



INSTITUTE FOR DEFENSE ANALYSES

Understanding Conflicts in a More Proliferated World

Andrew J. Coe
Victor A. Utgoff

December 2008

Approved for public release;
distribution unlimited.

IDA Paper P-4426

Log: H 09-000123

This work was conducted under contract DASW01-04-C-0003, Task BB-6-2926, for the Office of Net Assessment. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

© 2009 Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882 • (703) 845-2000.

This material may be reproduced by or for the U.S. Government pursuant to the copyright license under the clause at DFARS 252.227-7013 (NOV 95).

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-4426

**Understanding Conflicts in
a More Proliferated World**

Andrew J. Coe
Victor A. Utgoff

PREFACE

This analysis was performed for the Office of Net Assessment, in the Office of the Secretary of Defense (OSD) by the Institute for Defense Analyses (IDA) under the task entitled “Understanding Conflicts in a More Proliferated World – Part 2.” The authors greatly appreciate the support of Mr. Andrew W. Marshall, Director of Net Assessment, in funding the analysis and in providing many useful suggestions as the work progressed. We are also grateful for the painstaking reviews of successive drafts of this paper by Brad Roberts, Katy Oh Hassig, Jeffrey McKittrick, General Larry Welch USAF (ret.), David Hockaday and Heather Williams.

This page is intentionally left blank.

CONTENTS

Preface.....	iii
Summary	S-1
1. Introduction.....	1
2. A Model of Limited Nuclear Warfare	7
A Model of Limited Nuclear War on the Korean Peninsula	10
Optimal Strategies for Limited Nuclear War.....	14
3. Conflict between the DPRK and a Perfect Alliance	19
Example Settlements.....	19
Damage Thresholds	20
Nominal Evolution of Example Conflict	22
A Case for Assuming Modest Escalations.....	24
Potential Implications of US/ROK Ballistic Missile Defenses.....	25
Potential Implications of US/ROK Preemptive Strike Forces.....	27
Possible Limitations on DPRK Situational Awareness	30
Concluding Observations.....	31
4. From Perfect Alliance to Strategic Relations.....	33
Sincere Alliance with Disagreement on Damage Thresholds.....	34
A Less than Sincere Alliance.....	39
Strategic Relations Rather Than Alliance.....	43
5. Potential Intervention by both the US and China	49
Initial War between the DPRK and an Independent, Nuclear-Armed ROK	50
Best Strategies for Interventions by the US and China.....	52
Potential Effects of DPRK and ROK Anticipation of Interventions	57
6. Generalizing to a More Proliferated World	61
7. General Comments and Conclusions.....	69

FIGURES

2.1	Bounds to conflict derive from players' value schemes and beliefs about each other	16
3.1.	Example of a DPRK vs. US/ROK Conflict	20
3.2	A Potential DPRK vs. US/ROK Alliance Conflict	24
3.3	Examples of DPRK vs. Alliance Conflict with Missiles Defenses and Preemption	27
4.1	A Potential Conflict between the DPRK and a US/ROK Alliance with Individual Damage Thresholds	32
5.1	Nuclear War between DPRK and an Independent ROK	43
5.2	Potential Subsequent War between US and DPRK	45

SUMMARY

This paper employs a new model of nuclear conflict to explore the potential nature of a more proliferated world. A fundamental postulate of this model is that it takes more than an arsenal of nuclear weapons to win an interest or defend it. A state also needs the capacity to absorb retaliations and escalations until its adversary is no longer willing to take more damage either to win the interest at stake or to defend it. A second fundamental postulate is that states and alliances would be better able to discipline the evolution of a nuclear war to a tolerable outcome with the least possible damage if they think beforehand about what the potential outcomes might be and establish how much damage they are willing to take to pursue each plausible outcome.

A major challenge for strategy in a more proliferated world is that it will become more likely that a nuclear war could involve more than two nuclear states. To understand such possibilities the paper examines a series of case studies of potential nuclear war on the Korean Peninsula that involve first two and then progressively more than two independent states. The first case examines a perfect ROK/US alliance, i.e., a unitary actor defending – as in all our cases – against a desperate DPRK gambling that it could gain control of the entire Peninsula by force. Three additional cases examine how successively looser security relationships between the ROK and the US might affect the conflict. In the last of these, a nuclear-armed ROK and the US have dissolved their alliance and defend independently against the DPRK, thus making the war essentially trilateral. A final case explores a quadrilateral war in that after an initial war between an independent ROK and the DPRK, the two Korean states face potential independent interventions by both the US and China.

The paper goes on to generalize some of the results of these case studies, concluding that coalitions of small nuclear states can defeat a much larger nuclear state and thus may be able deter its attacks. As a group they would sum not only their capabilities to impose nuclear damage on a larger state, but also their capabilities to absorb the damage the larger state can impose on them and keep fighting. So long as the group members impose at least commensurate damage for the total imposed on them, the larger state suffers the sum of the damages done by the members while they suffer only a share of the damage it imposes. The incentives of smaller states to form such coalitions

against even a larger nuclear aggressor and live up to their "obligations" appear strong. The formation or even potential formation of such defensive coalitions should have a substantial stabilizing effect on a more highly proliferated world.

The analysis presented below should be understood as a preliminary exploration of the issues addressed, and of the analytic methodology employed. The Office of Net Assessment has funded a follow-on study with the authors, which is producing three papers expanding upon this study. The first will spell out the potential policy implications for nuclear deterrence of adopting the perspectives and methodology presented in this analysis. The second will explain in less abstract terms what this analysis has to say about the prospective nature of conflicts in a more proliferated world. The third will develop a more formal and rigorous explanation of the game theoretic methodology we have employed in this work.

1. INTRODUCTION

It is not too hard to imagine that the proliferation of nuclear weapons to states around the world will continue. On the supply side, the raw ingredients from which fissile material is manufactured occur naturally all over the world. Advances in technology and long-term economic growth imply that the technological prerequisites will steadily become cheaper and easier to develop. Expertise can be obtained illicitly from former and current nuclear states; even rudimentary designs for nuclear weapons are available.

On the demand side, the present crop of pariah states seems likely to persist for the foreseeable future. For these states, nuclear weapons provide independent means to guarantee their survival and perhaps underwrite future aggression. North Korea and Iran are but the most recent examples—others might follow their lead. These states may expand the black market for the necessary technology, machinery, and expertise, accelerating the decline in supply-side barriers to proliferation. Moreover, the acquisition of nuclear arms by these states strengthens the incentives of others to do so, if only to deter aggression from neighboring rogues. Thus proliferation can encourage more proliferation.

Of course, the rest of the world may not watch idly as proliferation spirals on. The United States and other responsible stakeholders in the international community have strong reasons to attempt to prevent further proliferation and also a variety of powerful tools at their disposal for pursuing this objective. But each of these states also has other important objectives, not all of which are consistent with stopping nuclear proliferation. Each has proven willing to compromise the goal of non-proliferation in order to address other threats to its security; some have even done so for simple economic gain.

Calculating the balance among these opposing forces, easing and constraining proliferation, is no mean feat.¹ Governments therefore find it very difficult to determine

¹ The analysis presented here, of nuclear conflicts in a more proliferated world, should not be taken to imply that the authors believe that further proliferation is inevitable, or even likely. On the contrary, we have little confidence in any prediction of future proliferation. However, even if one thinks more proliferation is unlikely, it still behooves the US and other responsible stakeholders to consider the possibility and to attempt to understand the potential consequences. It is to this preparation that we wish to contribute.

the likely consequences for further proliferation of allowing, say, North Korea to retain its nuclear capability. This uncertainty can make it easy for even the most responsible of leaders to trade an uncertain, long-term increase in proliferation risk for more immediate and tangible reduction in other threats to security.

Even if governments could predict the future of proliferation, there is little concrete understanding of just how dangerous a more proliferated world might be. A line of research in political science, initiated by Kenneth Waltz and subsequently labeled “proliferation optimism”, suggests that such a world would be quite stable and would offer a high level of security for most states. But this theory has its critics: namely, Scott Sagan and other “proliferation pessimists”. The pessimists paint a quite different picture of a more proliferated world: one riven by instability and subject to occasional spasms of nuclear violence.²

These projections rely upon highly abstracted models of strategic interaction among nuclear states. Waltz’s model assumes that bilateral stability is guaranteed by an inescapable logic of absolute nuclear deterrence. From this, he concludes that all bilateral nuclear relationships will be at least as stable as the US-USSR relationship was, and that the addition of more nuclear states to the fold will only fortify the already ironclad deterrence of aggression. In this model, the apparent complexity introduced by adding more nuclear powers is in fact rendered irrelevant by the foolproof dictates of deterrence.

The proliferation pessimists have identified a host of potential problems with Waltz’s model. These critiques employ an understanding of the sources of apparently “irrational” behavior on the part of states. Psychological flaws in leaders and organizational flaws in bureaucracies can lead to large errors in national decision-making, sufficient to violate even Waltz’s simple logic of nuclear deterrence. These problems are exacerbated by the complexity introduced by the spread of nuclear weapons, making occasional nuclear wars possible and creating dangerous insecurity for many states.

No consensus has arisen between proliferation optimists and pessimists. Even if one does, it should offer little comfort to actual planners and strategists in the US defense community. The US might someday face a more proliferated world, one in which the possibility of conflicts involving more than two nuclear-armed states cannot be ignored. Those responsible for preparing for these conflicts cannot afford to rely on the

² We rely here on the exposition of these theories given in Scott D. Sagan and Kenneth N. Waltz, *The Spread of Nuclear Weapons: A Debate Renewed*, New York: W.W. Norton (2002).

simplifications of Waltz's model, even if the pathologies identified by the pessimists are overblown in theory or mitigated in practice. Prudence demands analysts remain skeptical.

Our previous work suggests that such skepticism may be well founded, even in the case of purely bilateral conflict between "rational" (i.e., not prone to errors in decision-making or organization) states.³ In a study of conflicts between the US and China over the status of Taiwan, we found that there are plausible situations in which both states perceive vital interests to be at stake. However, we also found that there could be great uncertainty in each state's assessment of its opponent's interests, a possibility dismissed in Waltz's theory. This uncertainty derives not from flaws in each state's decision-making, but instead from each side's incentives to obfuscate or exaggerate its interests prior to the conflict. In such conflicts, war may be the only way for each state to credibly communicate its interests and reliably probe those of its opponent.⁴ If the stakes are high enough, such wars could escalate to the use of nuclear weapons.

If such combinations of high stakes and high uncertainty can be expected to occur, then adding more nuclear-armed states to the world only expands the consequences of this omission in Waltz's theory. It introduces the possibility of third-party nuclear intervention in a previously bilateral conflict. The interests of this third party may be unclear to the first two combatants—does this increase or decrease the likelihood of nuclear escalation? The third party might seek to aid an ally in the original conflict; to take advantage of a weakened enemy; or to pursue some non-aligned purpose. How do these possibilities affect the choices made by the first two combatants? If the war reveals the true interests of the combatants, might the perceived commonality of interests that underwrote a prior alliance be eroded? What if there were multiple potential third parties? Might these states compete to influence the outcome of the original conflict?

This paper offers an exploratory study of the conduct of nuclear warfare in a more proliferated world. In it, we use ideas drawn from game theory to analyze detailed case studies of plausible multi-state conflicts. We believe that this approach offers a number of advantages over further theorizing that abstracts over all states.

³ Andrew J. Coe and Victor A. Utgoff, "The New Triad: Potential Implications for a Nuclear War over Taiwan," Institute for Defense Analyses, P-3966 (December 2005).

⁴ James D. Fearon, 1995, "Rationalist Explanations for War," *International Organization*, Vol. 49, No. 3, pp. 379-414.

Game theory offers a rigorous platform for studying strategic interaction. Its use allows us to be clear in our assumptions about state behavior and to derive the implications of a range of assumed interests on the part of the combatants. It also provides for the possibility of formalizing intuitive models of nuclear war and ensuring their logical consistency.

