



PROJECT MUSE®

The Nuclear Balance Is What States Make of It

David C. Logan

International Security, Volume 46, Number 4, Spring 2022, pp. 172-215 (Article)

Published by The MIT Press



➔ For additional information about this article

<https://muse.jhu.edu/article/855435>

The Nuclear Balance Is What States Make of It

David C. Logan

In the second winter of his presidency, Donald Trump again traded nuclear-tinged warnings with North Korean leader Kim Jong Un.¹ After Kim reminded the U.S. president of North Korea's burgeoning nuclear capabilities, Trump tweeted, "North Korean Leader Kim Jong Un just stated that the 'Nuclear Button is on his desk at all times.' Will someone from his depleted and food starved regime please inform him that I too have a Nuclear Button, but it is a much bigger & more powerful one than his, and my Button works!"² With this afternoon mis-sive, Trump staked out his position in a debate that has raged among academics and policymakers since the advent of the nuclear age.

Central to the study of interstate conflict are questions of whether nuclear superiority offers political or military benefits and, importantly, whether these benefits accrue past the acquisition of a secure second-strike capability. These questions incorporate foundational debates about deterrence theory, relative versus absolute gains, and the role of material capabilities.³ The nu-

David C. Logan is a Ph.D. candidate at Princeton University and a Stanton Nuclear Security Fellow at the Massachusetts Institute of Technology. Starting in summer 2022 he will be Assistant Professor in National Security Affairs at the U.S. Naval War College.

For helpful comments on earlier drafts, the author thanks Mark Bell, Leyatt Betre, Tom Christensen, Chris Chyba, Bryan Early, Aaron Friedberg, Frank Gavin, Mindy Haas, Robert Jervis, Vipin Narang, Pavel Podvig, Travis Sharp, Eric Snyder, Lauren Sukin, Sam Winter-Levy, participants in both the Princeton University Security Studies Graduate Research Seminar and the Massachusetts Institute of Technology Security Studies Working Group, and the anonymous reviewers. Research for this article was supported by the Center for International Security Studies at Princeton University. The opinions expressed in this article do not represent the official position of the U.S. Navy or any other organization with which the author is or has been affiliated. An online appendix is available at <https://doi.org/10.7910/DVN/E8RJDK>.

1. Peter Baker and Michael Tackett, "Trump Says His 'Nuclear Button' Is 'Much Bigger' than North Korea's," *New York Times*, January 2, 2018.

2. Donald Trump (@realDonaldTrump), "North Korean leader Kim Jong Un just stated that the 'nuclear button is on his desk at all times.' Will someone from his depleted and food starved regime please inform him that I have a Nuclear Button, but it's a much bigger and more powerful one than his, and my Button works!" *Twitter*, January 2, 2018, 4:49 p.m., <https://web.archive.org/web/20180103005039/https://twitter.com/realDonaldTrump/status/94835557022420992>.

3. Christopher H. Achen and Duncan Snidal, "Rational Deterrence Theory and Comparative Case Studies," *World Politics*, Vol. 41, No. 2 (January 1989), pp. 143–169, <https://doi.org/10.2307/2010405>; Richard Ned Lebow and Janice Gross Stein, "Rational Deterrence Theory: I Think, Therefore I Deter," *World Politics*, Vol. 41, No. 2 (January 1989), pp. 208–224, <https://doi.org/10.2307/2010408>; and Robert Powell, "Absolute and Relative Gains in International Relations Theory,"

clear superiority issue has also had important policy implications, as U.S. policymakers have debated the bomb's political and military utility. For decades, much of the theoretical and qualitative scholarship has argued that securing a second-strike capability is more important than achieving nuclear parity or superiority,⁴ though some scholars have suggested that superiority might have influenced the outcomes of some historical cases.⁵

Scholars have recently begun analyzing whether nuclear superiority influences interstate politics by applying statistical methods to data on nuclear crises, interstate disputes, and compellent threats. This new body of work is small, and its findings are mixed, though it has advanced the field by theorizing mechanisms through which nuclear superiority can confer advantages and remedying some of its predecessors' methodological weaknesses by, for instance, better matching cases with theory.⁶ Despite these contributions, however, this literature's empirical investigations have at least two crucial weaknesses.

First, this statistical work relies on overly simplistic measures, which sometimes fail to capture either the actual nuclear balance (assuming such an objective balance even exists and is measurable) or, more importantly, decision-makers' perceptions of it. Analysts of nuclear weapons have long sought methods for assessing, measuring, and comparing nuclear forces. Throughout the Cold War, the U.S. government and independent analysts dedicated significant resources to measuring the strategic nuclear balance be-

American Political Science Review, Vol. 85, No. 4 (December 1991), pp. 1303–1320, <https://doi.org/10.2307/1963947>.

4. Robert Jervis, "Why Nuclear Superiority Doesn't Matter," *Political Science Quarterly*, Vol. 94, No. 4 (Winter 1979/80), pp. 617–633, <https://doi.org/10.2307/2149629>; Robert Jervis, *The Illogic of American Nuclear Strategy* (Ithaca, N.Y.: Cornell University Press, 1984); and Kenneth N. Waltz, "Nuclear Myths and Political Realities," *American Political Science Review*, Vol. 84, No. 3 (September 1990), p. 738, <https://doi.org/10.2307/1962764>.

5. For example, Glenn Herald Snyder and Paul Diesing, *Conflict among Nations: Bargaining, Decision Making, and System Structure in International Crises* (Princeton, N.J.: Princeton University Press, 1978), pp. 459–462; and Amardeep Athwal, "The United States and the Sino-Soviet Split: The Key Role of Nuclear Superiority," *Journal of Slavic Military Studies*, Vol. 17, No. 2 (2004), pp. 271–297, <https://doi.org/10.1080/13518040490450547>.

6. Matthew Kroenig, "Nuclear Superiority and the Balance of Resolve: Explaining Nuclear Crisis Outcomes," *International Organization*, Vol. 67, No. 1 (January 2013), pp. 141–171, <https://doi.org/10.1017/S0020818312000367>; Matthew Kroenig, *The Logic of American Nuclear Strategy: Why Strategic Superiority Matters* (New York: Oxford University Press, 2018); Todd S. Sechser and Matthew Fuhrmann, "Crisis Bargaining and Nuclear Blackmail," *International Organization*, Vol. 67, No. 1 (January 2013), pp. 173–195, <https://doi.org/10.1017/S0020818312000392>; Todd S. Sechser and Matthew Fuhrmann, *Nuclear Weapons and Coercive Diplomacy* (Cambridge: Cambridge University Press, 2017); and Matthew Kroenig, Miriam A. Krieger, and Hans Noel, "Dare to Fail: Nuclear Superiority, Threat Initiation, and Compellent Success," unpublished manuscript, Georgetown University, 2015.

tween the United States and the Soviet Union. They developed complex metrics such as static quantitative measures (e.g., counts of warheads and delivery vehicles), composite measures (e.g., counter-military potential and hard-target kill capability), and dynamic nuclear exchange models, as well as qualitative measures assessing the accuracy, throw-weight, and survivability of delivery systems.⁷ Yet recent quantitative scholarship investigating whether nuclear superiority affects interstate conflict and crisis dynamics relies on simple counts of the total warheads controlled by a state in a given year. In doing so, these works elide the myriad other factors that affect a state's nuclear capability.

Second, and relatedly, this work largely assumes that states share the same (accurate) information about the relative nuclear balance. The statistical work relies on nuclear capability measures built from independent organizations' data, published many years after the interstate interactions that form research observations. In doing so, scholars discount the importance of how the nuclear balance is perceived at the time of the interaction. Further, scholars sometimes analyze these data using symmetric state dyads, implicitly assuming that leaders share the same perceptions of the nuclear balance. Yet while states may possess accurate information about their own nuclear arsenals—though even this may not always be true for some recent entrants to the nuclear club—they may often have imperfect information about others' nuclear arsenals, whether due to normal intelligence failures or counterintelligence efforts.⁸

Here, I argue that these two assumptions of warhead prominence and complete information often do not hold. First, policymakers across different states and periods have used different metrics to assess the nuclear balance—metrics that recent scholarship does not capture. Using an original dataset on state nuclear capabilities covering nearly sixty years, I demonstrate that the results of empirical tests of nuclear superiority depend strongly on how the concept is operationalized. Second, by comparing data from declassified U.S. intelligence estimates of Soviet nuclear deployments to their actual numbers, I show that U.S. perceptions of the nuclear balance were often substantially incorrect. In short, scholars have not adequately coupled measures either with theories of nuclear superiority or with states' actual experiences in assessing the nuclear balance. Beyond investigating the impact of the nuclear balance, these

7. Sydell P. Gold, "Report of the Chairman: What Role Can Analysis Play in Illuminating Arms Control Objectives and Options?" in Rudolf Avenhaus, Reiner K. Huber, and John D. Kettelle, eds., *Modelling and Analysis in Arms Control* (New York: Springer, 1986), pp. 125–142; and Claire M. Levy, *Soviet Strategic Nuclear Measures of Effectiveness* (Santa Monica, Calif.: RAND, 1992).

8. Francis J. Gavin, "What We Talk about When We Talk about Nuclear Weapons: A Review Essay," *H-Diplo/ISSE*, Forum No. 2, June 15, 2014, pp. 11–36.

findings also have implications for other purportedly “objective” measures of state capabilities, such as the widely used Composite Index of National Capability scores.

I then propose and use a new measure of the perceived balance, states’ real-time estimates of relative damage in a nuclear exchange. Building on recent scholarship, original archival work, and declassified interviews with Soviet leaders, I show that official nuclear exchange fatality estimates were, at best, weakly correlated with dyadic crisis outcomes.⁹ Moreover, these fatality assessments varied considerably and were tenuously connected to the nuclear balance’s technical-military features. Ultimately, the findings reinforce a key theoretical tenet of the nuclear revolution: Once states have achieved a secure second-strike capability, nuclear superiority does not confer additional political benefits.

More significantly, the evidence presented here highlights a key but underappreciated feature of the relative nuclear balance: it depends on not only technical-military factors but also state perceptions and beliefs. In a nuclear exchange, factors such as megatonnage and throw-weight as well as the destruction that they wreak are tragically real. But in domains short of nuclear exchanges, such as the brinkmanship contests that have come to characterize much of the nuclear age, state beliefs are paramount.¹⁰ In short, the nuclear balance is often what states make of it.

I begin with an overview of the major theoretical and empirical works on nuclear superiority. Next, I examine the most recent quantitative work and highlight the literature’s two key assumptions: (1) warhead counts are an appropriate measure of the strategic nuclear balance, and (2) states have complete information about the balance. I demonstrate that these assumptions often fail to hold, and I replicate two prominent recent works to show that results are sensitive to how the balance is operationalized. I then demonstrate that perceptions of anticipated damage in a nuclear exchange are an improved alternative measure of the nuclear balance. Drawing on archival and interview

9. For recent scholarship, see Caroline Reilly Milne, “Hope Springs Eternal: Perceptions of Mutual Vulnerability between Nuclear Rivals,” Ph.D. dissertation, Princeton University, 2017, <http://arks.princeton.edu/ark:/88435/dsp01cf95jf13p>.

10. Jacques E. C. Hymans, *The Psychology of Nuclear Proliferation: Identity, Emotions, and Foreign Policy* (Cambridge: Cambridge University Press, 2006); Rachel Elizabeth Whitlark, “Nuclear Beliefs: A Leader-Focused Theory of Counter-Proliferation,” *Security Studies*, Vol. 26, No. 4 (2017), pp. 545–574, <https://doi.org/10.1080/09636412.2017.1331628>; Elizabeth N. Saunders, “The Domestic Politics of Nuclear Choices—A Review Essay,” *International Security*, Vol. 44, No. 2 (Fall 2019), pp. 146–184, https://doi.org/10.1162/isec_a_00361; and Rupal N. Mehta, *Delaying Doomsday: The Politics of Nuclear Reversal* (New York: Oxford University Press, 2020).

data, I conclude that within the U.S.-Soviet Cold War dyad, the nuclear balance and its meaning were often not viewed as objective, coherent facts. Finally, I argue that even individuals with access to similar information about the balance disagree about its meaning and how to respond to it. Although state military intelligence agencies may produce estimates that inform individuals' beliefs, people can—and often do—respond differently to the balance. The nuclear balance is, therefore, what they make of it.

Three Strands of Nuclear Superiority Research

There has long been a substantive back-and-forth in the academic and policy communities about whether, why, and how nuclear superiority matters. I divide this work into three categories. In the first category, early theoretical work on the impact of nuclear weapons on interstate relations mostly argues that nuclear superiority was irrelevant if the relevant states enjoyed a secure second strike. As the Cold War drew on, the interactions between a growing number of nuclear-armed states provided case studies against which to test these theories. Much of the ensuing qualitative work constitutes a second category and appears to confirm the theoretical arguments that nuclear superiority was largely irrelevant, though some findings suggest that the relative nuclear balance sometimes mattered in nuanced ways. Finally, and most recently, large-N statistical work attempts to determine whether there is a systematic relationship between nuclear superiority and dynamics in international politics. This literature, investigating the outcomes of interstate crises,¹¹ the initiation of interstate disputes,¹² and the success of coercive diplomacy,¹³ has produced somewhat mixed results.

The dominant theories largely reject the value of nuclear superiority. Writing just after the U.S. atomic bombings of Japan in 1945, Bernard Brodie dismissed the very notion of nuclear superiority: "Superiority in numbers of bombs is not in itself a guarantee of strategic superiority in atomic bomb warfare."¹⁴ For Brodie, it was more important for states to acquire and maintain a secure second strike than to achieve nuclear superiority.¹⁵ Two decades later, Thomas

11. Kroenig, *The Logic of American Nuclear Strategy*.

12. Erik Gartzke, Jeffrey M. Kaplow, and Rupal N. Mehta, "Nuclear Deterrence and the Structure of Nuclear Forces," unpublished manuscript, University of California, San Diego, 2015, http://dl.jkaplow.net/GKM_Deterrence_2014.pdf.

