of our own. More likely would be suggestions to abandon weapons that were provocatively exposed to the other side. Note how different in spirit this would be from the "ban the bomb" orientation. Whatever the propaganda implications of such a topic, it at least has the merit of viewing deterrence as something to be enhanced, not dismantled.
- At one extreme is the "pure" strike-back weapon: the relatively inaccurate vehicle with a super-dirty bomb that can kill just about everything in the enemy's country *except* a well-protected or well-hidden retaliatory force, and that itself is so well protected or well hidden as to be invulnerable to any weapons the other side might possess. Ideally, this weapon would suffer no disadvantage in waiting to strike second and gain no advantage in striking first. At the opposite extreme is a weapon that is itself so vulnerable that it could not survive to strike second, or a weapon so specialized for finding and destroying the enemy's retaliatory forces before they are launched that it would

[2]See Thomas C. Schelling, *The Strategy of Conflict*, Harvard University Press, Cambridge, Massachusetts, 1960, pp. 230–254.

[3]In all cases, the emphasis is Schelling's.

lose most of its usefulness if it were held until the other side has already started.

- The arms race does not necessarily lead to a more and more unstable situation, and it is *not* a foregone conclusion that disarmament, in the literal sense, leads to stability.

FIRST-STRIKE STABILITY AS A CRITERION FOR EVALUATING STRATEGIC FORCES

The United States now faces important decisions concerning strategic offensive forces, strategic defenses, and prospective U.S.-Soviet agreements to reduce strategic arms. These decisions have the potential to fundamentally alter the strategic landscape, and their consequences could affect the national security of the United States well into the twenty-first century.

Strategic issues under scrutiny include the modernization of the basing modes of U.S. intercontinental ballistic missiles (ICBMs), the development and deployment of nationwide strategic defenses, the development of a capability to hold Soviet mobile missiles at risk, and the planning for U.S. strategic nuclear forces constrained by a U.S.-Soviet strategic arms reduction treaty.

Decisionmakers must settle these issues, at least in part, by assessing alternative strategic forces in terms of the ability of these forces to underwrite U.S. national security objectives. The key national security objectives relevant to the evaluation of strategic forces are to

- Minimize Soviet influence and coercion short of war (mold political perceptions)
- Enhance arms race stability
- Eliminate any temptations for a Soviet leader to launch a nuclear first strike against the United States (provide central deterrence)
- Reduce the pressures facing either superpower leader in a crisis to launch a nuclear first strike against the homeland of the other (enhance first-strike stability)
- Deter aggression against friendly and allied nations and attack on U.S. forces abroad (provide extended deterrence)
- Should deterrence fail, limit damage and deny Soviet war aims to the extent possible and encourage conflict termination through negotiation while preserving U.S. security and sovereignty.

Several of these objectives conflict. The objective of limiting damage should deterrence fail, for example, might call for "decapitating" attacks on Soviet leadership and command and control elements to deny efficient execution of a Soviet counterstrike. Such attacks, however, by eliminating key Soviet leaders, might preclude the termination of hostilities through negotiation.

The most important conflict arises between the objectives of enhancing first-strike stability, on one hand, and extending deterrence and limiting damage, on the other; i.e., the more robust the Soviets believe first-strike stability to be in a deep crisis, the less they might hesitate to precipitate a deep crisis by engaging in serious aggression, for example, in Western Europe.

Balancing between first-strike stability and extended deterrence presents a problem in the planning of strategic forces. Those who would emphasize reducing the risk that a deep crisis might lead to nuclear war would give priority to enhancing first-strike stability. Conversely, those who would stress the goal of deterring the Soviets, through the presence of strategic nuclear forces, from provoking a deep crisis in the first place would give priority to strengthening extended deterrence.

Indeed, one might argue that an optimal amount of first-strike instability is possible: that is, enough to deter the Soviets from generating a major crisis (say, by invading Western Europe), but not enough to allow a major crisis to spiral out of control. Whether or not such an optimum actually exists, the concept provides the proper intellectual framework in which to think about the trade-off between first-strike stability and extended deterrence.

Official U.S. documents increasingly emphasize stability as a criterion for evaluating alternative strategic offensive force postures.[4] For example, the President's Commission on Strategic Forces, known as the Scowcroft Commission, noted the importance of stability in its April 1983 report, as follows:

> Stability should be the primary objective of our strategic forces and of our arms control proposals. . . . As we try to enhance stability . . . other objectives should be subordinated to the overall goal of permitting the United States to move—over time—toward more stable

---

[4]Though these documents refer simply to "stability" rather than to "first-strike stability," their descriptions of deployments and postures of forces demonstrate adherence to a common criterion, if not a common vernacular.

> deployments, and giving the Soviets the strong incentive to do the same.[5]

Congress recently called for an analysis of the effects of the prospective strategic arms reduction (START) treaty on first-strike stability. According to an amendment to the FY-1989 defense authorization bill, the President is required, prior to the conclusion of a treaty, to provide an analysis of alternative force postures for the United States permitted under an arms reduction agreement and "the damage-limitation capability, survivability, and retaliatory potential of each posture, and the implications for strategic stability, assessed with regard to the likely force postures of the Soviet Union and the first-strike potential of such Soviet force postures."[6]

METHODOLOGY FOR EXAMINING FIRST-STRIKE STABILITY

This report outlines a methodology for assessing the relative first-strike stability of alternative strategic offensive force postures. The methodology may be used to indicate which postures of strategic forces might act as a catalyst to instability in a crisis.

In Section II, we create the methodology that leads to a numerical index of first-strike stability. In Section III, we apply this methodology and calculate the stability index for current and postulated U.S. and Soviet strategic offensive force structures and the postures of forces within those structures.[7] Section III also suggests quantitative measures for some of the key national security objectives outlined above. Section IV presents concluding observations.

Appendix A contains tables outlining the current and postulated U.S. and Soviet strategic nuclear force structures used in the report. Appendix B describes postures of forces (primarily the degree to which they are targetable) within a single force structure. Appendix C elaborates on the concept of cost developed in Section II and treats alternative policies in assessing these costs. In Appendix D, we examine how the index of stability is affected by alternative curves that relate "potential damage to value" to "weapons available for attack of that value."

---

[5] *Report of the President's Commission on Strategic Forces,* Brent Scowcroft, Chairman, United States Government Printing Office, Washington, D.C., April 1983, p. 3.

[6] H.R. 4264, FY 89 Defense Authorization Bill, Amendments, Sec. 939, "Analysis of Alternative Strategic Nuclear Force Postures for the United States Under a Potential START Treaty," p. 198.

[7] We began the calculations for this report in 1988, using 1988 U.S. and Soviet force structures; these structures are considered current.

II. ASSESSING FIRST-STRIKE STABILITY

The proposed methodology for examining first-strike stability involves six steps:

1. Portray the first-order interaction of U.S. and Soviet strategic offensive forces with *draw-down curves* displaying *counterforce attack* options against specific force elements in the *weapons domain*.[1]
2. Define and quantify *cost* by developing lines of constant cost in the *damage domain*.
3. Link the damage and weapons domains with *damage curves* that relate the number of weapons available for attacking *value* to potential damage to that value.
4. Use the damage curves to transfer the lines of constant cost from the damage domain to the weapons domain.
5. Overlay draw-down curves representing alternative U.S. and Soviet *force postures* and read off costs to each side associated with striking first and second (four costs for each posture).
6. Define a numerical index of first-strike stability that combines the four costs; use this index to rank alternative force postures on the basis of first-strike stability.

DRAW-DOWN CURVES IN WEAPONS DOMAIN

To express the potential first-order interaction of U.S. and Soviet strategic forces, we display draw-down curves in the weapons domain. This format enables us to portray options available to each side for allocating weapons to underwrite its operational objectives of limiting damage to itself and inflicting damage on the adversary to deny him his war aims. The extent to which either side can achieve these objectives drives its costs associated with striking first or second.[2]

To limit damage to himself, an attacker conducts a counterforce attack to weaken the enemy's retaliatory blow. In a counterforce attack, the attacker strikes the enemy's strategic nuclear forces, including ICBMs in silos, ballistic missile submarines (SSBNs) in port, and

[1]Italicized words are defined in the Glossary, above.

[2]This format is described in detail in Glenn A. Kent, Randall J. DeValk, David E. Thaler, *A Calculus of First-Strike Stability (A Criterion for Evaluating Strategic Forces)*, The RAND Corporation, N-2526-AF, June 1988.

bombers at their bases. To deny the adversary his war aims, the attacker launches a *countervalue attack,* in which he strikes the adversary's theater projection forces, war-supporting facilities, and possibly leadership. Thus, in our vernacular, the destruction of value corresponds to the denial of war aims, rather than to the devastation of cities.

We assume that the retaliator cannot effectively target weapons that the first-striker does not use in his initial attack. We consider any weapons that the first-striker withholds as nontargetable because they are aboard SSBNs at sea; atop mobile ICBMs that have been dispersed; on bombers that are on strip alert; or, if on ICBMs at fixed locations, launchable under attack.

The draw-down curves in our format demonstrate the first-striker's options for counterforce attack and the weapons remaining to each side for countervalue attack. The first-striker has the option of terminating his counterforce attack at any point along his draw-down curve. The weapons available to the initiator for countervalue attack are those not used in the counterforce attack. The weapons available to the retaliator for countervalue attack are those that survive the initiator's counterforce attack.[3]

Figure 1 displays a generic weapons domain. The ordinate indicates the number of U.S. weapons available to attack Soviet value, while the abscissa indicates the number of Soviet weapons available to attack U.S. value.

In the upper left portion of the domain, many weapons are available to the United States for denying the war aims of the Soviet Union, while the Soviets retain few weapons with which to deny U.S. war aims. This situation is notionally depicted by the shaded area along the ordinate. In the lower right portion, represented by the shaded area along the abscissa, many Soviet weapons are available against U.S. value and few U.S. weapons are available against Soviet value. In the middle of the domain each side has many weapons available for countervalue attack against the other.

The shaded area in the lower right is larger than the shaded area in the upper left to reflect the assumption that the Soviet value base is larger and more dispersed than its U.S. counterpart. Further, a large portion of the U.S. arsenal consists of weapons carried by bombers that must pass through dense Soviet air defenses to reach their targets. The effects of these factors on first-strike stability are detailed below

[3]The initiator's countervalue attack might coincide with his counterforce attack rather than constituting an entirely separate phase.

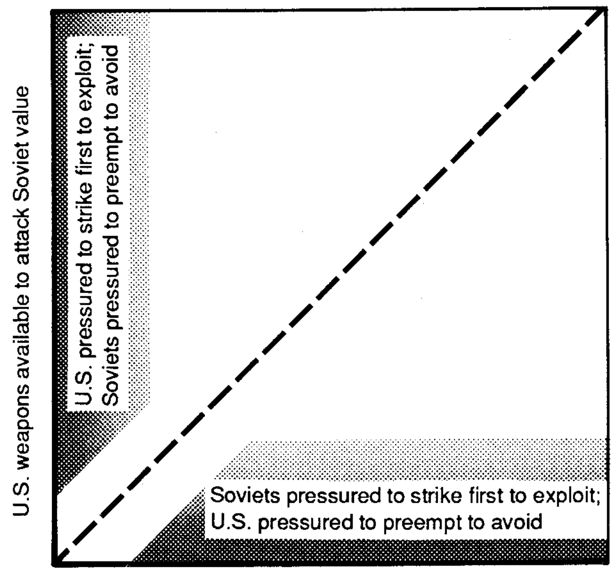


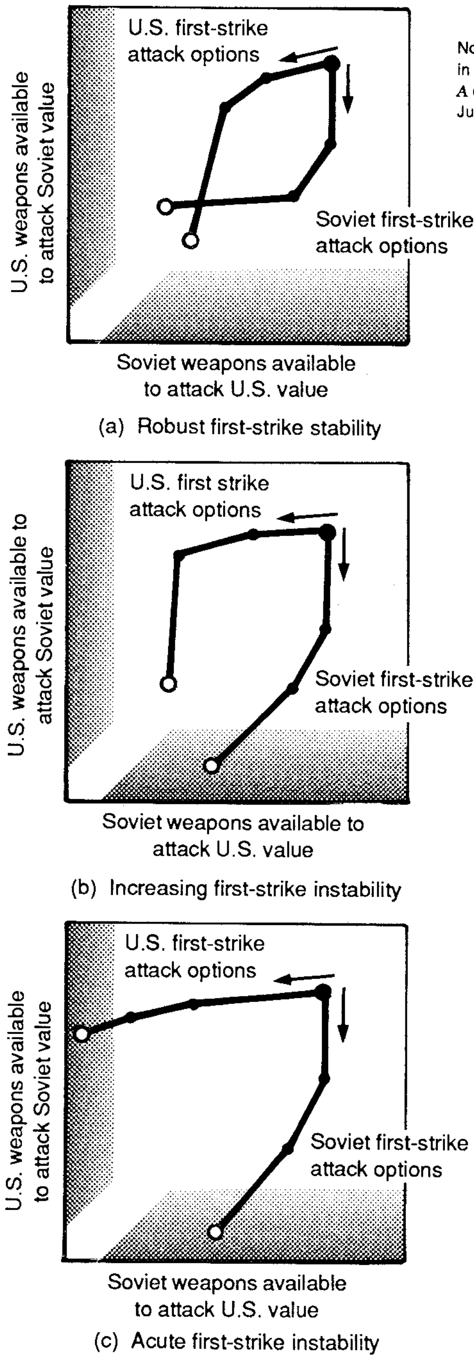
Fig. 1—Generic weapons domain

in the sections covering damage curves and lines of constant cost in the weapons domain.[4]

Figure 2 provides examples of draw-down curves representing notional force postures that give rise to various levels of stability or instability. First-strike stability is considered robust when neither side can substantially limit damage to itself by striking first, i.e., a counterforce first strike by either side results in points located toward the middle of the domain, as portrayed in Fig. 2(a). Each side maintains a large buffer of retaliatory forces that the other side cannot target.

The closer the draw-down curve depicting a Soviet counterforce first strike comes to the abscissa, as shown in Fig. 2(b), the more first-strike instability increases. First-strike instability is acute when both sides can substantially limit damage to themselves by launching a first strike and destroying the bulk of the other side's retaliatory forces, as depicted in Fig. 2(c). To maintain first-strike stability, therefore, the shaded areas must be avoided; each side, therefore, must maintain a sizable buffer of survivable forces.

[4]See also App. D.



Note: The format of these draw-down curves is explained in Glenn A. Kent, Randall J. DeValk, and David E. Thaler, *A Calculus of First-Strike Stability*, N-2526-AF, June 1988.

Fig. 2—Notional U.S. and Soviet first-strike attack options

A central purpose of our methodology is to define these shaded areas quantitatively. Such a methodology would allow us to evaluate more rigorously the relative first-strike stability of alternative force postures.

LINES OF CONSTANT COST IN DAMAGE DOMAIN[5]

To recapitulate, first-strike stability is assessed by comparing the cost to each side of striking first relative to its cost of incurring a first strike. A cost to each side is associated with each point in the weapons domain. Each point along a U.S. draw-down curve entails a cost to the United States of striking first (and suffering retaliation) and a cost to the Soviet Union of incurring a first strike and retaliating. Likewise, a cost to the Soviet Union of going first and a cost to the United States associated with going second apply to each point along a Soviet draw-down curve.

Because cost is the central issue here, we now define and quantify it, using the damage domain as the format. The damage domain is a useful tool for expressing decisionmakers' attitudes as to the trade-offs between limiting damage to one's own value and denying the adversary his war aims. These attitudes drive the quantification of cost. The damage domain shown in Fig. 3 depicts all possible combinations of damage to U.S. value and damage to Soviet value as a result of any nuclear exchange.

In an actual exchange, the overriding objective would be to limit damage to oneself. A secondary objective, given a nuclear exchange, would be to deny the enemy his war aims by inflicting damage on the enemy's value. Accordingly, the overall cost to a country involved in a nuclear exchange is a function of (1) the damage incurred and (2) the damage not inflicted (or war aims not denied). The cost to a country is driven primarily by damage suffered by that country; it increases to some extent when the country fails to deny the opponent his war aims.

Rationale for Construct of Cost in Damage Domain

To express cost in the damage domain, we assign a value of cost to each of the four corners of that domain (Fig. 3). We first examine the costs to the United States, beginning at point A in the upper left corner of the domain.

At this point, the United States suffers no damage to its own value and inflicts complete damage on Soviet value. The cost assigned to this exchange outcome (as far as the United States is concerned) is determined by the decisionmaker's view of U.S. objectives involved.

---

[5]See "line of constant cost" in the Glossary, above.

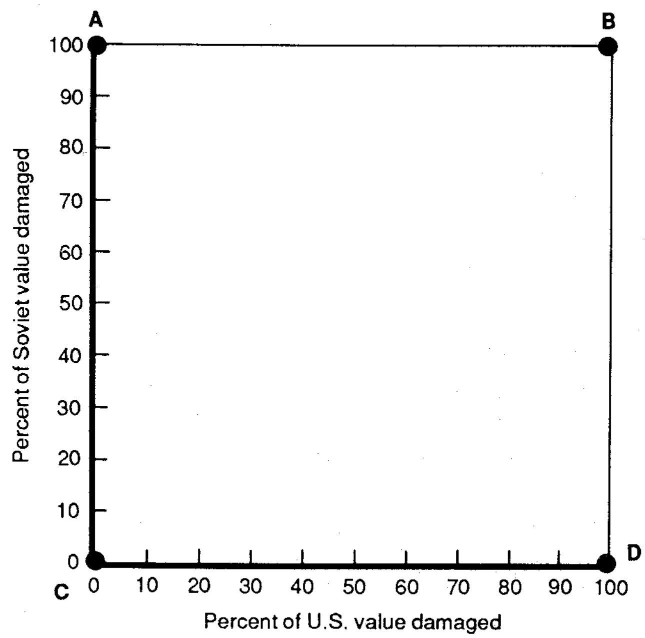


Fig. 3—Generic damage domain, given nuclear exchange

A leader might maintain, for instance, that the United States derives no benefit from a nuclear exchange with the Soviet Union—even if the United States achieves its objectives of limiting damage to itself and denying the enemy's war aims by inflicting severe damage on his value. In this case, the cost to the United States associated with point A in Fig. 3 logically cannot be less than zero.[6]

However, if a leader believes that the United States gains some benefit from a strategic nuclear exchange if his country receives no damage and achieves other objectives (such as preventing Soviet domination of Europe), the cost associated with point A would be less than zero (such as −0.2), i.e., a negative cost, or a benefit.