Furthermore, we can choose cases so as to maximize the plausibility and applicability of our results. By focusing on particular conflicts, it becomes feasible to incorporate many more variables into the analysis and thereby avoid having to contrive explanations for dismissing these variables out of hand. This in turn allows for more detailed and rigorous modeling. Of course, the goal of this approach is to generate new insights that pave the way for a more nuanced general theory.

Our cases are variants of a potential conflict on the Korean peninsula, in which imminent collapse of North Korea (hereon, the DPRK) leads it to initiate a war with South Korea (the ROK). We chose this setting according to three criteria.

First, the aspects of multi-state nuclear warfare that we wish to study might plausibly come into play in this conflict. The DPRK might anticipate its own collapse and precipitate a war as a last hope for regime survival, a war that might escalate to nuclear weapon use. The US might intervene on the ROK's behalf or to take advantage of a weakened DPRK. The US and ROK might have substantially different interests in the resolution of the conflict. The ROK might seek its own nuclear weaponry and/or end the alliance, prior to or during the conflict. China might also seek to influence the outcome, and in so doing may compete with the US.

Second, this setting offered less complexity and more tractability, for our purposes, than other plausible locales. One might instead examine potential multi-state nuclear warfare, say, in the Middle East. But the locale of the Middle East is, from our perspective, considerably more complex. The number of states that might be involved in a plausible conflict there is higher. A Middle East study would presumably need to consider the possibility of intervention by sub-state actors, for example suppressed ethnic groups, which might undermine the assumption of unitary state decision-making. It would also have to deal with the possible role of terrorist organizations with interests radically different than those of states. The Korean peninsula lacks these particular wrinkles.

Finally, this locale offered the opportunity to expand our earlier work on conflicts over Taiwan. The earlier study examined conflict between two major powers, both of

which had strong but not existential interests at stake, both too large and powerful to see their annihilation as a plausible outcome. The study at hand examines a conflict involving relatively weak states, for whom total destruction (at least of the regime) is a real possibility. This enabled us to extend the framework we have developed to analyze these conflicts. Section 2 describes this framework and the game-theoretic ideas on which it is based.

Section 3 presents our analysis of the simplest case considered here: a “perfect” US/ROK alliance where the allies approach the conflict as a unified actor, or nearly so. The perfect alliance case serves as a base case from which other cases will depart. We also use this simplified setting to analyze the implications of US missile defenses and preemptive capabilities for the alliance’s and the DPRK’s strategies.

Section 4 discusses three, more realistic, cases, where the US and the ROK approach the conflict as two actors with varying degrees of perceived and real commonality of interests. The first case assumes a sincere alliance between the US and the ROK and discusses the process by which the allies could establish a joint strategy for conflict, and the problems this poses for the DPRK. The second case examines the implications of a possibly insincere alliance, in which the allies may have incentives to deceive each other about their true interests. The third case assumes that the US and the ROK are unallied and interact strategically in pursuit of substantially different interests. This case enables an assessment of the incentives for an unallied ROK to acquire its own nuclear weapons.

Section 5 introduces the possibility of additional third-party intervention by China. It analyzes the possibilities for competition or collusion with the US or the DPRK and discusses the ways in which these possibilities can affect earlier stages of a conflict.

The results of these case studies can inform speculation about the general nature of a more proliferated world. In particular, they allow us to speculate about a number of possible mechanisms by which such a world might be stabilized. Section 6 discusses these mechanisms. Section 7 offers some tentative conclusions.

Before turning to the analysis, it is important to be explicit about the limitations of our work. First, taken as a study of potential conflicts on the Korean peninsula, our analysis is quite shallow. In particular, we essentially ignore the conventional aspects of any Korean conflict. We also ignore the many possible complications introduced by the DPRK’s chemical weapons, although their possible use is not altogether dismissed. Though we will eventually consider the involvement of both the US and China, we

ignore the possibility of intervention by other states (e.g., Japan). In so doing, we do not mean to suggest that these features are unimportant in understanding these conflicts. Rather, we introduce these simplifications both for tractability and also to focus the analysis on the particular questions we have posed.

Second, the models presented in this paper are not mathematically formalized. That is, although we draw extensively on the ideas of game theory, our models are not grounded in the detailed, rigorous specifications that typically accompany game-theoretic analysis. To explore our questions, we rely instead on informal arguments, derived from a less precise deductive logic. We recognize that an informal treatment of such a complex subject must ultimately be inadequate, and we reserve for future work the formalization of these models. For the present paper, we hope that these arguments, together with the illustrations and examples we offer, are sufficient to show the reader that our ideas are at least plausible, if not yet convincing.

Our goal in this paper is neither to provide a full development of a theory of nuclear war, nor to give a comprehensive treatment of potential nuclear conflicts on the Korean peninsula. Instead, we intend for it to serve as a profitable springboard for future work, for ourselves, and perhaps for others interested parties.

The importance of developing a more sophisticated understanding of proliferation's consequences is growing rapidly. With rapid changes in the balance of power, increasing competition in the global economy, and the specter of catastrophic terrorism, responsible nations cannot shirk making hard decisions about the relative importance of preventing proliferation. Though a highly proliferated world may seem distant, decisions made in the near future about relationships with allies, responses to the present proliferants, modifications to the nuclear non-proliferation regime, development of the US New Triad, and more, may have lasting implications for proliferation and the global order.

2. A MODEL OF LIMITED NUCLEAR WARFARE

How does one go about modeling nuclear warfare? There are no real-life examples of nuclear wars to be examined and generalized upon. It is not easy to find analogues in history. Situations where one state, group, or individual faced the possibility of annihilation by another are quite common. But in how many of those situations could the annihilated entity expect to remain able to inflict annihilation upon its destroyer, after its own demise? The potential for “revenge from beyond the grave” is a novel element in the strategic interaction of nuclear war

Given this vacuum, any model intended to frame the problem of nuclear warfare is inherently speculative. A good model should exhibit logical self-consistency; it should appeal to intuition or have a convincing rationale for any counter-intuitive aspects; and it should be cognitively realistic, in the sense of imposing reasonable demands on the capacity of combatants to comprehend nuclear conflict and choose strategies. Unfortunately – for planners, at least – these criteria leave ample room for a variety of plausible models.

Looking back to the Cold War, commentary and research by US nuclear strategists focused on two models, which we term the “escalation-by-risk” and “escalation-by-damage” models.⁵ These models share a number of features. Both assume that there are exactly two ways to “win” a nuclear war. First, a combatant can win outright by exhausting his opponent’s nuclear forces, through some combination of suffering strikes and preempting them, before his own forces run out. Second, a combatant can win by coercion. That is, the victor wins by convincing his opponent that conceding victory would be preferable to continuing the fight. Given the sizes and survivability of the US and USSR nuclear forces during most of the Cold War, strategists typically assumed that outright victory was impossible, since the exhaustion of either combatant’s nuclear forces would mean the annihilation of the other. Thus, victory could only be achieved by coercion.

5 Robert Powell, 1985, “The Theoretical Foundations of Strategic Nuclear Deterrence,” *Political Science Quarterly*, Vol. 100, No. 1, pp. 75-96.

Both models also assume that coercion would be pursued through *escalation*, or discrete increases in the intensity of the conflict. These escalations are a means for each combatant to credibly communicate its will to win. They serve to increase the opponent's estimate of the costs of victory, pushing him to reconsider his own commitment to victory. Given that at least one combatant puts only finite value on winning, the less committed of the two would eventually come to prefer conceding victory over enduring further escalations.

The principal difference between the two models lies in the presumed vehicle for escalation. In the "escalation-by-risk" model, an escalation consists of a combatant imposing an increase in the probability of a general, all-out nuclear exchange in which both combatants could expect to be completely destroyed. A combatant quits when the value of winning is no longer sufficient to justify enduring the increasing risk of an all-out exchange. This model requires that each combatant be able to exercise control over the probability of all-out exchange, but not over its actual occurrence. Returning to the Cold War precedent, the large, diverse nuclear forces of the US and USSR offered numerous escalation options: submarines could be put to sea; bombers could be put on-alert or launched on patrols; and ground-based missiles could be fueled and, if mobile, deployed to launch locations. Presumably, each increase in the alert posture of these forces implied an increase in the probability of an undirected engagement, which might rapidly and uncontrollably precipitate a massive exchange. Moreover, the US and USSR could also be reasonably confident that each could observe the other's escalations, given their persistent and detailed surveillance of each other's forces.

Alternatively, the "escalation-by-damage" model assumes that each side would escalate by employing limited nuclear strikes to impose damage on its opponent. A combatant quits when the value of winning no longer justifies the costs of strikes suffered and potentially still to come. In this model, each combatant must be able to exercise good control over the use of its nuclear weapons. However, the combatants' forces need be neither large nor diverse. Even if a combatant possesses a small number of identical weapons, he can devise a variety of escalation options by varying the number used, the targets attacked, and so on. Furthermore, to observe the opponent's escalations, a combatant need only be able to assess the damage caused by the opponent's strikes. Thus this model poses less demanding requirements on each combatant than the escalation-by-risk model.

The strategy community as a whole has reached no consensus on which is the "correct" or "better" model, or even how to judge these qualities. In retrospect, the

evidence suggests that the military strategists of both the US and USSR favored the escalation-by-risk model and structured their nuclear forces and plans accordingly during the Cold War. In contrast, the academic strategists in the US came to favor the escalation-by-damage model. The complexities introduced by new proliferants such as the DPRK have only muddied the already murky water.

However, it does not really matter which model is “correct” in an absolute sense. What is more important is that each combatant knows to which model his opponent will be playing. If both combatants play according to the escalation-by-risk model, then each can plan his sequence of increasingly risky deployments and employ them when the time comes. But if either combatant chooses to play by the escalation-by-damage model, the other will have little choice but to follow. A combatant that had initially chosen to play according to escalation-by-risk would be literally and figuratively struck by an escalation-by-damage opponent. His bluff called by the opponent’s strike, the first combatant would have three choices: a limited retaliation and subsequent play according to escalation-by-damage; an all-out strike according to escalation-by-risk, thereby putting himself entirely at the mercy of his (still) well-armed opponent; or quitting the conflict. Obviously, if the first combatant wishes to continue the fight, he will have to play according to escalation-by-damage. In our case study of war on the Korean peninsula, this implies that, should the DPRK choose the escalation-by-damage model, the US would do best to do the same.

In fact, there are several reasons to think that the DPRK would conceptualize nuclear war in accordance with the escalation-by-damage model. Our scenario presumes that the DPRK finds itself in desperate times, and therefore, is willing to employ desperate measures, including a war aimed at salvaging what it can from a final settlement of the Korean conflict. Under these circumstances, the DPRK regime would be searching for some theory—any reasonable model—that would give it some hope that it might survive and perhaps even prosper as a result of the war.

The DPRK would surely realize its inferiority to the US and the ROK in most measures of military capability. In the face of the Alliance’s much stronger conventional forces, perhaps soon to be advancing into its territory, the DPRK could hardly afford to engage in the long, drawn-out sequence of escalatory force alerts and deployments called for in the escalation-by-risk model. It might have little choice but to initiate the use of nuclear weapons just to halt the allied advance. In so doing, it would be foregoing the possibility of playing according to the escalation-by-risk model, having already stepped across the nuclear threshold.

Even if the DPRK managed to stave off Alliance forces without using nuclear weapons, its modest nuclear forces would not be wholly sufficient to threaten the annihilation of its enemies. If it attempted to escalate by risk, it would be raising the danger of severe damage faced by its opponents, but there would be no risk of total destruction equivalent to that with which the US could threaten the DPRK. Moreover, its small, homogeneous forces would give it relatively few options for risk-increasing escalations. So in choosing to escalate-by-risk, the DPRK would be fighting according to a model for which the US was better-prepared and better-equipped.

Finally, choosing to escalate-by-risk might also endanger the survivability of DPRK nuclear forces. To be effective, such an escalation must not only increase the risk of nuclear exchange, but must also credibly communicate this increased risk to the opponent. To ensure that its escalations were believed, the DPRK would have to make it possible for the US to observe its forces moving to higher states of readiness for use. However, in the absence of sophisticated surveillance on its part, the DPRK would have difficulty assuring itself that, by revealing its forces' changed posture, it was not also exposing them to preemption by US forces.