13. Sechser and Fuhrmann, *Nuclear Weapons and Coercive Diplomacy*.

14. Bernard Brodie, ed., *The Absolute Weapon: Atomic Power and World Order* (New York: Harcourt, 1946), p. 46.

15. Bernard Brodie, "The Development of Nuclear Strategy," *International Security*, Vol. 2, No. 4

Schelling noted the degree to which nuclear weapons and strategic delivery systems had compressed the speed with which states could inflict violence, arguing that modern interstate conflict would be characterized not by competitions of military strength but rather by “competitions in risk taking.”¹⁶ Because of the speed and assurance with which they can deliver destruction, “nuclear weapons make it possible to do monstrous violence to the enemy without first achieving victory.”¹⁷ Within this world, the relative nuclear balance was insignificant—all states were vulnerable. Refining these arguments, Robert Jervis demonstrated that superiority might only matter in a war of attrition, which seemed very unlikely in the nuclear era.¹⁸ Continuing in this tradition, Kenneth Waltz similarly rebuffed the notion that nuclear superiority could convey advantages: “so long as two or more countries have second-strike forces, to compare them is pointless.”¹⁹ He acknowledged that some U.S. policymakers did not share this view but argued this was a failure to appreciate the changes wrought by the nuclear revolution.²⁰

Proponents of the benefits of nuclear superiority, however, argue that superiority strengthens coercive bargaining leverage, reduces damage in prospective exchanges, and bolsters extended deterrence commitments.²¹ All things being equal, the state that expects to suffer less damage is more likely to press its claims, whereas the state that expects to suffer more damage is more likely to capitulate. Nuclear superiority, it is argued, might decrease expected damage in one of two ways. First, it might permit the nuclear-superior state to launch a successful counterforce strike and eliminate much or all the adversary’s nuclear forces, thereby insulating itself from a retaliatory nuclear strike. Second, in a purely countervalue exchange, the nuclear-superior state could inflict more damage than its adversary.²²

The theory of the nuclear revolution and the relative unimportance of the

(Spring 1978), pp. 65–66, <https://doi.org/10.2307/2538458>. Elsewhere, Bernard Brodie illustrates the concept of deterrence by proposing that a hypothetical state, though conventionally inferior to the Soviet Union, could nonetheless enjoy significant deterrent power by threatening Moscow with “a single thermonuclear bomb.” See Bernard Brodie, *Strategy in the Missile Age* (Princeton, N.J.: Princeton University Press, 1959), p. 275.

16. Thomas C. Schelling, *Arms and Influence* (New Haven, Conn.: Yale University Press, 1966), pp. 92–125.

17. *Ibid.*, p. 22.

18. Jervis, “Why Nuclear Superiority Doesn’t Matter”; and Jervis, *The Illogic of American Nuclear Strategy*.

19. Waltz, “Nuclear Myths and Political Realities,” p. 738.

20. *Ibid.*, p. 741.

21. Kroenig, “Nuclear Superiority and the Balance of Resolve,” pp. 146–152.

22. Robert Jastrow, “Why Strategic Superiority Matters,” *Commentary*, March 1983, <https://www.commentary.org/articles/robert-jastrow/why-strategic-superiority-matters/>.

nuclear balance appears to have been largely supported by case studies conducted during the Cold War. Analyzing the Cuban Missile Crisis, Marc Trachtenberg argues that nuclear superiority did not influence U.S. decisions and that fears of escalation and general nuclear war predominated in U.S. discussions.²³ Richard Betts reviewed nearly a dozen examples of nuclear coercion to conclude that nuclear superiority was at times an ancillary factor influencing state behavior, but that “in crisis decisions the United States exhibited a proclivity toward nuclear coercion that was not strongly governed by the nuclear balance of power.”²⁴ In their examination of threatened and actual uses of force since World War II, Barry Blechman and Stephen Kaplan conclude that the relative nuclear balance between the United States and the Soviet Union had no impact on state crisis behavior.²⁵ Statements from senior U.S. officials also appear to cast doubt on arguments about the importance of the relative nuclear balance. Writing twenty years after the Cuban Missile Crisis, several former John F. Kennedy administration officials rejected the notion that U.S. nuclear superiority—then at its all-time greatest, as measured in warheads—had played any role in the outcome.²⁶ Former Secretary of Defense Robert McNamara wrote of a “recognition by U.S. civilian and military officials that NATO’s vastly superior nuclear capabilities, measured in terms of numbers of weapons, did not translate into usable military power.”²⁷ In reviewing the U.S. record of “atomic diplomacy,” former National Security Advisor McGeorge Bundy concluded that nuclear superiority was not a significant factor.²⁸

Still, other scholars point to some episodes as potentially demonstrating the importance of nuclear superiority. Glenn Herald Snyder argues that, on the

23. Marc Trachtenberg, “The Influence of Nuclear Weapons in the Cuban Missile Crisis,” *International Security*, Vol. 10, No. 1 (Summer 1985), pp. 147–156, <https://doi.org/10.2307/2538793>. Marc Trachtenberg also argues that there is some limited evidence to support the idea of asymmetric effects of nuclear superiority (*ibid.*, pp. 156–161). That is, although U.S. officials were not emboldened, Soviet officials may have been intimidated by concerns of their relative nuclear inferiority. But his argument about the Soviet motivations for not alerting their nuclear forces is largely speculation and does not, likely due to lack of sources, point to any evidence that the nuclear balance definitively influenced Soviet thinking.

24. Richard K. Betts, *Nuclear Blackmail and Nuclear Balance* (Washington, D.C.: Brookings Institution Press, 1987), p. 213.

25. Barry M. Blechman and Stephen S. Kaplan, *Force without War: U.S. Armed Forces as a Political Instrument* (Washington, D.C.: Brookings Institution Press, 1978), p. 132.

26. Dean Rusk et al., “Essay: The Lessons of the Cuban Missile Crisis,” *Time*, September 27, 1982, pp. 89–92.

27. Robert S. McNamara, “The Military Role of Nuclear Weapons: Perceptions and Misperceptions,” *Foreign Affairs*, Vol. 62, No. 1 (Fall 1983), p. 64.

28. McGeorge Bundy, “The Unimpressive Record of Atomic Diplomacy,” in Gwyn Prins, ed., *The Choice: Nuclear Weapons versus Security* (London: Chatto and Windus, 1984), pp. 44–47.

question of whether nuclear superiority matters, “pure logic gives a clear negative to this question,” but that “real world experience does not quite follow this logic.”²⁹ But in a footnote at the end of the discussion on nuclear superiority, his coauthor Paul Diesing rebuts Snyder’s analysis, arguing that “the empirical evidence on whether nuclear superiority confers any bargaining advantage is so weak that no conclusions can be reached.”³⁰ As noted, both Trachtenberg and Betts acknowledge that nuclear superiority (or inferiority, in the case of Trachtenberg) may matter, though they circumscribe when and how.

More recently, a handful of studies investigate whether nuclear superiority matters by subjecting large-N datasets to statistical regressions. Matthew Kroenig, for instance, analyzes a set of interstate crises between nuclear-armed states and finds that states that possess a greater proportion of nuclear weapons in the dyad are more likely to achieve their goals.³¹ Using a dataset on militarized compellent threats, Todd Sechser and Matthew Fuhrmann find no evidence that nuclear superiority influences the success of states’ coercive threats.³² Erik Gartzke, Jeffrey Kaplow, and Rupal Mehta find that nuclear superiority is not associated with changes in the propensity of conflict between states.³³ Finally, Matthew Kroenig, Miriam Krieger, and Hans Noel replicate and rebut Sechser and Fuhrmann’s work, arguing that nuclear-superior states are more likely to succeed with compellent threats in part because they are more willing to issue them.³⁴

This recent wave of statistical work offers notable contributions. Kroenig formalizes the logic connecting nuclear superiority to relative damage assessments to interstate crisis outcomes. Sechser and Fuhrmann identify and resolve two weaknesses of earlier work: indeterminate research designs (largely through selection on the dependent variable) and inappropriate quantitative data (including observations that do not qualify as coercive threats). Gartzke, Kaplow, and Mehta introduce a measure of nuclear force diversification, which they find affects interstate conflict dynamics. These contributions have

29. Snyder and Diesing, *Conflict among Nations*, pp. 459–462. The section analyzing nuclear superiority rightly notes that, on questions of the nuclear balance, decision-makers’ beliefs may matter more than the objective balance.

30. *Ibid.*, p. 462.

31. Kroenig, “Nuclear Superiority and the Balance of Resolve.” For a fuller development of the argument, including qualitative case studies and policy implications, see Kroenig, *The Logic of American Nuclear Strategy*.

32. Sechser and Fuhrmann, “Crisis Bargaining and Nuclear Blackmail.”

33. Gartzke, Kaplow, and Mehta, “Nuclear Deterrence and the Structure of Nuclear Forces.”

34. Kroenig, Krieger, and Noel, “Dare to Fail.”

added to our understanding of how to conceptualize, measure, and test nuclear superiority.

Warhead Prominence and Complete Information

Despite these strengths, the recent statistical work generally suffers from two related weaknesses.³⁵ First, it assumes that total warhead counts are an appropriate measure of nuclear capabilities. All the statistical analyses discussed in the previous section use a simple total of each state's nuclear warheads. But the ability to inflict or limit damage in a nuclear exchange does not accrue simply in proportion to the number of warheads. Those warheads must be deployed and made operationally available to permit strikes. Even a modest damage limitation capability requires that a state's nuclear arsenal meet a slew of demanding technical and operational criteria, from miniaturized warheads with sufficiently high yield-to-weight ratios to strategic launchers with highly accurate guidance systems. Total warhead counts do not consider variation in whether and how those warheads are deployed or even in the characteristics of the warheads themselves, to say nothing of the variation in the intelligence, surveillance, and reconnaissance (ISR), targeting, and command-and-control capabilities required for damage limitation strikes.³⁶

Second, the work assumes that states and their leaders possess complete (or at least symmetric) information about other states' nuclear arsenals. Yet states may often have imperfect information about the size and configuration of other states' nuclear forces. In addition to the general risk of intelligence failures, states often have incentives to obscure or misrepresent their military capabilities, especially in the nuclear realm. What matters in determining state crisis behavior is not the size and capability of the adversary's nuclear forces but, rather, the state *perceptions* of those forces. Without demonstrating that a state possesses reasonably accurate information about the adversary's nuclear arsenal, the quantitative literature on nuclear superiority may fall short.

Throughout my analysis, I focus on the nuclear arsenals of the United States and the Soviet Union throughout the Cold War. This focus presents theoretical, evidentiary, and methodological advantages. First, the Cold War superpower dyad represents a hard test of my claim that the impact of the nuclear balance

35. In analyzing the recent quantitative literature on nuclear superiority, I focus my discussion on three representative works: Kroenig, "Nuclear Superiority and the Balance of Resolve"; Kroenig, *The Logic of American Nuclear Strategy*; and Sechser and Fuhrmann, "Crisis Bargaining and Nuclear Blackmail." Of note, all the works cited above use the same operationalization.

36. I thank Eric Snyder for this point.

is mediated by state beliefs and information rather than a direct result of nuclear arsenals' technical features. These two states invested considerable resources in both knowing the characteristics of each other's nuclear forces and developing concepts for measuring and responding to those forces. If any states could claim to meaningfully know, conceptualize, and respond to the nuclear balance, they would be the Cold War superpowers. Second, in evidentiary terms, the declassification of U.S. government documents from the Cold War provides a rich source of data on the nuclear forces of the two states. Third, methodologically, these two states represent a large proportion of the observations in the statistical studies of nuclear superiority.³⁷ In Kroenig's analysis of nuclear crisis outcomes, either the Soviet Union or the United States participated in sixteen of the twenty nuclear crises in the dataset. Both countries appeared in thirteen of the crises. In Sechser and Fuhrmann's analysis, either country appears in half of the forty-eight cases in which a nuclear-armed state issued a compellent threat.

ASSUMPTION OF WARHEAD PROMINENCE

Recent works use a state's number of nuclear warheads to measure the nuclear balance. In assessing the impact that nuclear superiority has on the success of compellent threats, Sechser and Fuhrmann employ three variables measuring nuclear superiority: a dichotomous variable (measuring whether the state possessed more nuclear warheads than its target), a nuclear ratio variable (measuring the proportion of state-controlled nuclear warheads in the dyad), and a variable measuring the difference in the number of warheads possessed by each state.³⁸ In measuring the impact of nuclear superiority on crisis outcomes between nuclear-armed states, Kroenig employs two measures of nuclear superiority: a dichotomous variable measuring whether the state possessed more nuclear warheads than its counterpart, and a second variable measuring the proportion of state-controlled nuclear warheads in the dyad. Kroenig, Krieger, and Noel, in assessing whether nuclear superiority affects the success of compellent threats, again use measures of total warhead stockpiles.³⁹

All these measures calculate the state's total nuclear warhead stockpile. While this approach may present advantages in terms of measurement and conceptual simplicity, I argue that it is an inappropriate operationalization of nuclear capability for at least four reasons. First, counting total stockpiles in-

37. This large U.S. and Soviet representation in the data may, itself, be analytically problematic.

38. Sechser and Fuhrmann, "Crisis Bargaining and Nuclear Blackmail," pp. 185–186.

39. Though not discussed in detail here, Erik Gartzke, Jeffrey Kaplow, and Rupal Mehta use the

cludes tactical warheads that often have little, if any, relevance to the overall strategic balance. Tactical nuclear weapons might not be used in a direct attack on the adversary's home territory and could not be used in a counterforce strike. Stockpile-based nuclear superiority measures will therefore systematically misrepresent the nuclear arsenals of states that deploy significant numbers of tactical nuclear weapons (e.g., the United States, the Soviet Union, France, and Pakistan). For example, the Soviet Union's nuclear warhead stockpile surpassed that of the United States in 1978, when the U.S. share of nuclear weapons in the U.S.-Soviet dyad dropped to 48 percent. Yet the United States still enjoyed significant strategic nuclear superiority over the Soviet Union because it possessed 72 percent of the strategic bombs within the dyad.⁴⁰

The second reason why stockpiles are an inappropriate operationalization of nuclear capability is because they include both deployed and non-deployed warheads. Non-deployed weapons are unlikely to influence a state's behavior in a short-term crisis. There can be large differences between total warhead stockpile and deployed (or loaded) weapons. In 1966, for instance, the United States had an estimated 11,232 strategic warheads, barely half of which were assessed as loaded on delivery vehicles. Counting all warheads thus results in a doubling of the United States' relevant nuclear arsenal in 1966. Similar discrepancies would arise for states with a "catalytic posture," such as Israel and South Africa, which kept their nuclear weapons undeployed.⁴¹

Third, stockpile-based measures do not account for weapon yield. If, as suggested by theories of nuclear superiority, nuclear-superior states are more successful in interstate negotiations because they anticipate suffering less damage, then states should care greatly about the destructive power of the adversary's nuclear arsenal. An arsenal composed of low-yield fission weapons could theoretically confer weaker bargaining advantages than one composed of multi-megaton thermonuclear bombs. Stockpile-based measures elide this potentially significant difference by treating all warheads the same. For exam-

same binary and ratio measures of total stockpiles as Matthew Kroenig to operationalize nuclear capability. Gartzke, Kaplow, and Mehta, "Nuclear Deterrence and the Structure of Nuclear Forces."

40. Robert S. Norris and Thomas B. Cochran, *U.S.-USSR/Russian Strategic Offensive Nuclear Forces, 1945-1996*, Nuclear Weapons Databook Working Paper 97-1 (Washington, D.C.: Natural Resources Defense Council, January 1997), p. 12. This 1997 Databook and the Archive of Nuclear Data resources largely agree. If there are differences, then I use the data in the Archive of Nuclear data that was compiled and published more recently. "Archive of Nuclear Data," *Natural Resources Defense Council*, <https://web.archive.org/web/20061031132859/http://www.nrdc.org///nuclear/nudb/datainx.asp>.