Whatever the cost assigned to point A, the United States loses 100 percent of its value in moving from the left side of the damage domain

[6]In fact, though point A represents the outcome most favorable to the United States, a U.S. leader might associate some cost with that outcome (i.e., a cost greater than zero, zero being indifference between war and peace). In the 1950s, for example, the United States may have had the capability to devastate the Soviet Union in first strike while receiving no damage in return. U.S. leaders did not attack, however, because of the high political and moral cost of such an action.

to the right side. If we assume the cost associated with point A to be zero, then point B—where both the United States and the Soviet Union suffer complete devastation—represents a U.S. cost of 1.0. By the same logic, point B describes a cost to the United States of 0.8 if we assign the benefit (negative cost) of −0.2 to point A.

The question now arises as to the cost associated with point C relative to the costs assigned to points A and B. Again, the answer to this question depends on the relative priority a leader assigns to his nation's objectives. One could argue, for example, that the only U.S. objective in a U.S.-Soviet strategic nuclear exchange is that of limiting damage. If so, the cost to the United States associated with point C equals that represented by point A, and whether the Soviets achieve their war aims does not concern the U.S. leader.

Conversely, a U.S. leader might have objectives other than limiting damage; he might want to inflict damage on the opponent to deny him his war aims. In this case, the U.S. cost assigned to point C in Fig. 3 would exceed the cost assigned to point A because the United States at point C failed to deny Soviet war aims. Additionally, since limiting damage to oneself is more important than inflicting damage on the adversary, the U.S. cost at point C would be less than that at point B.

In Fig. 4, we assign the U.S. costs to each corner of the damage domain. The assignments are based on the following assumptions.[7]

First, the cost to each side of no nuclear exchange, or of both sides waiting, is zero. (The cost to each of both sides waiting is not represented in the damage domain, because the damage domain implies a strategic nuclear exchange.) Second, participation in a strategic nuclear war can never have a benefit, or negative cost. Accordingly, we assign a U.S. cost of zero to the point in the upper left corner.

Third, in any nuclear exchange, a U.S. leader sees some utility in inflicting damage on his adversary; i.e., denying the Soviets their war aims matters. Accordingly, we assign a higher cost to the United States at the point in the lower left corner than at the point in the upper left corner. This represents a surcharge of 0.3 for participating in a nuclear exchange. By failing to deny Soviet war aims, the United States incurs a cost of 0.3 despite receiving no damage to its own value.[8]

Finally, a U.S. cost of 1.3 obtains if U.S. value is completely destroyed and Soviet value remains intact (the point in the lower right corner of the domain in Fig. 4). If the United States overcomes the

[7]Alternative criteria, including those cited above, are further discussed in App. C. The values assigned to the four corners are negotiable; the methodology responds to any criteria chosen.

[8]The surcharge of 0.3 is not entirely arbitrary. See App. C.

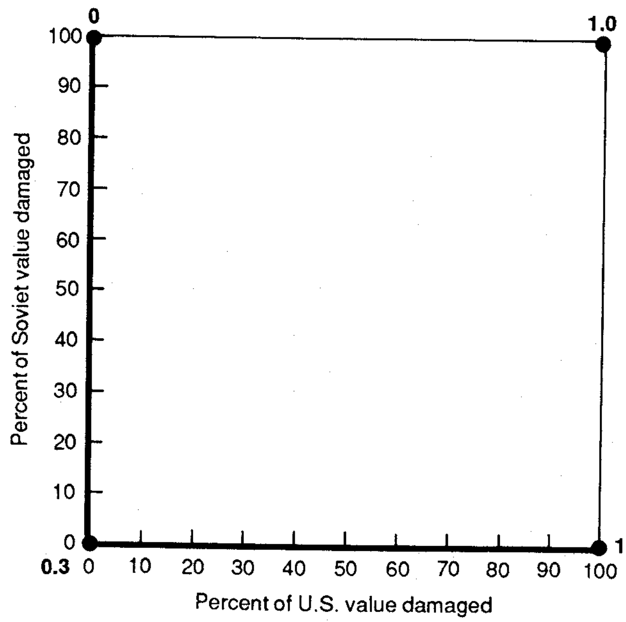


Fig. 4—U.S. costs at damage domain corners, given nuclear exchange

surcharge by completely destroying Soviet value and thwarting Soviet war aims, but is completely devastated itself, the U.S. cost is 1.0 (the point in the upper right corner).

Expressing Trade-offs with Lines of Constant Cost

With values assigned to the four corners of the damage domain, we now demonstrate the various trade-offs—in damage to U.S. value versus damage to Soviet value—that a U.S. leader might accept in different regions of that domain. To depict these trade-offs, or marginal rates of substitution, we set three additional criteria for drawing lines of constant cost within the domain.

First, the lines of constant cost must meet the boundary conditions set forth above. As displayed in Fig. 5, the United States would incur a cost of zero if it suffered no damage to itself and denied Soviet war aims. The United States would incur a cost of 1.0 by moving along the domain's boundaries from left to right, and a further cost of 0.3 along the boundaries from top to bottom.

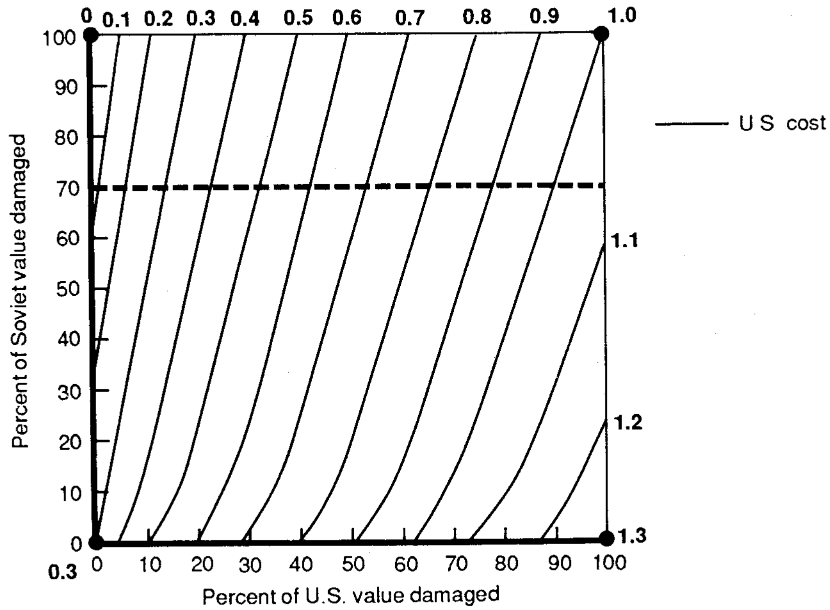


Fig. 5—U.S. lines of constant cost in damage domain, given nuclear exchange

Second, the cost lines in Fig. 5 are concave upward to reflect our perception of *diminishing marginal returns*; i.e., marginal substitution rates vary throughout the damage domain. When damage to Soviet value is high and damage to U.S. value is low, a U.S. leader sees little utility (according to our intuition) in inflicting an additional increment of damage on the Soviet Union, but great utility in further limiting damage to the United States—hence the steep slopes in the upper left area of Fig. 5. Here, Soviet war aims are all but thwarted. In the U.S. scheme of marginal returns, the trade-off swings even more toward limiting damage to U.S. value than inflicting damage on Soviet value.

Conversely, when damage to U.S. value is high and damage to Soviet value is low—as in the lower right portion of the domain in Fig. 5—the slopes are less steep. The United States has had little or no effect on Soviet war aims, but has sustained near devastation to its own value. The U.S. leader, under these circumstances, is willing to accept a higher substitution rate in damage suffered by the United States for damage inflicted on the Soviet Union.

Finally, we compress the lines of constant cost (i.e., we generate gradients that are steeper) as one moves in Fig. 5 from high damage to U.S. value to low damage (right to left). For example, at 70 percent Soviet value damaged (along the dashed horizontal line), limiting damage to U.S. value from ten percent damage to no damage is of greater utility to a U.S. leader than limiting it from 90 percent to 80 percent. Whereas from 10 percent to no damage, the United States crosses two of its cost lines (substantially reducing U.S. cost), it crosses no cost line between 90 percent and 80 percent damage to its value.

This criterion further reflects our sense of diminishing marginal returns: The U.S. leader gives his country more credit for incrementally limiting damage at low levels of damage to U.S. value than at high levels. Another way of thinking about this compression is as follows: the physical difference in damage levels is more readily distinguishable between 10 percent and no damage than between 90 percent and 80 percent damage.

We draw Soviet lines of constant cost in Fig. 6 by inverting the U.S. lines of constant cost. The Soviet cost of 0 is represented by the point in the lower right corner of the domain, and the cost of 1.3 is in the upper left corner. Though we have assumed similar U.S. and Soviet attitudes for the purposes of this discussion, Soviet cost lines need not mirror U.S. cost lines.[9]

The cost matrix in Fig. 6 reflects our judgment as to the costs and the marginal substitution rates between limiting damage and denying war aims by inflicting damage. This matrix represents our sense of policy and attitudes. The numerical values assigned to the corners can be changed and the gradients and curvatures of the cost lines can be exaggerated or de-emphasized. However, as we demonstrate in Appendix C, such changes affect neither our basic methodology nor the conclusions we reach from applying the methodology.

To make our cost matrix useful, we should relate it to the drawdown curves in the weapons domain that depict a first-order interaction of strategic forces. Thus, we must replicate in the weapons domain the lines of constant cost that we portrayed in the damage domain in Fig. 6.

---

[9]The leader or decisionmaker might create cost lines incorporating perceived differences in U.S. and Soviet cost functions. See App. C.

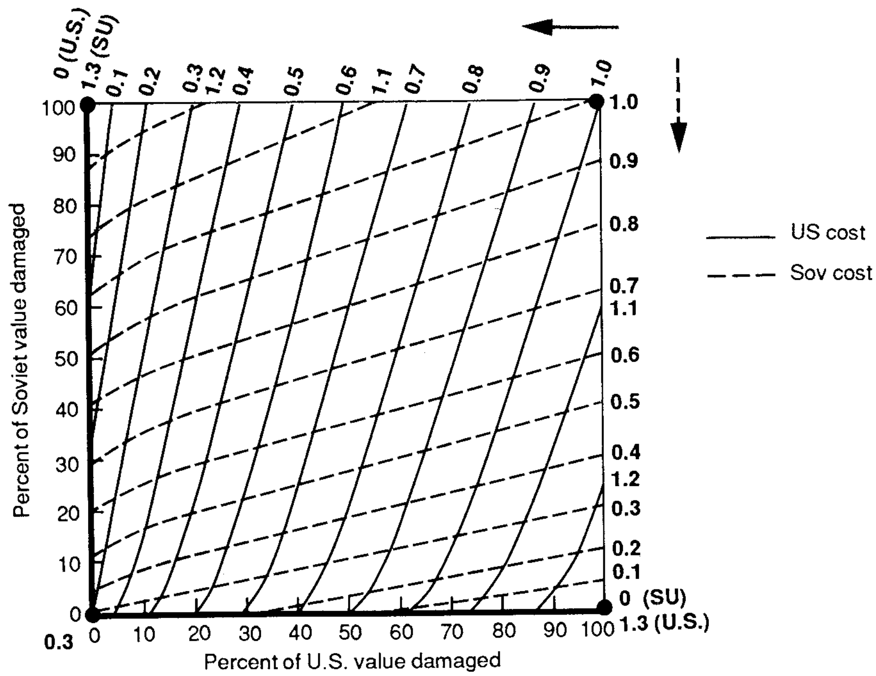


Fig. 6—U.S. and Soviet lines of constant cost in damage domain, given nuclear exchange

DAMAGE CURVES RELATING "POTENTIAL DAMAGE TO VALUE" TO "NUMBER OF WEAPONS AVAILABLE TO ATTACK THAT VALUE"

To replicate our lines of constant cost in the weapons domain, we draw a pair of *damage curves*, one U.S. and one Soviet, that expresses the relationship between denying enemy war aims and the weapons available for that purpose. These damage curves, displayed in Fig. 7, provide the potential countervalue destructive capacity of any given number of U.S. or Soviet weapons.[10]

[10]Altering the slopes of the curves does not affect the basic methodology. However, altering the damage curves affects the measure of stability in significant, yet predictable, ways. In App. D, we conduct an excursion in which a new Soviet curve resembles the present U.S. curve, while a new U.S. curve rises more rapidly. The excursion demonstrates that using steeper damage curves reinforces our concluding observations.

For other sources that discuss "value," see Michael M. May, George F. Bing, and John D. Steinbruner, "Strategic Arsenals After START: The Implications of Deep

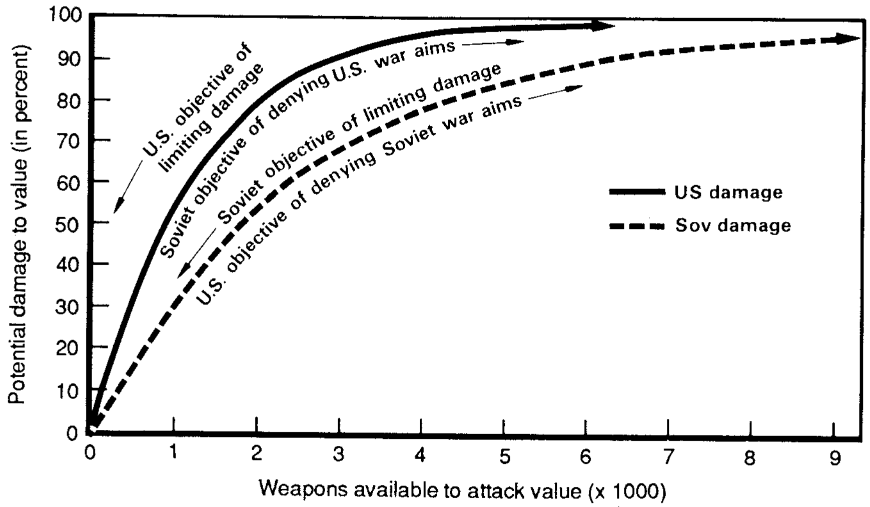


Fig. 7—U.S. and Soviet damage curves

In drawing these damage curves, we express crisis decisionmaking in a macro sense. Our intention is to address the following questions, which the President likely would ask his advisers in a deep crisis in deciding whether to strike first or wait:

- What is the likelihood that the Soviets will attack us if we wait?
- What fraction of our national value could the Soviets destroy if they launched first?
- What portion of our strategic retaliatory forces could the Soviets believe that they could catch in a vulnerable posture if they attacked now?
- How much better off would we be if we preempted their attack?

The President's advisers could be fairly certain as to the number of residual U.S. and Soviet weapons available (in first strike or in retaliation) for countervalue attack. In this regard, the inputs are relatively

Cuts," *International Security*, Vol. 13, No. 1, Summer 1988, p. 115; Desmond Ball, "Targeting for Strategic Deterrence," *Adelphi Papers*, No. 185, International Institute for Strategic Studies, London; and Ashton B. Carter, "BMD Applications: Performance and Limitations," *Ballistic Missile Defense*, edited by Ashton B. Carter and David N. Schwartz, The Brookings Institution, Washington, D.C., 1984, p. 168.

straightforward, involving such factors as the *price to attack* each side's strategic retaliatory forces and the number of weapons invulnerable to attack.

However, these advisers could only speculate about the President's first question as to the likelihood of a Soviet attack if he waits. Also, they could not meaningfully estimate the damage that the United States would incur in striking first or incurring a first strike; i.e., they could not predict the Soviet targeting doctrine with any reliability. Without knowing how the Soviet Union would use its residual weapons, the advisers could only inform the President as to the damage a given number of Soviet weapons *could* inflict on the United States. On this crucial point, the U.S. (or the Soviet) President must base his decision upon imperfect knowledge.

Both curves rise rather sharply from the origin and become asymptotic (flat) as the number of weapons targeted against value increases. This reflects the concept of diminishing marginal returns in that each additional increment of weapons yields a declining percentage of additional value destroyed. When the number of weapons on value is high, significant changes in that number only marginally affect damage levels, since we read these levels off the asymptotic portion of the damage curves. Conversely, when the number of weapons is low, we read damage levels off the steep part of the damage curves; small changes in the number of weapons cause substantial variations in damage levels.

Three factors cause the U.S. damage curve to rise more sharply than the Soviet damage curve: We assume that (1) Soviet value is larger, more dispersed, and more hardened against nuclear effects than U.S. value; (2) Soviet bomber defenses are more effective than U.S. bomber defenses; and (3) the average Soviet weapon generally is larger than the average U.S. weapon.

According to Fig. 7, both leaders perceive that 2000 Soviet weapons could destroy some 80 percent of U.S. value. The same number of U.S. weapons potentially could damage only about 55 percent of Soviet value.[11]

[11] In effect, we assume that an objective observer with perfect knowledge of each side's target base and countervalue attack capabilities has drawn these damage curves. Indeed, four curves would be more realistic, given that each leader maintains a perception of the effectiveness of his countervalue attack that differs from his opponent's perception of that attack. Thus, for instance, there could be two Soviet curves: one depicting the Soviet perception of U.S. countervalue attack capability and one depicting the U.S. perception of that capability. We believe, however, that our depiction of the U.S. and Soviet damage curves is adequate. Drawing four curves entails a sense of precision that is not warranted.

LINES OF CONSTANT COST IN WEAPONS DOMAIN

Using the damage curves described above, we now undertake to transfer U.S. and Soviet lines of constant cost from the damage domain to the weapons domain. Taking a point along a U.S. cost line in the damage domain in Fig. 6, we establish U.S. value damaged and Soviet value damaged. We enter these values onto the U.S. and Soviet damage curves introduced in Fig. 7, above, and determine the number of U.S. and Soviet weapons required to inflict these levels of damage.

The number of U.S. weapons and Soviet weapons represent coordinates that describe one point on that particular U.S. cost line in the weapons domain. Repeating this process for several points along that same U.S. cost line allows us to mechanically replicate the entire line in the weapons domain. We transfer the other U.S. and Soviet cost lines in the same manner.