Escalation-by-damage eliminates these liabilities. The DPRK's first escalation-by-damage might well be a limited nuclear strike to stop allied forces penetrating its territory. In moving towards higher levels of escalation, the DPRK would have plenty of options for varying the types and numbers of targets struck, and its inventory would suffice to exact a terrible, if not total, price on its enemies. And it need only expose those forces of which it was about to make use, leaving the rest in hiding to maximize their survivability.

Thus, it seems likely that, of the two models, escalation-by-damage offers the DPRK the better hope for success in the war. Given its presumed desperate situation, the DPRK could be expected to seize upon this possibility, and choose its strategy accordingly. Thus, the US would do best to assume that the DPRK would escalate-by-damage and choose a strategy appropriate for that model.

A MODEL OF LIMITED NUCLEAR WAR ON THE KOREAN PENINSULA

Answering the kinds of questions we pose in this paper requires a considerably more complex version of the escalation-by-damage model than has been previously posited. Some of the added features ought to be common to most or all real limited nuclear wars and hence allow us to derive more specific predictions about such wars from the model. Other features enrich the model's depiction of the particulars of wars on the

Korean peninsula, thereby increasing the relevance of the analysis to the various policy dilemmas posed by that region. Of course the general aim in introducing these complications is to raise the model's overall level of fidelity to real world nuclear wars, with the hope that the attendant increase in analytical difficulty is outweighed by the improved results.

We begin with a discussion of the fundamental features of the model, before turning to an explanation of how these features give rise to the model dynamics. We first assume that the ultimate objective of the conflict, for all combatants, is to determine the political structure of the Korean peninsula for the long-term. That is, the combatants seek to arrive at a final or nearly final resolution to the original, 50-year-old Korean War. Obviously, there are many dimensions to such a resolution: the degree to which North and South will remain politically separate entities; if they unite, the arrangement by which the northern and southern leaderships will share power, or not; the final status of either regime, if removed from power; the final disposition of US, DPRK, and ROK forces, including nuclear forces; and so on. Our model of the conflict thus assumes there are a number of "intermediate" settlements, in addition to the usual modeled outcomes of complete victory for one side and total defeat for the other. In principle, there is a multitude of conceivable settlements; in practice, we will typically deal with only a handful for the sake of expositional clarity.

The ultimate sources of conflict in our model are the differences in the preferences each combatant has over the possible final outcomes. For example, the DPRK might prefer its total victory over all other outcomes and prefer unions where it retains substantial power. ROK preferences seem likely to be just the opposite. For its part, the US might be most concerned that any final outcome should include nuclear disarmament of the DPRK, and possibly also of the ROK. Given these differences, the selection of any particular final outcome involves compromise by at least one of the combatants—therein lies the conflict.

It is helpful to introduce some terminology that will be used throughout the paper. Each combatant's preferences are represented by a "value scheme". The value scheme describes which outcomes a combatant prefers to others, and *by how much*. For each possible outcome, the value scheme associates to that outcome a "damage threshold", defined as follows. Denote the combatant's most favored outcome (usually total victory over his opponent(s)) by V . The *damage threshold* for any outcome X is the amount of damage a combatant would willingly suffer in order to receive outcome V over outcome X . Intuitively, the damage threshold for outcome X is the amount the combatant would

pay in order to achieve total victory instead of X. Put another way, the damage threshold is the amount the combatant would be willing to pay in order to avoid outcome X. Thus, less preferred outcomes have higher damage thresholds (the combatant would be willing to suffer more damage to avoid them).

In particular, a combatant's *maximum damage threshold* is the damage threshold associated with the player's least favored outcome. Typically, this is the total defeat of the player, so the maximum damage threshold is the damage a combatant would willingly suffer in order to avoid total defeat. Note that this threshold may be quite large, perhaps even infinite (i.e., total annihilation). If a combatant's maximum damage threshold is infinite, then it has an "existential" interest at stake in the conflict: an interest (e.g., survival) so important that the combatant would suffer any amount of damage to defend it. The DPRK might be taken to have an infinite maximum damage threshold (i.e., an existential interest) in our hypothesized conflict.

A combatant may also have a *minimum damage threshold*: the level of damage that a combatant must suffer before it becomes willing to exit the conflict (by surrendering, quitting, or accepting some intermediate settlement). That is, even if a combatant knows that it will lose the conflict, it may continue to fight until its minimum damage threshold has been reached. This threshold may arise from the need of a state's leadership to satisfy domestic or international audiences that it had tried "hard enough" to defend its interests and/or uphold its commitments. So, for instance, even if the DPRK leadership were convinced that it could and would end the war by accepting a settlement very favorable to the US and the ROK, it might continue to fight until enough damage had been suffered to satisfy the North's elites that the leadership tried earnestly to achieve a better outcome.

Violence in the model arises from the uncertainty each combatant has about the other combatants' value schemes (that is, the value of the opponent's minimum and maximum damage thresholds). To see why this is true, suppose that each combatant had certain knowledge of all the other combatants' value schemes. Then, each combatant could easily predict the outcome of a war, because each could confidently assess how much damage the other combatants were willing to suffer to avoid each possible outcome. Each combatant could thus see the point at which a particular outcome would be accepted by all the combatants. Since each has the same information, all will come to the same conclusion about the final outcome of a war. But since the war itself is costly, all the combatants would be better off if they simply agreed to the predicted outcome without fighting a war. Thus, there would be no violence.

Of course, strategically-interacting combatants would never have certain knowledge of each others' value schemes. This is easy enough to believe in the case of a war on the Korean peninsula. (How much damage would the DPRK be willing to suffer to avoid, say, a long-term confederation between North and South? How much for the ROK or the US?) But this should also be true for any set of combatants in this model. The reason is that, if all the combatants thought that they knew all the others' value schemes, each would face strong incentives to misrepresent his value scheme. By doing so, the deceiving combatant can shift in his favor the final outcome predicted and hence accepted by the others. Because this reasoning applies to all combatants, they all have incentives to lie about their value schemes. And because peace offers the combatants no way to prove their claims about their value schemes, they must find an alternative mechanism by which to credibly communicate the strength of their interests.

The infliction and suffering of damage (here, nuclear strikes) by the combatants provides this mechanism. Each enters the war with uncertain estimates of the others' value schemes. As a combatant stays in the game, the others revise their estimates of its value scheme to account for its continued persistence. This occurs because, by staying in, the combatant exposes itself to additional strikes by the others. It thus demonstrates its continued willingness to accept damage while holding out for a more favorable settlement.

Combatants can also offer settlements to their opponents at any given time in the conflict. These offers are assumed to be binding, if accepted. That is, if for example the alliance offers the DPRK a long-term confederation between North and South, and the DPRK accepts, it is assumed that the war ends, and the offer is implemented. Moreover, all the combatants are assumed to know that any settlement offer is credible and binding.⁶

As the war continues, the damage suffered by all those contesting it increases. The combatants continue to revise their estimates of the strength of their opponents, and become willing to accept less and less favorable settlements. Though it is not actually required, we assume for expository simplicity that the combatants take turns in this process. During its turn, a combatant can accept an offered settlement (including surrender), make a counter-offer, and/or launch additional strikes.

⁶ This is a substantial simplification. The DPRK in particular has a long history of renegeing on past agreements. Nevertheless, this assumption allows us to focus on the war rather than on the political maneuvering that might follow it.

One can thus think of conflict as being a process of estimation. By inflicting damage on their opponents, combatants reliably probe their opponents' value schemes. However, in order to continue probing, they must stay in the conflict, and therefore they are also exposed to being probed. Hence, they are "paying" in damage for reduced uncertainty about their opponents' value schemes. These intuitions are helpful in exploring the best strategies for the combatants in the kind of conflict we are considering.

OPTIMAL STRATEGIES FOR LIMITED NUCLEAR WAR

With these elements of our model in place, we can now turn to characterizing the optimal strategies for combatants. We first describe the form taken by strategies in this model and define what is meant by "optimal". We then present the intuition behind optimal strategy, and analyze the role of risk tolerance or confidence in setting bounds on the optimal strategies. Within these bounds, optimal strategies are given by solutions to the problem of choosing the best "sizes" for a combatant's escalations and settlement offers. This paper will not provide solutions to the sizing problem, and so does not fully specify the optimal strategies. However, the incomplete characterization offered here is sufficient for present purposes.

Strategies in this game consist of a set of rules. These rules tell the combatants when to offer or refrain from offering a particular settlement, and when to accept or reject an opponent's offer. They also tell a combatant how to choose the size of its escalations and how to choose which settlements to offer. These rules will depend on the combatant's value scheme, its initial and evolving estimates of its opponent's value scheme, the range of escalatory options available to the combatant, and the range of settlements it could potentially offer.

We take the best, or "optimal", strategies to be those that maximize a combatant's expected utility from the conflict, *given the information that is available to it*. We also consider below the case of risk non-neutrality, where combatants are assumed to maximize risk-adjusted utility. This definition of optimality enables a particularly appealing intuition about optimal play.

First consider the following naïve strategy. A combatant should seek its most preferred outcome initially and offer each alternative settlement, from most to least preferred, as its damage threshold for each settlement is reached (or accept that settlement if the opponent has offered it). In other words, the combatant simply follows along until a new damage threshold is reached, concludes that holding out for its previously tabled

settlement is no longer worth it, and offers (or accepts if already offered) its next most favored settlement (i.e., the one associated with the new damage threshold).

It may not be obvious why this strategy is naïve. After all, passing a damage threshold without offering the associated settlement might be expected to result in the combatant “overpaying” for that settlement, if it were accepted later in the conflict. Similarly, offering a settlement before its associated damage threshold was reached might be expected to result in the combatant getting stuck with that settlement when a superior one was possible.

However, there are circumstances where deviation from the naïve strategy is beneficial. These circumstances derive from wise use of the estimate each combatant maintains of the value scheme of his opponent. For example, if a combatant reached a new damage threshold, but was relatively confident that its opponent could be coerced into a more favorable settlement later in the game, it might beneficially (in expectation) choose not to offer the associated settlement. Similarly, if a combatant were convinced that its opponent had generally higher damage thresholds, it might beneficially (in expectation) choose to offer a settlement before its associated damage threshold had passed. This happens because the combatant might judge from its estimate of its opponent’s damage thresholds that a better offer is not likely to be made before it reaches that damage threshold. It might therefore seize the opportunity to end the conflict at the current lower damage level.

These examples highlight the importance of a combatant’s confidence in its estimate of its opponent’s value scheme, as well as its attitude toward risk. A risk-neutral combatant with mildly uncertain estimates would generally be inclined to stick with naïve strategy. By contrast, a risk-acceptant combatant, especially one with confidence in low estimates of its opponent’s value scheme, would be inclined to let damage thresholds pass, gambling on the likelihood of a better deal ahead. On the other hand, a risk-averse player, especially one with confident high estimates of his opponent’s value scheme, would be inclined to offer settlements early, averting the risk of unnecessarily prolonging the war.

Figure 2.1 shows the bounds on optimal strategies formed by the four possible combinations of risk-acceptant combatants with confident low estimates and risk-averse combatants with confident high estimates. In general, the more risk-acceptant and confident in low estimates of their opponent’s damage thresholds two combatants are, the more destructive the war is likely to be, and the converse is also true.

These bounds delimit the freedom combatants have to employ the intuition discussed above for optimal play. It remains to determine the actual sequence of escalations and offers prescribed by the optimal strategies. In other words, an optimal strategy should specify the size of the escalation employed by a combatant in response to any history of its own and its opponent's actions, as well as the particular deal the combatant should offer at each such point in time.

In essence, the problem, then, is to determine a combatant's cost-minimizing path by which to "search" the opponent's value scheme. Finding the solution to this "sizing problem" requires complex calculations. Fortunately for us, our case study analysis requires only that we demonstrate that the optimal sizes are neither too small nor too big.