41. Vipin Narang, *Nuclear Strategy in the Modern Era: Regional Powers and International Conflict* (Princeton, N.J.: Princeton University Press, 2014). I thank Bryan R. Early for this point.

ple, in 1969, the United States possessed 16,000 more nuclear warheads than the Soviet Union, but the total megatonnage of the Soviet arsenal surpassed that of the U.S. force.

Fourth, stockpile-based measures do not account for the delivery vehicles necessary to execute an attack. If states are concerned about their relative ability to limit the damage that they will suffer in a prospective nuclear exchange, then they may care more about the relative balance of delivery vehicles, not nuclear warheads. In 1972, the United States controlled nearly two-thirds of all the warheads possessed between it and the Soviet Union. But that same year was also the first in which the Soviet Union deployed more strategic delivery vehicles than the United States.

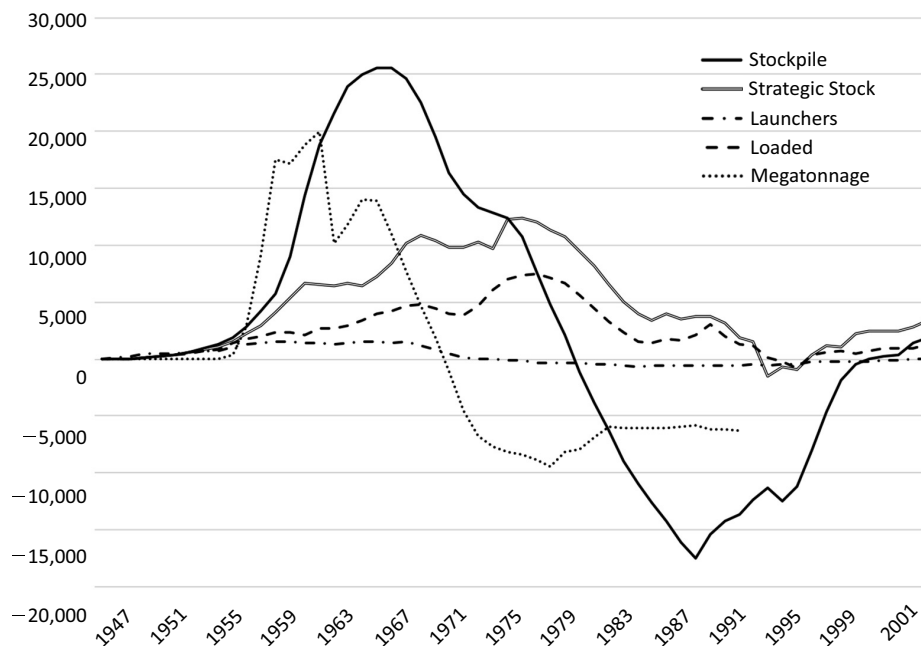
Some of the recent statistical work recognizes that nuclear capabilities may be measured in other ways but ultimately concludes that these measures are unimportant, inaccessible, or redundant. For example, Kroenig observes that “nuclear analysts often consider additional factors when calculating the nuclear balance between states including: total megatonnage, numbers and accuracy of delivery vehicles, and the ability of command-and-control systems to execute war plans in a crisis.” But he argues that using a stockpile-based measure is nonetheless appropriate both because “detailed information on these variables is not available for every nuclear weapon state in every year” and because “there is good reason to believe that simple warhead counts and more complicated assessments of nuclear capabilities are highly correlated.”⁴² Similarly, Gartzke, Kaplow, and Mehta acknowledge that “these simple measures do not capture the full complexity of nuclear superiority calculations made by U.S. and Soviet nuclear strategists during the Cold War, which implicate such varied topics as geography, missile accuracy and yield, targeting strategies, and nuclear fratricide.” They conclude, however, that “absent a more comprehensive measure of the nuclear balance, we agree with Kroenig that these simpler measures of nuclear arsenal size are at least likely to be strongly correlated with more complex calculations. A full empirical test of this assumption awaits future data collection.”⁴³ This assumption of correlation likely holds true across many nuclear-armed state dyads in many years. For example, there is no year in which China could be said to have enjoyed material nuclear superiority over the United States or the Soviet Union.

The assumption of correlation may fail to hold in many important instances, however, especially for the U.S.-Soviet dyad. To illustrate this, I constructed a

42. Kroenig, “Nuclear Superiority and the Balance of Resolve,” p. 155.

43. Gartzke, Kaplow, and Mehta, “Nuclear Deterrence and the Structure of Nuclear Forces,” p. 19.

Figure 1. U.S.-Soviet Union/Russia Absolute Nuclear Balance, 1945–2002



NOTE: The higher the line is above zero, the greater the U.S. superiority in that measure; the lower the line is below zero, the greater the Soviet/Russian superiority in that measure.

new dataset, the Nuclear Capabilities Dataset.⁴⁴ This dataset records nuclear capabilities for all nuclear-weapon states from 1945 to 2002 across five measures: nuclear warheads stockpile (stockpile), strategic warheads stockpile (strategic stock), strategic delivery vehicles (launchers), warheads loaded on strategic delivery vehicles (loaded), and total megatonnage (megatonnage). Figure 1 plots the difference between the number of U.S. capabilities for a given measure and the number of Soviet capabilities. The higher the line above zero, the greater the U.S. superiority is in that measure; the lower the line is below zero, the greater the Soviet superiority in that measure.

As shown in figure 1, there appears to be some co-movement in the relative balances of stockpile sizes and strategic delivery vehicles. Yet the relationship

44. For more on the dataset, see the Nuclear Capabilities Dataset Codebook in the online appendix.

between the overall stockpile and other measures is more tenuous. The balance in overall megatonnage reversed a full decade earlier than the reversal in overall stockpile size. Conversely, the United States consistently maintained superiority in both the size of its strategic warhead stockpile and the number of warheads loaded onto strategic delivery vehicles. Throughout this period, the outcomes of crises and threats between the United States and the Soviet Union would have been overdetermined by U.S. nuclear superiority in strategic warheads and loaded strategic warheads.

To further illustrate the degree to which stockpile-based measures can distort assessments of the nuclear balance, I also constructed a measure of the relative nuclear balance across each of these five measures. This measure represents the proportion of capabilities possessed by the United States each year. Values closer to 1 indicate higher degrees of U.S. nuclear superiority, while values closer to 0 indicate higher degrees of Soviet nuclear superiority. A value of 0.5 indicates perfect parity. In figure A1 in the online appendix, I plotted the relative nuclear balance along each of the five measures of nuclear capabilities. Figure A1 shows that depictions of the U.S.-Soviet nuclear balance often depend on the measure that is used. Though the trend of each measure is generally toward greater parity over the course of the Cold War, the measures do not always move in the same direction. Most significantly, even when the measures do move together, they move at dramatically different rates. By the mid- to late-1960s, assessments of the nuclear balance would have depended heavily on what measure was used. This disparity only increased with time. In 1976, for instance, a stockpile-based measure reported a value of 0.55, indicating almost perfect parity. In the same year, however, the United States had more than three-quarters of all strategic warheads and strategic warheads loaded onto delivery vehicles, while the Soviet Union had the majority of strategic launchers and nearly three-quarters of the total megatonnage. It is true that, until the early 1960s, the United States enjoyed significant superiority across all five measures. But observations of interstate relations under conditions of such extreme superiority raise questions about whether it is nuclear superiority that matters or whether the inferior state no longer possesses a secure second-strike capability.⁴⁵

45. By “nuclear superiority,” I mean numerical superiority in warheads. For other works that find survivability (and diversification) more important than superiority, see Gartzke, Kaplow, and Mehta, “Nuclear Deterrence and the Structure of Nuclear Forces”; and Dana Higgins, Connor Huff, and Anton Strezhnev, “Survivability Not Superiority: A Critique of Kroenig,” unpublished paper, Harvard University, 2013, https://robobees.seas.harvard.edu/files/danahiggins/files/higgins_et_al_-_survivability_not_superiority.pdf.

Throughout the Cold War, U.S. analysts and policymakers emphasized more comprehensive, nuanced measures than just warhead counts. First, early warnings in the United States of an impending “bomber gap” and a later “missile gap” demonstrate that U.S. analysts and policymakers focused on delivery systems, not on warheads. From the late 1940s to the mid-1950s, U.S. officials became (mistakenly) concerned that the Soviet Union had gained a significant strategic advantage over the United States by deploying large numbers of jet-powered, long-range, nuclear-capable bombers.⁴⁶ In the late 1950s, analogous fears emerged of a “missile gap” following the Soviet launches of the SS-6 ICBM (intercontinental ballistic missile) and the Sputnik satellite, which were seen as signaling Soviet technological superiority.⁴⁷ Those fears were further heightened by the publication of the Gaither Report, which warned that “by 1959, the USSR may be able to launch an attack with ICBMs carrying megaton warheads, against which [Strategic Air Command] will be almost completely vulnerable under present programs.”⁴⁸

Throughout both episodes, U.S. officials were much more concerned with a relative inferiority of delivery vehicles (bombers and, later, missiles) than with nuclear warheads. The 1948 Finletter Report warning of an impending bomber gap, for instance, focused on the development and production of nuclear-capable aircraft, with little reference to the bombs that they would deliver.⁴⁹ The Gaither Report recommended measures to ensure the survivability of the U.S. deterrent that were focused on delivery systems, including guidance to increase bomber alert status, to harden and disperse Strategic Air Command facilities, and to ramp up production of ICBMs.⁵⁰ Similarly, concerns in the United States in the 1970s of an alleged “window of vulnerability” centered around Soviet strategic delivery vehicles’ estimated throw-weight and accuracy.⁵¹

Second, National Intelligence Estimates (NIEs) on Soviet nuclear capabilities produced throughout the Cold War focused almost exclusively on delivery

46. Luke Benjamin Wells, “The ‘Bomber Gap’: British Intelligence and an American Delusion,” *Journal of Strategic Studies*, Vol. 40, No. 7 (2017), pp. 963–989, <https://doi.org/10.1080/01402390.2016.1267006>.

47. Greg Thielmann, “Looking Back: The Missile Gap Myth and Its Progeny,” *Arms Control Today*, Vol. 41, No. 4 (May 2011), p. 44, <http://www.jstor.org/stable/23629042>.

48. U.S. Office of Defense Mobilization, *Deterrence and Survival in the Nuclear Age* (Washington, D.C.: Government Printing Office [GPO], 1957), p. 14 (hereafter, *Deterrence and Survival in the Nuclear Age*).

49. Thomas Knight Finletter, *Survival in the Air Age: A Report by the President’s Air Policy Commission* (Washington, D.C.: GPO, 1948).

50. *Deterrence and Survival in the Nuclear Age*, pp. 6–7.

51. Pavel Podvig, “The Window of Vulnerability That Wasn’t: Soviet Military Buildup in the 1970s—A Research Note,” *International Security*, Vol. 33, No. 1 (Summer 2008), pp. 118–138, <https://doi.org/10.1162/isec.2008.33.1.118>.

systems. The U.S. intelligence community consistently produced two series of NIEs addressing the Soviet Union's nuclear capabilities.⁵² One, the 11-2 series, reported on developments in Soviet atomic energy, including progress on nuclear reactors, nuclear materials production, nuclear weapons development, and atomic aid provided to other states. Estimates of the Soviet nuclear warhead stockpile did not feature prominently in these analyses. For example, the 1959 NIE on the Soviet atomic energy program noted, "we lack sufficient evidence to support a firm estimate of the Soviet weapons stockpile by number, by type, by mission, or otherwise."⁵³ Similarly, the 1963 NIE estimated cumulative Soviet U-235 production at 130,000 kilograms but, considering significant margins of error, noted that the true number could have fallen anywhere between 80,000 and 180,000 kilograms.⁵⁴ Though the 1959 report presents "illustrative allocations" of Soviet nuclear forces, it cautions that "in the light of the range and complexity of the factors discussed above, we do not believe it is possible or desirable to arrive at any single 'most probable' estimate of the Soviet weapons stockpile at selected periods."⁵⁵ These NIEs instead focused largely on qualitative changes in the Soviet nuclear weapons program⁵⁶ and do not appear to have provided any usable estimate of the Soviet stockpile size until 1964, and even these were presented with significant caveats.⁵⁷

52. For a brief history of the origins of National Intelligence Estimates, see Donald P. Steury, ed., *Intentions and Capabilities: Estimates on Soviet Strategic Forces, 1950–1983* (Washington, D.C.: Central Intelligence Agency, 1996), pp. xi–xxii. For a summary of how contemporary NIEs are written and their role in the policy process, see Greg Bruno and Sharon Otterman, *National Intelligence Estimates* (New York: Council on Foreign Relations, 2008), <https://www.cfr.org/background/national-intelligence-estimates>.

53. See document A1, pg. 59, in the online qualitative appendix for the active citation. For more on active citation and increasing transparency in qualitative international relations and security studies scholarship, see Andrew Moravcsik, "Active Citation: A Precondition for Replicable Qualitative Research," *PS: Political Science & Politics*, Vol. 43, No. 1 (January 2010), pp. 29–35, <https://doi.org/10.1017/S1049096510990781>; Andrew Moravcsik, "Active Citation and Qualitative Political Science," *Qualitative & Multi-Method Research*, Vol. 10, No. 1 (Spring 2012), pp. 33–37, https://www.princeton.edu/~amoravcs/library/active_citation.pdf; Andrew Moravcsik, "Trust, but Verify: The Transparency Revolution and Qualitative International Relations," *Security Studies*, Vol. 23, No. 4 (2014), pp. 663–688, <https://www.doi.org/10.1080/09636412.2014.970846>; Diana Kapiszewski and Dessislava Kirilova, "Transparency in Qualitative Security Studies Research: Standards, Benefits, and Challenges," *Security Studies*, Vol. 23, No. 4 (2014), pp. 699–707, <https://doi.org/10.1080/09636412.2014.970408>; and Elizabeth N. Saunders, "Transparency without Tears: A Pragmatic Approach to Transparent Security Studies Research," *Security Studies*, Vol. 23, No. 4 (2014), pp. 689–698, <https://doi.org/10.1080/09636412.2014.970405>.

54. See document A2, p. 37, in the online qualitative appendix.

55. Document A1, p. 62, in the online qualitative appendix.

56. Niccolò Petrelli and Giordana Pulcini, "Nuclear Superiority in the Age of Parity: U.S. Planning, Intelligence Analysis, Weapons Innovation, and the Search for a Qualitative Edge, 1969–1976," *International History Review*, Vol. 40, No. 5 (2018), pp. 1191–1209, <https://doi.org/10.1080/07075332.2017.1420675>.

57. References to the estimates are made in document A3, p. 29, in the online qualitative appen-

A second set of annual NIEs, the 11-8 series, focused on Soviet nuclear forces that could be used to attack the continental United States. The substantive focus of these NIEs on delivery vehicles is demonstrated by the sections into which they are divided, with each estimate reporting Soviet developments on, in order: ICBMs, space-based capabilities, submarine-launched ballistic missiles, and long-range nuclear-capable bombers. In fact, by late 1955, Director of Central Intelligence Allen Dulles declared that collecting reliable intelligence on Soviet ICBMs was “of the highest priority, probably of even greater ultimate importance to our national security than atomic energy intelligence.”⁵⁸ In short, U.S. intelligence estimates on Soviet nuclear capabilities generally discounted warhead counts because of both the difficulty of developing such estimates and their relative unimportance compared with delivery systems estimates.