Figure 8 depicts lines of constant cost in the weapons domain for both the United States and the Soviet Union. The plot of the lines is not intuitively obvious. The regions where changes in the number of weapons cause significant changes in cost generally resemble the shaded areas shown in Fig. 1, above. At this point, we can describe these shaded areas in detail.

At point A, for example, U.S. cost is relatively low primarily because few Soviet weapons are available to attack U.S. value. Thus, the number of available Soviet weapons on U.S. value associated with point A in Fig. 8 provides a point on the steep portion of the U.S. damage curve in Fig. 7. Any increase or decrease in the number of weapons substantially changes potential damage to U.S. value (and thus greatly affects U.S. cost). Conversely, the number of U.S. weapons available to attack Soviet value represents a point along the asymptotic portion of the Soviet damage curve. Any decrease in U.S. weapons available, therefore, changes only marginally the percentage of Soviet value damaged (and thus has negligible effect on U.S. cost).

The U.S. and Soviet lines of constant cost in this region of the weapons domain are parallel to the ordinate. The steep gradients between U.S. cost lines in this region suggest that the United States could lower its cost significantly if its counterforce attacks could reduce residual Soviet weapons to fewer than about 2000 weapons. Additionally, the farther below 2000 weapons a U.S. counterforce attack reduces Soviet retaliatory forces, the faster the United States diminishes its cost.

At point B in Fig. 8, many Soviet weapons are available to attack U.S. value and few U.S. weapons are available to attack Soviet value. Thus, damage to Soviet value is relatively low and damage to U.S.

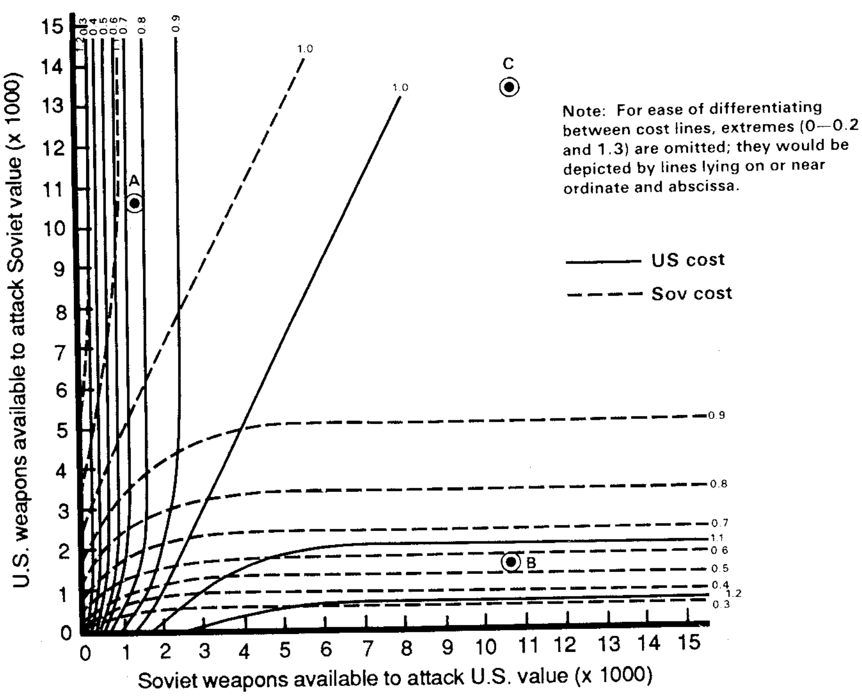


Fig. 8—U.S. and Soviet lines of constant cost in weapons domain

value is relatively high. Additionally, the gradients between the Soviet cost lines in Fig. 8 in the region of point B are less steep than those of U.S. cost lines in the region of point A because the Soviet damage curve (Fig. 7) does not rise as steeply as the U.S. damage curve.

Point C in Fig. 8 is located in an area where many weapons are available to each side for attack of the opponent's value. Gradients in this quadrant are present, but slight. Even if their counterforce attacks were more effective, neither the United States nor the Soviet Union could significantly reduce the cost (limit the damage) to themselves. To get from point C to the steep gradients along the ordinate, the United States would require a capability to draw down Soviet weapons from 11,000 to approximately 1500. Likewise, a Soviet counterforce attack must reduce U.S. weapons by some 10,000 to reach the steep gradients along the abscissa from point C.

COUPLING LINES OF CONSTANT COST WITH DRAW-DOWN CURVES

We can use the cost lines in the weapons domain repeatedly and without recalculation to examine the relative stability of any force posture. Accordingly, we couple U.S. and Soviet lines of constant cost in the weapons domain with the draw-down curves that represent the potential first-order interaction of strategic forces. Figure 9 depicts notional draw-down curves for explanatory purposes; in Sec. III, we evaluate actual and postulated U.S. and Soviet force structures and the various postures within those force structures.

In Fig. 9(a), we portray a notional U.S. and Soviet force posture that exhibits relatively robust first-strike stability. Neither side could significantly limit damage to itself by striking first because each side maintains a sizable buffer of survivable forces. This is shown by their failure to reach their own areas of steep gradient; each could reduce cost only marginally. Under these circumstances, therefore, there is little difference between the cost to each side of launching a first strike and the cost of waiting and incurring a first strike.

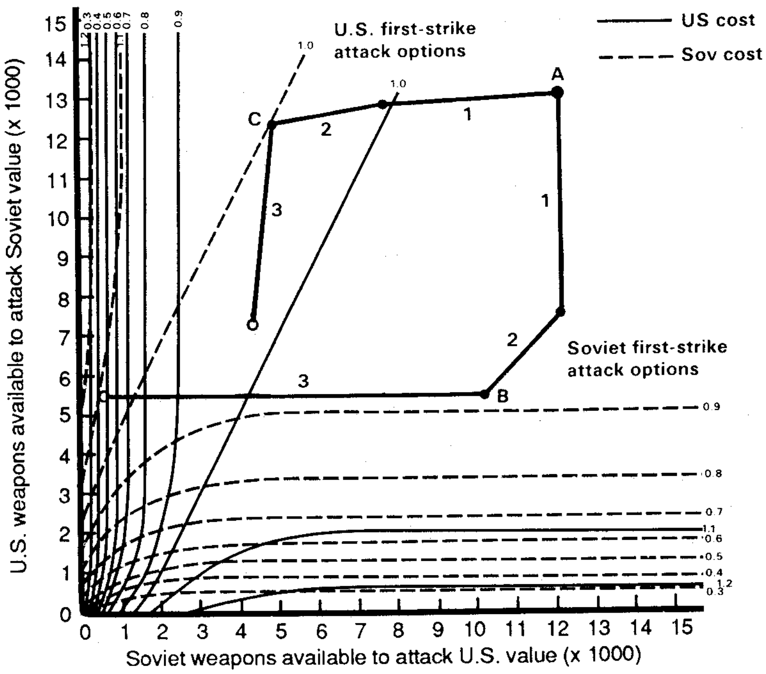
The cost lines aid in determining the point at which each leader would probably terminate his nation's counterforce attack, with the remaining weapons available for countervalue attack. The first priority of the initiator would be to minimize his own cost. If the first-striker's cost remained constant (i.e., the cost had been reduced to the extent possible), his second priority would be to maximize the enemy's cost—without, of course, increasing his own.[12]

In terms of the notional Soviet attack curve in Fig. 9(a), starting from point A, the Soviet leadership would conduct the counterforce attacks represented by segments 1 and 2 of that curve. Such attacks would reduce Soviet cost somewhat without a large expenditure of Soviet weapons. Segment 3, however, runs parallel to the Soviet 0.9 cost line, suggesting an inefficient counterforce attack, in which the Soviets would expend many weapons but destroy few U.S. weapons.

By following segment 3, the Soviets soon would increase their own cost and reduce the cost to the United States. They would begin disarming themselves, thereby undermining their secondary objective of denying U.S. war aims. Thus, using our formula, the Soviets likely would terminate their counterforce attack at point B, leaving over 10,000 Soviet weapons available against U.S. value and 5500 U.S. weapons available against Soviet value.

[12]Clearly, if a leader were interested only in limiting damage to his own country, he would conduct inefficient counterforce attacks as long as these attacks yielded even a small number of the adversary's weapons destroyed. See App. C.

(a)



(b)

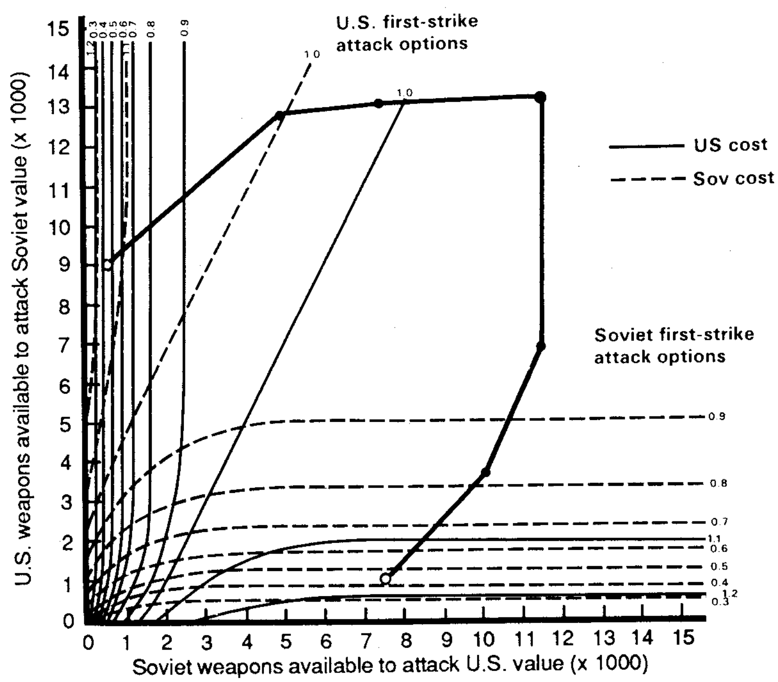


Fig. 9—Notional U.S. and Soviet first-strike attack options with lines of constant cost

If the United States struck first, its counterforce attack would end at point C after the strikes represented by segments 1 and 2 on its attack curve. Further counterforce attacks (segment 3) would increase U.S. cost and decrease Soviet cost. Following a U.S. counterforce first strike, 12,000 U.S. and 5000 Soviet weapons would be available for countervalue attack.

We derive four costs from each pair of draw-down curves. At point C on the U.S. draw-down curve in Fig. 9(a), the U.S. cost of going first is about 0.97 and the Soviet cost of incurring a first strike, 1.0. At point B on the Soviet curve, the Soviet cost of going first is 0.9 and the U.S. cost of incurring a first strike is 1.05.

The range between the Soviet costs associated with going first and second is small—0.9 compared with 1.0. For the United States, this is even smaller—0.97 compared with 1.05. Since there is little difference to either side between going first and incurring a first strike, first-strike stability is robust.

Figure 9(b) portrays a situation in which first-strike instability is acute because each country maintains a large portion of its forces in a vulnerable posture. By striking first, each side could drive its attack curve into the steep gradients along its own axis. Under these circumstances, each side could substantially reduce the cost (limit the damage) to itself by striking first.

We can gauge the extent of relative instability by establishing the costs to each side. The range of the costs to each of initiating and incurring a first strike is considerable: for the United States, 0.5 compared with 1.18, and for the Soviets, 0.45 compared with 1.15. The postures represented in Fig. 9(b), therefore, exhibit a much greater level of first-strike instability than those of Fig. 9(a).

INDEX OF FIRST-STRIKE STABILITY

We now extend the methodology to create a numerical *index of first-strike stability*. The index will provide a practical and reliable way to rank alternative force postures in terms of the relative first-strike stability inherent in each. As described above, the output of our methodology so far, for each posture of U.S. and Soviet strategic nuclear forces, is a set of four costs. We present below a method for collapsing these four costs into one numerical index.

Our guideline in deriving a first-strike stability index is that it must relate *solely* to force structures and the various postures of forces within those force structures. We eschew the use of the index to predict the probability of war, given a deep crisis (crisis instability).

The decision to strike first or wait is taken in an emotionally volatile atmosphere, in which a leader seemingly faces desperate alternatives and ambiguous (if not wrong) information; such a decision depends largely on nonquantifiable factors.

The stability index serves simply as an indicator of the relative stability or instability of alternative postures of forces. When the index indicates a decrease in stability, we seek the underlying causes by returning to the draw-down curves atop the cost lines in the weapons domain.

Decision Tree[13]

We derive the stability index from decision theory, portrayed by the decision tree in Fig. 10. At any time, the leader of the United States (and the leader of the Soviet Union) has two fundamental choices: to initiate a strategic nuclear attack on the Soviet Union (United States) or not to, i.e., to wait. These two choices, initiating or waiting, are represented by the two main branches of the U.S. decision tree.

A U.S. decision to launch a first strike leads down the "initiate" branch. In this situation, the cost relevant to the United States is its cost of striking first (C_1^{US}). Submerged in this cost is the assumption of certainty (as perceived by the U.S. leadership) that the Soviets would retaliate if the United States struck first.

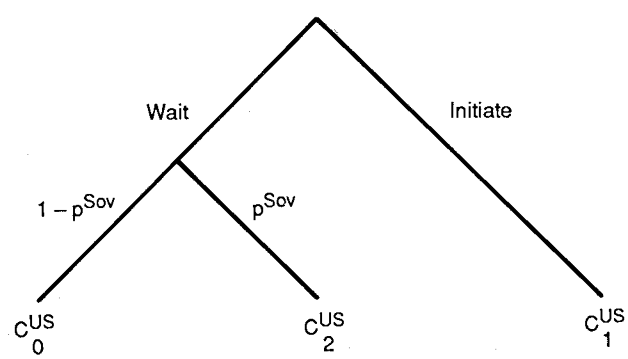


Fig. 10—U.S. decision tree

[13]Variations of the decision tree have been demonstrated by Albert Carnesale to the senior author and by Dean Wilkening and Kenneth Watman in *Strategic Defenses and First-Strike Stability*, The RAND Corporation, R-3412-FF/RC, November 1986.

The "wait" branch signifies a U.S. decision not to launch a first strike. The "wait" branch has two branches: one denoting a Soviet decision to strike first, the other a Soviet decision to wait. If, by waiting, the U.S. leader incurs a Soviet first strike, the U.S. cost of going second (C_2^{US}) applies. Associated with the expected value of the cost to the United States of incurring a first strike is the probability that the Soviets will strike first (p) if the U.S. leader chooses to wait.

If the Soviets also wait, the United States incurs no cost, as we equate the cost of no nuclear exchange (C_0^{US}) to 0. Associated with the U.S. cost of no exchange is the likelihood that the Soviet leader will wait ($1-p$) if the U.S. leader waits.

If the U.S. leader—in terms of expected value—perceives the same utility (or disutility) to initiating and waiting, i.e., he "prefers" neither branch, then the expected values of the two branches are equal. Mathematically, we express this equality by

$$C_0^{US}(1 - p) + C_2^{US}(p) = C_1^{US}$$

Eliminating the first term (since $C_0^{US} = 0$) reduces the equation to

$$C_2^{US}(p) = C_1^{US}$$

Solving for p yields a ratio of the U.S. cost of striking first to the U.S. cost of incurring a first strike and striking second, or

$$p = \frac{C_1^{US}}{C_2^{US}}$$

We interpret this equation as follows: If the U.S. leader believed that the probability of a Soviet first strike (if he waited) was greater than the U.S. ratio of expected costs, he would perceive an advantage in initiating. If, on the other hand, he believed that the probability of a Soviet first strike was less than the U.S. ratio of expected costs, he would perceive an advantage in waiting.

We develop the Soviet decision tree in the same fashion. There is also a ratio of the Soviet cost of striking first to the Soviet cost of incurring a first strike and retaliating. The same logic, therefore, applies to the Soviet perception of their advantage in initiating or waiting. By definition, C_1 is less than or equal to C_2 for both countries.

Probability Domain

Coupling the U.S. and Soviet ratios of costs is the next step in the process of establishing an index of first-strike stability. Figure 11 is a generic *probability domain* that provides graphically any combination of probabilities, actual or perceived, of U.S. and Soviet initiation. The ordinate and the abscissa in Fig. 11 are drawn for heuristic purposes only; we do not calculate or assume these probabilities or perceptions in this methodology.

The legend on the ordinate involves two factors: (1) the U.S. likelihood of initiating and (2) the Soviet perception of that likelihood. The actual and perceived probability of a first strike in a given situation would differ. For instance, the Soviets might perceive the likelihood of U.S. initiation as high, while the actual likelihood of U.S. initiation might approach zero. The abscissa pertains to the Soviet probability of initiating and the U.S. perception of that probability. Because they depict probabilities, the ordinate and the abscissa range between 0 and 1.0.

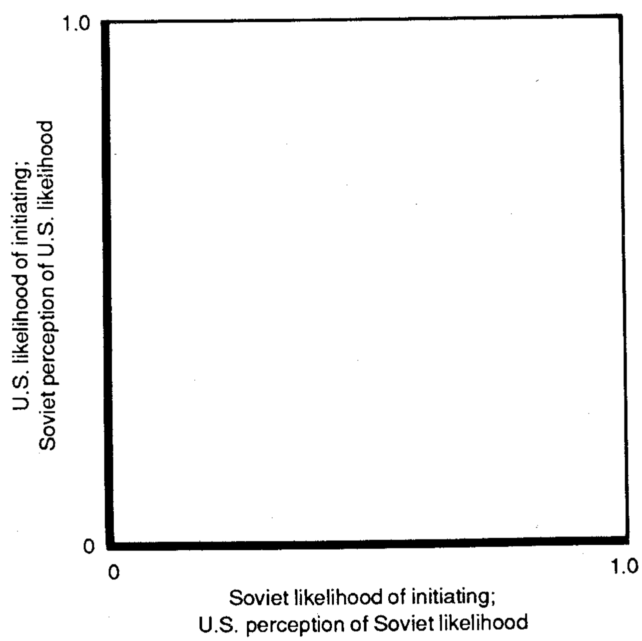


Fig. 11—Generic probability domain

In Fig. 12, we discard the legends, as they do not pertain to our calculations. Instead, we plot a point along the abscissa that represents the ratio of U.S. costs for some notional force posture (determined by superimposing the associated draw-down curves in the weapons domain atop the lines of constant cost). From that point, we draw a vertical line. If the U.S. leader perceives that the Soviet likelihood of initiating is to the right of that line, decision theory suggests an advantage to the United States in preempting; to the left, a perceived U.S. advantage in waiting.