That escalations and changes in offered settlements cannot be too small follows from the strong desires on the part of all combatants to limit the time during which their militaries and populations are subject to the stresses and terrors of nuclear war. Combatant leaders cannot afford to dilly-dally while probing the will of their opponents. In any case, there is nothing small about the use of even one nuclear weapon, especially since the DPRK and perhaps the ROK will presumably have limited inventories.

That escalations should not be too large follows from the combatants' need to avoid wanton destruction and the disproportionate retaliation it invites. All of the combatants should be looking to pay the minimum price possible for their favored settlement; large escalations risk overpaying. Similarly, combatants would not want to "give away the farm" by being overly generous with changes to their offered settlements. Thus, we assume that the escalations chosen and the changes from settlement to settlement will be modest in size.

Bounds to conflict derive from players' value schemes and beliefs about each other.

US/ROK DPRK	Sees <i>Itself</i> More Committed to Complete Victory	Sees <i>DPRK</i> More Committed to Complete Victory
	Sees <i>Itself</i> More Committed to Complete Victory	Sees <i>US/ROK</i> More Committed to Complete Victory
Sees <i>Itself</i> More Committed to Complete Victory	US/ROK wins at DPRK maximum damage threshold if its maximum damage threshold is higher; loses at its own maximum damage tolerance otherwise.	DPRK wins at US/ROK minimum damage threshold.
Sees <i>US/ROK</i> More Committed to Complete Victory	US/ROK wins at DPRK minimum damage threshold.	DPRK wins at US/ROK minimum damage threshold if its minimum is greater; loses at its own minimum otherwise.

Figure 2.1

One last requirement for optimal strategies must be imposed. The model presented here assumes that the postulated war is the only important interaction for the involved combatants. Of course, in the real world, the ROK, the DPRK, and especially the US and China, would all be concerned with the effects of their behavior in this Korean war on their reputations, with each other and with other nations. These reputations condition expectations between these and other nations for other conflicts at other times. So, each of the combatants would attempt to ensure that it did not appear to be taken advantage of by its opponents. Even if it chose to end the conflict by conceding, a combatant would like to be seen as exacting a fair price for its concession. And in particular, a combatant would prefer to be seen as “giving as good as it got”. Thus, optimal strategies would require that combatants always respond to escalation with comparable or greater retaliation.

The description of our model and its optimal strategies given here is neither comprehensive, nor fully rigorous. But we hope that it communicates the intuition that has guided our approach to this research and that animates our case studies. This picture of limited nuclear war—with combatants characterized by value schemes and uncertain estimates of each other, choosing modest escalations and offers of settlement to dispel this uncertainty—offers the possibility of an examination of the nature of a more proliferated world. We begin with the simplest variant of our hypothetical Korean war.

This page is intentionally left blank.

3. CONFLICT BETWEEN THE DPRK AND A PERFECT ALLIANCE

The first of our four investigations considers conflict between the DPRK and a “perfect” alliance of the US and the ROK. By “perfect” we mean that the allies would agree on all important matters of strategy and fight as one.⁷ We consider this idealization in this section only, relaxing it in subsequent sections.

In this first investigation, as in all that follow, we assume that the DPRK has found itself in desperate straits and decided to gamble that an attack on the Alliance might defeat it, or at least result in a settlement of some advantage well short of the DPRK’s total defeat. It is certainly possible that a conflict between a desperate nuclear-armed DPRK and the US/ROK Alliance could escalate to nuclear warfare. Anticipating this, the Alliance, and perhaps soon the DPRK as well, could be most interested in bringing the war to a tolerable close with as little damage as possible. A wide spectrum of potential settlements can be imagined that would combine important factors such as the nature of the post-war political arrangements for governing the Korean Peninsula, economic aid, post-war security arrangements, etc.

Thus, the US would want to ensure that any settlement ruled out the possibility of another war on the Peninsula, especially a nuclear war. The ROK would want to minimize the damage to both Koreas, in part because of its identification with the citizens of the DPRK, but also because it might expect to bear most of the costs of recovery for the entire Peninsula. The ROK would expect to emerge with at least as much political power as it had enjoyed prewar, and it too would want to rule out the possibility of a future war. Finally, once it knew it could not win control of the ROK, the DPRK would want a settlement that would preserve as much of its autonomy and power as possible.

EXAMPLE SETTLEMENTS

To keep our investigation transparent yet rich enough to develop some of the more important implications of our model, we look at a specific example that assumes only three feasible settlements to the conflict.

⁷ Obviously, an alliance that drew the US into a nuclear war could not be deemed perfect in all senses.

The first settlement in our example is long-term confederation (LTC) of the two Koreas with substantial political autonomy for each of the confederated states. The US would want the confederation to be dominated by the South and would not want it to have nuclear weapons. The second settlement in our example is disarmament and development (D/D) for the DPRK. This would involve multinational security guarantees to the DPRK, elimination of all its weapons of mass destruction (WMD), and deep reductions in its conventional forces, in return for strong economic aid and help in development. It would leave the DPRK government in place.

The third settlement would demand regime change for the DPRK, but with immunity for the senior leadership (RC/I). This settlement constitutes total defeat of the DPRK, but offers the senior leadership a comfortable exile if, from the point at which they receive the offer, they forgo any further use of nuclear weapons. Of course, this offer would not be made if the Alliance knew that the DPRK had no more deliverable nuclear weapons. Under this agreement, the US would expect the ROK to occupy the North, hand over all the remaining DPRK nuclear weapons, dismantle any surviving DPRK nuclear weapons-related facilities, and continue to honor its promises under the NPT.

Of course, none of these settlements is totally satisfactory for any of the combatants, but that is to be expected. In view of the great destruction and the high emotions the two sides would experience in the course of a nuclear war, if a settlement is to halt the war well short of its full destructive potential, all the political leaderships involved must see benefits in it.

DAMAGE THRESHOLDS

As discussed in the previous section, the damage level at which a given side would plan on accepting or offering any given settlement – assuming that substantially more damage were still possible – is its damage threshold for that settlement. Figure 3.1 illustrates a nominal evolution of this greatly simplified conflict, and we will see how the damage thresholds assumed for each side can determine the war's outcome. Each axis corresponds to the accumulated damage each side suffers as the conflict progresses. The damage thresholds each side attaches to each of the settlements are shown on its axis. The diagonal line represents commensurate damage to both combatants.

As discussed, the lowest of each side's damage thresholds is assumed to exceed a minimum damage, below which it would not offer any settlement. This minimum corresponds to the possible needs of each leadership to put up a respectable initial fight before offering or accepting any settlement.

Finally, note that if the Allies are to fight as one, jointly offering or accepting any given settlement, they must agree on joint damage thresholds. Negotiations on this subject would surely be difficult. Both allies could be expected initially to propose exaggerated damage thresholds to pull the joint thresholds in their desired direction. For strategic reasons at least, the proceedings and the outcome of such negotiations must be kept secret. For this first investigation, consistent with our assumption of a “perfect” alliance, we assume that agreed joint thresholds have been reached and would not be subject to renegotiation or reneged on by either ally during crisis or war. We loosen this assumption in Section 4.

Note that joint thresholds imply that the Alliance’s decisions depend only on the total damage the Allies have suffered, not on how damage is allocated between them. And since by changing this damage allocation the DPRK cannot affect Alliance decisions regarding the offering or accepting of settlements or defeat, its decision making on these matters is also indifferent to this allocation. Acting as one also implies that the ROK would not be concerned that it did not have nuclear weapons, as it could count on the US to use its nuclear weapons in ways that would be consistent with joint strategic interests.

Example of a DPRK vs. US/ROK Conflict

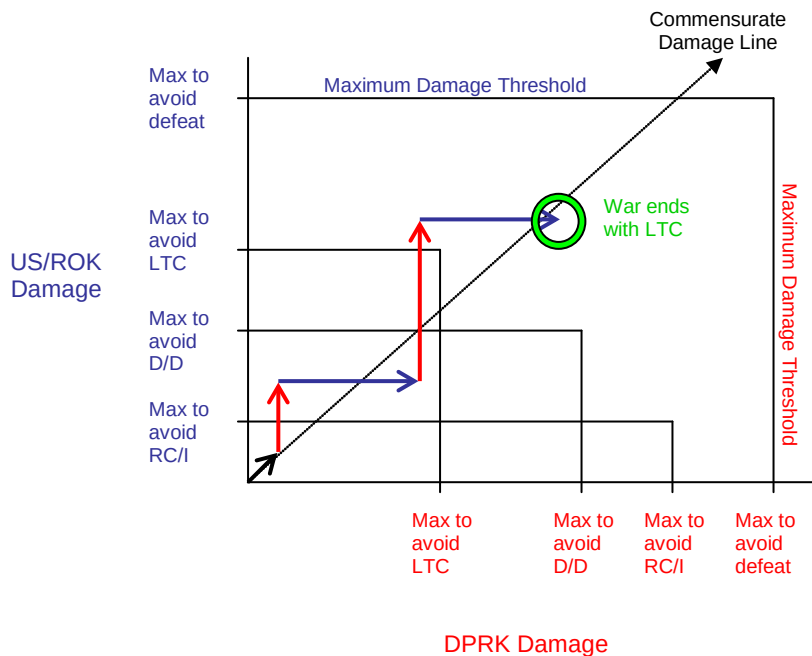


Figure 3.1

NOMINAL EVOLUTION OF EXAMPLE CONFLICT

The nominal evolution of our example conflict begins with initial attacks by the DPRK answered with at least commensurate responses by the Alliance. As argued in the previous section, at least commensurate responses should always be expected and executed, at least as long as the adversaries see the possibility of continued conflict.

This phase of the conflict is assumed to involve conventional forces possibly supported by modest DPRK use of chemical weapons against Alliance forces and bases to suppress their efforts to mobilize and fight. Thus the violence employed initially is modest compared to what is to come, and the damage done to both sides can be expected to closely follow the commensurate damage line indicated by the first arrow coming from the origin in the figure.

We might then suppose that the Alliance defense comes together reasonably well and signs begin to appear that DPRK attacks into the ROK are faltering. In this event, the DPRK would presumably have to escalate if it wants to preserve any prospect of winning. In this nominal conflict evolution, we assume the DPRK escalation involves substantial use of chemical weapons against Seoul resulting in perhaps thousands of casualties. The presumed aim of the DPRK would be to create mass panic that could force the ROK government to consider whether it can continue. It is also possible that the DPRK might use a nuclear weapon for this purpose. This escalation is depicted by the first vertical arrow, which carries the Alliance above its lowest damage threshold, that for regime change with immunity (RC/I).

In our nominal evolution, the Alliance is then assumed to retaliate strongly, driving the total damage to the DPRK well past the commensurate damage line. Attacks by the Alliance's conventional airpower over several days or more might accomplish this if the DPRK had only used chemical weapons in its prior escalation. Nuclear weapons could be used instead, which would seem reasonably likely if the DPRK had already initiated nuclear warfare. Immediately after its response to the original DPRK escalation, the Alliance would offer the RC/I settlement corresponding to the damage threshold it had exceeded.

The DPRK would presumably be very disturbed by this turn of events. But as shown in Figure 3.1, it would not yet have reached the point of taking or offering any settlement, and is nowhere near ready to accept so unfavorable a settlement as RC/I. Its response, depicted in the figure, is assumed to involve the use of a small number of

nuclear weapons. This response drives the damage suffered by the Alliance substantially above that so far suffered by the DPRK. The chart shows that this escalation would exceed the damage thresholds of the Alliance for both the disarmament and development (D/D) and the long-term confederation (LTC) settlements.

The illustrated response of the Alliance is then to first to retaliate at least commensurately for the damage it has just suffered and then presumably to offer both of the remaining settlements. As depicted, the Alliance response would drive the accumulated damage to the DPRK above its damage threshold for offering LTC, or in this case accepting it. Thus, the conflict ends with LTC, as this is the first settlement both are willing to accept to avoid further damage.