Even the five measures captured in the Nuclear Capabilities Dataset fail to encompass the universe of nuclear forces measurements. As early as the 1960s, observers were attempting to measure and compare strategic nuclear forces’ capabilities to predict how the balance might change under future arms control agreements.⁵⁹ Over the years, analysts have developed numerous metrics, including equivalent megatons, lethal area potential, counter-military potential, hard-target kill capability, equivalent weapons, joint throw-weight and warhead measures, standard weapon stations, and distinct blasts index.⁶⁰ Cold War analysts modeled the economic consequences of nuclear exchanges, the

dix. Though the estimates themselves appear to have been excised from the declassified document, it is possible that similar estimates were made in earlier NIEs and that they had merely been excised from the declassified versions, though there is no indication of this. Some earlier NIEs included “illustrative allocations,” though the authors were careful to emphasize that these were notional and not meant to provide any meaningful estimate of Soviet nuclear forces.

58. As quoted in Steury, *Intentions and Capabilities*, p. 55.

59. See, for example, Glenn A. Kent, *On the Interaction of Opposing Forces under Possible Arms Agreements* (Cambridge, Mass.: Center for International Affairs, Harvard University, 1963), p. 5; Thomas A. Brown, “U.S. and Soviet Strategic Force Levels: Problems of Assessment and Measurement,” *Annals of the American Academy of Political and Social Science*, Vol. 457 (1981), pp. 18–27, <http://www.jstor.org/stable/1044185>; Jeffrey T. Richelson, “Evaluating the Strategic Balance,” *American Journal of Political Science*, Vol. 24, No. 4 (November 1980), pp. 779–803, <https://doi.org/10.2307/2110959>; Jeffrey T. Richelson, “Static Indicators and the Ranking of Strategic Forces,” *Journal of Conflict Resolution*, Vol. 26, No. 2 (1982), pp. 265–282, <https://doi.org/10.1177/0022002782026002004>; Lowell Bruce Anderson and Frederic A. Miercort, “On Weapons Scores and Force Strengths,” *Naval Research Logistics*, Vol. 42, No. 3 (April 1995), pp. 375–395, [https://doi.org/10.1002/1520-6750\(199504\)42:3<375::AID-NAV3220420305>3.0.CO;2-0](https://doi.org/10.1002/1520-6750(199504)42:3<375::AID-NAV3220420305>3.0.CO;2-0); and Fred A. Payne, “The Strategic Nuclear Balance: A New Measure,” *Survival*, Vol. 19, No. 3 (1977), pp. 107–110, <https://doi.org/10.1080/00396337708441677>.

60. For a review of most of these measures, many of which enjoy their own acronyms, see Gold, “Report of the Chairman,” pp. 132–135. The Soviet Union developed its own measures to evaluate the effectiveness of nuclear forces. See Levy, *Soviet Strategic Nuclear Measures of Effectiveness*.

role of strategic defense systems, and the importance of command and control.⁶¹ Scholars have recently conducted detailed simulations of exchanges involving contemporary China, Russia, and the United States.⁶² The five measures of nuclear forces discussed above may appear crude by comparison. Further, applying these complex measures to the same nuclear forces often produces very different assessments of the nuclear balance, and some observers may cherry-pick their measure to lobby for a preferred policy.⁶³

It is difficult to say *ex ante* which measure is most appropriate. The metrics that matter may depend on what kind of nuclear exchange is anticipated. In a counterforce nuclear exchange, perhaps the number of strategic delivery vehicles matters most since it better represents the force's survivability.⁶⁴ In a countervalue exchange, however, loaded warheads or total megatonnage may matter more, since they better measure the state force's destructive capability. Different metrics' importance may also vary with changes in technology. For instance, in the early years of the Cold War, when delivery vehicles could only carry a single warhead, the number of strategic launchers may have been most significant. But following the development of multiple independently targetable reentry vehicles, warhead loadings may have gained prominence. Different metrics might also shift in salience with changes in political or military leadership. For instance, for political leaders with relatively "unsophisticated" understandings of nuclear weapons, "pre-attack" indicators such as the number of deployed warheads may be most salient, whereas experts may

61. Leonard J. Bickley, Jane-Ring F. Crane, and Edward S. Pearsall, *Estimates of the Potential of the U.S. Economy following a Strategic Attack in 1975* (Alexandria, Va.: Institute for Defense Analyses, 1967); Colin S. Gray, "Nuclear Strategy: The Case for a Theory of Victory," *International Security*, Vol. 4, No. 1 (Summer 1979), pp. 54–87, <https://doi.org/10.2307/2626784>; Michael F. Altfeld and Stephen J. Cimbala, "Closing the Window of Vulnerability: Peacekeeper and Point Defense," *Comparative Strategy*, Vol. 5, No. 4 (1986), pp. 375–393, <https://doi.org/10.1080/01495938608402697>; and Daniel Shuchman, "Nuclear Strategy and the Problem of Command and Control," *Survival*, Vol. 29, No. 4 (1987), pp. 336–359, <https://doi.org/10.1080/00396338708442368>.

62. Keir A. Lieber and Daryl G. Press, "The End of MAD? The Nuclear Dimension of U.S. Primacy," *International Security*, Vol. 30, No. 4 (Spring 2006), pp. 7–44, <https://doi.org/10.1162/isec.2006.30.4.7>; Keir A. Lieber and Daryl G. Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence," *International Security*, Vol. 41, No. 4 (Spring 2017), pp. 9–49, https://doi.org/10.1162/ISEC_a_00273; and Wu Riqiang, "Living with Uncertainty: Modeling China's Nuclear Survivability," *International Security*, Vol. 44, No. 4 (Spring 2020), pp. 84–118, https://doi.org/10.1162/isec_a_00376.

63. Richelson, "Static Indicators and the Ranking of Strategic Forces."

64. Assumptions about other variables, such as the performance characteristics of the delivery vehicles and the nature of the nuclear strike, would also significantly change assessments of the strategic nuclear balance. For a demonstration of how predicted outcomes of a nuclear exchange change along some of these variables, see John D. Steinbruner and Thomas M. Garwin, "Strategic Vulnerability: The Balance between Prudence and Paranoia," *International Security*, Vol. 1, No. 1 (Summer 1976), pp. 138–181, <https://doi.org/10.2307/2538581>.

emphasize more complex exchange models.⁶⁵ But without first interrogating the historical record, it is difficult to say which measures will be most salient for decision-makers and, therefore, difficult to construct a plausible test.

Official U.S. documents suggest that analysts have been interested in all these and other nuclear capability measures. Government assessments suggest substantial nuance in considering Soviet nuclear forces, incorporating measures of delivery vehicles, throw-weight, accuracy, and megatonnage. At times, these assessments seemed to implicitly reject the stockpile-based approach when assessing the relative nuclear balance and its strategic implications. For example, 1969 was the first year in which Soviet strategic delivery vehicles outnumbered those of the United States.⁶⁶ Yet the NIE distributed that year was careful to acknowledge the diversity within this force and, significantly, its inability to execute a counterforce strike against the United States, noting that “the SS-9 is the only [Soviet] ICBM with the combination of payload and accuracy to attack hard targets effectively, but in its present numbers with single warheads it could attack no more than a small percent of the U.S. ICBM force.”⁶⁷ This kind of nuanced analysis of a state’s actual nuclear operational capability is elided by stockpile measures.

To investigate the extent to which empirical findings vary with regard to how the balance is measured, I replicated work by Kroenig and by Sechser and Fuhrmann.⁶⁸ These two works represent recent prominent and sophisticated attempts to statistically investigate the impact of the nuclear balance.

Using these new data, I reestimated the full models in both analyses. In total, I ran ten models for the Kroenig analysis, two for each of the five measures of nuclear balance: one measuring superiority as a binary variable, and another measuring superiority as the ratio of nuclear capabilities possessed by one state divided by the total nuclear capabilities possessed by both states in the dyad. I ran fifteen models for the Sechser and Fuhrmann analysis: the ten described above plus five additional models measuring superiority as the logged difference between each of the five capabilities. The results of the mod-

65. Richelson, “Evaluating the Strategic Balance,” pp. 781, 801; and Snyder and Diesing, “Conflict among Nations,” p. 459.

66. I include ICBMs, submarine-launched ballistic missiles, and long-range bombers.

67. See document A4, p. 3, in the online qualitative appendix.

68. In replicating Kroenig’s work, following Todd Sechser and Matthew Fuhrmann, I correct the identifier used for clustering standard errors by crisis dyad. Matthew Fuhrmann and Todd S. Sechser, “Debating the Benefits of Nuclear Superiority, Part III,” *Duck of Minerva*, March 28, 2013, <https://www.duckofminerva.com/2013/03/debating-the-benefits-of-nuclear-superiority-part-iii.html>.

Table 1. Logit Analyses of Impact of Various Forms of Nuclear Superiority on Success in Interstate Crises (Kroenig replication)

	Total Stockpile	Strategic Stockpile	Strategic Delivery Vehicles	Strategic Loadings	Total Megatonnage
Nuclear Superiority (binary)	2.005** (0.953)	0.763 (0.739)	1.124 (0.715)	0.763 (0.739)	0.458 (0.634)
Nuclear Ratio	4.252** (2.160)	2.268 (1.526)	1.376 (1.441)	2.672* (1.551)	1.315 (1.293)

Standard errors in parentheses. Control variables not shown. *p < 0.1, **p < 0.05, ***p < 0.01

Table 2. Logit Analyses of Impact of Various Forms of Nuclear Superiority on Success in Interstate Crises (Sechser and Fuhrmann replication)

	Total Stockpile	Strategic Stockpile	Strategic Delivery Vehicles	Strategic Loadings	Total Megatonnage
Nuclear Superiority (binary)	-0.274 (0.251)	-0.134 (0.248)	0.002 (0.257)	0.067 (0.255)	-0.191 (0.219)
Nuclear Ratio	-0.537 (0.464)	-0.179 (0.498)	-0.270 (0.488)	-0.079 (0.496)	-0.467 (0.456)
Nuclear Difference (difference)	0.001 (0.035)	0.000 (0.038)	-0.016 (0.044)	0.005 (0.038)	-0.012 (0.032)

Standard errors in parentheses. Control variables not shown. *p < 0.1, **p < 0.05, ***p < 0.01

els for the variable of interest (nuclear superiority) are reported in table 1 and table 2.⁶⁹

I begin with the Kroenig replication, the results of which motivate two conclusions. First, the results depend on how superiority is measured. In the original two models, in which the nuclear balance is determined by total nuclear stockpiles, the results are statistically significant at the 0.05 level. In the model employing the loaded strategic warhead ratio, the results are significant at the 0.10 level. In the remaining seven models, however, the results are not statistically significant at any conventional threshold. Without a well-formulated argument about which indicator is most relevant to states' actual understandings of the nuclear balance, it is not clear how to interpret these results. Sec-

69. Full results are available in the online quantitative appendix.

ond, and relatedly, the overall results suggest that nuclear superiority was not a significant factor in crisis outcomes. In most models, the nuclear superiority measure is statistically insignificant, including most of those which, a priori, would seem the most relevant for state perceptions, such as strategic delivery vehicles and strategic warhead loadings. Again, this implicates the role of beliefs about which components matter most in determining and comparing nuclear capabilities.

The Sechser and Fuhrmann replication results largely confirm their original findings about the irrelevance of the nuclear balance, despite the new operationalizations. None of the fifteen models are statistically significant, and over half the models report negative coefficients. Together, the two replication results suggest both that results are sensitive to the measure and that, overall, there is little evidence suggesting a role for the nuclear balance in the outcomes of either interstate crises or compellent threats.

ASSUMPTION OF COMPLETE INFORMATION

The second major weakness of the recent wave of statistical work is its insufficient attention to the role of information and beliefs. All the quantitative work discussed here uses data on nuclear warheads published by nongovernmental sources years or decades after the fact,⁷⁰ and in doing so, makes two crucial implicit assumptions.

First, the quantitative work assumes that states have accurate information, and thus assumes that what matters is the “actual” nuclear balance and not what states subjectively perceive. But each state possesses, at best, perfect information only about its own arsenal, and imperfect information about the adversary’s arsenal (though some states, especially those with relatively new nuclear forces, may not even have complete information about their own capabilities). In this way, both states in a dyad might simultaneously believe that they are in a state of nuclear superiority or inferiority, depending on their information about the nuclear forces of the adversary and their beliefs about what determines the nuclear balance. For some more balanced dyads that appear frequently in data (e.g., U.S.-Soviet Union and India-Pakistan), small errors in intelligence assessments could produce important differences in perceptions of the balance.

Second, this work uses data that may not even accurately represent states’

70. All these works rely on data published by either the Natural Resources Defense Council or the Bulletin of the Atomic Scientists. In both instances, these independent analyses draw on unclassified and declassified U.S. government sources.

nuclear arsenals. All the work uses data on warhead counts published by non-governmental sources. But much of these data are, fundamentally, compiled from U.S. government estimates. This means that while data on the U.S. nuclear arsenal are likely accurate, data on other states' nuclear arsenals are, at best, composite estimates.