Likewise, we draw a horizontal line from the point on the ordinate representing the ratio of Soviet costs. According to decision theory, the Soviet Union would see an advantage in preempting if it perceived the likelihood of a U.S. first strike as above the line (i.e., greater than their ratio of costs). Similarly, the Soviets would see an advantage in waiting if they perceived the U.S. likelihood as below this line.

The probability domain in Fig. 12 now contains four boxes: (A) both leaders perceive an advantage in waiting, (B) the U.S. leader

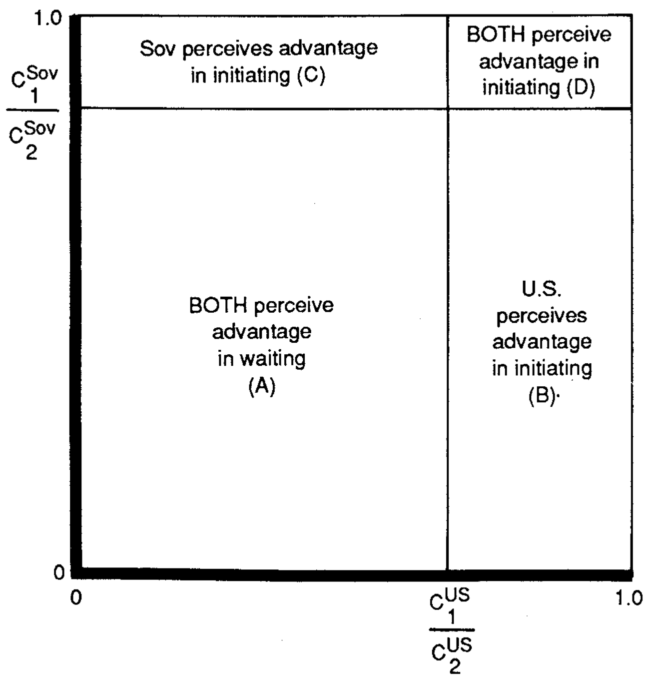


Fig. 12—Probability domain with notional ratios of U.S. and Soviet costs

perceives an advantage in initiating, (C) the Soviet leader perceives an advantage in initiating, and (D) both leaders perceive an advantage in initiating.

We define the index of first-strike stability as the ratio of the area of box A, where both leaders perceive an advantage in waiting, to the total area of the domain (A + B + C + D). Arithmetically, then, we express the stability index as

$$\left(\frac{C_1^{US}}{C_2^{US}}\right)\left(\frac{C_1^{Sov}}{C_2^{Sov}}\right)$$

This is simply the height of box A multiplied by its width since, by definition, the index is bounded by 0 and 1.0.[14]

The larger the area of box A, the higher the level of first-strike stability. If box A fills the whole area of the domain—there is no difference to either country between going first or second—the index is 1.0 and conditions are completely stable, at least with regard to the posture of forces. If the index is 0 (one or both sides can, in a first strike but not in a second strike, devastate the enemy and receive no damage in return), neither leader perceives an advantage in waiting and conditions are completely unstable.

Using the index, we can rank alternative force postures on the basis of the first-strike stability criterion. In Fig. 9(a), above, we demonstrated draw-down curves representing notional force postures. We determined that the U.S. costs associated with going first and second were 0.97 and 1.05, respectively, and that the Soviet costs were 0.90 and 1.0. According to our formula, the stability index in this situation would be

$$\left(\frac{C_1^{US}}{C_2^{US}}\right)\left(\frac{C_1^{Sov}}{C_2^{Sov}}\right) = \left(\frac{0.97}{1.05}\right)\left(\frac{0.90}{1.0}\right) = (0.92)(0.90) = 0.83$$

Under the notional force posture portrayed in Fig. 9(b), the United States would incur a cost of 0.5 if it struck first and 1.18 if it suffered a Soviet first strike. A Soviet first strike would result in a Soviet cost of 0.45, while the Soviet cost associated with striking second would be 1.15. These four costs generate a stability index of 0.16. Thus, the posture in Fig. 9(b) exhibits much greater first-strike instability than the posture in Fig. 9(a).

[14]The authors are indebted to Jonathan K. Cave of RAND for suggesting that the index derives from the ratio of the area of box A to the total area of the domain.

III. APPLYING THE METHODOLOGY

FORCE STRUCTURES AND POSTURES TO BE EVALUATED

Having established a methodology for examining first-strike stability, we now apply the methodology to assess the relative first-strike stability of current and postulated U.S. and Soviet strategic offensive force postures. We examine four cases, each representing a pair of U.S. and Soviet strategic nuclear force structures. Within each force structure, we examine several postures—states of *generation*—of these forces.

In analyzing various postures of forces, we purposely avoid the constructions of surprise attack and "bolt out of the blue" or "gray." Following such constructions leads some analysts to argue that examinations of day-to-day postures in the strategic realm are unrealistic. They contend that either side would have enough strategic warning in a deep crisis to generate forces to a more survivable posture than had obtained on a day-to-day basis.

The decisionmakers do not always react to signs of impending attack, however, and the attacker has an incentive to catch the victim in a vulnerable posture and strives to do so. For instance, historians still debate whether the Japanese attack in 1941 against U.S. airfields in the Philippines was a "surprise attack." It could hardly be argued that the attack was a complete surprise to some U.S. officials at the time. They knew that the Japanese fleet was at sea, and that a portion of that fleet had already attacked Pearl Harbor. Still, the Japanese caught the U.S. forces in an unprepared posture.

We contend that the posture of strategic nuclear forces is not immune to decisionmakers' failure to act. It is possible for the attacker to perceive that he could catch the adversary's strategic nuclear forces in a vulnerable posture. Analyses where the attacker strives to catch the adversary in a less survivable day-to-day posture, therefore, are highly relevant.

We examine three postures of forces: A, B, and C.[1] Posture A is equivalent to current day-to-day U.S. and Soviet force postures (their lowest state of generation). This posture represents a situation in which the first-striker attacks before the retaliator generates additional forces.

[1]See App. B.

We assume in posture B a medium state of generation, in which each side has generated more of its forces before the other side launches a first strike. Each side has placed more bombers on strip alert, put more SSBNs to sea, and dispersed more mobile ICBMs than in posture A. Overall, forces operating in posture B are less vulnerable to attack than those in posture A.

Posture C signifies a state of maximum generation, in which all bombers in good repair have been placed on strip alert, all SSBNs not in overhaul have put to sea, and all mobile ICBMs have been dispersed. Variations in postures generally apply to the side incurring a first strike. We assume that the initiating side always operates from posture C.

For each force posture within each force structure, we gauge the effect on first-strike stability of another action rendering forces nontargetable: *prompt retaliatory launch* (PRL). PRL, also known as launch under attack, involves the retaliator launching his ICBMs on tactical warning (i.e., upon confirmation that an attack has begun) before the first-striker's warheads can destroy them in their silos.

The force structures that we analyze are characterized as follows:

- Case I: current U.S. and Soviet forces
- Case II: postulated, modernized U.S. and Soviet forces unconstrained by an arms reduction treaty
- Case III: postulated, modernized U.S. and Soviet forces constrained by a strategic arms reduction (START) treaty; U.S. MX deployed in rail garrison mode
- Case IV: postulated, modernized U.S. and Soviet forces constrained by a START treaty; U.S. road-mobile Midgetman deployed.

In Case I, silo-based reentry vehicles (RVs) constitute 60 percent of the Soviet strategic nuclear weapons inventory. U.S. RVs on silo-based ICBMs make up only 18 percent of the U.S. arsenal. The United States possesses a little over 1800 warheads—all on ICBMs—capable of effectively attacking over 1300 hardened Soviet silos. Over 5000 Soviet ICBM RVs have hard-target-kill capability against the 1000 U.S. silos. Bombers and submarine-launched ballistic missiles (SLBMs) carry most of the U.S. inventory. Notably, bomber weapons make up less than 10 percent of the Soviet force structure.

Case II is characterized by a considerable expansion of superpower strategic nuclear inventories in a world of unconstrained competition. Both sides have increased manifold their hard-target-kill-capable RVs.

Case III represents notional forces under a U.S.-Soviet START treaty limiting each to 6000 accountable weapons and 4900 ballistic missile RVs. However, each side maintains an actual inventory of over 9000 weapons by exploiting an assumed counting rule for bomber weapons in which bombers not designated as air-launched cruise missile (ALCM) carriers count as one weapon. Because we assume that the Soviets choose to take advantage of the rule, bomber weapons now comprise 45 percent of the Soviet strategic nuclear inventory.

In addition, a greater proportion of Soviet ICBMs are mobile; still, over half (1540) of all Soviet ICBM RVs remain in fixed sites and in only 154 silos. The United States deploys its 50 MX ICBMs in a rail-garrison mode while maintaining the Minuteman IV, a single-RV, silo-based ICBM. Further, the United States greatly enhances its hard-target-kill capability with the deployment of D-5 SLBMs.

Case IV is a variant of Case III: Leaving the 50 MX ICBMs in silos, the United States deploys 500 single-RV Midgetman ICBMs (also known as the small ICBM, or SICBM) on hardened, road-mobile launchers. All other assumptions concerning force structure in Case III pertain as well to Case IV.[2]

RESULTS

The bar charts in Figs. 13 and 14 illustrate the results of applying our methodology of first-strike stability to the force structures and force postures described above. Figure 13 depicts the initiator's perception of "ride-out"—absorbing an attack before reacting—by the retaliator's ICBMs, and Fig. 14 portrays the initiator's perception that the retaliator would conduct a PRL of his ICBMs. Each case (bar) represents a force structure, and the shadings within each bar in Figs. 13 and 14 signify the value of the index for various force postures: darkened for posture A, crosshatched for posture B, and dotted for posture C.

Case I: Current U.S. and Soviet Strategic Offensive Forces

Figure 15 displays the interaction of current U.S. and Soviet forces (Case I) operating at posture A. For the purpose of illustration, we focus on the Soviet draw-down curve portraying Soviet first-strike attack options.

---

[2]Appendix A provides details of current and postulated U.S. and Soviet force structures: Tables A.1 and A.2 describe Case I, Tables A.3 and A.4 describe Case II, Tables A.5 and A.6 describe Case III, and Table A.7 describes Case IV.

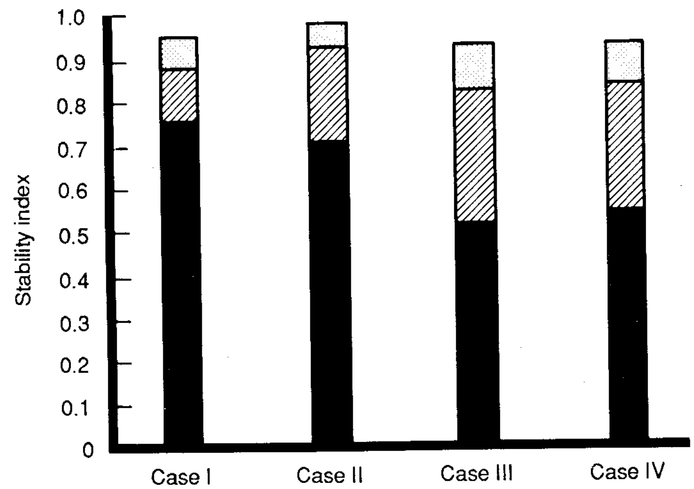


Fig. 13—First-strike stability index of force postures in alternative U.S. and Soviet strategic offensive force structures —perception of "ride-out"

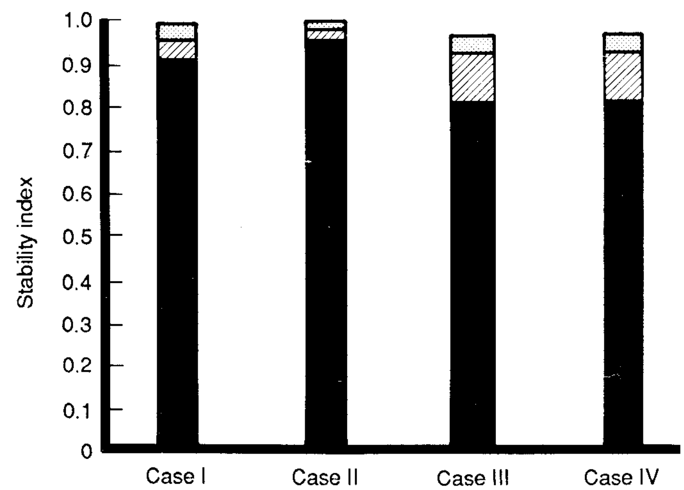


Fig. 14—First-strike stability index of force postures in alternative U.S. and Soviet strategic offensive force structures —perception of prompt retaliatory launch

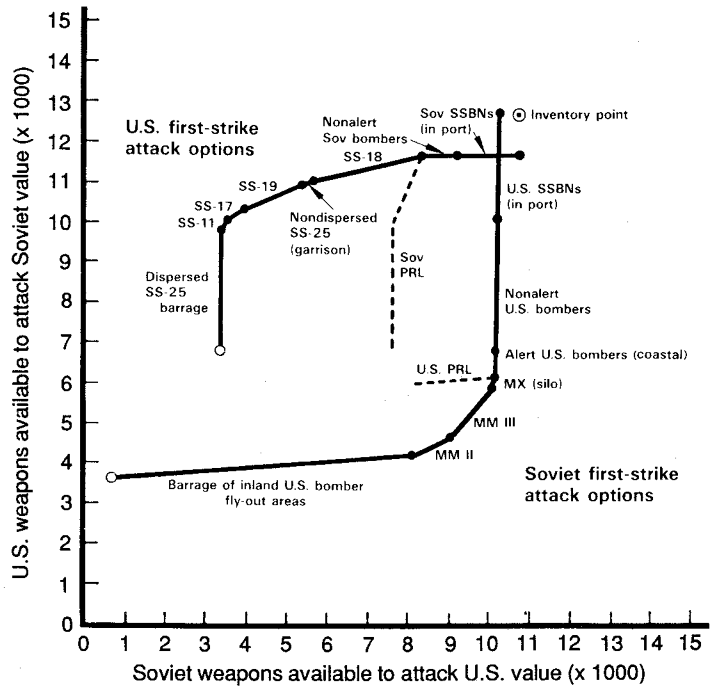


Fig. 15—U.S. and Soviet first-strike attack options for Case I forces operating in posture A (without cost lines)

The inventory point represents the total strategic nuclear arsenal of each side (roughly 12,600 U.S. weapons and 10,600 Soviet weapons).[3] We assume that in striking first, each side can muster up to 90 percent of its SSBN and bomber forces in addition to its entire ICBM force. The remaining 10 percent of these forces—consisting of submarines in overhaul and bombers under repair—do not play a role in the nuclear exchange. Thus, the Soviet draw-down curve begins to the left of the inventory point and the U.S. curve begins below.

If the Soviets, in striking first, used all available forces (comprising some 10,100 weapons) for denying U.S. war aims, they would operate at the top of the Soviet draw-down curve. While causing maximum damage to the United States, the Soviets would have done nothing to limit damage to themselves—12,600 U.S. weapons would be available (eventually) for U.S. retaliation against Soviet value.

[3]See Tables A.1 and A.2, App. A, for current U.S. and Soviet force structures.

In an attempt to limit damage to the Soviet Union, the Soviet leader would attack U.S. strategic forces, represented by movement down the Soviet draw-down curve. The Soviet curve in Fig. 15 is arranged according to efficiency of attack (not according to a time scale); thus, attacking U.S. SSBNs in port is most efficient, then bombers on the ground (nonalert bombers at all bases and alert bombers at some coastal bases), MX missiles in silos, and finally, Minuteman missiles in silos. Barrage of fly-out areas of alert U.S. bombers at inland bases yields few U.S. weapons destroyed for many Soviet weapons expended, and thus represents an inefficient Soviet counterforce attack option. The U.S. draw-down curve is arranged in similar fashion.

The lines of PRL emanating from the U.S. and Soviet draw-down curves represent the initiator's perception that the side incurring a first strike would launch most of its ICBMs out from under the attack. Therefore, the first-striker expends weapons against largely empty silos and destroys few of the adversary's ICBMs.[4]

The leader who initiates has the option of stopping the counterforce portion of his nation's attack anywhere along its draw-down curve. In the counterforce attack, the initiating side attempts to draw down opposing forces and reduce its cost to the extent possible, i.e., drive its first-strike attack curve as far as possible toward its own axis without wasting weapons on inefficient counterforce attacks—weapons that could be used more effectively in countervalue attacks.

Thus, one may assume that the Soviet leader probably would not barrage U.S. bomber fly-out areas in a first strike. Likewise, the U.S. leader probably would terminate his country's counterforce attack after striking Soviet SSBNs in port, nonalert bombers, and fixed ICBMs. Barraging SS-25 deployment areas, at least for the location uncertainty we have postulated, would be a highly inefficient attack option.

The superimposition of these draw-down curves atop the lines of constant cost in the weapons domain, as shown in Fig. 16, supports our assumptions. Using our formula of minimizing one's own cost and, within that constraint, maximizing that of the adversary, a Soviet counterforce attack would end following strikes against U.S. SSBNs in port, nonalert bombers, alert bombers at coastal bases, and silo-based ICBMs. A U.S. first strike would involve attacks against Soviet SSBNs in port, nonalert bombers, silo-based ICBMs, and mobile ICBMs in garrison.

The Soviets, therefore, probably would not barrage U.S. bomber fly-out areas. Indeed, a barrage would increase the Soviets' own cost.

[4]We assume that PRL would enable either side to launch 90 percent of its ICBMs out from under the attack.