Note that both sides would be making and steadily updating estimates of the other side's damage thresholds. As indicated in Section 2, how these estimates would evolve would depend upon each side's initial estimates of how these values for its opponent's damage thresholds might be distributed, and then upon the adjustments to the estimates that were made as the conflict evolved and the sides learned more about their opponent's damage thresholds.

The moves assumed to be made here led to LTC, but at a much higher damage level than the two sides need to have suffered. Had they chosen responses that had not so substantially exceeded commensurate retaliation, and smaller escalations, the DPRK might have offered LTC first. Had that happened, the Alliance might have accepted the offer even though below its own damage threshold for LTC.

It would be especially willing to do so if its evolving estimate of the DPRK damage thresholds had led it to believe that neither of the two settlements it had already offered was likely to be accepted before it reached its own damage threshold for LTC. The Alliance might even press the DPRK to sweeten its LTC offer given that the Alliance damage threshold for this settlement is higher.

Even if the Alliance had continued the conflict until it had reached its own damage threshold for LTC, had it done so with more modest retaliations and escalations, the total damage suffered by both sides at the end of the war would have been considerably lower than that indicated in the figure. In short, both sides would have been better off to advance the conflict in smaller steps.

A CASE FOR ASSUMING MODEST ESCALATIONS

Of course this is just a single evolution of a conflict that could develop in a continuum of different ways. As stated in Section 2, absent a solution of the sizing problem, we can't specify a single best strategy for the two sides or be confident of which settlement might end the conflict. We can only specify a range within which the optimal strategy must lie. Retaliate, and escalate—boldly or conservatively—that is the question.

As shown in our nominal conflict, taking big steps can risk substantially more damage than necessary to find a mutually acceptable settlement. In addition, more time to reach any given damage level is more time for both sides to think about what they are doing. On the other hand, by going too slowly a combatant suggests reluctance to continue, which could lead its opponent to reduce its evolving estimates of the combatants' damage thresholds, thus encouraging it to think the combatant was closer to making or accepting a settlement offer, or perhaps even conceding defeat. As pointed out in Section 2, there are situations where this kind of encouragement can work to the combatant's disadvantage.

Finally, going slowly would also increase the time it takes to reach the damage level at which the two sides would agree on a settlement. And time is important when states are suffering the levels of damage that would come with the use of nuclear weapons. Citizens and government functionaries will be increasingly anxious for the war to end and for all effort to be focused on aid to the suffering and recovery in general. And under the extreme pressures of nuclear war, the temptations of subordinate elites to subvert the government or citizens to revolt could grow quickly.

We are persuaded that in high stakes conflict where there is great uncertainty for both sides about how the other side values potential settlements and its own defeat, the combatants seem likely to be relatively risk averse and, on balance, most concerned to avoid unnecessary suffering to reach a settlement. Thus, the Alliance seems most likely to retaliate at least commensurately for the damage imposed on it but to escalate modestly as it reaches for the possibility of a commonly acceptable settlement. And the DPRK would be wise to do the same.

If the sides escalate modestly, and if neither side starts the war somehow highly confident of very wrong estimates of its opponent's damage thresholds, figures of the type shown above tell us which of the settlements would be agreed upon, and provide a reasonable upper bound on the accumulated damage for the sides. We suggest that the reader assume modest escalation while interpreting the analyses presented below.

Nonetheless, the larger conclusions we will draw below do not depend upon this assumption, only upon the fact that some settlement would be reached under broad and straightforward conditions.

POTENTIAL IMPLICATIONS OF US/ROK BALLISTIC MISSILE DEFENSES

Consider again our example conflict between the DPRK and the “perfect” US/ROK Alliance. For simplicity, assume that the only WMD strike systems the DPRK has are nuclear-armed ballistic missiles. Assuming for the moment that the Alliance does not have a substantial strike capability that might preempt against the DPRK missile force, what effect would Alliance missile defenses have on the potential conflict?

Clearly, useful defenses would increase the numbers of missiles the DPRK would need to cause any given amount of damage. But the DPRK might be considerably uncertain of how good the Alliance missile defenses and especially US missile defenses are. How might it deal with such uncertainties?

Testing the effectiveness of Alliance defenses against DPRK missiles requires shooting some missiles at the Alliance. Until the DPRK knows its missiles can penetrate in numbers that are likely adequate to its strategic needs, it would not want to impose and suffer great damage. Thus, it might probe the Alliance defenses by arming with nuclear replica warheads or conventional warheads some of the types of missiles it uses for nuclear strikes, and launching them at targets the Alliance would likely defend. Such probes would allow the DPRK a useful if crude assessment of whether it has the capability to cause enough damage to meet its potential strategic needs.

Alliance missile defenses might not concern the DPRK if, despite them, it estimates that it can still drive the Alliance above even the most conservative estimate of its damage threshold for conceding defeat – while holding enough damage capability in reserve to deter a disproportionate Alliance response. To the extent that Alliance defenses led our desperate DPRK to doubt that it can meet this criterion for pursuing defeat of the Alliance, it would have to hope it can meet such a criterion for a tolerable settlement that the Allies would prove willing to offer.

The DPRK could also find it useful to build a large enough missile force to impose a very conservative estimate of the Alliance’s maximum damage threshold – despite a very conservative estimate of the effectiveness of Alliance defenses. But with its uncertainties multiplying in this way, the DPRK could find that very costly. Finally,

note that the DPRK could find it useful to keep the size of its nuclear missile forces secret.

Figure 3.2 describes the situation the combatants could find themselves in when the Alliance has strong defenses. We assume for illustration that the DPRK has discovered that the Alliance defenses are so capable that it cannot impose more damage than that depicted by the dashed green line. If the Alliance has also discovered this limit, both sides would know the DPRK could be defeated at less cost to the Alliance than the damage level indicated by the line. In contrast, the DPRK would be vulnerable to as much damage as the Alliance cared to impose.

If during the conflict both sides were to discover this state of affairs, it would not be surprising if the Alliance immediately offered RC/I in order to take advantage of these circumstances, even though neither it nor the DPRK had yet imposed or suffered anywhere near their damage thresholds for that settlement. The Alliance might find this situation too advantageous not to exploit to the fullest. Of course, the Alliance would have to run the risk that the DPRK might still choose to test the Alliance's will to the extent it could.

Note that even a considerably less capable defense could have substantial value, by ruling out the potential opponents' ability to force the US and its allies to accept defeat, or even to offer some of the settlements they see as most undesirable.

A Potential DPRK vs. US/ROK Alliance Conflict

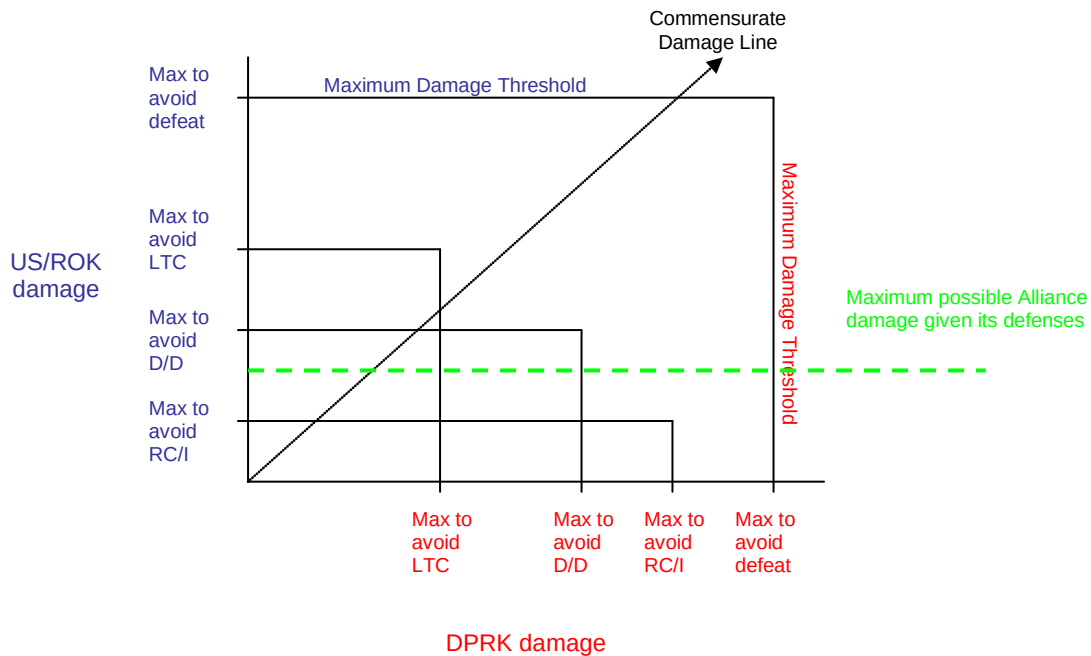


Figure 3.2

POTENTIAL IMPLICATIONS OF US/ROK PREEMPTIVE STRIKE FORCES

The potential strategic implications of US/ROK capabilities to attack DPRK nuclear forces depend upon their effectiveness, how well the DPRK knows their effectiveness, whether the DPRK can detect incoming preemptive attacks and launch the threatened missiles before they are destroyed, and the effectiveness of Alliance defenses. For purposes of this discussion, assume that Alliance defenses cannot be destroyed or exhausted by DPRK attacks.

Consider first a situation where the DPRK has the capability to detect incoming preemptive attacks and to launch its threatened missiles quickly enough to allow their escape. In this case, if the DPRK knows beforehand that the potential effectiveness of Alliance preemptive attacks is low, it might choose to ride out the attacks. A few missiles might be lost, but if they represent a small fraction of the DPRK capability, the total damage it could inflict would not change much.

Suppose the DPRK knows beforehand that Alliance preemption capabilities are good enough to pose a moderate risk that if none of its missiles are launched under attack

it might not be able to test the Alliance's willingness to resist defeat, or even to resist offering some of the better settlements the DPRK seeks. In this case, the DPRK could decide to launch some of its missiles under attack. This might constitute an escalation that, while perhaps larger than the attack that the DPRK would otherwise have chosen at that point, would still be reasonable. This tactic would allow the DPRK to stretch its strike capability and reduce the risk that it might not have enough.

Now, suppose that the Alliance preemptive capability is near-perfect and again the DPRK knows that beforehand, and can detect the attack and launch its missiles before they are destroyed. If such a preemptive attack were detected, the DPRK would know that, one way or the other, its last chance to do great damage to the Alliance is about to disappear. What might it do?

If the incoming Alliance attack were known beforehand as certain to impose enormous damage on the DPRK, its leadership might choose to launch all its missiles under attack, seeking to impose rough justice on the Alliance. It might ignore the possibility of suffering further nuclear attacks by the Alliance, which it would have no way of deterring. This choice seems plausible as the leadership might see itself as having nothing left to lose, save the possibility for revenge.

A second choice for the DPRK would be to allow all its missiles to be destroyed on the ground. This might make sense if the leadership recognized that the Alliance preemptive attack would lead to its quick defeat in any case, that there is little sense in risking more suffering by its people, and that it wanted the Alliance to be as motivated and as capable as possible of aiding the recovery of the people of the DPRK. Standing down might also make sense if the DPRK leadership were somehow to calculate that it had a chance of riding international sympathy to a better outcome by depicting the sacrifice of its nuclear forces as a humanitarian act and the US as the reckless party. However logical it might be for the DPRK leadership to forego launching some or all of its missiles when it sees their destruction as imminent, the current DPRK leadership as we know it seems unlikely to make either choice. In fact, it is far from clear that the US could make such a choice were it ever to find itself in a similar situation.

Would it make any difference if the Alliance capabilities to preempt employed either conventional or very low yield nuclear weapons that would cause relatively little collateral damage? In the cases considered above, we doubt that it would. The logic that seems likeliest to drive the DPRK leadership's decisions would not change. The deterrent values of promising retaliation before conflict starts and in prewar crises, and the wartime

satisfaction that the DPRK leadership might find in imposing a punishing revenge should it find itself facing total defeat seem likely to dominate other considerations.