I now assess the accuracy of these assumptions. First, I examine the extent to which states' real-time perceptions of the nuclear balance reflected reality. Determining those perceptions' accuracy requires information both on what states believed and a "true" baseline against which to compare those beliefs. Here, I examine the real-time assessments of the United States by compiling data on the U.S. intelligence community's estimates of the Soviet strategic nuclear force size and composition from 1960 to 1977.⁷¹ I draw from the 11-8 series of the NIEs that, in each year, provided estimates of the number of operational Soviet ICBMs, submarine-launched ballistic missiles, and long-range nuclear-capable bombers. This provides the most authoritative view of the U.S. intelligence community.⁷²

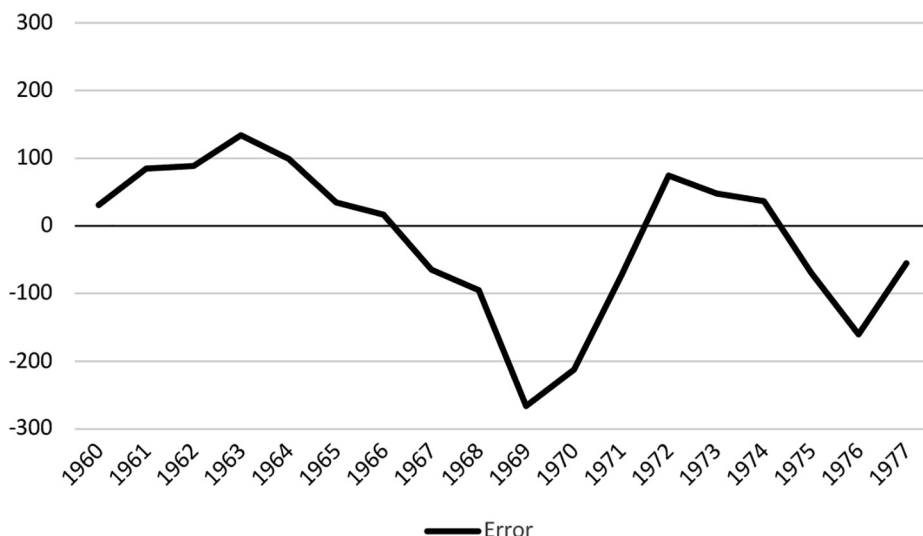
Next, I construct a "true" baseline of data on Soviet nuclear forces from the same period. Past efforts to assess the accuracy of U.S. intelligence estimates of Soviet nuclear capabilities have often suffered from a lack of authoritative data on the Soviet nuclear arsenal and have been forced to rely on "independent" assessments, such as those produced by the Natural Resources Defense Council (NRDC).⁷³ The NIE and NRDC both produced estimates of the total number of Soviet strategic delivery vehicles. Here, I construct two measures of U.S. intelligence errors using the NRDC data as a baseline: one reporting the error as the percent of the actual number of delivery vehicles that the NRDC reported, and the second reporting the error as the difference in the number of delivery vehicles between the NRDC and NIE estimates. The results are presented in figure A2 in the online appendix and in figure 2. The figures illustrate how U.S. intelligence estimates of Soviet strategic delivery vehicles exhibited nontrivial errors throughout the 1960s and 1970s, averaging an error of roughly 15 percent (without regard to sign) and 100 weapons (without re-

71. Unfortunately, it does not appear possible to construct similar datasets for other measures of nuclear balance, such as warhead stockpile, strategic warhead stockpile, or total megatonnage. For the NIEs produced during this period, estimates of these characteristics are either not provided or they have been excised or redacted in the declassified versions.

72. Richard A. Best Jr., *Intelligence Estimates: How Useful to Congress?* (Washington, D.C.: Congressional Research Service, January 6, 2011), p. 1.

73. This is the same data that I analyze in the earlier section examining the different ways of conceptualizing and measuring the strategic nuclear balance. See Norris and Cochran, *U.S.-USSR/Russian Strategic Offensive Nuclear Forces, 1945-1996*.

Figure 2. Errors in Delivery Vehicles Estimate as Difference in Number of Estimated and Actual Delivery Vehicles, 1960–1977



NOTE: Absolute error of estimation, computed as the difference between estimated number of delivery vehicles and the actual number of delivery vehicles.

gard to direction). Though these errors were significantly reduced in the early 1970s, they at times exceeded 40 percent and 250 delivery vehicles.

Unfortunately, however, the NRDC data make for an imperfect baseline for two reasons. First, little if any of the data come directly from Soviet sources; rather, much of it is built from U.S. government and nongovernmental sources. Thus, the numbers represent merely a best estimate rather than an actual baseline. Second, and most significantly, both the baseline and the estimates rely on some of the same sources.⁷⁴ For instance, the baseline relies, in part, on the declassified NIEs.⁷⁵ Because of this source duplication, the estimates and the baseline will naturally exhibit a higher-than-“natural” degree of correlation.⁷⁶ To avoid these problems, I rely on data published by

74. *Ibid.*, pp. 10–11.

75. See, for example, Daryl G. Press, *Calculating Credibility: How Leaders Assess Military Threats* (Ithaca, N.Y.: Cornell University Press, 2005), pp. 91–94. To construct his “true baseline” of Soviet nuclear forces, Daryl Press relies on data published in Pavel Podvig, ed., *Russian Strategic Nuclear Forces* (Cambridge, Mass.: MIT Press, 2001). The relevant data in this source comes from Norris and Cochran, *U.S.-USSR/Russian Strategic Offensive Nuclear Forces, 1945–1996*, which relies in part on U.S. government estimates.

76. Some analysts apparently fail to realize the extent to which their own “true baselines” are built

Pavel Podvig that draw from the personal papers of a senior Soviet defense official.⁷⁷ These papers include data on the number of Soviet ICBMs deployed in each year from 1970 to 1990.

To address the second weakness discussed above, I also examine the degree to which data from independent estimates, such as those compiled by the NRDC, reflect the actual size and composition of nuclear arsenals. I do this by comparing NRDC estimates for Soviet ICBMs to the Podvig data. As before, I again calculated these estimates' absolute and relative errors. I plotted the absolute (figure A3 in the online appendix) and relative (figure A4 in the online appendix) estimation errors of both the NIE and NRDC estimates.

The figures motivate several conclusions. First, over the period for which comparisons are possible, U.S. intelligence on Soviet ICBM estimates exhibited significant errors, sometimes overestimating the force by as much as 20 percent and underestimating it by as much as 10 percent. Without regard to sign, the average U.S. Soviet ICBM estimate was off by more than 10 percent. Though these errors appear smaller than the differences between the NIE and NRDC estimates, they are still significant for the "easy" case of ICBM estimates. Second, data from independent sources largely track the NIE data and exhibit similar estimation errors. Finally, by the 1980s, estimation errors had been significantly reduced to under 50 ICBMs, or about 5 percent of the actual size of the Soviet arsenal.

Though existing data only permit a direct assessment of U.S. intelligence on Soviet ICBMs from 1970 to 1977, there is reason to believe that U.S. assessments about other aspects of the Soviet nuclear arsenal, especially warhead counts, were subject to even greater inaccuracies. First, and most importantly, the error estimates reported above may be considered minimum expected error rates compared to errors in other types of estimates. It is likely easier to count silo-based delivery vehicles than either other delivery vehicles or warheads, and it is easier to evaluate quantitative rather than qualitative aspects of

from the same data sources as U.S. intelligence estimates, though they do sometimes supplement their analyses with additional sources. I thank J. Robert Logan for this point.

77. Podvig, "The Window of Vulnerability That Wasn't," p. 124. Podvig's data come from the personal papers of Vitalii Leonidovich Kataev, who served as senior adviser to the secretary for the defense industry of the Central Committee of the Communist Party in the Soviet Union from 1974 to 1990. Kataev, who was previously an aviation engineer, kept meticulous notes and the collection, held at the Hoover Institution, includes "reports, memoranda, correspondence, diaries, notes, meeting and conference materials, electronic documents, and sound recordings relating to Soviet military policy, the Soviet defense industry, Soviet nuclear weapons, and Soviet participation in arms limitation negotiations." See "Register of the Vitalii Leonidovich Kataev Papers," Online Archive of California, Hoover Institution Library and Archives, 2007, <http://www.oac.cdlib.org/findaid/ark:/13030/kt900039p6/>.

nuclear arsenals.⁷⁸ Warheads are smaller and more easily concealed and transported than their associated launchers. Indeed, the relative ease of counting launchers compared to warheads is a primary reason why arms control agreements have largely restricted the former and not the latter. Second, as discussed earlier, U.S. intelligence assessments themselves directly acknowledged the difficulty in counting warheads and at times either declined to provide such an estimate or provided very wide ranges.⁷⁹ Analysts seemed reluctant to provide an estimate of the stockpile size given the inherent difficulties.

Finally, other evidence suggests that U.S. estimates of the warhead stockpile and other attributes of the Soviet nuclear arsenal exhibited even greater inaccuracies. For example, in 1993, the then head of the Russian Ministry of Atomic Energy stated that the Soviet Union's nuclear warhead stockpile had peaked in 1986 at about 45,000 warheads, as much as 17,000 more bombs or more than 60 percent higher than the U.S. government had estimated at the time.⁸⁰ U.S. estimates of the Soviet forces' technical features were similarly flawed. U.S. estimates from the middle and late Cold War incorrectly placed the accuracy of key Soviet ICBMs at almost half (250 meters) of their true level (400 meters or more).⁸¹ Other reports overestimated the hardness of Soviet ICBM silos by as much as a factor of ten.⁸²

These same challenges confront independent analyses of nuclear arsenals, especially to the extent that they rely on U.S. government documents. These estimation errors suggest that U.S. decision-makers frequently did not have accurate information about key aspects of the Soviet nuclear arsenal. As detailed information on Soviet intelligence estimates has yet to be released, it is impossible to similarly analyze Soviet perceptions. Yet there is little reason to believe that Soviet intelligence products would have been significantly more accurate than those of the United States.

When the differences between nuclear arsenals are clearly large, assump-

78. Steve Fetter, "A Comprehensive Transparency Regime for Warheads and Fissile Materials," *Arms Control Today*, Vol. 29, No. 1 (January/February 1999), pp. 3–7, <https://www.jstor.org/stable/23626040>.

79. See, for example, document A1 and document A2 in the qualitative online appendix.

80. William J. Broad, "Russian Says Soviet Atom Arsenal Was Larger than West Estimated," *New York Times*, September 26, 1993, <https://www.nytimes.com/1993/09/26/world/russian-says-soviet-atom-arsenal-was-larger-than-west-estimated.html>; and Houston T. Hawkins, "Rethinking the Unthinkable," *National Security Science*, Los Alamos National Laboratory, December 1, 2014, p. 10, https://www.lanl.gov/discover/publications/national-security-science/2014-december/rethinking_the_unthinkable.php.

81. Podvig, "The Window of Vulnerability That Wasn't," pp. 125–129.

82. The U.S. estimated hardness rating was 15,000 to 25,000 pounds per square inch, whereas the actual maximum Soviet hardness rating was 1,500 pounds per square inch. *Ibid.*, pp. 129–132.

tions of complete information may not necessarily bias the analysis. These systematic biases will be most significant when the two states' nuclear capabilities are more similar and when they possess relatively smaller arsenals. At least two sets of cases stand out as potentially more vulnerable to this bias. The first set is a series of interactions between the United States and the Soviet Union from the late 1960s to the late 1970s, when the two states' nuclear arsenals approached parity and the Soviet Union began to obtain superiority across some nuclear measures.⁸³ The errors during this period were likely large enough to misrepresent even which side possessed superiority. In 1972, for example, just before the crisis over the Yom Kippur War, the U.S. NIE assessment of strategic nuclear delivery vehicles was that the Soviet Union possessed 2,238 compared to 2,167 for the United States. It is still unclear exactly how many strategic delivery vehicles the Soviets deployed. But retrospective calculations by the NRDC estimate only 2,164 delivery vehicles, which is less than the United States deployed. Assuming that total strategic delivery vehicle estimates suffered from (merely) the same error rates as the ICBM estimates, the Soviets may have only deployed 1,895 delivery vehicles, nearly 300 fewer than the United States, meaning that both countries may have believed that they possessed superiority in this dimension.

The second set includes the several crises involving India and Pakistan from 1990 to 2001.⁸⁴ In each of these observations, estimates of the two nuclear arsenals' sizes are almost identical, differing by, at most, only two warheads.⁸⁵ It is unlikely that decision-makers at the time were aware of, much less sensitive to, differences between the two arsenals' sizes. As Francis Gavin observes, "it

83. The Sechser and Fuhrmann dataset includes the Cienfuegos base incident in Cuba as an instance of coercion by the United States directed against the Soviet Union, while Kroenig's dataset includes crises over the Cienfuegos submarine base (1970), the Yom Kippur War (1973), the war in Angola (1975), and the invasion of Afghanistan (1979). It is not clear whether nuclear weapons featured prominently in each of these cases but, for the sake of replication, I include them in the analysis. Kroenig, *The Logic of American Nuclear Strategy*; and Sechser and Fuhrmann, *Nuclear Weapons and Coercive Diplomacy*. For an argument that nuclear weapons did not actually feature very prominently in some of these nuclear crises, see Mathew Fuhrmann and Diane Labrosse, eds., "ISSF Roundtable 10-25 on the Logic of American Nuclear Strategy: Why Strategic Superiority Matters," *H-Diplo*, ISSF Roundtable, Vol. 10, No. 25 (March 29, 2019), <https://issforum.org/roundtables/10-25-nuclear>.

84. The Kroenig dataset includes four crises involving India and Pakistan (1990, 1998, 1999, and 2001), whereas the Sechser and Fuhrmann data include the 2001 attempt by India to coerce Pakistan into controlling terrorist organizations. Kroenig, *The Logic of American Nuclear Strategy*; and Sechser and Fuhrmann, *Nuclear Weapons and Coercive Diplomacy*.

85. These data are taken from Hans M. Kristensen and Robert S. Norris, "Global Nuclear Weapons Inventories, 1945–2013," *Bulletin of the Atomic Scientists*, Vol. 69, No. 5 (September 2013), pp. 75–81, <https://doi.org/10.1177%2F0096340213501363>. The Bulletin's data only identify India and Pakistan as initially possessing nuclear arsenals in 1998.

is not even clear that Indian and Pakistani civilian leaders knew the size of their *own* nuclear arsenals during this period" (italics added).⁸⁶

Toward Superior Nuclear Superiority Research

In this section, I propose two ways to improve the literature and offer some provisional evidence that, at least for the United States and the Soviet Union, the nuclear balance did not matter and, indeed, was often both elusive and subjective.

First, to incorporate state perceptions, research should use data from real-time intelligence assessments and employ them within directed dyads. As discussed, work in this area has relied on third-party estimates and assumed that each state in a dyad shares the same (accurate) information about the relative nuclear balance. Employing asymmetric directed dyads would better model real-world dynamics because perceptions of the nuclear balance may generate asymmetric effects, depending on which side of the balance a state believes it is. That is, a state's response to a perceived condition of nuclear superiority may not simply be the mirror of its response to a perceived condition of nuclear inferiority.⁸⁷ Research could reasonably assume that states possess perfect information about their own arsenals and imperfect information about other states' arsenals. Scholars should also employ sensitivity tests to ensure that their findings are robust to misperceptions generated by intelligence errors.

Second, to address the problems associated with using total warhead counts as a proxy for nuclear capabilities, scholars can develop more nuanced measures of nuclear capabilities. This should, if possible, reflect decision-makers' beliefs and perceptions at the time. In short, nuclear superiority will generate political effects only when and how leaders believe it does.

Perceptions of the U.S.-Soviet dyad balance can illustrate these points. To investigate nuclear superiority's potentially asymmetric effects, I employ an asymmetric directed dyad approach, examining first the perceptions of U.S. decision-makers and then those of their Soviet counterparts. To best capture how states perceive nuclear superiority, I eschew static mechanical measures of nuclear arsenals and instead attempt to examine expectations of the results of a nuclear exchange.