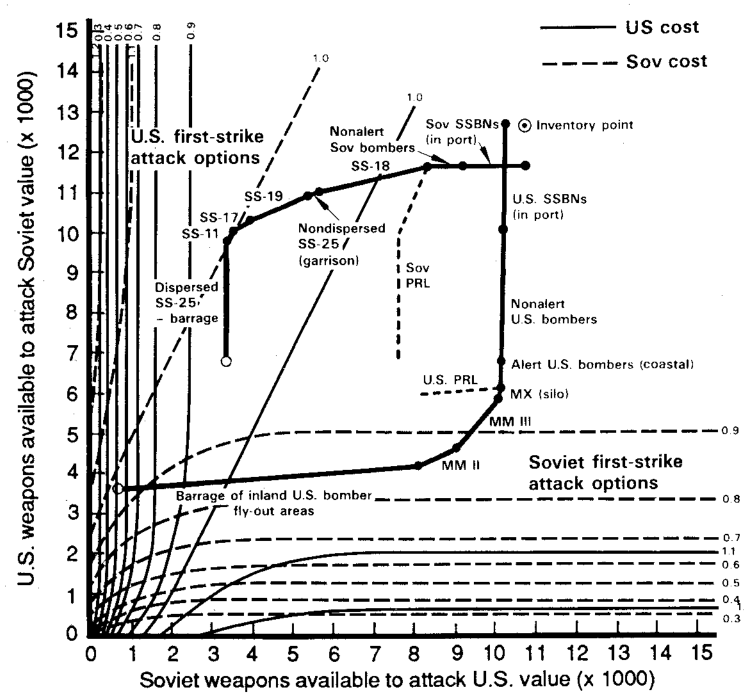


Fig. 16—U.S. and Soviet first-strike attack options for Case I forces operating in posture A

Without further limiting damage to themselves, barraging fly-out areas would detract from the secondary Soviet objective of denying U.S. war aims by inflicting damage on U.S. value.

Using this methodology, we can gauge the significance to each leader of the weapons remaining by reading the costs to each of striking first or incurring a first strike. The United States would incur a cost of about 0.94 if it struck first and 1.07 if it waited and suffered a Soviet first strike. Initiating would cost the Soviets 0.86, and incurring a first strike would cost them 1.0. Multiplying the U.S. and Soviet ratios of costs gives a stability index of 0.76.

Because the costs of striking first and striking second do not vary greatly for either side, current forces, even under posture A, exhibit a relatively high level of first-strike stability. Each side maintains a buffer of nontargetable forces large enough to prevent the adversary from successfully limiting damage through a counterforce attack.

If the initiator believed that the opponent's forces would operate at a higher state of generation (Fig. 13, above) and/or would conduct PRL (Fig. 14, above) by the time the counterforce attack arrived, these buffers of nontargetable forces would increase and the stability index would rise. It could not rise significantly, as each side already would have a robust retaliatory capability under posture A and without PRL.

According to Fig. 16, if either side believed that the other would launch its ICBMs under attack, the initiating side would not cross even one of its own cost lines—even if the retaliator's forces operated at posture A. In this instance, the U.S. costs associated with going first and second would be about 1.0 and 1.04, respectively, and the Soviet costs would be 0.92 and 0.97. The index of first-strike stability would be 0.91, as shown in Fig. 14, above.

Obviously, we do not examine methodologically the many drawbacks associated with PRL. One such drawback might be a greater likelihood of accidental launch by one side in a crisis if that side actually sustained a doctrine of launch under attack. However, one should not ignore the utility of creating the uncertainty as to whether the retaliator would conduct PRL if attacked; i.e., neither side could be confident that the retaliator would not launch his missiles on tactical warning. This uncertainty would dampen each leader's expectation that he could reduce the cost to his country by striking first.

Cases II, III, and IV: Postulated U.S. and Soviet Strategic Offensive Forces

We now examine draw-down curves generated by force postures of greatly expanded and modernized U.S. and Soviet force structures. Figure 17 depicts Case II, with forces operating at posture A. In this case, the inventory point is located far in the northeast corner of the weapons domain.[5]

This quadrant of the domain contains only slight gradients between the lines of constant cost. Though the attacker destroys many of the opponent's weapons as a result of a massive counterforce first strike, the initiator accomplishes little in terms of reducing the cost to himself, i.e., there is little movement over the cost lines. Thus, despite a substantial increase in counterforce capabilities, the larger arsenal on each side, while maintaining the same generation rates (indicating greater number of SSBNs at sea, bombers on strip alert in the U.S. case, and dispersed mobile missiles), precludes any significant capability to limit damage by striking first. The erosion in first-strike stability is marginal relative to current conditions.

[5]See Tables A.3 and A.4.

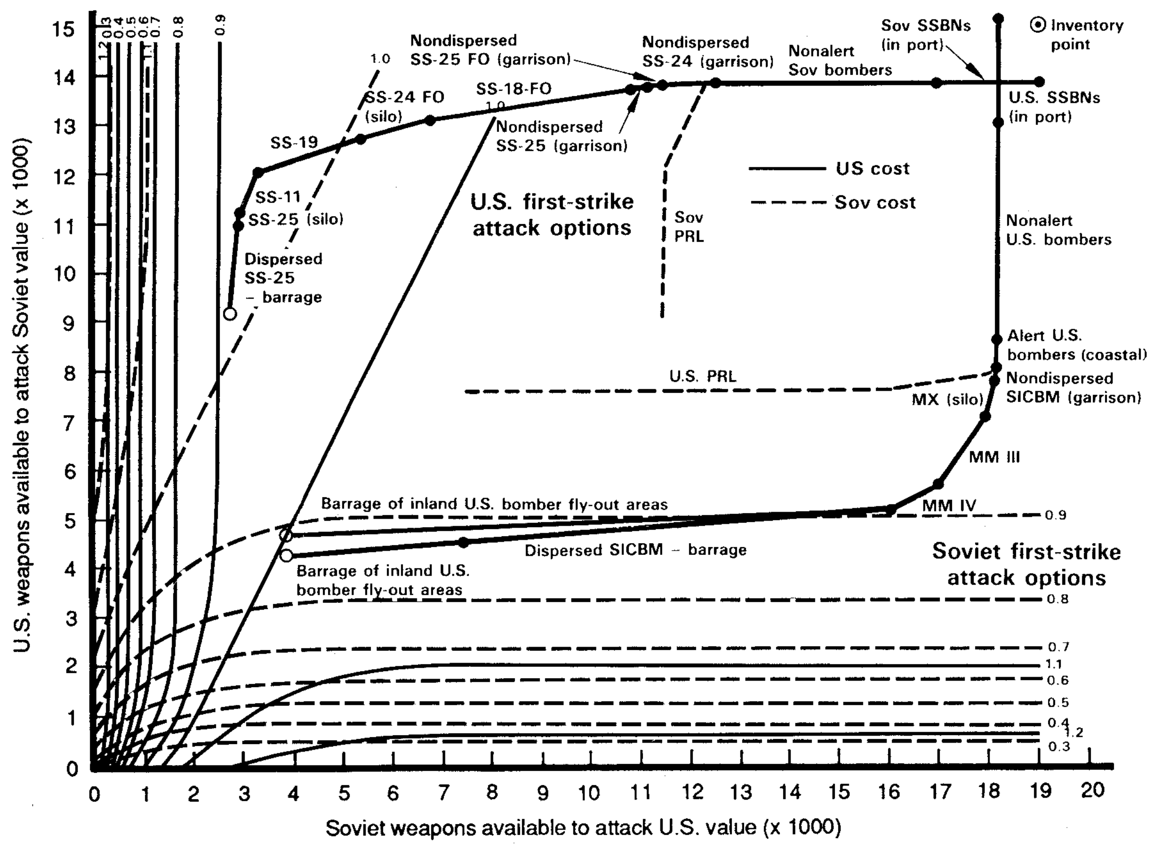


Fig. 17—U.S. and Soviet first-strike attack options
for Case II forces operating in posture A

Figure 18 portrays Case III—postulated U.S. and Soviet forces that have been modernized within the constraints of a START agreement. The major changes in force structure include the enhancement of U.S. hard-target kill (with the deployment of the D-5 SLBM) and the substantial expansion of the Soviet bomber force in proportion to RVs atop ICBMs and SLBMs. Furthermore, the United States has deployed 50 MX ICBMs in a rail-garrison mode.[6]

The new inventory point reflects smaller arsenals overall. Under posture A (the same posture as that assumed for current forces in Fig. 16, above), by definition, fewer weapons on each side are nontargetable. Each side begins to drive its draw-down curve into its own region of steep gradient in terms of declining cost. This phenomenon appears most clearly in the U.S. first-strike attack curve.

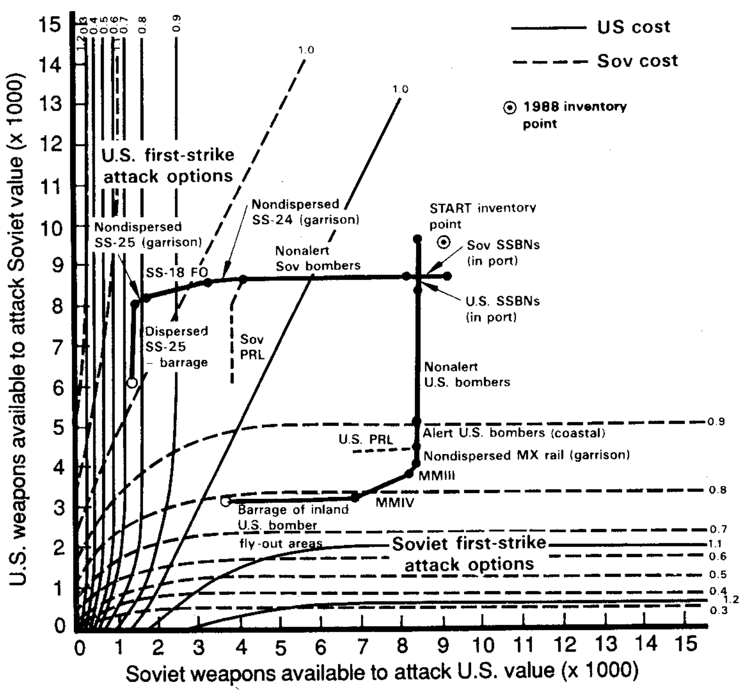


Fig. 18—U.S. and Soviet first-strike attack options for Case III forces operating in posture A

[6]See Tables A.5 and A.6.

Soviet bombers destroyed on the ground constitute a much greater portion of Soviet strategic offensive forces, illustrated by the long segment on the U.S. curve. Moreover, hard-target-kill-capable RVs on D-5 and MX missiles eliminate most of the silo-based Soviet ICBMs. The lower index results especially from the U.S. leader's belief that if he preempted, he could catch all Soviet bombers before they could escape and many fixed ICBMs in silos and mobile ICBMs in garrison.

The modernized, START-constrained forces in Case III, when operating at posture A, demonstrate a higher level of first-strike instability than current forces—an index of 0.52 compared with 0.76 (see Fig. 13, above). Stability could be enhanced if each leader perceived that the side under attack would conduct PRL. The stability index with PRL as shown in Fig. 14, above, would be 0.80, a level only slightly above the relative stability exhibited by current forces operating at posture A without PRL.

The introduction of accurate, effective RVs on the MX ICBM and D-5 SLBM provides the United States the capability to successfully attack the Soviet ICBM force. Although some would attribute the lower index to the U.S. deployment of the MX and D-5, a more strategic assessment would ascribe it to the Soviet posture of forces. Neither side could prevent development of increasingly accurate RVs. Rather, each must undertake unilateral measures to make its own forces unattackable, i.e., to make the the adversary's price to attack one's own strategic retaliatory forces so high that he would find such attacks counterproductive.

We now demonstrate how generating the forces in Case III to less vulnerable postures raises the stability index substantially. Figure 19 shows the Case III force structure generated to posture B, in which the side incurring the first strike maintains a higher proportion of bombers on alert, SSBNs at sea, and mobile ICBMs dispersed out of garrison.

Putting greater numbers of SSBNs to sea and bombers on strip alert shortens the segments in each draw-down curve that portray attacks against those elements. Doing so also pulls the attack curves out of the regions in which steep gradients exist. The first-strike stability index for this posture and structure resembles that associated with current forces operating at posture A.

A comparison of Case III (posture A) in Fig. 18, above, and Case IV in Fig. 20 provides an example of the effect on first-strike stability of instituting certain ICBM modernization options facing the United States in 1989. The Soviet force structure and the U.S. SLBM and bomber force elements in Case IV remain the same as in Case III. In Case IV, the United States deploys 500 single-RV, road-mobile small

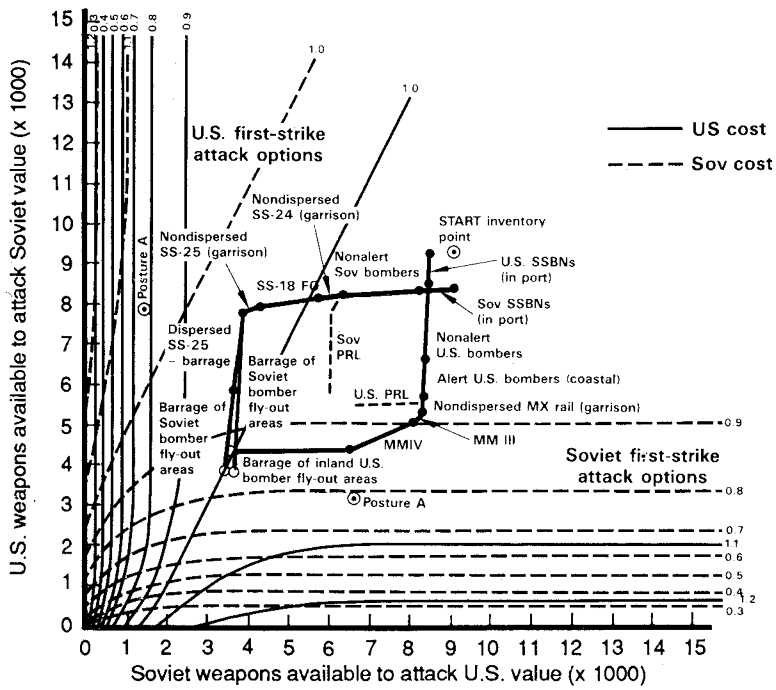


Fig. 19—U.S. and Soviet first-strike attack options for Case III forces operating in posture B

ICBMs rather than the 50 ten-RV, rail-garrison MX missiles as in Case III. In Case IV, the 50 MX ICBMs remain in 50 silos.[7]

A comparison of Fig. 20 with Fig. 18 demonstrates no change in the U.S. draw-down curve, as U.S. counterforce capabilities are equal in both cases. However, the Soviet attack curve is shorter in Fig. 20 (Case IV, posture A) than in Fig. 18 (Case III, posture A).

The change in the Soviet curve results from our assumptions as to how the rail-mobile MX and road-mobile SICBM operate. The survivability of the rail-garrison MX depends upon a decision by a designated authority to disperse them onto the U.S. rail system, i.e., the rail-garrison MX is not dispersed on a day-to-day basis.

Under posture A in Case III, we assume that the Soviet decision-maker believes that he could catch all rail-mobile MX ICBMs in a few

[7]See Tables A.5 and A.7 for U.S. force structures and Table A.6 for the Soviet force structure.

vulnerable garrisons in a first strike. Under the alternative force structure in Case IV, conversely, we postulate that 90 percent of the SICBM force is dispersed under day-to-day conditions; SICBM survivability does not depend upon a decision by a designated authority.

The Soviet decisionmaker faces the prospect of expending a prohibitive number of weapons to barrage the SICBM dispersal area in an inefficient attack. Given the virtually flat slope of the line in Fig. 20 depicting that attack, he probably would not opt to barrage the dispersal area. Thus, 450 more weapons—all SICBM RVs—would survive under posture A in Case IV than would survive under posture A in Case III.

The stability index changes little, however, as a consequence of altering the U.S. force structure; it increases under posture A from 0.52 to 0.55 as shown in Fig. 13, above. In Case III, the Soviet counterforce first strike would end in an area of relatively shallow gradient in terms of the U.S. and Soviet cost lines. Taking into account the

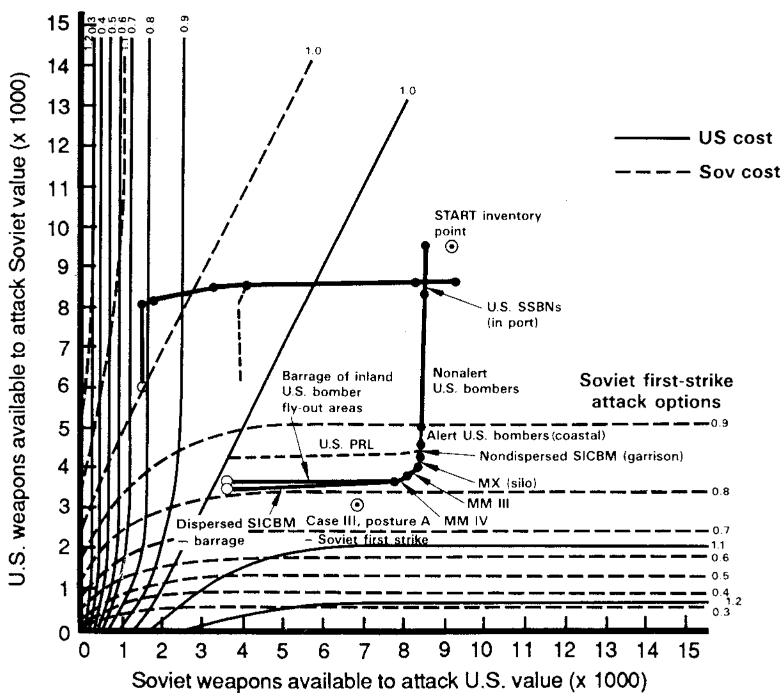


Fig. 20—U.S. and Soviet first-strike attack options for Case IV forces operating in posture A

nontargetable U.S. weapons at sea and on alert bombers at inland bases, 450 more U.S. retaliatory weapons would not cause much fluctuation in U.S. or Soviet cost.

If, however, the survivability of SSBNs came into serious doubt, the burden of maintaining a robust retaliatory capability would fall on the surviving ICBMs and bombers. In the event SSBNs became vulnerable, the significance of the 450 surviving SICBM RVs in denying a Soviet damage-limiting capability would greatly increase.

The differences in the stability index between Cases III and IV (without PRL) would diminish further if forces operated at posture B (when assumed dispersal rates of the rail-mobile MX and the road-mobile SICBM are 50 percent and 90 percent, respectively). They would disappear at posture C (when assumed dispersal rates of both approach 100 percent).