But there are other cases to be considered. Suppose that the DPRK knows that it faces near-perfect Alliance preemption capabilities that would cause little collateral damage. And suppose that the DPRK cannot detect an Alliance preemptive attack early enough to launch its missiles before they can be destroyed. Facing such limitations, the DPRK would have to decide when to use its nuclear missiles based only on knowledge of its general strategic situation. If the Alliance were aware of these limitations, it could feel strong motivations to preempt before any DPRK use of nuclear weapons and could be expected to attack immediately upon detecting DPRK preparations to strike. In this case, the Alliance would expect to succeed with its preemptive attack and the only damage involved would be that caused by the preemption, making low collateral damage preemption capabilities of great value.

Note that if it were aware that it faces near-perfect preemption capabilities, the DPRK would understand that it has at most only one chance to use nuclear weapons. It should also see great danger in trying to probe the uncertain effectiveness of Alliance defenses. Of course, it might not be totally aware of its situation.

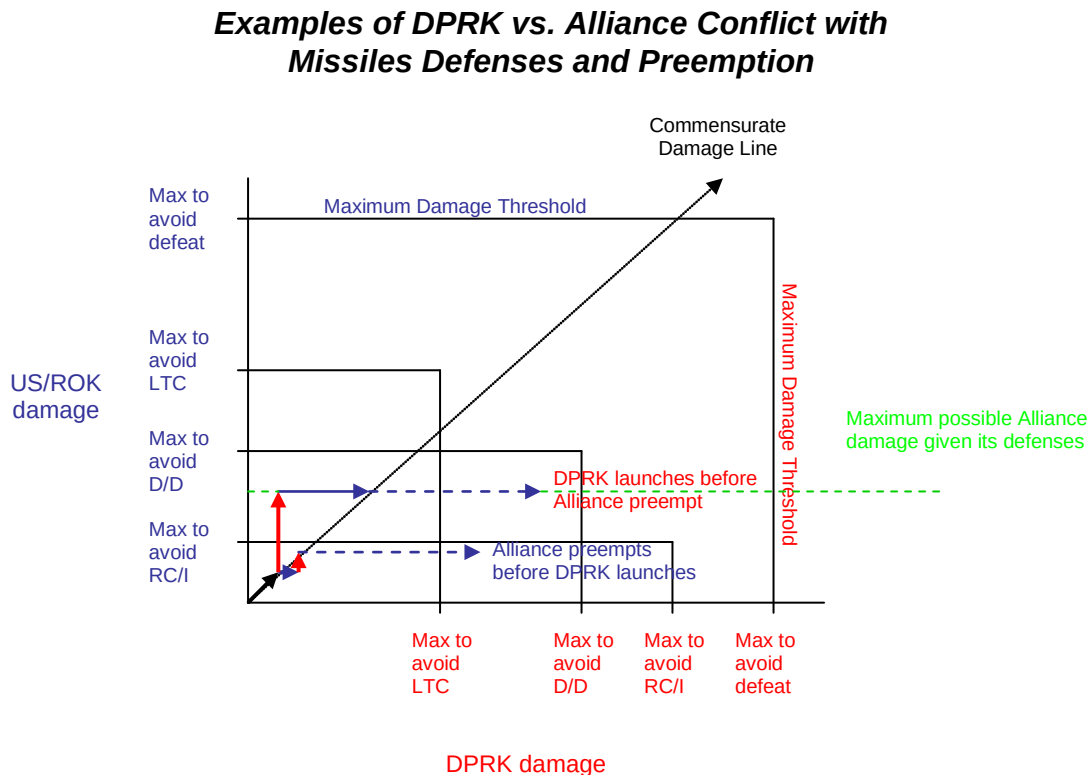
If the Alliance has near-perfect, low collateral damage preemption capabilities, the potential damage trajectories defining our example war would take either of two simple forms. In both cases, the trajectory would initially remain relatively close to the commensurate damage line as long as the two sides employ only conventional weapons. If the Alliance were to preempt before the DPRK escalated to nuclear warfare, the damage trajectory would then move horizontally to the right a distance corresponding to the collateral damage caused by the preemptive attack. The trajectory might then move vertically to the extent the DPRK had any surviving weapons, launched them, and succeeded in penetrating Alliance defenses. From that point on, only the DPRK could suffer further major damage.

The other possible form for the damage trajectory would depart from the commensurate damage line vertically if the DPRK were the first to use nuclear weapons. The damage done would be determined by the number of missiles the DPRK chose to launch knowing that whatever it did not launch would quickly be destroyed, and by the effectiveness of the Alliance defenses. After this first and only nuclear attack by the DPRK, any further major damage would be imposed by the Alliance and represented by a horizontal line to the right. The Alliance might choose to impose nuclear damage to

coerce a quicker DPRK surrender, or alternatively, complete its defeat solely by conventional means.

Examples of these two trajectories for a conflict where the Alliance has both strong missile defenses and preemption capabilities are shown in Figure 3.3. Note that no trajectory can exceed whatever maximum defense line is implied by the Alliance defenses.

Of course, once the DPRK is unable to make further nuclear attacks, the damage thresholds for both sides would cease to have meaning. Both would know that the Alliance could not be coerced by the threat of additional nuclear strikes into offering or accepting any settlement. The dashed horizontal blue arrows represent the potential choices the Alliance would have to coerce the DPRK to accept its defeat.



POSSIBLE LIMITATIONS ON DPRK SITUATIONAL AWARENESS

The situational awareness of the DPRK is clearly a key factor for both sides. Three kinds of limitations on situational awareness come readily to mind. First, the

DPRK might be unable to receive or transmit messages concerning settlements that are being offered or accepted. This would clearly be a serious problem for both sides.

A second conceivable difficulty is that the DPRK might be ignorant or slow to learn of the results of its attacks on the Alliance. This could have both good and bad effects for the Alliance. For example, if the DPRK were as a result to doubt the success of attacks it had made that had in fact succeeded, it might wrongly think that additional attacks were needed and make them. On the other hand, the Alliance might see an advantage in suppressing reports of damage to the extent it can in order to reduce the fears of its own citizens or perhaps to suggest to the DPRK that Alliance defenses are stronger than they really are.

A third conceivable difficulty for DPRK situational awareness is that it might be slow to learn of the damage that the Alliance is doing to it. The Alliance might try to manipulate this kind of DPRK awareness. If it could, it would presumably want to delay DPRK awareness that a preemptive attack is being launched against it, or that it is being invaded. Alternatively, the Alliance might find it important for the DPRK to know that such things are not happening when they are not.

In summary, the Alliance would want the DPRK to have good situational awareness in some circumstances but not in others, and it would want good two-way communications in order to exchange settlement offers and acceptances. Once the conflict reaches the nuclear level, these capabilities could be in doubt, especially given the possibilities for electromagnetic-pulse damage to the required communications systems.

CONCLUDING OBSERVATIONS

Obviously, there are many other questions about the “perfect” alliance one can explore using this model and the general point of view it provides; but, of course, we cannot explore them all in this paper. We close this section, then, with a few final points.

First, it is clear that the deployment by the Alliance of a near-perfect capability to destroy DPRK nuclear strike capabilities before they can be used could be of great advantage. Correspondingly, such an Alliance capability can pose great problems for the DPRK by making its nuclear forces usable only if used first. In such circumstances, the DPRK could at most look to its nuclear forces to deter a war, not to win one.

Note that this kind of crisis instability arises only when at least one side sees its continued existence brought into question if a war were to start. In this case, the DPRK

must know it is most unlikely that it would be able to establish a nuclear capability to bring the existence of the US into question. In contrast, the continued existence of the DPRK in a nuclear war against the Alliance would be at the latter's discretion. Clearly, it takes more than big weapons to be a strong state. The strategic depth to absorb damage and keep going is also needed, and because it lacks such depth, the DPRK is fundamentally a weak state and will continue to be one. Standing alone, the ROK would also be a weak state in this sense, but as part of a reliable Alliance with the US it is not standing alone.

Note also that if the Alliance deploys strong defense and preemption capabilities, and the DPRK accurately recognizes its resulting situation, both sides should be more deterred from challenging each other, given the dangers of doing so. But of course, challenges are not always seen as free choices, and the challenged party may see the dangers and decide to retreat.

Third, it is certainly possible that the US and the ROK might have different views of the importance of retaliating against a disarmed and collapsing DPRK that had caused great damage by launching its missiles under attack. The US could be concerned at the precedent that would be set for confrontations with future nuclear adversaries if the DPRK were not subjected to some nuclear retaliation. The ROK on the other hand would be less concerned with setting the right precedents if the outcome of the war has already been determined than with saving as many Korean lives as possible.

Of course, our speculations about this particular example do not tell us how an actual nuclear war between the US/ROK Alliance and the DPRK might work out. They are meant only to illustrate how a two-sided conflict can be characterized in terms of relative damage thresholds, how some settlement short of defeat for either side can exist, and how such a settlement might be reached – despite the fact that both sides would likely start such a conflict with substantial uncertainties about the possibilities for settlements and about the damage thresholds the other side might associate with possible settlements.

And, of course, it is possible that a side might be significantly uncertain of its own damage thresholds even if it had thought carefully about them in advance. How confident could an American president be of how much damage the US could take before beginning to disintegrate politically?

4. FROM PERFECT ALLIANCE TO STRATEGIC RELATIONS

In the previous section, we idealized the alliance of the US and the ROK as that of unitary actor. The Alliance was thus assumed to be “perfect” and the conflict between it and the DPRK could be considered purely bilateral. In the case studies that follow, we consider three weaker security relationships between the US and the ROK that assume reduced agreement and cooperation between the two.

The first of these weaker relations assumes, as does the perfect alliance case, that the US and the ROK have a strong and enduring alliance and can rely on each other to act jointly based on whatever agreements they reach. Both see the long-term value of their Alliance as greater than any gain they might make in the immediate conflict by reneging on their agreements or misleading one another. Thus, they will act sincerely in all their dealings with one another.

But sincerity does not require that they see eye to eye on everything. In this security relationship, each of the two leaderships recognizes that the endurance and will of its state to persevere in the conflict would be more strongly driven by the damage it suffers than by that suffered by its ally. Thus, each leadership wants to be sure that the Alliance does not require its own country to accept more damage than it believes resisting any given settlement is worth. So in this relationship, the Allies agree to jointly offer or accept any given settlement at the point where either one of them reaches the individual damage threshold it had set for that settlement. And, of course, given our assumption of sincere dealings with each other, each ally honestly informs the other of its individual damage thresholds prior to the war.

The second of these weaker relationships also assumes the two allies agree to act jointly, and it further assumes that they agree on joint damage thresholds determining when the Alliance will offer or accept any given settlement. But in this case, the Allies are not so convinced of the long-term value of the Alliance and are thus willing to risk their reputations with one another in order to shape to their advantage the agreements under which the immediate conflict would be fought. Thus, they may mislead each other about how much damage they are willing to take to resist offering or accepting any given agreement.

The third security relationship assumes that the alliance between the US and the ROK has been dissolved and that the two states would make independent strategic decisions on how to defend against the DPRK attack on the ROK. Two variants of the relationship are examined, one in which the ROK does not have its own nuclear weapons and one in which it does.

The best strategies for all three of the participants in our assumed Korean conflict are very different for each of the three alternative security relationships described above and for the perfect alliance case discussed in the last section. Note also that, in moving from the perfect alliance to each of these three alternatives in succession, we are stepping toward a trilateral war in that the US and the ROK are acting more independently in resisting the DPRK attack.

Finally, observe that with strong common opposition to the DPRK campaign in all of the security relationships described, we do not encounter any reasons for the ROK and the US to ever impose damage on one another. Some possibilities for that will show up in Section 5.⁸

The following subsections provide observations about the strategies of the combatants for each of the three security relationships described above.

SINCERE ALLIANCE WITH DISAGREEMENT ON DAMAGE THRESHOLDS

In Section 3, we assumed that the Allies could settle on and adhere to joint damage thresholds based on the *total* damage to the Alliance. Of course, the Allies could differ substantially in the amounts of damage to their own country that they would judge as worth suffering to resist offering some given settlement. Thus it seems more realistic to assume that the Allies would want these differences to be accommodated in joint decision-making.