I propose using states' expectations about the outcome of a nuclear ex-

86. Gavin, "What We Talk about When We Talk about Nuclear Weapons."

87. For some evidence of this, see Trachtenberg, "The Influence of Nuclear Weapons in the Cuban Missile Crisis," pp. 147–161.

change, especially expected fatality counts. Arguments about the significance of nuclear superiority rest on state assumptions about future nuclear exchanges and the relative damage that each side would suffer. In this way, the best measure of how states perceive the nuclear balance is not raw weapon counts but, rather, the anticipated effects of using them.⁸⁸ Historically, the expected damage from suffering a nuclear strike was typically measured in fatalities and damage to industrial production. Yet these kinds of fatality and damage estimates are not easily constructed solely from warhead counts. Fatality estimates can vary widely depending on assumptions about force postures, exchange details, weather patterns, population density, missile accuracy, and effectiveness of civil defense measures, among others.⁸⁹ In fact, analysts have argued that even with roughly symmetrical strategic nuclear forces, these climatic, socioeconomic, and policy factors provided the Soviet Union an enduring advantage in its strategic competition with the United States.⁹⁰

U.S. NUCLEAR EXCHANGE FATALITY EXPECTATIONS

Throughout the Cold War, U.S. government agencies produced dozens of assessments estimating the damage that would be suffered by the United States and the Soviet Union in the event of a nuclear exchange. Many of the early estimates were prepared by the National Security Council's Net Evaluation Subcommittee (NESC), which was tasked with estimating the outcomes of nuclear attacks.⁹¹ Similar estimates were produced by other government organizations, including the Defense Department's Weapons Systems Evaluation Group, the Arms Control and Disarmament Agency, the Central Intelligence Agency's Office of Research and Estimates, the U.S. Air Force, and the Joint Chiefs of Staff.⁹² These estimates accounted for many of the other technologi-

88. Even using raw fatality estimates may oversimplify and distort state perceptions. For instance, would states care more about the number of fatalities inflicted, or the proportion of the state's population those fatalities represent? During the Cold War, U.S. officials often defined "unacceptable damage" to the Soviet Union in terms of a percentage of population and lost economic activity. U.S. nuclear exchange models often highlighted the "recovery time" of a society as a key variable impacting the relative strategic balance. See examples in the online qualitative appendix.

89. Desmond Ball, *Can Nuclear War Be Controlled?* (London: International Institute for Strategic Studies, 1981), p. 27.

90. Francis P. Hoerber, "How Little Is Enough?" *International Security*, Vol. 3, No. 3 (Winter 1978/79), p. 56, <https://www.jstor.org/stable/2626718>.

91. "National Security Council Directive," National Security Council 5511, February 14, 1955, *Foreign Relations of the United States [FRUS], 1950-1955: The Intelligence Community, 1950-1955* (Washington, D.C.: GPO, 2007), doc. 207, <https://history.state.gov/historicaldocuments/frus1950-55Intel/d207>.

92. For an excellent review of the Net Evaluation Subcommittee (NESC) and other government reports estimating the damage of prospective nuclear exchanges, including the impact of such reports on the thinking and decisions of policymakers, see Milne, "Hope Springs Eternal." For a discussion of the NESC and its reports, see William Burr, "Studies by Once Top Secret Government

cal, strategic, meteorological, and demographic variables that would otherwise be difficult to incorporate. They were often detailed examinations using different assumptions about how many weapons each side deployed, the weapons' capabilities, the targets to which they were assigned, and how they were launched. More significantly, from an analytical perspective, they provide a clear measure of the variable that should matter in theories of nuclear superiority: expected damage, typically in the form of expected fatalities. These assessments were more than just bureaucratic exercises. They reached and influenced the highest levels of the U.S. government. The results were regularly briefed to the president and senior officials, including the secretary of defense and national security adviser. Personal memos, meeting minutes, and private diaries are peppered with references to these reports' gruesome conclusions.⁹³ These estimates may not have always modeled the fatality or damage assessments for every specific scenario, but they nonetheless represent the most detailed predictions produced by the U.S. government about what would happen in a nuclear exchange, and there is evidence that they informed the views of U.S. officials.

Based on an analysis of declassified U.S. documents and recent work by Caroline Reilly Milne, I compiled data on expected U.S. and Soviet fatalities.⁹⁴ I began with Milne's data and supplemented these with additional estimates from a variety of sources, including the *Foreign Relations of the United States*, materials available through the presidential libraries, the Central Intelligence Agency's Freedom of Information Act Reading Room, and documents released through the National Security Archive. Table A1 in the online appendix shows the number of fatalities in a nuclear exchange between the United States and the Soviet Union as estimated by U.S. government assessments.⁹⁵ For each assessment, I recorded the year it was published and the estimated range of U.S. and Soviet fatalities. For assessments that include estimates of both U.S. and Soviet fatalities, I then recorded—at the low, mean, and high estimates—which of the two states was predicted to “win” a large-scale nuclear exchange by suffering the fewest casualties.⁹⁶

Entity Portrayed Terrible Costs of Nuclear War,” Electronic Briefing Book No. 480, July 22, 2014, National Security Archive, George Washington University, Washington, D.C., <https://nsarchive2.gwu.edu/nukevault/ebb480/>.

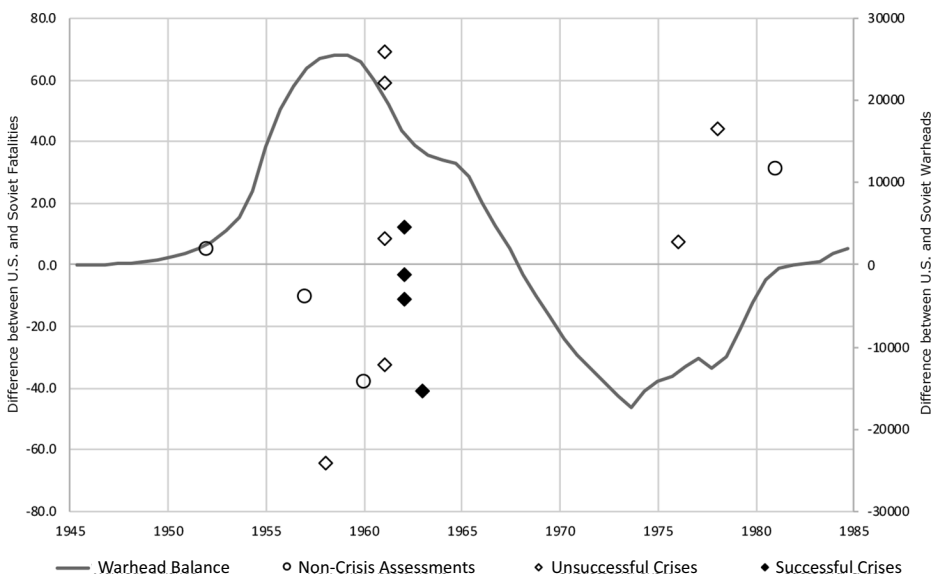
93. Milne, “Hope Springs Eternal,” pp. 24–121.

94. For illustrative excerpts from some of these reports, especially those not captured in Milne, “Hope Springs Eternal,” see documents B1 through B5 in the online qualitative appendix.

95. These data are taken from Milne, “Hope Springs Eternal,” with some supplementation from primary sources. See the online qualitative appendix.

96. Simply reporting the range of the fatality estimates necessarily simplifies the predictions of fa-

Figure 3. U.S.-Soviet Crisis Outcomes and Nuclear Exchange Fatality Estimates



NOTE: For each nuclear exchange report that produced estimates of both U.S. and Soviet fatalities, figure 3 plots the difference between the mean expected number of U.S. fatalities and Soviet fatalities. Icons indicate if a report was conducted in a year adjacent to a year in which the U.S. and Soviet Union experienced a crisis (diamonds), or adjacent to a year in which they did not experience a crisis, according to the Kroenig dataset (circles). Figure further depicts whether the United States succeeded (solid diamonds) or failed (hollow diamonds) to achieve its goals in each crisis. Figure also plots the difference between the number of warheads in the U.S. and Soviet arsenals.

Next, I assessed whether crisis outcomes were associated with the nuclear balance or fatality estimates. For each assessment with both U.S. and Soviet fatality estimates, I calculated the difference between mean U.S. and Soviet fatalities.⁹⁷ I then plotted each estimate in figure 3. For each estimate, I also indicated whether a crisis with the Soviet Union occurred in the same year or

talities. When estimates provide a range of possible fatalities, I produce a “mean” estimate from the highest and lowest estimates. For reports that only produce one estimate of fatalities (e.g., Soviet fatalities in the 1957 NESC assessment), I use the same number for low, mean, and high figures. Even the fatality estimate ranges are imperfect measures because they sometimes report results across different simulated exchanges, which might make some inappropriate comparisons. Both Milne and many of the original estimates do not provide detailed breakdowns of the different exchanges and report overall ranges. Still, within a given estimate, different scenarios reported different outcomes about which state would suffer more. For the 1957 NESC assessment, see Milne, “Hope Springs Eternal,” p. 45.

97. I use the mean values provided by the estimates.

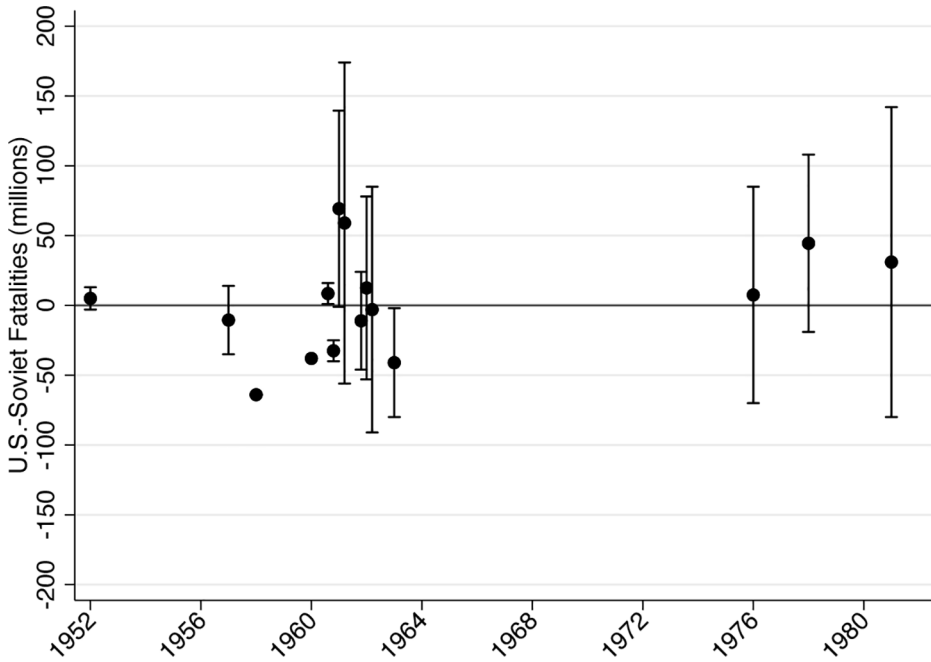
the year following the assessment. For assessments that were proximate to nuclear crises, I then marked whether the United States succeeded (solid diamonds) or failed (hollow diamonds) to achieve its goals in the crisis.⁹⁸ On the same figure, I also plotted the difference between the number of warheads in the U.S. and Soviet arsenals.

Two findings emerge. First, there is some evidence of a relationship between expected damage and crisis outcomes, though the data is sparse. The United States succeeded in two of the three crises in which assessments suggested it would suffer fewer fatalities than the Soviet Union. By comparison, it succeeded in none of the three crises in which proximate assessments predicted it would suffer more fatalities than the Soviet Union. These data might be interpreted as weak evidence in favor of the underlying assumption of nuclear superiority, which is that states expected to suffer less damage will be more likely to prevail in interstate bargaining.

Second, and most significantly, however, there appears to be little if any relationship between the warhead balance and expectations of relative damage. For example, the late 1950s to the mid-1960s was a period of considerable stockpile superiority for the United States, during which it possessed at least 20,000 more nuclear weapons than the Soviet Union. Yet the estimates produced during this period show little agreement on which side would suffer more fatalities, as several estimates even predicted that the United States would suffer more than the Soviet Union. In the two years prior to the Cuban Missile Crisis, the U.S. government created at least six separate assessments of nuclear exchange outcomes. In 1960, when the United States possessed roughly 25,000 more warheads than the Soviet Union, the NESAC estimated that in such an exchange, the United States would suffer 61 million fatalities compared to 99 million Soviet Union fatalities. The next year, five assessments were produced in the U.S. government, one by the NESAC, two by the Joint Chiefs of Staff, and two by the Office of the Assistant Secretary of Defense for International Security Affairs (ISA). These reports produced widely varied damage estimates. The most optimistic of these assessments, produced by the Joint Chiefs of Staff, estimated that the United States would suffer 32 million fewer fatalities than the Soviet Union. By contrast, the two ISA assessments produced that year estimated that the United States would suffer 60–70 million more fatalities than the Soviet Union. Fatality estimates varied significantly even within individual assessments, with one ISA assessment predicting that

98. As determined by the International Crisis Behavior dataset. Michael Brecher and Jonathan Wilkenfeld, *A Study of Crisis* (Ann Arbor: University of Michigan Press, 2000).

Figure 4. Difference between U.S. and Soviet Nuclear Exchange Fatality Estimates



NOTE: For each of the fifteen nuclear exchange reports that produced estimates of both U.S. and Soviet fatalities, this figure plots the difference between the median expected number of U.S. fatalities and Soviet fatalities. Error bars extend to the highest expected differences favorable to the United States (lower on the figure) and the Soviet Union (higher on the figure).

U.S. fatalities could range anywhere from “negligible” to 140 million.⁹⁹ These estimates, issued within nearly a year of one another, all analyzed an essentially unchanging material balance. And yet they predicted dramatically different outcomes.

Another way of appreciating the variability in U.S. expectations about a nuclear exchange is to directly consider the variation within and between estimates. For each of the fifteen nuclear exchange reports that produced estimates of both U.S. and Soviet fatalities, figure 4 plots the mean ex-

99. I assume that “negligible” means close to or at zero. Other assessments offered similarly low predictions. A separate Department of Defense, International Security Affairs assessment conducted in 1961 estimated that U.S. fatalities might be as “low” as 12 million, while a Defense Department report conducted the following year estimated there might be “only” 4 million. Office of the Assistant Secretary of Defense for International Security Affairs (1961), August 14, Series 08,

pected number of U.S. fatalities minus the mean number of expected Soviet fatalities.¹⁰⁰ Error bars extend to the highest expected differences favorable to the United States (lower on the figure) and the Soviet Union (higher on the figure). As the figure illustrates, the expected relative fatalities varied significantly at the levels of both individual assessment and the year in which assessments were produced. In many cases, it was not even clear which side would suffer more fatalities.

This variation is driven by a range of assumptions about how such a nuclear exchange would occur: which side would strike first, the size and configuration of each side's nuclear forces, targeting approaches, the degree to which nuclear forces are hardened, the effectiveness of civil defense procedures, the accuracy of early warning capabilities, the proximity of population centers, and (literally) which direction the winds will blow. All these factors are subject to uncertainty, and many are hidden by the adversary. These gaps in knowledge are filled by the assumptions, perceptions, and beliefs of leaders. The entire process is then compounded by parallel processes taking place on the other side. In this way, the nuclear balance and its meaning were often what observers made of them.