Interestingly, the "carry-hard" concept, under which U.S. ICBMs would be hidden among an array of shelters in a shell game, would yield approximately the same stability index as the SICBM if 500 RVs were combined with about 2000 shelters under a START regime. As with the SICBM, the survivability of "carry-hard" ICBMs would not depend upon a decision by a designated authority in a deep crisis. ICBMs in a "carry-hard" deployment would be nontargetable in the sense that the Soviets likely would not attack them because such an attack would not measurably reduce their cost.

In sum, one may use these cost lines in the weapons domain to evaluate alternative postures of forces. If one disagrees with our construct of cost in the damage domain or our assumptions as to the steepness of the damage curves, one may develop one's own. Nonetheless, armed with the approach to first-strike stability presented in this report, one can gauge the relative stability inherent in any number of alternative force structures and force postures within those structures.

MEASURES OF NATIONAL SECURITY OBJECTIVES

Strategic nuclear forces, as noted in Sec. I, should be evaluated on their ability to underwrite key national security objectives. Figure 21 and the text that follows list these objectives and the measures of each. Some of the objectives conflict; the text elaborates on some of the tensions.

- <u>First-strike stability</u>

  U.S. and USSR maximize $\left(\frac{C_1^{US}}{C_2^{US}}\right)\left(\frac{C_1^{Sov}}{C_2^{Sov}}\right)$

- <u>Central deterrence</u>

  U.S. maximizes C_1^{Sov}

  USSR maximizes C_1^{US}

- <u>Extended deterrence</u>

  U.S. minimizes C_1^{US}

- <u>If deterrence fails</u>

  — <u>Limit damage, deny enemy war aims</u>

  U.S. minimizes $C_{1,2}^{US}$

  USSR minimizes $C_{1,2}^{Sov}$

  — <u>Terminate conflict on acceptable terms</u>

  After U.S. counterforce attack, U.S. maximizes $\frac{W_1^{US}}{W_2^{Sov}}$

  After Soviet counterforce attack, USSR maximizes $\frac{W_1^{Sov}}{W_2^{US}}$

  (where W = weapons)

Fig. 21—Measures of national security objectives

Enhancing First-Strike Stability

The goal of enhancing first-strike stability is to reduce the pressures—stemming from the posture of strategic forces—facing either superpower leader in a deep crisis to launch a first nuclear strike against the homeland of the other. First-strike stability is robust when the cost to each side of striking first is not much less than its cost of waiting and potentially incurring a first strike.

First-strike stability differs from most other U.S. national security objectives in that the Soviet Union shares it. Both countries strive to ensure that the posture of forces does not act as a catalyst in a deep crisis to a strategic nuclear war, i.e., that the stability index is maximized. This confluence of goals does not occur with regard to any other national security objective except arms race stability. Survivable postures of U.S. and Soviet strategic forces—in the form of bombers on alert, SSBNs at sea, mobile ICBMs in dispersal areas, and the perception that ICBMs stand ready to be launched under attack—serve to raise the stability index.

Providing Central Deterrence

To provide central deterrence, the United States strives to achieve force postures that the Soviets perceive as maximizing their losses and minimizing their gains if they launch a first strike. The United States seeks to avoid any perception by the Soviets that aggression against U.S. territory could be considered successful by any conceivable measure of merit. The Soviet Union's cost of going first should greatly exceed its cost of avoiding a strategic nuclear war so that Soviet leaders are never tempted to strike first. Maintaining an effective U.S. retaliatory force that can hold at risk a wide array of Soviet assets enhances central deterrence; hence, this objective can be compatible with first-strike stability.

Strengthening Extended Deterrence

In strengthening extended deterrence, the United States strives to create a Soviet perception that it would be extremely dangerous for the USSR to provoke a crisis in the first place. To achieve this objective, the United States seeks to minimize its cost of going first. Holding at risk a large portion of Soviet strategic nuclear forces (and fostering the Soviet perception of a U.S. damage-limiting capability) represents a means to this end.

Deploying weapons effective for counterforce first strikes in vulnerable basing modes serves to enhance extended deterrence. With such deployments, the United States creates the perception that a major crisis might escalate to strategic nuclear war. Obviously, the objective of extended deterrence contradicts that of first-strike stability.

If Deterrence Fails

If deterrence fails, the United States must attempt to limit damage and deny Soviet war aims. Thus, the United States endeavors to minimize its cost of striking second, as well as its cost of striking first. Achieving the capability to limit damage in a first strike conflicts with first-strike stability. In contrast, maintaining the retaliatory capability to deny Soviet war aims by inflicting devastating damage on Soviet value complements first-strike stability as well as central deterrence.

The ratio of weapons remaining on each side after a purely counterforce first strike may be associated in some way with war termination on acceptable terms. Some analysts, for instance, contend that by achieving superiority in forces remaining after a counterforce first strike, the attacker has some advantage in intrawar negotiations.

According to this measure, in conducting a counterforce first strike the Soviet Union strives to maximize the ratio of Soviet weapons not used to U.S. weapons surviving, i.e., to destroy many U.S. weapons with as few Soviet weapons as possible. The United States, in striking first, endeavors to maximize the reverse of this ratio. Accordingly, each side benefits further by maintaining a large weapons inventory in peacetime.

To summarize, the national security objectives that strategic forces are designed to support exhibit important tensions that must be reconciled by decisionmakers as they plan strategic forces and formulate arms control proposals. The most troublesome tensions occur between first-strike stability, on one hand, and extended deterrence and limiting damage on the other. In light of these tensions, one might argue that an optimal amount of instability for the United States exists: enough to deter the Soviets from precipitating a crisis, but not enough to cause a crisis to spiral out of control should it occur. In this context, decisionmakers must define a metric of optimality in determining the trade-offs between first-strike stability and extended deterrence.

IV. CONCLUDING REMARKS

The methodology presented above demonstrates which structures and postures of U.S. and Soviet strategic nuclear forces enhance first-strike stability and which do not, and it relates cause and effect. The application of the methodology by policymakers would promote perceptive and disciplined discourse on matters surrounding first-strike stability in several ways.

First, the definition and quantification of cost, as required by the use of the damage domain, would force policymakers to analyze logically their views as to each side's operational objectives. Moreover, use of the methodology would compel policymakers to think about the vulnerability or invulnerability of the entire force structure rather than of one element of it. Though U.S. land-based ballistic missiles are vulnerable to Soviet attack, for example, draw-down curves demonstrate clearly that many U.S. retaliatory weapons would remain after a Soviet first strike.

The methodology also would foster an understanding of what information is relevant and available to decisionmakers in a crisis. It would compel the leader to focus on the pertinent questions of potential exchange outcomes and the difference between launching first and incurring a first strike, rather than on how well his forces could attack a target list generated by his own subordinates.

Finally, the methodology would provide planners a means for gauging how the implementation of particular modernization programs, with or without arms control constraints, would affect relative first-strike stability. Planners might also use it to assess the implications of breakout from an arms reduction treaty.

Our analysis of alternative postures of strategic forces based on this methodology suggests the following observations:

- **The index of first-strike stability applies equally to the United States and the Soviet Union.**

There is no U.S. index separate and distinct from the Soviet index. If the index suggests relative stability or instability, that stability or instability applies equally to each side. The United States and the Soviet Union share the national security objective of first-strike stability, and maintaining it requires joint efforts and cooperation.

Because first-strike stability is a shared goal, each country should spend its resources to make its own strategic nuclear forces

invulnerable, rather than to make the other's strategic nuclear forces vulnerable. Thus, U.S. policymakers should seek to increase the survivability of U.S. strategic forces, rather than to erode the survivability of Soviet strategic forces.

- **The attacker's perception of the posture of the adversary when the attacker's weapons arrive drives the requirement for each side to maintain a buffer of nontargetable forces on day-to-day alert.**

Whether or not an attack is a surprise is irrelevant. Decisionmakers might not react to impending signs of attack (so-called strategic warning) by generating forces. Thus, forces that depend on such reactions for survivability may not contribute to first-strike stability. An attacker might decide to preempt earlier rather than later to ensure that his weapons catch the adversary in a vulnerable posture. Accordingly, each side should sustain a day-to-day posture to guarantee that, although the side incurring the first strike may be surprised, it will also be prepared.

- **First-strike stability under current conditions is relatively robust.**

Even with peacetime states of generation and without PRL, the present cost to each side of striking first is not much less than its cost of incurring a first strike and retaliating. Each country could deny the other's war aims by inflicting devastating damage. Under current force postures, therefore, neither leader would be likely to perceive the other as pressured in a crisis by the posture of forces to strike first. First-strike stability is high today because neither the United States nor the Soviet Union has been able to achieve the capability to significantly limit damage by a first strike.

- **Increasing the size of U.S. and Soviet strategic nuclear inventories does not necessarily erode first-strike stability.**

A world with larger numbers of U.S. and Soviet strategic nuclear weapons would drive the inventory point in the weapons domain farther into the northeast corner of that domain. Even in the face of massive counterforce attacks, each side would maintain a sizable buffer of nontargetable forces. Of course, an unconstrained arms buildup might undercut such objectives as arms race stability and crisis stability.

- **Reducing the size of U.S. and Soviet strategic nuclear inventories does not necessarily enhance first-strike stability.**

Without proper attention to the posture of strategic nuclear forces, force reductions may erode first-strike stability. As inventories are reduced, forces usually would require higher day-to-day generation rates to maintain the current level of first-strike stability. Any agreement on reductions, therefore, should encourage unilateral survivability measures that render forces nontargetable or that at least significantly increase the price to attack these forces. An arms reduction regime must account as much for postures of forces as for numbers of weapons. The banning of mobile missiles, of course, would contravene the objective of enhancing first-strike stability.

We have demonstrated through our methodology how the superpowers can achieve an arms reduction treaty without jeopardizing first-strike stability. In so doing, we keep in mind that, aside from the issue of first-strike stability, both sides have persuasive reasons for concluding a treaty. For example, such a treaty could help to slow the arms race and to ease the confrontational relationship between the United States and the Soviet Union.

- **The United States and the Soviet Union should realize the importance of both leaders' ordering the generation of forces early in a crisis.**

The most favorable situation obtains when no generation from peacetime rates is required to sustain an adequate number of weapons in a survivable posture, thereby ensuring that neither side could limit damage significantly by striking first. If additional generation is required, however, such generation could be viewed in the context of first-strike stability rather than confrontation.

Many analysts perceive the act of generating forces in a crisis—placing more bombers on strip alert, putting more SSBNs to sea—as a confrontational move designed to show resolve. This fuels the perception that such acts send the wrong signal.

We have demonstrated that such acts actually increase the first-strike stability inherent in the posture of forces. Accordingly, we urge, in the interest of improving crisis management and enhancing first-strike stability, that the United States and the Soviet Union both come to realize that generating forces in a crisis can be a confidence-building measure. Joint U.S.-Soviet risk reduction centers could be established to manage cooperative generation of forces; barring this, each side

might consider unilateral generation by either side as alleviating tensions rather than increasing them.

- **Our index of first-strike stability does not predict the likelihood of either leader's launching a first strike in a deepening crisis (crisis instability).**

We can make reasonably definitive statements concerning the relationship of the index to the probability of war only at the extreme values of the index. If the stability index approaches 1.0, neither leader is likely to initiate a preemptive strike in a crisis. Conversely, if the stability index approaches zero, a deepening crisis is more likely to result in a nuclear exchange inasmuch as one of the leaders could feel considerable pressure from the posture of forces to strike first, i.e., the posture of forces could become a catalyst to instability in a crisis.

- **The United States has key national security objectives other than first-strike stability, and the objective of first-strike stability conflicts with the objectives of limiting damage and extended deterrence.**

As we state frequently in this study, strategic nuclear forces should be evaluated on the basis of their ability to underwrite key national security objectives. The difficulty in balancing these complicated, and in some cases, opposing objectives poses a profound problem in planning strategic forces.

In particular, first-strike stability conflicts with extended deterrence and damage limitation. Enhancing first-strike stability relieves pressures on both superpower leaders—based on the posture of forces—to strike first in a deep crisis. Conversely, extended deterrence implies that the Soviets are deterred from taking actions severely detrimental to U.S. interests because they perceive a grave danger of unwanted and uncontrollable escalation in a crisis.

Thus, as they modernize U.S. forces and formulate arms control proposals, U.S. policymakers must consider the trade-offs between these competing national security objectives. At the same time, U.S.-Soviet cooperation to enhance first-strike stability with respect to strategic nuclear forces mandates the creation of a stable environment in central Europe. Efforts by NATO to provide a more robust forward defense through increased capabilities of conventional forces and/or conventional arms control represent preferred means of enhancing stability in this region.

Appendix A

U.S. AND SOVIET STRATEGIC OFFENSIVE FORCE STRUCTURES

The U.S. and Soviet strategic offensive force structures presented in Tables A.1 and A.2 are based on figures in *Modernizing U.S. Strategic Offensive Forces: Costs, Effects, and Alternatives*, Congressional Budget Office, November 1987; *Jane's All the World's Aircraft 1987–88*, Jane's Publishing Company Limited, London, 1987; and Edward L. Warner III and David A. Ochmanek, *Next Moves: An Arms Control Agenda for the 1990s*, Council on Foreign Relations, New York, 1988. The postulated U.S. and Soviet force structures in the remaining tables in this appendix (A.3–A.7) are used only for the purpose of illustration.

Table A.1

CASE I—CURRENT U.S. STRATEGIC OFFENSIVE FORCE STRUCTURE

Force	Number	Throwweight (1000 kg) per Missile	Throwweight (1000 kg) Total	RVs per Missile	Total RVs
ICBMs					
Minuteman II	450	0.7	315	1	450
Minuteman III	523	1.1	575	3	1569
MX/Peacekeeper (silo)	27	3.6	97	10	270
Total	1000		987		2289
SLBMs					
C-3	256	1.5	384	10	2560
C-4	384	1.4+	538	8	3072
Total	640		922		5632
Total ballistic missiles	1640		1909		7921
		Take-off Gross Weight (1000 kg) per Bomber	Take-off Gross Weight (1000 kg) Total	Weapons per Bomber	Total Weapons
Bombers					
B-52G/ALCM	98	222	21756	12	1176
B-52G	69	222	15318	8	552
B-52H/ALCM	49	222	10878	20	980
B-52H	49	222	10878	8	392
B-1B	100	217	21700	16	1600
Total	365		80530		4700
Total missiles and bombers	2005				12621

Table A.2

CASE I—CURRENT SOVIET STRATEGIC OFFENSIVE FORCE STRUCTURE

Force	Number	Throwweight (1000 kg) per Missile	Throwweight (1000 kg) Total	RVs per Missile	Total RVs
ICBMs					
SS-11	440	1.0	440	1	440
SS-13	60	0.6	36	1	60
SS-17	150	2.7	405	4	600
SS-18	308	7.6	2341	10+	3080
SS-19	360	3.4	1224	6	2160
SS-25 (road)	100	1.4	140	1	100
Total	1418		4586		6440
SLBMs					
SS-N-6	272	0.7	190	1	272
SS-N-8	292	0.8	234	1	292
SS-N-18	224	1.1	246	7	1568
SS-N-20	100	2.3	230	9	900
SS-N-23	64	2.3	147	4	256
Total	952		1047		3288
Total ballistic missiles	2370		5633		9728
		Take-off Gross Weight (1000 kg) per Bomber	Take-off Gross Weight (1000 kg) Total	Weapons per Bomber	Total Weapons
Bombers					
Bear H/ALCM	50	188	9400	8	400
Bear	100	188	18800	4	400
Bison	15	158	2370	4	60
Total	165		30570		860
Total missiles and bombers	2535				10588

NOTE: SS-25 is mounted on road-mobile launchers and, when dispersed, is deployed over an area of 20,000 square nautical miles.

Table A.3

CASE II—POSTULATED U.S. STRATEGIC OFFENSIVE FORCE STRUCTURE: NO TREATY CONSTRAINTS, MODERNIZED FORCES, AND IMPROVED BASING MODES

Force	Number	Throwweight (1000 kg) per Missile	Throwweight (1000 kg) Total	RVs per Missile	Total RVs
ICBMs					
Minuteman III	500	1.1	550	3	1500
Minuteman IV	450	1.1	495	1	450
MX/Peacekeeper (silo)	100	3.6	360	10	1000
SICBM/Midgetman (road)	500	0.6	300	1	500
Total	1550		1705		3450
SLBMs					
C-4	192	1.4+	269	8	1536
D-5	480	2.4	1152	8	3840
Total	672		1421		5376
Total ballistic missiles	2222		3126		8826
		Take-off Gross Weight (1000 kg) per Bomber	**Take-off Gross Weight (1000 kg) Total**	**Weapons per Bomber**	**Total Weapons**
Bombers					
B-52/ALCM	96	222	21312	20	1920
B-52	69	222	15318	8	552
B-1B/ALCM	100	217	21700	20	2000
ATB/B-2	132	?	?	12	1584
Total	397		?		6056
Total missiles and bombers	2619				14882

NOTES: SICBM is mounted on hardened, road-mobile launchers and, when dispersed, is deployed over an area of 10,000 square nautical miles. Data on B-2 advanced-technology bomber are notional.

Table A.4

CASE II—POSTULATED SOVIET STRATEGIC OFFENSIVE FORCE STRUCTURE: NO TREATY CONSTRAINTS, MODERNIZED FORCES, AND IMPROVED BASING MODES

Force	Number	Throwweight (1000 kg) per Missile	Throwweight (1000 kg) Total	RVs per Missile	Total RVs
ICBMs					
SS-11	400	1.0	400	1	400
SS-18FO	308	7.6	2341	14	4312
SS-19	360	3.4	1224	6	2160
SS-24FO (silo)	150	3.6	540	10	1500
SS-24 (rail)	150	3.6	540	10	1500
SS-25 (silo)	60	1.4	84	1	60
SS-25 (road)	275	1.4	385	1	275
SS-25FO (road)	100	1.4	140	3	300
Total	1803		5654		10507
SLBMs					
SS-N-6	64	0.7	45	1	64
SS-N-8	88	0.8	70	1	88
SS-N-18	128	1.1	141	7	896
SS-N-20	200	2.3	460	10	2000
SS-N-23	224	2.3	515	4	896
Total	704		1231		3944
Total ballistic missiles	2507		6885		14451

		Take-off Gross Weight (1000 kg) per Bomber	Take-off Gross Weight (1000 kg) Total	Weapons per Bomber	Total Weapons
Bombers					
Bear H/ALCM	100	188	18800	12	1200
Blackjack	200	188	37600	16	3200
Total	300		56400		4400
Total missiles and bombers	2807				18851

NOTES: When dispersed, rail-mobile SS-24 is assumed to be nontargetable. SS-25 is mounted on road-mobile launchers and, when dispersed, is deployed over an area of 20,000 square nautical miles. Data on Blackjack are notional.