US willingness to endure damage in order to win what it sees as one of the more preferable settlements would be influenced by its global responsibilities and interests, many of which the ROK does not share. And, because defeat of the Alliance would mean the immediate end of the ROK, but not of the US, the Allies seem likely to see substantially different individual damage thresholds as appropriate for this outcome. Further, as mentioned earlier, because the ROK identifies with the citizens of the DPRK and because it could see itself paying most of the costs of post-war recovery of the two

⁸ In our view, it is not plausible, now or in the foreseeable future, that the US or ROK would impose on one another the kind of damage we discuss in this paper, in pursuit of whatever gain. That said, this possibility is not directly proscribed by the model, so we mention it here for completeness.

Koreas, it might be generally less willing to impose and accept damage to resist settlements other than defeat than the US.

The importance of somehow accommodating such differences in joint decision-making by the Alliance is well illustrated by the current situation of the Alliance. As the US is reasonably confident that the DPRK would be able to land few if any nuclear weapons on the US homeland (and hopes to keep it that way), it knows that the ROK would suffer most of the damage if the war continues long enough. This consideration might make it more willing than the ROK to propose higher damage thresholds for the settlements the US would least prefer.

If Alliance decision-making is based on the individual damage thresholds of the two allies as done in this first alternative security arrangement, their different views on how hard to resist any given settlement can be reasonably accommodated. Note, however, that this accommodation does not come free. The ally with the lower damage thresholds might somehow be selected by the DPRK as its preferred victim (we'll say more about this later), while the ally with the higher thresholds may have good reasons to want to resist more strongly. Clearly, radical differences in the Allies' individual damage thresholds would be a serious challenge to the solidarity of the Alliance that we have assumed in this case. We will discuss this in the next subsection.

The nature of the combatants' best strategies given this first approach to accommodating differences in the Allies' damage thresholds can be more readily understood with the aid of Figure 4.1. It depicts the damage level for each of the two allies on separate axes and shows each ally's individual damage thresholds. Note that in this potential conflict the allies have the same preference ordering for the settlements, though all their damage thresholds are significantly different save for that they assign to regime change with immunity. One would not expect very different preference orderings, but for reasons described above, we should not be surprised if one of the allies has damage thresholds that are significantly higher than those of the other.

A Potential Conflict between the DPRK and a US/ROK Alliance with Individual Damage Thresholds

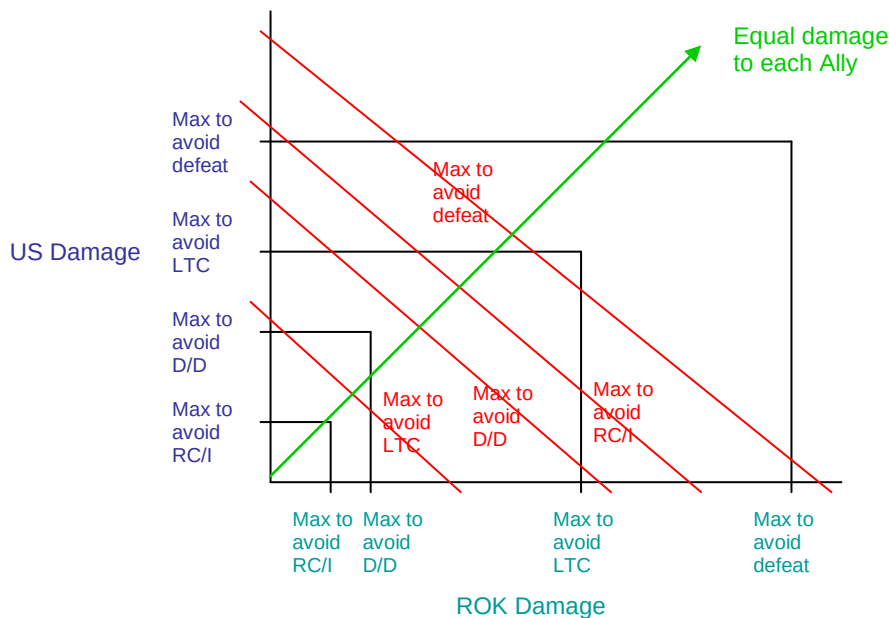


Figure 4.1

In this example, the US is significantly less willing than the ROK to settle the war with disarmament and development of the DPRK. This difference might make sense if the US strongly valued putting an end to the decades of trouble that it has had with the DPRK but is skeptical that D/D, which would leave an autonomous DPRK, would provide it. The difference in damage thresholds for D/D might also make sense if the ROK were particularly concerned about paying the bills for the DPRK and saw a disarmed and developed but autonomous DPRK as likely to be less burdensome than if the DPRK becomes its responsibility.

The example damage thresholds also show that the US is significantly more willing than the ROK to settle the war with long term confederation of the two Koreas. This difference might make sense if the US saw a well-constructed LTC as likely to end its long term troubles with the DPRK. The higher ROK damage threshold for LTC might make sense if the ROK were particularly sensitive to the infringement of its autonomy implied by sharing political power in confederation with the DPRK.

We have already noted that there is little difference between the Allies' damage thresholds for regime change with immunity. The figure also reflects our premise that

the ROK should be expected to fear defeat of the Alliance more than the US would, and thus have a higher damage threshold for this worst of all outcomes.

Consider now the damage thresholds of the DPRK. Note that it is assumed to prefer ending the war with long term confederation than with its disarmament and development. This preference would make sense if the DPRK would rather be part of a strong and prosperous confederation with the ROK than weak and surrounded by strong military powers. The DPRK might also fancy its chances to build increased political power within a long-term confederation even if the ROK were initially much the stronger party.

The reader should have little trouble imagining other plausible explanations for the relative values of example damage thresholds assumed for the Allies and even for the DPRK. In any case, the example damage thresholds are not important for the general points made below.

In this example, only the US and the DPRK have an independent nuclear strike capability. Thus, nuclear retaliation for DPRK attacks would be done either by the US, or by the ROK using US nuclear weapons released for this purpose. And because the security relationship we are examining is presumed sincere, the Alliance would agree on the need and means of nuclear retaliation and fulfill their obligations.

Points on the chart representing any given total of the damages done to the Allies are represented by straight lines at 45-degrees to the damage axes. The chart shows in red the specific lines corresponding to the commensurate damage imposed by the Alliance that would be equal to each of the damage thresholds assumed for the DPRK.

In projecting how this example conflict might end, we assume again that the DPRK and the Alliance choose modest escalations and retaliations in order not to impose and suffer substantially more damage than they must in order to reach an agreed settlement. Given this assumption, the settlement ending our previous “perfect alliance” example conflict did not depend on how much of the total damage to the Alliance is allocated to each ally.

Not so, now. If the DPRK were to focus all its damage on either of the Allies, the conflict would be settled by agreeing to long-term confederation. If, however, the DPRK were to allocate half of the total damage it imposes to each ally, the conflict would be settled by agreeing on disarmament and development. Since in this example the DPRK prefers LTC to D/D, it would be better off to concentrate all its damage on one or the other of the allies. And because the US would not fight as hard to resist LTC as would the

ROK, the DPRK would be best off if it chose to concentrate all its damage on the US, all else being equal. Of course, it would not know which ally to concentrate on, unless it somehow had found out which had the lower damage threshold for accepting LTC.

We should expect the DPRK to do its best to form a good estimate of which Ally has the lower damage thresholds. Even before the war, it should try to assess which Ally is less committed to resisting possible settlements. Observation of the Allies' behavior in bargaining over the DPRK nuclear program could provide some clues, as could prewar estimates of the relative costs and benefits each ally might see for each of the plausible settlements. Should a war start, the DPRK also could probe the wills of both Allies with its initial attacks. If one ally were to offer a settlement early, and better yet were quick to offer an even better one, that ally might seem the best target for any further DPRK attacks.

In general, if the Alliance were to operate under an agreement to offer a settlement as soon as either Ally reaches its damage threshold for that settlement and that settlement turns out to end the war, all the damage that was imposed on the other ally and all the corresponding commensurate damage imposed on the DPRK is wasted suffering. Thus, the best course of action for the DPRK would be to choose quickly, if not from the outset, a single ally as the preferred victim for its nuclear attacks.

The DPRK would face a complex dilemma in making such a choice. Choosing the ROK as the sole victim would require that the DPRK not shoot back at the US when it retaliates as agreed for the nuclear strikes made on the ROK. Further, the DPRK might reasonably assume that the ROK has the higher damage threshold for resisting defeat of the Alliance. If this were true, as we have assumed in our example, defeating the Alliance would require the DPRK to absorb more damage than defeating it by driving the US out of the war.

On the other hand, making the US the preferred victim for nuclear strikes risks especially grave consequences for the DPRK. While the DPRK might reasonably assume that the US has the lower damage thresholds, it couldn't be sure. It is possible that the US attaches very high value to preserving its global reputation as a reliable ally and to protecting its non-proliferation interests by preventing any gain from aggression backed by nuclear weapons. This could lead it to attach a very high damage threshold to resisting the defeat of the Alliance.

Finally, note the extra uncertainty that may be facing our desperate DPRK when it decides to gamble it can defeat the Alliance. It still faces the uncertainties that we have

discussed previously, about factors such as the potential effectiveness of Alliance missile defenses, the effectiveness of its preemption capabilities, and the values of Alliance damage thresholds. But it may also be unsure that the Alliance is operating under an agreement where either Ally can cause the Alliance to jointly offer or accept a settlement or even defeat when it reaches its corresponding individual damage threshold. What if it is wrong and, after defeating the ROK as its preferred victim, it finds that the US chooses to fight on?

A LESS THAN SINCERE ALLIANCE

Whether or not the Allies judge that their long term interests in preserving a reliable Alliance dictate that they deal with each other sincerely, they would likely see incentives to misrepresent their interests to one another. In particular, if their security arrangements require them to establish joint damage thresholds, and if they can anticipate how their true damage thresholds differ from those of their ally, each could see an advantage in exaggerating the differences between them in order to pull the final compromises closer to its true values.

As another example of temptations to misrepresent its damage thresholds, if the US wants to discourage ROK nuclear proliferation, it might represent its thresholds as not being meaningfully different from those of the ROK, when in fact they might be. To the extent the ROK sees such representations as sincere, it would be more confident that the US would use its nuclear weapons as the ROK would wish, and it wouldn't need its own.

Given such incentives, the US and the ROK cannot be certain of each other's true damage thresholds. Still, if both see the long-term value of their Alliance as dominating any interests they might have in seeking near-term advantage, they would want to establish joint damage thresholds that they could count on to guide the decisions of the Alliance through the entire war. This should lead them toward rationalizing their own damage thresholds to minimize differences with those of their Ally and to find a way to agree on joint damage thresholds to which they could commit.

There are many ways they might try to arrive at such joint damage thresholds. Perhaps splitting the difference between their individual damage thresholds for each settlement might seem fair. Perhaps the Allies might find an agreed way to weight each of their proposed damage thresholds with the damage each ally actually suffers so that the thresholds of the ally taking greater punishment would count for more. However they might arrive at joint damage thresholds, if the Alliance is going to continue to act as

originally agreed throughout the conflict, the differences between the thresholds the Allies have agreed to and their true damage thresholds cannot be too great.

Now, suppose that the Allies do not believe their Alliance is of fundamental and enduring importance for the long term, or do not understand the dangers of misrepresenting their true damage thresholds. In this case, they might feel free to seek whatever advantage for the immediate conflict they think they can derive from misrepresenting their damage thresholds. And even if both allies are both convinced that their Alliance is of fundamental and enduring value and were thus inclined to represent their damage thresholds honestly, it might still be possible that one or both might worry that the other Ally is not thus convinced and honest. An ally that has such doubts might then think it needs to misrepresent its damage thresholds to compensate for the misrepresentations that its Ally might be making. Finally, it is also possible that the Allies might both be convinced and honest, but that one or both would discover in the course of the conflict that its true damage thresholds are very different from what it had originally believed.