POLITICAL OBSTACLES TO SOVIET NUCLEAR ASSESSMENTS

Analysis within the Soviet Union of the relative nuclear balance was similarly shaped by its leaders' perceptions and beliefs, first, as political ideology suppressed efforts to conduct "objective" studies of the results of nuclear war and, later, as beliefs about the fragility of command-and-control structures prevented the Soviets from enjoying the fruits of their quantitative superiority. As shown previously, the U.S. intelligence and defense communities had developed various (and sophisticated) models for measuring and comparing nuclear capabilities. By comparison, independent assessments of the strategic balance did not appear to affect Soviet decision-making until decades into the Cold War. Until the late 1960s and early 1970s, the Soviet Union possessed practically no means of systematically measuring the relative nuclear balance.¹⁰¹ Throughout this period, nuclear exchange models were often discounted

Carl Kaysen Files, Box 374, Military Policy, SAC, NSF, John F. Kennedy Presidential Library, cited in Milne, "Hope Springs Eternal," p. 77; and Summary of Population Fatalities from Nuclear War in 1966 (in Millions), U.S. Department of Defense, February 17, 1962.

100. To facilitate discrimination, data points for reports produced in the same year are slightly offset horizontally.

101. Levy, *Soviet Strategic Nuclear Measures of Effectiveness*, pp. 1–3; and Milne, "Hope Springs Eternal."

because to accept their conclusions of a Soviet defeat would have been to undermine the prevailing ideology at the time.¹⁰² According to Marshal Serge F. Akhromeyev, who rose through the Soviet General Staff and eventually served as Mikhail Gorbachev's national security adviser, "modeling and analysis did contribute" to operational and strategic planning, but that this "was more true in the mid-1970s and later. Many other factors, however, went into such decisions."¹⁰³ According to Milne, whose recent work examines Soviet perceptions of the strategic balance, "credible calculations of the effects associated with large-scale nuclear exchanges did not make an impression on the decision-making process until the early 1980s; prior to this point models of nuclear war were either nonexistent or misrepresented to make the costs seem more palatable."¹⁰⁴

When the Soviets eventually did formally assess the balance, their assessments incorporated more than just warhead totals. According to one senior Soviet official intimately involved with nuclear weapons issues, "nuclear power [*iadernaia moshch*] in our assessments is a function of yield, nuclear weapons, and accuracy."¹⁰⁵ Similarly, a RAND report produced at the end of the Cold War identified four distinct Soviet measures of effectiveness for assessing nuclear capabilities.¹⁰⁶ The first, Quantitative Correlation of Nuclear Forces, measured the relative numbers of similar types of weapons in states' arsenals.¹⁰⁷ The second, Equivalent TNT Correlation of Nuclear Forces, adjusted arsenal size to account for the yield of the weapons.¹⁰⁸ A third, the Anureyev Correlation of Nuclear Forces, incorporated the arsenals' technical features as well as assumptions about how they would be used.¹⁰⁹ Finally, the Destruction Potential Correlation of Nuclear Forces sought to explicitly model the effects of a nuclear exchange with the United States.¹¹⁰ Colonel General Andrian A. Danilevich, who led the Soviet group in charge of strategic and operational planning from 1977 to 1986, recognized the poverty of warhead-

102. Milne, "Hope Springs Eternal," pp. 122–161; and John G. Hines, Ellis M. Mishulovich, and John F. Shull, *Soviet Intentions 1965–1985: Volume I: An Analytical Comparison of U.S.-Soviet Assessments during the Cold War* (McLean, Va.: BDM Federal, 1995), pp. 25–27.

103. Interview with Sergei F. Akhromeev in Hines, Mishulovich, and Shull, *Soviet Intentions 1965–1985*, p. 4. See document C1 in the qualitative appendix.

104. Milne, "Hope Springs Eternal," p. 122.

105. Interview with Vitalii Leonidovich Kataev in Hines, Mishulovich, and Shull, *Soviet Intentions 1965–1985*, p. 100. See document C4 in the online qualitative appendix.

106. Levy, *Soviet Strategic Nuclear Measures of Effectiveness*, pp. v–vii. See document C5 in the online qualitative appendix.

107. *Ibid.*, pp. 19–20.

108. *Ibid.*, pp. 21–23.

109. *Ibid.*, pp. 23–27.

110. *Ibid.*, pp. 29–31.

counting, saying, “analysis of quantity alone provides only half of the analytical picture. Because of qualitative deficiencies, one side could have a tenfold quantitative advantage and still be behind.”¹¹¹

Unfortunately, there are few details on the findings of those damage assessments. But at least two strands of evidence provide some insight into Soviet assessments and how they impacted decision-making. Together, they show that, similar to the U.S. experience, Soviet perceptions and reactions to the material balance were highly subjective.

First, even when Soviet assessments of the relative nuclear balance concluded that the Soviet Union enjoyed numerical superiority across various dimensions, Soviet officials still did not believe that they enjoyed overall superiority because of the weakness of their command-and-control arrangements. The Soviets actively sought strategic superiority throughout much of the Cold War.¹¹² The Soviet Union eventually did perceive superiority in several areas, including number of launchers, silo protection, yield of warheads, and range and power of missiles.¹¹³ Yet this numerical superiority did not translate into any meaningful strategic superiority from the Soviet perspective. For instance, a report from the Soviet General Staff concluded that command-and-control infrastructure was so fragile that “after sustaining an all-out nuclear strike the Soviets would be able to launch only 2% of their missiles.”¹¹⁴ As late as 1982, Soviet officials were still bemoaning their inability to validate the survivability of either their nuclear forces or their related command-and-control systems. A March 1982 report from the head of the Central Committee’s Defense Industry Department to the Chief of the General Staff and other senior defense officials argues that a combination of bureaucratic hassles, financial shortfalls, and technical challenges meant that “notwithstanding the importance of carrying out the verification of the stability of various forms of military equipment, until now the general five-year [verification] plan has not worked out.”¹¹⁵ A separate high-level report issued in February that year by the Central Committee’s Defense Industry Department observed the Soviets’

111. Interview with Andrian A. Danilevich in Hines, Mishulovich, and Shull, *Soviet Intentions 1965–1985*, p. 22. See document C3 in the online qualitative appendix.

112. Hines, Mishulovich, and Shull, *Soviet Intentions 1965–1985*, pp. 2–13.

113. Interview with A.S. Kalashnikov in Hines, Mishulovich, and Shull, *Soviet Intentions 1965–1985*, pp. 89–90. See document C2 in the online qualitative appendix.

114. *Ibid.*

115. “On Work Checking the Resistance of Military Equipment to the Destructive Factors of a Nuclear Explosion,” Vitalii Leonidovich Kataev Papers, February 1982, document 1, folder 8, box 5, Hoover Institution, Stanford University, Stanford, California (hereafter, Hoover Institution). See document C6 in the online qualitative appendix.

failure to evaluate and ensure their nuclear command-and-control systems, noting that “a comprehensive comparative assessment of the necessity and sufficiency of the work ongoing in the country with the goal of ensuring control of nuclear forces in war (especially in conditions of nuclear strike) has not yet been conducted. This does not make it possible to identify weak links in the systems being created or to eliminate shortcomings.”¹¹⁶

Without the ability to launch the weapons, their greater numbers were useless. According to one former senior Soviet defense official, “Soviet superiority in the number of launchers did not give them any real advantage. This numerical superiority reflected a mechanistic, wasteful approach to force building.”¹¹⁷ This sense of vulnerability persisted late into the Cold War. By 1988, the Soviet Union enjoyed superiority over the United States in total stockpile, megatonnage, and strategic launchers. Yet a high-level memo within the defense industry noted that “current Strategic Rocket Forces are capable of striking 80 enemy targets in a retaliatory strike, by 1995, 100 targets and by 2000, 150 targets, slightly below the calculated level necessary to fulfill the assigned mission of a retaliatory strike-200 targets.”¹¹⁸ Though the memo later concludes that “mobile missiles are an effective means for a retaliatory strike from now until 2000,” its conclusions illustrate the weaknesses of only examining the nuclear balance’s static indicators and not considering how those indicators were interpreted by states themselves.

Second, even when predicting that the United States would suffer relatively more damage in a nuclear exchange, Soviet officials did not embrace the supposed benefits of nuclear superiority. The Soviet Union conducted three exercises in the early 1970s aimed at predicting the consequences of a large-scale nuclear exchange with the United States.¹¹⁹ At the time, the Soviet Union had obtained a moderate lead in overall megatonnage, had reached complete parity in number of strategic launchers, and faced moderate inferiority in total number of nuclear weapons. Despite the somewhat mixed strategic balance, the models depicted devastating consequences for the Soviet Union. Andrian

116. “On the Assessment of the Capabilities of Controlling the Country’s Nuclear Forces in Hostile Conditions,” Vitalii Leonidovich Kataev Papers, February 1982, document 8, folder 8, box 5, Hoover Institution. See document C7 in the online qualitative appendix.

117. Interview with A.S. Kalashnikov in Hines, Mishulovich, and Shull, *Soviet Intentions 1965–1985*, p. 90 (document C2).

118. “Reference: On Mobile Basing of Missile Systems,” Vitalii Leonidovich Kataev Papers, n.d., folder 8, box 5, Hoover Institution. The memo is undated and unsigned but was likely produced around 1987–1988 by either the Central Research Institute of Machine Building or the 4th Institute of the Ministry of Defense. I am grateful to Pavel Podvig for his help with this source.

119. Interview with Andrian A. Danilevich in Hines, Mishulovich, and Shull, *Soviet Intentions 1965–1985*, p. 27 (document C3).

Danilevich, who helped oversee strategic nuclear planning in the Soviet Union at that time, described the results of the final exercise: “We explained our conclusions that after the strike the Armed Forces would be reduced to 1/1,000 of their previous strength; 80 million citizens would be dead; 85% of the industrial capability of the Soviet Union would be destroyed; the European part of the USSR would be contaminated by radiation at extremely lethal levels of 3,000 roentgens. Given all of this, *the consequences of a retaliatory strike against the U.S. would be even more lethal to that country*” (italics added).¹²⁰

But despite the apparent assumption that the Soviets enjoyed relative “superiority” in the expected damage from a large-scale exchange with the United States, the assessment shocked senior officials. According to Danilevich, “Brezhnev and Kosygin were visibly terrified by what they heard. During [an] exercise three launches of ICBMs with dummy warheads were scheduled. Brezhnev was provided with a button in the exercise and was to ‘push the button’ at the appropriate time When the time came to push the button, Brezhnev was visibly shaken and pale and his hand trembled and he asked Grechko several times for assurances that the action would not have any real-world consequences.”¹²¹ Testimonial evidence from a range of Soviet officials confirms that from at least the late 1960s onward, the Soviet Union concluded that it could not meaningfully “win” a nuclear war.¹²² Despite at times believing that the Soviet Union would emerge from a large-scale nuclear exchange relatively better off than the United States, officials were still loathe to invite the catastrophic damage of a nuclear war.

THE NUCLEAR BALANCE IS WHAT LEADERS MAKE OF IT

To assert that the nuclear balance matters for interstate politics is to make three underlying claims. First, states must have sufficiently accurate (or at least symmetric) information about how the various technical-military components that make up the nuclear balance are distributed. Second, assuming accurate (or symmetric) information about the material distribution, they must have shared beliefs about which of those components matter and how they aggregate to construct the overall nuclear balance. Third, assuming states agree on the nuclear balance’s overall condition, they must have shared beliefs about what that balance means strategically and politically.

As research strands on intelligence failures and the role of beliefs have

120. Ibid.

121. Ibid.

122. Hines, Mishulovich, and Shull, *Soviet Intentions 1965–1985*, p. 26.

demonstrated, each of these claims may be undermined in several ways. Bureaucratic pathologies, inflexible standard operating procedures, or normal intelligence failures may confound efforts to gain accurate or symmetric information about nuclear capabilities.¹²³ Once state intelligence agencies develop nuclear capability estimates, those estimates must be interpreted. Those interpretations can be shaped significantly by individual- or organizational-level traits such as military experience, political ideology, standard operating procedures, institutional biases, and others.¹²⁴

A rich body of research demonstrates how beliefs matter, including in the nuclear domain. Elizabeth Saunders, for instance, develops a framework that explains how the salience of leader beliefs can vary with both the perceived external threat and the nature of the state's domestic politics.¹²⁵ Rachel Elizabeth Whitlark shows how decisions to "consider and use preventive force rests not only on material factors but more importantly on a leader's prior beliefs about nuclear proliferation and the threat posed by a specific adversary" (italics added).¹²⁶ In survey data, Scott Sagan and Benjamin Valentino show how the U.S. public's preferences for a nuclear strike vary with partisan affiliation, gender, and stance on the death penalty.¹²⁷ Similarly, Joseph Ripberger, Hank Jenkins-Smith, and Kerry Herron show that cultural orientations are associated with systematically different beliefs about the importance of retaining nuclear weapons.¹²⁸ Rose McDermott, Anthony Lopez, and Peter Hatemi reveal how the very functioning of deterrence depends on the psychology of revenge.¹²⁹ Fiona Cunningham and M. Taylor Fravel provide evidence that

123. Robert Jervis, *Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War* (Ithaca, N.Y.: Cornell University Press, 2010); and Joshua Rovner, *Fixing the Facts: National Security and the Politics of Intelligence* (Ithaca, N.Y.: Cornell University Press, 2011).

124. Graham T. Allison, "Conceptual Models and the Cuban Missile Crisis," *American Political Science Review*, Vol. 63, No. 3 (September 1969), pp. 689–718, <https://doi.org/10.2307/1954423>; Roger Z. George and James B. Bruce, eds., *Analyzing Intelligence: Origins, Obstacles, and Innovations* (Washington, D.C.: Georgetown University Press, 2008); Richard K. Betts, *Enemies of Intelligence: Knowledge and Power in American National Security* (New York: Columbia University Press, 2009); and Jervis, *Why Intelligence Fails*.

125. Saunders, "The Domestic Politics of Nuclear Choices," pp. 162–163.

126. Whitlark, "Nuclear Beliefs," p. 545.

127. Scott D. Sagan and Benjamin A. Valentino, "Revisiting Hiroshima in Iran: What Americans Really Think about Using Nuclear Weapons and Killing Noncombatants," *International Security*, Vol. 42, No. 1 (Summer 2017), pp. 62–67, https://doi.org/10.1162/ISEC_a_00284.

128. Joseph T. Ripberger, Hank C. Jenkins-Smith, and Kerry G. Herron, "How Cultural Orientations Create Shifting National Security Coalitions on Nuclear Weapons and Terrorist Threats in the American Public," *PS: Political Science and Politics*, Vol. 44, No. 4 (October 2011), pp. 715–719, <https://www.jstor.org/stable/41319956>.

129. Rose McDermott, Anthony C. Lopez, and Peter K. Hatemi, "Blunt Not the Heart, Enrage It': The Psychology of Revenge and Deterrence," *Texas National Security Review*, Vol. 1, No. 1 (November 2017), <http://hdl.handle.net/2152/63934>.