Table A.5

CASE III—POSTULATED U.S. STRATEGIC OFFENSIVE FORCE STRUCTURE: START CONSTRAINTS, MODERNIZED FORCES, AND IMPROVED BASING MODES

Force	Number	Throwweight (1000 kg) per Missile	Throwweight (1000 kg) Total	RVs per Missile		Total RVs
ICBMs						
Minuteman III	85	1.1	93	3		255
Minuteman IV (silo)	681	1.1	749	1		681
MX/Peacekeeper (rail)	50	3.6	180	10		500
Total	816		1022			1436
SLBMs						
D-5	408	2.4	979	8		3264
Total ballistic missiles	1224		2001			4700
		Take-off Gross Weight (1000 kg) per Bomber	Take-off Gross Weight (1000 kg) Total	Weapons per Bomber Counted/ Actual	Total Weapons Counted	Total Actual Weapons
Bombers						
B-52/ALCM	84	222	18648	12/12	1008	1008
B-52	60	222	13320	1/8	60	480
B-1B	100	217	21700	1/16	100	1600
ATB/B-2	132	?	?	1/12	132	1584
Total	376		?		1300	4672
Total missiles and bombers	1600				6000	9372

NOTES: When dispersed, rail-mobile MX is assumed to be nontargetable. Data on B-2 are notional.

Table A.6

CASE III—POSTULATED SOVIET STRATEGIC OFFENSIVE FORCE STRUCTURE: START CONSTRAINTS, MODERNIZED FORCES, AND IMPROVED BASING MODES

Force	Number	Throwweight (1000 kg) per Missile	Throwweight (1000 kg) Total	RVs per Missile		Total RVs
ICBMs						
SS-18FO	154	7.6	1170	10+		1540
SS-24 (rail)	112	3.6	403	10		1120
SS-25 (road)	344	1.4	482	1		344
Total	610		2055			3004
SLBMs						
SS-N-20	100	2.3	230	10		1000
SS-N-23	224	2.3	515	4		896
Total	324		745			1896
Total ballistic missiles	934		2800			4900
		Take-off Gross Weight (1000 kg) per Bomber	Take-off Gross Weight (1000 kg) Total	Weapons per Bomber Counted/ Actual	Total Weapons Counted	Total Actual Weapons
Bombers						
Bear H/ALCM	75	188	14100	12/12	900	900
Blackjack	200	?	?	1/16	200	3200
Total	275		?		1100	4100
Total missiles and bombers	1209				6000	9000

NOTES: When dispersed, rail-mobile SS-24 is assumed to be nontargetable. SS-25 is mounted on road-mobile launchers and, when dispersed, is deployed over an area of 20,000 square nautical miles. Data on Blackjack are notional.

Table A.7

CASE IV—ALTERNATIVE POSTULATED U.S. AND SOVIET STRATEGIC OFFENSIVE FORCE STRUCTURES: START CONSTRAINTS, MODERNIZED FORCES, AND IMPROVED BASING MODES

Force	Number	Throwweight (1000 kg) per Missile	Throwweight (1000 kg) Total	RVs per Missile	Total RVs
U.S. Strategic Offensive Force Structure					
ICBMs					
Minuteman III	85	1.1	93	3	255
Minuteman IV	181	1.1	199	1	181
MX/Peacekeeper (silo)	50	3.6	180	10	500
SICBM/Midgetman (road)	500	0.6	300	1	500
Total	816		772		1436
SLBMs		(Same as Table A.5)			
Bombers		(Same as Table A.5)			
Soviet Strategic Offensive Force Structure					
ICBMs		(Same as Table A.6)			
SLBMs		(Same as Table A.6)			
Bombers		(Same as Table A.6)			

NOTE: SICBM is mounted on hardened, road-mobile launchers and, when dispersed, is deployed over an area of 10,000 square nautical miles.

Appendix B

U.S. AND SOVIET STRATEGIC OFFENSIVE FORCE POSTURES

Throughout this study, we refer to force structures and the force postures within those force structures. We can describe force *structure* as (1) the systems that constitute the force (missiles, bombers, submarines); (2) the capabilities of those systems (probability of kill, reliability, stealth, hardening against nuclear effects); and (3) the basing mode (mobility, redundancy).

Force *posture* usually signifies the proportion of forces that cannot be targetable. Bombers on strip alert (except U.S. bombers at coastal bases), SSBNs at sea, and mobile ICBMs when dispersed are considered nontargetable because of the location uncertainty that we have postulated for them. We assume that the attacker would consider attacks on these elements counterproductive.[1]

Although silo-based ICBMs maintain a constant alert rate that approaches 100 percent, they are, of course, targetable. They become nontargetable if the attacker believes that his opponent will launch them under attack (based on tactical warning).

In this study, each force structure contains three basic force postures: A, B, and C, with and without PRL. Table B.1 shows a breakdown of postures postulated for our force-exchange calculations.

---

[1]In our assumptions, Soviet SSBNs patrolling off the Pacific and Atlantic coasts are able to destroy U.S. bombers on strip alert at coastal bases. Limited warning time is assumed to prevent these bombers from escaping. However, these bombers are still included in the alert rates of the U.S. bomber force.

Table B.1

PERCENTAGE OF U.S. AND SOVIET FORCES CONSIDERED NONTARGETABLE FOR THREE POSTULATED OFFENSIVE FORCE POSTURES

	Posture A	Posture B	Posture C
United States			
SSBNs	60	75	90
Bombers	30	60	90
MX (rail)	0	50	100
SICBM (road)	90	90	100
Soviet Union			
SSBNs	30	60	90
Bombers	0	50	90
SS-24 (rail)	25	50	100
SS-25 (road)	25	50	100

Appendix C

ALTERNATIVE COST FUNCTIONS

This study uses cost as the measure of merit that drives the criterion of first-strike stability. The definition of cost is an integral part of our methodology. We find the damage domain to be a most effective aid for thinking about this term.

The portrayal of cost in the damage domain reflects our judgment of policy and attitudes. The costs assigned to the corners of that domain relate to each leader's objectives of limiting damage to his own value and inflicting damage on that of the adversary to deny the adversary his war aims. Deciding the shapes of the lines of constant cost and the gradients between them compels us to think about the various trade-offs throughout the domain in terms of damage to one's own country versus damage to the opponent.

Cost Function Used in Report

Once determined by judgment, the lines of constant cost can be represented mathematically. The function we chose to represent the relationship between cost (C) and the fraction of value damaged (D) is as follows:

$$C^{US} = [(D^{US})^x] + \lambda[1 - (D^{Sov})^y]$$

$$C^{Sov} = [(D^{Sov})^r] + \gamma[1 - (D^{US})^s] \ ,$$

where $x = y = r = s = 0.75$ and $\lambda = \gamma = 0.3$. This function is consistent with our treatment of cost, which is damage incurred (D) plus the damage not inflicted ($1-D$) discounted by some factor (λ or γ). We formulate this function simply to produce lines that reflect our judgment as to the trade-offs within the damage domain. If our judgment differed, we would generate an alternative function to reflect the change.

Cost Functions Reflecting Alternative Policies

We could, for example, assume that the Soviet leader is more warlike than his U.S. counterpart, i.e., that the Soviet leader sees greater utility than the U.S. leader in inflicting damage on the adversary. Figure C.1 portrays graphically our new assumption.

Maintaining the U.S. cost function presented above and depicted in Fig. 5 (above), we now assume that if the Soviets could devastate the United States but received no damage in return, the Soviet leader would define his cost as −0.3 (a benefit). We assign this Soviet cost to the lower right corner of the damage domain in Fig. C.1. The negative cost connotes the Soviets' willingness to suffer a certain amount of damage to their value in return for achieving some objective, e.g., the domination of Europe.

Next, we assume that the Soviet leader places greater emphasis than we postulated above on denying U.S. war aims. To reflect this assumption, the Soviet lines of constant cost in Fig. C.1 have a steeper slope than those depicted in Fig. 6, above. Thus, in this instance, the Soviet

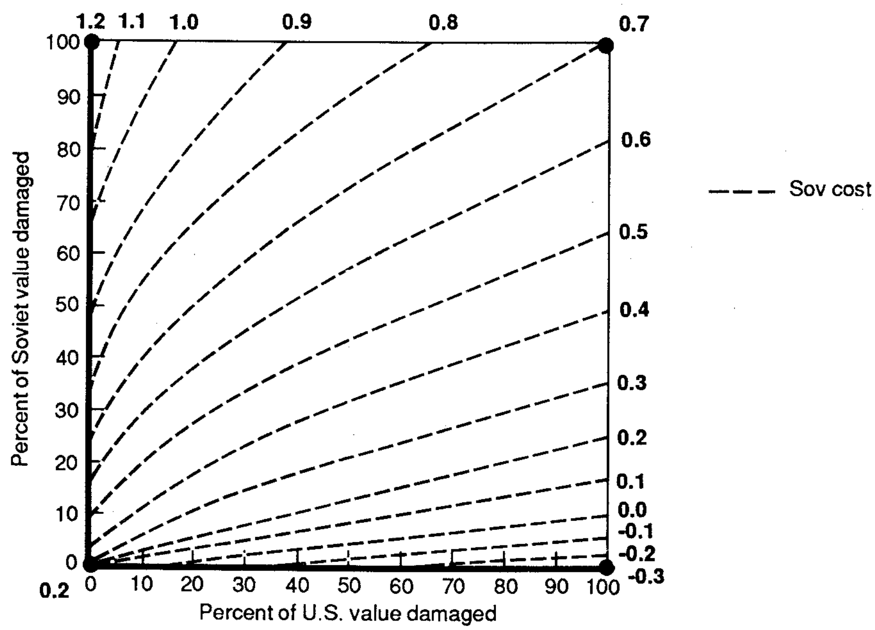


Fig. C.1—Soviet lines of constant cost in damage domain, given nuclear exchange and warlike Soviet leader

leader would be more inclined to accept damage to Soviet value in return for inflicting damage on U.S. value.

To express consistently the new lines of constant cost in the damage domain, we provide an alternative Soviet cost function by altering r, s, and γ in the equation

$$C^{Sov} = [(D^{Sov})^r] - \gamma[(D^{US})^s] + a \quad ,$$

where a is a constant. To create cost lines that reflect our new assumption, we set $r = s = 0.5$, $\gamma = 0.5$, and $a = 0.2$.

Figure C.2 is the result of transferring the lines of constant cost—along with the assumption of a more warlike Soviet leader—to the weapons domain. When we compare the Soviet cost lines in Fig. C.2 with those in Fig. 8 (above), the consequences of our new assumptions become readily apparent. (The U.S. lines of constant cost are those originally depicted in Fig. 8.)[1]

The Soviet lines of constant cost in Fig. C.2 have shifted up and to the left in comparison with the original Soviet cost lines in Fig. 8 (above). Based on these new cost lines, the costs to the Soviet Union of participating in a strategic nuclear exchange now are predictably lower for each force structure and posture than we had postulated originally.

The overall effect on the stability index of our new assumption of a more warlike Soviet leader is marginal. Using the notional force posture illustrated in Fig. 9(a), above, we note a small change in the stability index from 0.83 to 0.82. Likewise, the stability index of the force posture in Fig. 9(b) diminishes from 0.16 to 0.13. The range between the index of each posture expands from 0.67 to 0.69.

Figure C.3 demonstrates the cost lines in the damage domain if we assume that a leader views limiting damage to his country's value as his only objective in a strategic nuclear exchange. Since neither country would lower its cost by denying the other its war aims, U.S. lines of constant cost would be vertical and Soviet lines would be horizontal. We use this construct for illustrative purposes only.

The U.S. cost assigned to the point in the lower left corner is the same as the U.S. cost associated with the upper left corner (0 according to our original assignment). Further, the U.S. cost is 1.0 in the upper right corner of the domain. The cost is also 1.0 in the lower right

[1]To illustrate the effects of our new assumptions on the cost lines in the weapons domain, we transfer the lines through our original damage curves (Fig. 7, above). See App. D for a discussion of the consequences of altering these damage curves.

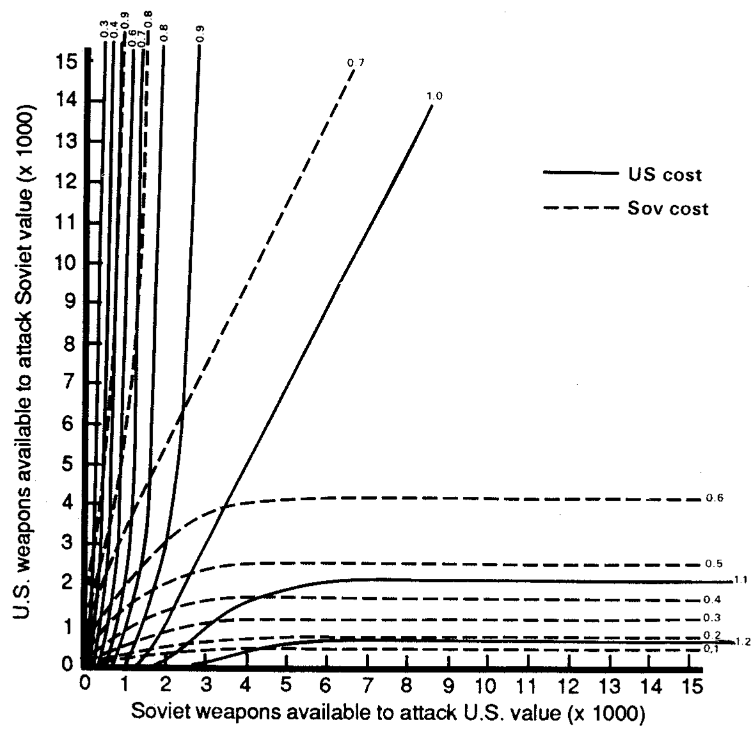


Fig. C.2—U.S. and Soviet lines of constant cost in weapons domain, given warlike Soviet leader

corner of the domain. For the United States, therefore, the cost function is

$$C^{US} = D^{US}$$

Each line of constant cost corresponds exactly with damage incurred.

Transferring the lines of constant cost from the damage domain, through the original damage curves, to the weapons domain results in Fig. C.4. Since the number of U.S. weapons available to attack Soviet value is not a factor in achieving U.S. objectives and thus does not affect U.S. cost, the U.S. cost lines are vertical.

The gradients between the lines of constant cost follow precisely the curvature of the damage curves (see Fig. 7, above). The U.S. cost lines are compressed near the ordinate because low numbers of Soviet

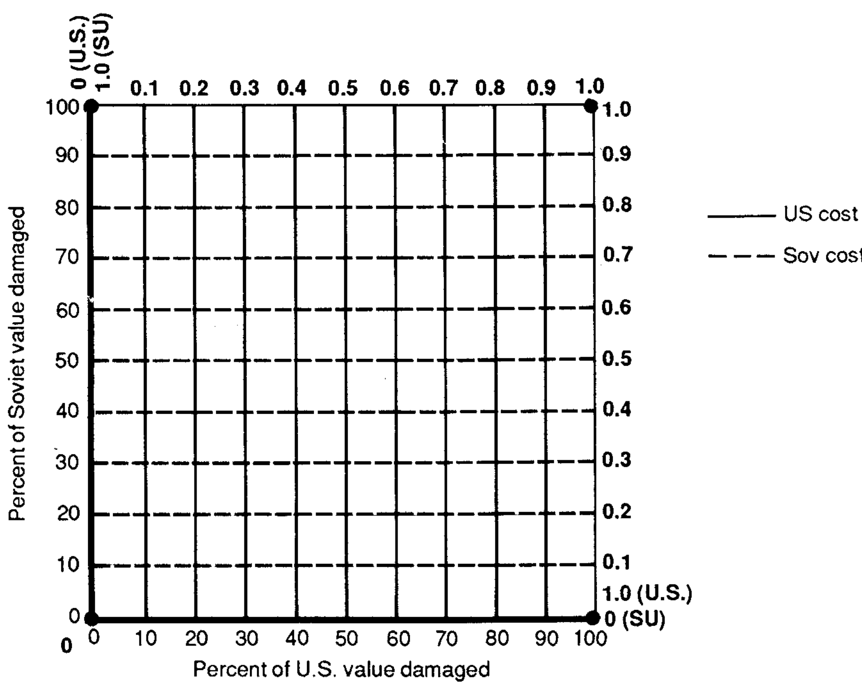


Fig. C.3—U.S. and Soviet lines of constant cost in damage domain, given nuclear exchange—damage limitation as only operational objective

weapons cause solutions along the steep part of the U.S. damage curve. The gradients progressively become less steep as the addition of Soviet weapons leads to solutions on the asymptotic portion of the U.S. damage curve. The U.S. 1.0 cost line lies far to the right of the U.S. 0.9 cost line.

The Soviet lines of constant cost are less compressed than their U.S. counterparts because the Soviet damage curve rises less steeply than the U.S. damage curve. The Soviet 1.0 cost line is located beyond the limits of Fig. C.4.

Using the draw-down curves from Fig. 9 (above), we find that the stability index changes only marginally if we use the cost lines in Fig. C.3, above, rather than the original cost lines in Fig. 8, above. A slight increase in the index from 0.83 to 0.86 occurs for the draw-down curves in Fig. 9(a). For the curves in Fig. 9(b), the index rises from 0.16 to 0.19.

Our methodology of first-strike stability, therefore, adapts to varying views of the costs to each side of a strategic nuclear exchange. The

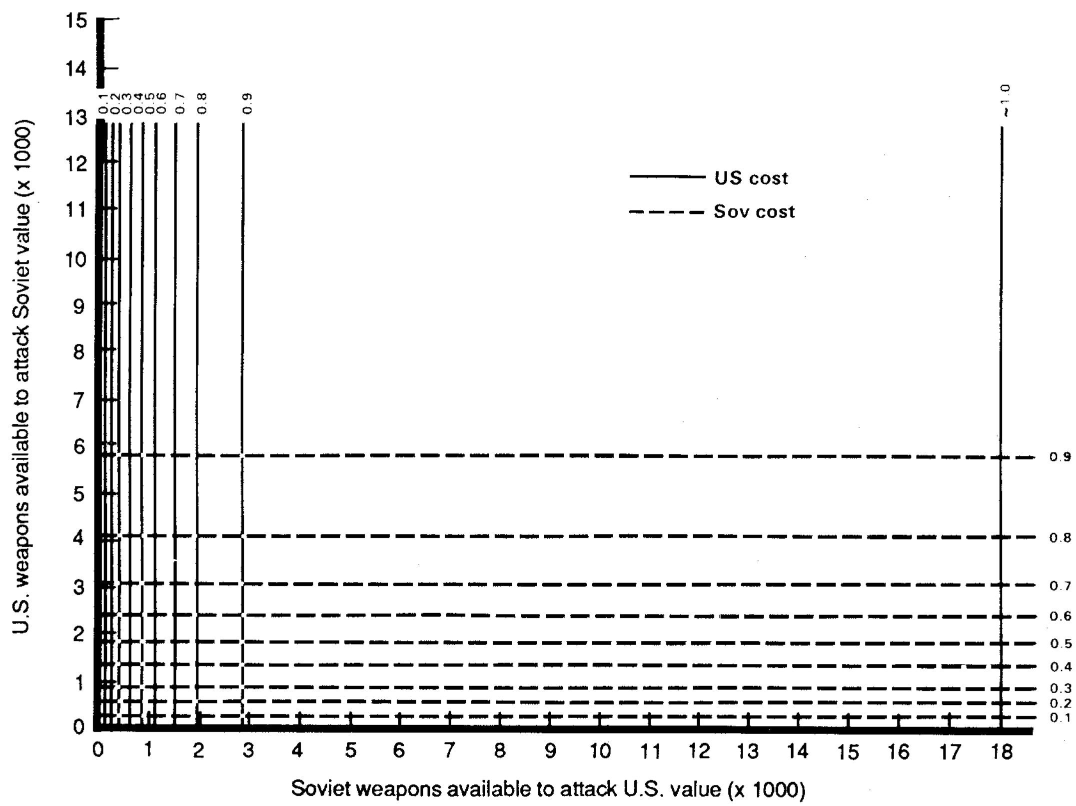


Fig. C.4—U.S. and Soviet lines of constant cost in weapons domain —damage limitation as only operational objective

damage domain expresses these views. Notably, however, the assumption of alternative policies only marginally alters the outcome in terms of the absolute values of the index and the range between the values. Most important, different assumptions as to the trade-offs between damage suffered and damage inflicted do not alter the relative stability of postures of forces.

Other Measures of Merit and Cost Functions

Other analysts have offered different measures of merit and cost functions for evaluating postures of strategic forces. We illustrate four of these measures and functions in the following four figures.

Some—for example, Major General I. Anureyev of the Soviet General Staff Academy—have portrayed ratios of weapons remaining after counterforce first strikes by either side.[2] We portray this measure graphically in Fig. C.5.

Using ratios of weapons—or any other surrogate of destructive capacity, such as throwweight or equivalent megatons—as the measure of merit, if accepted at face value, assumes that either side would perceive constant utility along a line of constant ratio of weapons remaining. According to this measure, a U.S. leader would see the same utility (or disutility) in outcomes that result in 10,000 weapons on each side as compared with 100 weapons on each side—as long as the ratio of residual weapons was always 1 to 1. We believe that such a construct of policy is incomplete and potentially misleading.

Another form of this measure of merit, depicted in Fig. C.6, is to plot ratios of damage in the damage domain as equal cost lines. While tracking damage might be better than tracking weapons, we assert that a country is not indifferent to whether 90 percent or 10 percent of each country is destroyed. Although the ratio is the same in both cases, the leader of each country surely would prefer the 10 percent outcome.

One could express cost in the damage domain as a constant difference, portrayed in Fig. C.7. In using constant difference, one must argue that a U.S. or Soviet leader would hold damage suffered by his own nation equal to damage suffered by the enemy. This produces U.S. and Soviet cost lines at a constant slope of 45 degrees. We would argue that the leader of each country would place far more emphasis on limiting damage than inflicting damage.

A more realistic definition of cost would be to discount damage to the opponent's value in relation to damage to one's own value. The

[2]See Peter Almquist and Stephen M. Meyer, *Insights from Mathematical Modeling in Soviet Mission Analysis (Part II),* Research Report No. 86-6, Center for International Studies, Massachusetts Institute of Technology, April 1985.

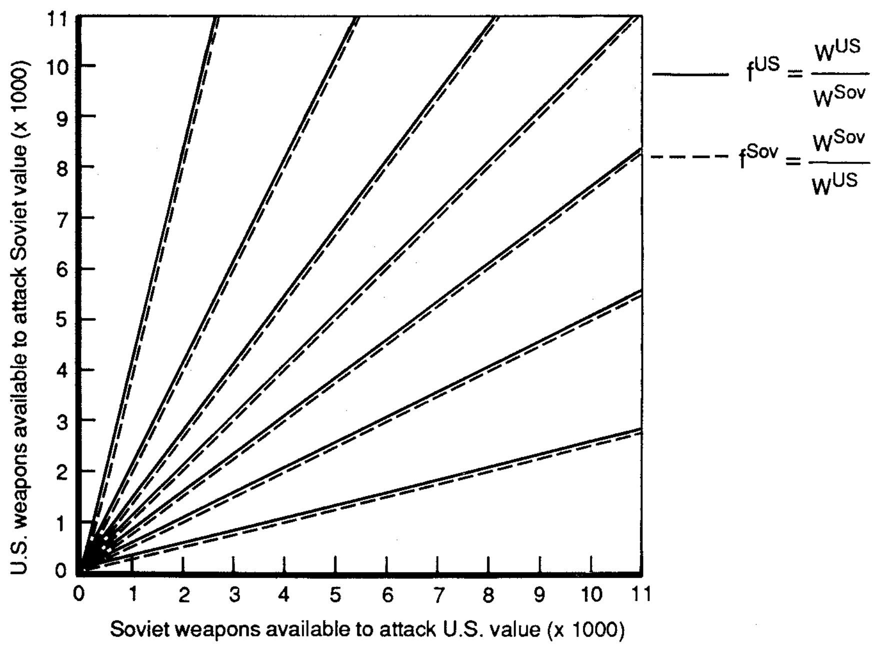


Fig. C.5—Function based on ratios in weapons domain

result, shown in Fig. C.8, consists of lines of constant cost with slopes equal to the discount; here, each leader values inflicting damage on the adversary's country 0.3 as much as limiting damage to his own (i.e., $\lambda = \gamma = 0.3$). Such curves are similar to those developed earlier in this study. The difference here is that the trade-off in damage to oneself versus damage to the opponent remains the same throughout the domain.

In summary, by drawing U.S. and Soviet lines of constant cost in the damage domain, we express cost in as forthright a manner as possible. One is not obliged to adopt our cost function. One is obliged, however, to present one's own view of the costs of a nuclear exchange in a way that encourages intellectual discourse and stands the test of common sense and informed judgment.

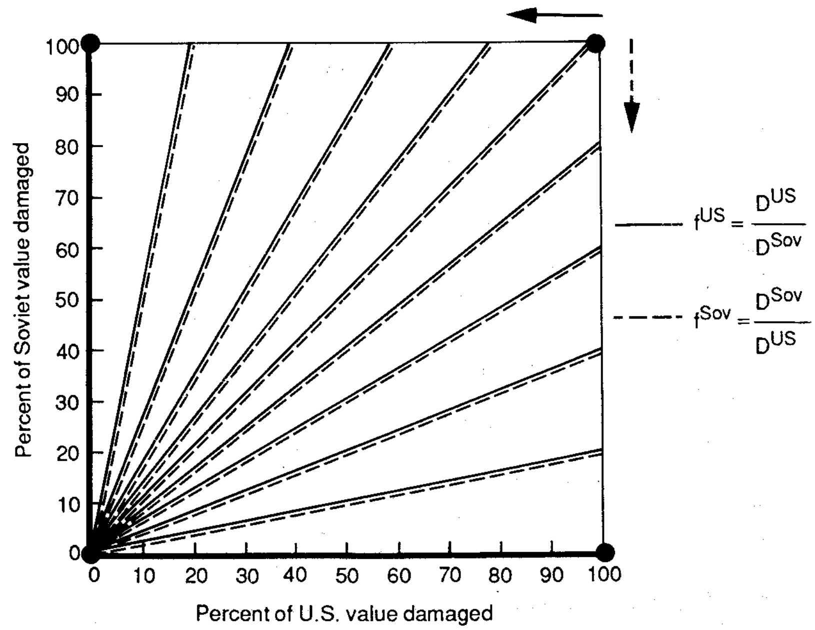


Fig. C.6—Function based on ratios in damage domain

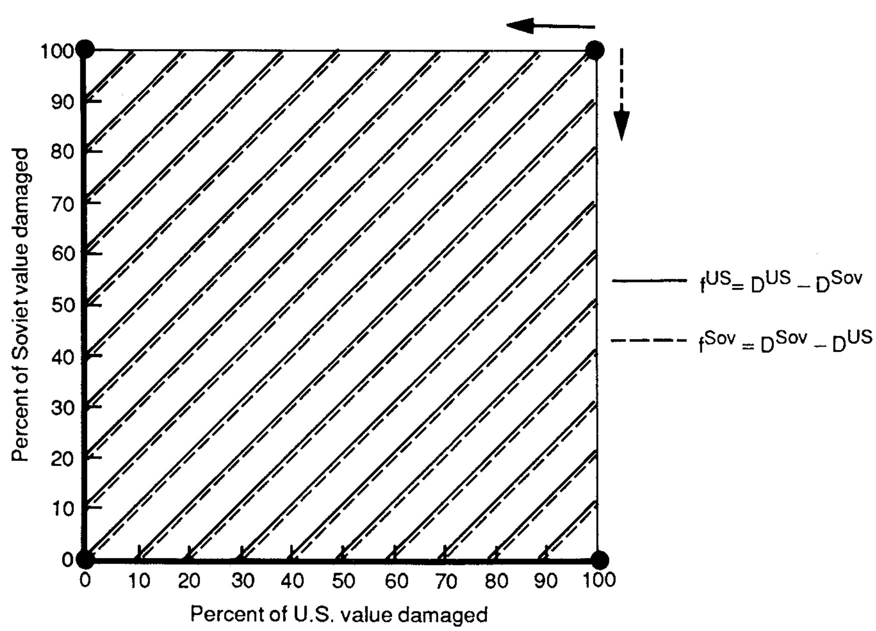


Fig. C.7—Function based on constant difference in damage domain

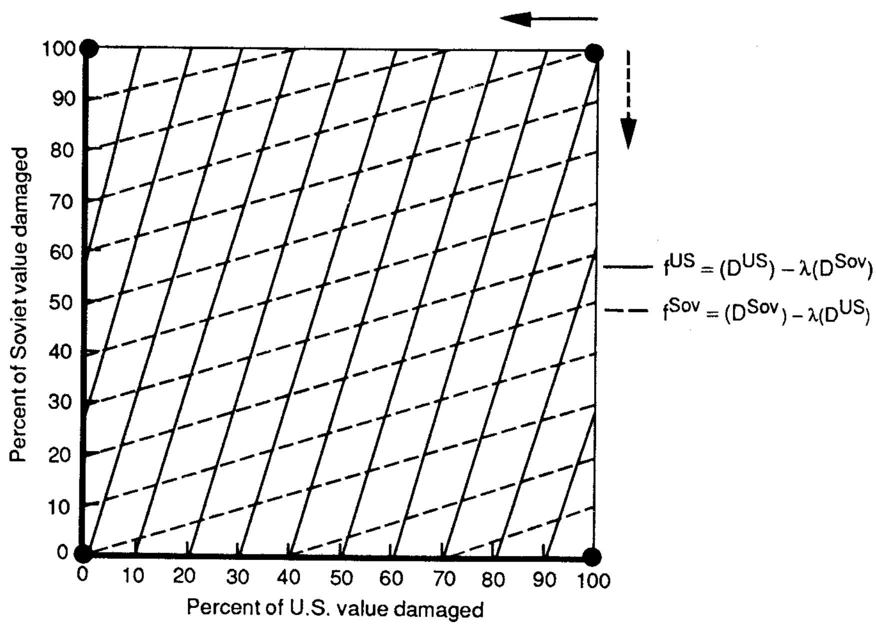


Fig. C.8—Function based on discounted difference in damage domain

Appendix D

ALTERNATIVE DAMAGE CURVES

This appendix describes the effects of altering the link between the damage domain and the weapons domain, i.e., the damage curves that express the relationship between potential damage to value and weapons available to attack that value. If analysts were to argue for changing the slopes of the damage curves we have presented, we believe that many would contend that the curves drawn in Fig. 7 (above) should rise more steeply. In Fig. D.1 we examine a new Soviet damage curve similar to the original U.S. damage curve in Fig. 7 and a new U.S. curve that rises more steeply than the original U.S. curve.

According to these new damage curves, 2000 U.S. weapons potentially could damage 80 percent of Soviet value. The same number of Soviet weapons potentially could destroy some 92 percent of U.S. value.

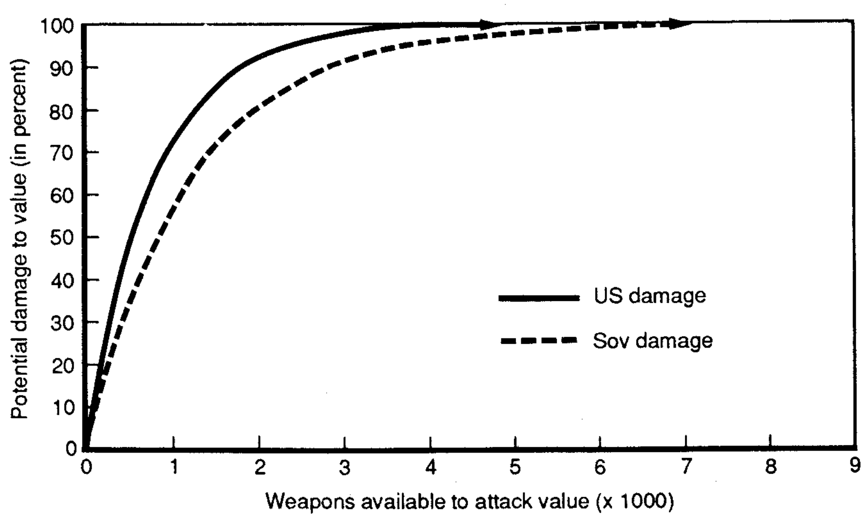


Fig. D.1—Alternative U.S. and Soviet damage curves

To portray the effect of damage curves that rise more rapidly, we establish U.S. and Soviet lines of constant cost in the weapons domain by transferring the original lines from the damage domain in Fig. 6 (above) through these new curves. Figure D.2 depicts the results.

The lines of constant cost in Fig. D.2 are more compressed along the axes than the original lines in Fig. 8 (above). Under the new assumptions, counterforce attacks by either side must be more effective than in the original case for each to enter its own areas of steep gradient, i.e., move from the asymptotic to the steep portion of the damage curve. Since the lines are so compressed, the new damage curves cause the stability index of each force posture to increase in comparison with the index that uses damage curves that rise less steeply.

The northeast quadrant of the domain represents a cost to the United States and the Soviet Union of 1.0; we portray this phenomenon by the widely dashed terminals of the 1.0 lines in Fig. D.2. Whereas some measurable gradient exists in this quadrant in Fig. 8, no gradient appears in Fig. D.2.

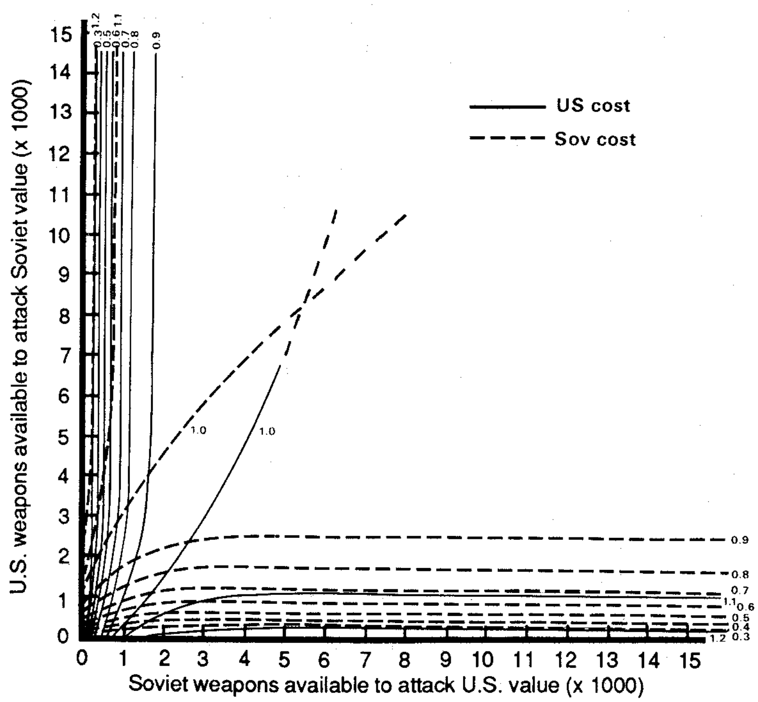


Fig. D.2—U.S. and Soviet lines of constant cost in weapons domain, given steeper damage curves

Superimposing the draw-down curves in Fig. 9(a), above, atop the cost lines in Fig. D.2 yields a stability index of 0.90, an increase over the original index of 0.83. The index for the posture in Fig. 9(b) exhibits a new index of 0.27; the original was of 0.16.

Causing the damage curves to rise more steeply increases the stability index for each posture. This change in the damage curves, moreover, reduces the range between the indexes for the two postures. However, the ranking of the postures in terms of first-strike stability remains the same.

RAND/R-3765-AF