Chinese and U.S. strategists appear to hold systematically different views about the likelihood of nuclear use in a conflict or crisis.¹³⁰ Scholars identify leader experiences and psychology as key variables in explaining both nuclear proliferation and nuclear reversal.¹³¹

Debates about the influence of the nuclear balance on the Cuban Missile Crisis illustrate how individual reactions to the same perceived balance can vary significantly. Even within the highest levels of the United States government, individuals possessing the same information about the nuclear balance and occupying similar positions reported different views of what the balance meant.¹³² In the midst of the crisis, Secretary of State Dean Rusk suggested superiority might matter, saying that “one thing Mr. Khrushchev may have in mind is that he knows that we have a substantial nuclear superiority, but he also knows that we don’t really live under fear of his nuclear weapons to the extent that he has to live under fear of ours.”¹³³ Similarly, Chairman of the Joint Chiefs of Staff General Maxwell Taylor argued in a memo that “we have the strategic advantage in our general war capabilities. . . . This is no time to run scared.”¹³⁴ Others, however, voiced concerns about escalation that belied much faith in the benefits of superiority. For instance, responding to claims about the advantages of U.S. superiority, Secretary of Defense Robert McNamara replied, “Look, you’re probably right that if we had to fight a war with the Soviet Union, we’d have fewer casualties today than if we had to do it later. But it’s not clear that we have to fight them. So for God’s sake, let’s try to avoid it.”¹³⁵ Kennedy himself seemed to express different views at different times. At some points in the crisis, he appeared to endorse the logic of preemption using U.S. superiority, but he later seemed to doubt the supposed benefits

130. Fiona S. Cunningham and M. Taylor Fravel, “Dangerous Confidence? Chinese Views on Nuclear Escalation,” *International Security*, Vol. 44, No. 2 (Fall 2019), pp. 61–109, https://doi.org/10.1162/isec_a_00359.

131. Hymans, *The Psychology of Nuclear Proliferation*; Matthew Fuhrmann and Michael C. Horowitz, “When Leaders Matter: Rebel Experience and Nuclear Proliferation,” *Journal of Politics*, Vol. 77, No. 1 (January 2015), pp. 72–87, <https://doi.org/10.1086/678308>; and Mehta, *Delaying Doomsday*.

132. For recent competing interpretations, see Mark S. Bell and Julia Macdonald, “How to Think about Nuclear Crises,” *Texas National Security Review*, Vol. 2, No. 2 (February 2019), <http://dx.doi.org/10.26153/tsw/1944>; Kroenig, *The Logic of American Nuclear Strategy*, pp. 81–113; and Brendan Rittenhouse Green et al., “Contrasting Views on How to Code a Nuclear Crisis,” *Texas National Security Review*, Vol. 2, No. 4 (August 2019), <http://dx.doi.org/10.26153/tsw/6663>.

133. “White House Tapes and Minutes of the Cuban Missile Crisis: Introduction to Documents,” *International Security*, Vol. 10, No. 1 (Summer 1985), p. 177, <https://doi.org/10.2307/2538794>.

134. John Lewis Gaddis, *Strategies of Containment: A Critical Appraisal of American National Security Policy during the Cold War*, rev. ed. (New York: Oxford University Press, 2005), p. 228n.

135. David Talbot, *Brothers: The Hidden History of the Kennedy Years* (New York: Free Press, 2007), p. 13.

of superiority by saying, “What difference does it make? They’ve got enough to blow us up now anyway.”¹³⁶

It remains unclear what shapes individual beliefs about the nuclear balance or how those beliefs aggregate to state action. In the Cuban Missile Crisis, it appears that President Kennedy’s own skepticism about the relevance of the balance won out in the end.¹³⁷ But what, if anything, accounts for Kennedy’s and other senior leaders’ varying views? And why did Kennedy’s superiority skepticism ultimately prevail?¹³⁸

The literature suggests some propositions about when different views matter more. In general, as in the Cuban Missile Crisis, the top leader’s views are likely to dominate. Nuclear use of authority is typically concentrated in only the head of state or, at most, a few high-level leaders.¹³⁹ But the importance of different elite beliefs may vary with state-level characteristics, such as nuclear command-and-control arrangements. For instance, in states with highly delegative command-and-control arrangements, the views of a wider range of military leaders will be more important.¹⁴⁰ Conversely, in states with highly assertive arrangements, civilian views are likely to dominate. In more democratic states, public opinion about nuclear use might serve as either a constraint or push to state policy, which might not be present in less democratic states.¹⁴¹ In states with highly experienced leaders, only the head of state’s views may win out, while for states with relatively inexperienced leaders, senior advisers’ beliefs would likely carry more weight.¹⁴²

At the individual level, these beliefs may vary with demographic characteristics. For example, leaders with rebel experience are more likely to “value the potential benefits of possessing nuclear weapons” and might, therefore, be more likely to use them under conditions of superiority.¹⁴³ Similarly, surveys

136. “White House Tapes and Minutes of the Cuban Missile Crisis,” 1985, p. 148.

137. Bell and Macdonald, “How to Think about Nuclear Crises.”

138. The ultimate dominance of Kennedy’s views might be explained by his position of authority or by the president’s experience relative to his advisers. See Elizabeth N. Saunders, “No Substitute for Experience: Presidents, Advisers, and Information in Group Decision Making,” *International Organization*, Vol. 71, No. S1 (2017), pp. S227, S242–S243, <https://doi.org/10.1017/S002081831600045X>. For more on the “aggregation problem” in international relations, see Ross James Gildea, “Psychology and Aggregation in International Relations,” *European Journal of International Relations*, Vol. 26, No. S1 (September 2020), pp. 166–183, <https://doi.org/10.1177%2F1354066120938830>.

139. See Alex Wellerstein, “NC3 Decision Making: Individual versus Group Process,” *NAPSNet*, August 8, 2019, <https://nautilus.org/napsnet/napsnet-special-reports/nc3-decision-making-individual-versus-group-process/>.

140. Giles David Arceneaux, “Beyond the Rubicon: Command and Control in Regional Nuclear Powers,” Ph.D. dissertation, Syracuse University, New York, 2019.

141. Sagan and Valentino, “Revisiting Hiroshima in Iran.”

142. Saunders, “No Substitute for Experience.”

143. Fuhrmann and Horowitz, “When Leaders Matter,” p. 74.

have shown that military experience is correlated with elite views about the effectiveness of military force, support for placing political and diplomatic constraints on the use of military power, and sensitivity to casualties.¹⁴⁴ These views might suggest that leaders with military experience are more likely to consider nuclear use and may view the nuclear balance in ways comparable to the conventional balance.¹⁴⁵ Together, these observations suggest some propositions about which beliefs matter when and point to areas for future research.

Conclusion

There has been a long and rich debate concerning the importance of nuclear superiority, including our ability to measure it. Recent contributions have helped advance that debate, but there are still areas for improvement. To claim that the nuclear balance matters is to claim that states have a meaningful, coherent, and stable conception of what constitutes the balance, how it can be measured, and why it matters. States must have a clear and shared understanding of the balance's component features, which components matter, and which do not. They must have accurate information about how those components are distributed between themselves and the adversary, and they must have a coherent theory of how that distribution generates political effects. But if states conceive of, perceive, measure, and respond to the nuclear balance in ways that are not fully captured by the "objective" technical-military features, analyses that rely on those features will necessarily fall short.

The argument presented here—that there is a crucial subjective component to how states perceive and respond to the balance—is not entirely new. Though he may not have employed the language of constructivism, Jervis recognized how the meaning of the nuclear revolution could be an intersubjective creation of nuclear-armed states.¹⁴⁶ He observed that propositions about the

144. Peter D. Feaver and Christopher Gelpi, *Choosing Your Battles: American Civil-Military Relations and the Use of Force* (Princeton, N.J.: Princeton University Press, 2011), pp. 21–63, 95–148. Peter Feaver and Christopher Gelpi are careful to note that, although there are differences between civilian-military and military-elite views about the use of force, "they are counterbalanced by a consensus that emerges if one looks only at absolute numbers" (p. 49).

145. For example, historians of the Cuban Missile Crisis note that Curtis LeMay, Chief of Staff of the U.S. Air Force, "rejected the notion that [nuclear weapons] were somehow special, morally or otherwise. 'The assumption seems to be,' he wrote, 'that it is much more wicked to kill people with a nuclear bomb, than to kill people by busting their heads with rocks,'" as quoted in Ernest R. May and Philip D. Zelikow, eds., *The Kennedy Tapes: Inside the White House during the Cuban Missile Crisis* (New York: W.W. Norton, 2002).

146. Nina Tannenwald, "The Meaning of the Nuclear Revolution: 75 Years of Non-Use," *Book Re-*

utility of nuclear superiority “rest on decision makers’ beliefs—beliefs, furthermore that can be strongly influenced by American policy and American statements. Although the Russians stress war-fighting ability, they have not contended that marginal increases in strategic forces bring political gains; any attempt to do so could be rendered less effective by an American assertion that this is nonsense.”¹⁴⁷

The material balance matters. If the bombs go off, the destruction will be painfully real. Even prior to nuclear use, the nature of the balance may constrain the role of information and beliefs, suggesting scope conditions for the argument presented here. That is, when the material balance is either highly symmetric or highly uncertain, there will be more space for information and beliefs. Conversely, when the balance is asymmetric and certain, information and beliefs may matter less.¹⁴⁸ Similarly, the importance of incomplete information may vary, gaining more significance when the material balance is roughly equal, when states configure their nuclear arsenals in ways that obscure their features, or when states suffer from serious intelligence failures. Even when the material balance appears highly and definitively asymmetric, however, beliefs may still influence perceptions of whether the inferior state possesses a survivable second strike. And, regardless of the perceived balance, leaders and states may still hold different beliefs about whether and how that balance matters. In the end, the material balance is not everything. Through their assumptions, perceptions, and beliefs, states create and shape both the balance and, crucially, its meaning.

These findings have important theoretical and methodological implications. Theoretically, they suggestively support one of the central tenets of the “nuclear revolution,” that nuclear superiority does not convey political benefits. This is significant because much recent work challenges some of the nuclear revolution’s principles.¹⁴⁹ It may also be the case that, as Mark Bell and Julia

view Roundtable: *The Meaning of the Nuclear Revolution 30 Years Later*, *Texas National Security Review*, April 30, 2020, <https://tnsr.org/roundtable/book-review-roundtable-the-meaning-of-the-nuclear-revolution-30-years-later/>.

147. Robert Jervis, “Cooperation under the Security Dilemma,” *World Politics*, Vol. 30, No. 2 (January 1978), pp. 209–210, <https://doi.org/10.2307/2009958>.

148. Saunders, “The Domestic Politics of Nuclear Choices,” pp. 162–164.

149. See, for example, Austin Long and Brendan Rittenhouse Green, “Stalking the Secure Second Strike: Intelligence, Counterforce, and Nuclear Strategy,” *Journal of Strategic Studies*, Vol. 38, Nos. 1–2 (2015), pp. 38–73, <https://doi.org/10.1080/01402390.2014.958150>; Brendan Rittenhouse Green and Austin Long, “The MAD Who Wasn’t There: Soviet Reactions to the Late Cold War Nuclear Balance,” *Security Studies*, Vol. 26, No. 4 (2017), pp. 606–641, <https://doi.org/10.1080/09636412.2017.1331639>; and Mark S. Bell, “Nuclear Opportunism: A Theory of How States Use Nuclear Weapons in International Politics,” *Journal of Strategic Studies*, Vol. 42, No. 1 (2019), pp. 3–28, <https://doi.org/10.1080/01402390.2017.1389722>.

Macdonald argue, the inconclusive empirical findings about nuclear superiority's political benefits emerge from a failure to appreciate the variation in nuclear crises.¹⁵⁰ This may be particularly true if states are more likely to select into certain nuclear crises based on the risks that they present. Future research may also investigate, at either the individual or the state-level, potential systematic variation in beliefs about how to assess the nuclear balance and the balance's meaning.

Methodologically, it suggests avenues for improving research on the implications of the nuclear balance. The impact of the nuclear balance is likely to be highly contingent on the information available to states in times of crisis and their beliefs about what that information means for their odds of success.¹⁵¹ If state beliefs about nuclear superiority's importance (or unimportance) can be learned and unlearned, this may not be captured by large-N methods.¹⁵² If state responses to nuclear superiority vary with the state's political leadership, this may not be captured.¹⁵³ In these ways, the nuclear balance may be, to a large extent, what states make of it.¹⁵⁴ This principle applies to not only strategic nuclear balance measures but also any attempts at "objectively" measuring capability and power.¹⁵⁵ Quantitative work can more fully establish that the operationalizations scholars choose best model the dynamics that they are investigating or that the results are robust across operationalizations. Future work might use the measures captured in the Nuclear Capabilities Dataset. Scholars might also use qualitative historical work to inform their quantitative coding decisions. Typically, this might be unreasonable with large-N work given the large number and variety of observations. But because of the relatively small number of observations involved in studies of nuclear weapons states, it may be possible to use more qualitative work to construct data on the nuclear balance. This is an area in which mixed methodological approaches may be fruitful.

Research on nuclear superiority also has important policy implications. As

150. Bell and Macdonald, "How to Think about Nuclear Crises."

151. Francis J. Gavin, *Nuclear Weapons and American Grand Strategy* (Washington, D.C.: Brookings Institution Press, 2020), pp. 49–74.

152. Joseph S. Nye Jr., "Nuclear Learning and U.S.-Soviet Security Regimes," *International Organization*, Vol. 41, No. 3 (Summer 1987), pp. 371–402, <https://www.jstor.org/stable/2706750>.

153. Campbell Craig and Sergey Radchenko, "MAD, not Marx: Khrushchev and the Nuclear Revolution," *Journal of Strategic Studies*, Vol. 41, Nos. 1–2 (2018), pp. 208–233, <https://doi.org/10.1080/01402390.2017.1330683>.

154. Jervis, "Cooperation under the Security Dilemma," pp. 208–210.

155. As William Wohlforth notes, "the relation of perceived power to material resources can be capricious." William C. Wohlforth, *The Elusive Balance: Power and Perceptions during the Cold War* (Ithaca, N.Y.: Cornell University Press, 1993), p. 306.

China expands and modernizes its nuclear forces, observers have debated whether U.S. nuclear superiority will matter vis-à-vis China and the value and prospects of arms control between the two states.¹⁵⁶ Some scholars have seized on evidence that nuclear superiority matters to lobby for a further expansion and modernization of the U.S. nuclear arsenal.¹⁵⁷ The recent body of research purporting to find the benefits of nuclear superiority is too small and too indeterminate. Further, as the replications presented here show, recent evidence for the significance of nuclear superiority remains tenuous. We will need more numerous and compelling works demonstrating the benefits of nuclear superiority before casting aside long-standing tenets of international relations theory and potentially placing U.S. policy on the path to arms racing once again.

156. David C. Logan, *Trilateral Arms Control: A Realistic Assessment of Chinese Participation* (Washington, D.C.: Stimson Center, 2021), <https://www.stimson.org/2021/trilateral-arms-control-a-realistic-assessment-of-chinese-participation/>.

157. Kroenig, *The Logic of American Nuclear Strategy*.