All three possibilities could put the Allies in a position where they would risk discovering major differences *in wartime* between each ally's true thresholds and the joint thresholds they had originally agreed upon. What might be the consequences?

Clearly, substantial differences could be the source of great stress on Alliance cohesion as the damage mounts, and in particular when the conflict escalates to the nuclear level. This would be especially so if the Allies did not originally see, or came to question, the fundamental and enduring value of their Alliance.

As the damage mounts, one or both Allies could begin to see the joint thresholds as too high or too low. This unease could manifest itself in the form of arguments as to why some particular settlement might better be offered or accepted at a different damage level than had been agreed upon. If the Allies were to agree that their joint damage had been revealed to be too high or too low, they should be able to agree quickly on adjustments. Still, adjustments could invite further adjustments, and pose the risk that the joint damage thresholds would lose their meaning. They could thus fail to discipline the evolution of the war, so that a tolerable agreed outcome was achieved with the least possible damage.

Of course, the most dangerous problem posed by differences between the true damage thresholds of the Allies and their agreed joint damage thresholds would arise when the high stresses of the conflict lead the Allies to want to change the joint

thresholds in opposite directions. For example, the US might have higher true damage thresholds for the potential settlements than the joint thresholds it had agreed upon, and come to see winning an advantageous settlement as even more important for its long-term global interests than it had originally thought. Correspondingly, the ROK might have had substantially lower true damage thresholds than the agreed joint damage thresholds but also see its citizens suffering far more than it had anticipated as the damage spiraled upward. Thus it could want to revise the joint thresholds downward.

This kind of divergence could seriously compromise Alliance cohesion and possibly even split the Alliance completely. Strains could initially manifest themselves simply as arguments made by both sides in favor of the opposing revisions they want. Such fundamental and diverging differences could seem irreconcilable. Fearing an open fission in the Alliance, the ROK and the US might continue to go through most of the motions of being good allies but start acting more independently.

The ROK might reduce its offensive efforts against the DPRK and secretly investigate the possibility of making a separate settlement with the DPRK. In contrast, the US might boost its efforts to cause the DPRK to submit. US retaliations and escalations for the damage the DPRK would still be imposing on the Alliance could begin to exceed the upper limits of what the ROK had agreed to or has come to see as appropriate.

If the Allies cannot find a way to bridge their differences, their actions might diverge completely. Continuing with the above example, the ROK might offer a separate settlement to the DPRK that it would find favorable enough to accept. The ROK might even capitulate if the DPRK would accept nothing less, and the ROK decided it could not tolerate further damage.

If the Alliance were to break up in this way, the US would have to decide whether or not to continue the conflict by itself. Defeating the DPRK or even winning a better settlement without the help of the ROK would be far more difficult. Allowing the DPRK to win control of the ROK as a result of its having used nuclear weapons would be a particularly difficult outcome for the US to accept, especially given its non-proliferation interests. Proliferation might be expected to accelerate if the DPRK manages to win an important gain as a result of its use of nuclear weapons, especially if that gain had been strongly defended by the US. Still, it is far from obvious what the US might decide to do in this unprecedented situation.

Finally, note that if the US somehow chose to leave the Alliance during the war, the ROK would have no choice but to capitulate given its dependence on US nuclear weapons. Given the US global security interests and the disastrous effect that abandoning an ally to nuclear defeat would have on the many other US alliances, it seems most implausible that the US would do this.

In any case, this example illustrates the dangers posed to the Alliance if the Allies have individual damage thresholds that differ greatly. It further illustrates the great value for both Allies of being convinced that their Alliance is of fundamental and enduring value for both. This conviction can help to motivate the sincere behavior required if the Alliance is to have its best chances of staying together despite the stresses of limited nuclear exchanges with the DPRK, and of achieving a tolerable if not preferable outcome to the war.

Still, the Allies may really have very different individual damage thresholds for the possible outcomes to such a war. They may know this before the war, or suspect it if the Alliance is not particularly sincere in its internal dealings. And they may anticipate that individual damage thresholds honestly exchanged before a nuclear war might change under the great stresses it would impose. How might they deal with such possibilities?

Sorting out how to best engage in such a conflict when major differences in the Allies' individual damage thresholds are suspected but not evident is an extraordinarily complex analytical challenge. Still, we can make two relevant observations. First, if the Allies are suspicious of one another, they would want to hedge against collapse of the Alliance to the extent they can. We will address how the ROK might hedge below. Second, if the DPRK suspects that the Allies might have very different damage thresholds from one another, it will have additional encouragement to attempt to determine which Ally has the lower thresholds and to concentrate its attacks on that Ally.

If the Allies know or strongly suspect before the war that they have irreconcilably different individual damage thresholds, they should consider a radically different security relationship. Thus, if the ROK knew or strongly suspected that the US has substantially lower damage thresholds for the various plausible settlements than it does, it would see advantage in getting its own nuclear forces in order to enforce its higher thresholds. But it would still have good reasons to preserve the Alliance. Beyond the many useful military capabilities of the US, the ROK would see substantial value in making the DPRK worry about the possible need to defeat both Allies. We will develop this last idea in more detail below.

Alternatively, if the ROK saw the US as having damage thresholds that are substantially greater than its own, it would again see advantage in getting its own nuclear weapons, but in this case would see strong incentives to end the Alliance, so long as it sees itself as having comparable or greater military strength than the DPRK. The ROK would be choosing to defend independently in order to be able to end the war with less damage than alliance with the US and the resulting higher joint damage thresholds would call for.

The US would seem likely to be strongly opposed to any security arrangements with the ROK that are greatly different from the long-standing Alliance between the two. It seems most likely to continue to oppose strongly any nuclear proliferation by the ROK. US general nonproliferation interests would call for continued opposition as would specific US concerns about the proliferation dynamics that might be set in motion across Asia. In addition, the US would worry that circumstances might arise in which a nuclear armed ROK would feel compelled to initiate the use of nuclear weapons despite strong opposition by the US – opposition based on its judgments that a successful conventional defense of the ROK was still possible and that the burden of starting a nuclear war should be left on the DPRK.

But if the ROK is not to have nuclear weapons, it must remain in the Alliance. Thus the US should see strong incentives to somehow guarantee that its individual damage thresholds are not significantly different from those of the ROK. That, of course, may simply be impossible, in which case, the ROK could see strong incentives to get its own nuclear weapons and perhaps to end the Alliance as well. This takes us to the third and last of our alternative security arrangements between the US and the ROK.

STRATEGIC RELATIONS RATHER THAN ALLIANCE

Suppose then that the US/ROK Alliance has been completely dissolved. We assume that the two states have made clear to each other that there are no longer any specific security obligations either is expected to honor. The two might share information when they see it to be in their interests, but each might also misrepresent its intentions if it sees any advantage to doing so. In these circumstances, the US and ROK security relationship would be purely strategic.

Nonetheless, the US would presumably still have strong reasons to oppose any DPRK attempt to gain control of the ROK by force. Thus, we will again assume that US objectives would be to see the DPRK totally defeated, or if not that, subjected to a settlement of the war under which it would be disarmed and effectively dominated by the

ROK. But the US would not be formally obligated to contribute to these ends, nor to pursue them in any particular way.

We further assume that the US would not want to risk the extra costs and difficulties of seeing the ROK defeated and then having to intervene to reverse this outcome. Thus, the US enters the war as soon as the need for its help is clearly evident. We also assume that the US chooses and implements its contributions well so that the ROK has little serious motivation to turn against the US, except possibly if it were to find itself in circumstances where it wants to accept an outcome to the war that the US opposes.

Finally, for purposes of this case study, we assume that China would have no interest in getting involved in a war on the Peninsula that seems likely to escalate to the use of nuclear weapons. Thus, it does all it can to discourage the war, including making it clear to all, especially to our desperate and aggressive DPRK, that it will stay on the sidelines once the shooting starts. In the next section, we will drop this assumption to examine the implications of both of the major powers being willing to intervene.

In order to illuminate the differences that ROK nuclear weapons could make, we will assume initially that the ROK does not have its own weapons. In such circumstances, the ROK would find itself essentially an instrument for the other combatants.

The most important contribution the US would need to make to the defense is nuclear protection for the ROK. Nuclear retaliation against the DPRK and perhaps even some escalation would be necessary to enable the ROK to stay in the war, and to drive the DPRK toward defeat. Note that the US could feel very uncomfortable carrying out nuclear strikes on behalf of a state no longer close enough to be an ally.

Keeping the ROK in the war could greatly reduce the costs to the US of ensuring a decisive victory over the DPRK, which would likely require that the DPRK be invaded and occupied. The possible burdens for the US of implementing the other possible settlements would also be far lighter if the ROK remains capable of playing a strong role.

The ROK could also be important instrumentally to a DPRK that is collapsing for economic reasons, if it can be captured reasonably intact. Alternatively, if its capture proves impossible, a largely intact ROK might offer a settlement to the DPRK that is superior to what the US might offer and press the US to accept it. Finally, if an enlightened DPRK were to find itself being totally defeated, it could see a more intact ROK as more capable of helping the DPRK recover, and might prefer occupation by the ROK to whatever the US might do in its stead.

Under these circumstances, as soon as the US enters the war, the DPRK might make a nuclear strike on it to see if it can be quickly driven out of the war, or at least forced to accept an outcome that the DPRK views as relatively favorable. The DPRK should expect at least commensurate retaliation. Still, it might hope the US would be so shocked by being struck with a nuclear weapon that, once it had retaliated, it would decide that the defense of a ROK, with which it was no longer allied, was not worth any further suffering. If the US withdraws, the DPRK can freely demonstrate its willingness to use nuclear weapons against the ROK, if necessary. Facing an opponent willing to use nuclear weapons against it, an ROK unprotected by likely retaliation in kind would see its defeat as assured and soon cease to resist.

If the US chose instead to continue to oppose a DPRK victory, as seems more likely, the DPRK would face a more difficult choice of how to allocate its damage. It might hope that a very few nuclear strikes against the ROK could create great pressure on its government to end the war immediately, even at the cost of accepting defeat, and to call on the US to cease all further military action. Given the lack of influence the ROK is likely to have (since it would lack nuclear weapons and no longer be an ally of the US), this DPRK strategy seems unlikely to succeed.

The DPRK might also consider that by attacking major ROK military capabilities it could render the ROK incapable of invading and occupying the North, and by doing so greatly raise the costs to the US of achieving decisive victory. It might hope that these increased costs would lead the US to support settlements the DPRK finds more attractive.

The paradox is that the more the DPRK attacks the ROK civil structure, the more it reduces the value of the prize it is trying to win. And as mentioned above, if the DPRK is going to lose in the end, it might prefer occupation and rule by the ROK to the arrangements the US might make if the ROK military had been damaged so badly that it was no longer capable of performing these missions.

Note that given its purely strategic relationship with the US, how the US might help to defend the ROK would depend solely upon US values and choices of which the ROK could be substantially uncertain. In particular, large differences between US and ROK damage thresholds could be costly for the ROK and could lead it to seek to escape from its instrumental role in the conflict. The obvious solution if it wants to or has dissolved the Alliance would be to acquire its own nuclear weapons.

Let us assume then that the ROK has acquired its own nuclear forces. We further assume that the ROK military would have a reasonably good chance of defending

successfully if the war were fought only with conventional forces. As before the US would have good reasons to want to ensure the defeat of the DPRK and could be expected to take military action toward that end if needed. Now, however, the US would see the ROK as capable of carrying much more of the burden of defending itself.

In these circumstances, the US and the ROK would each prefer that the other lead the campaign against the DPRK and bear the greater share of the damage. But the US can assure an advantageous division of the burdens as long as it wants, since it is assumed to have only a strategic relationship with the ROK, the ROK can retaliate independently for DPRK nuclear strikes, and the ROK has no choice but to defend itself.

Note that because it has its own nuclear forces, the ROK would be in a far better position to achieve the war outcomes it prefers. If the US damage thresholds were lower than those of the ROK, the ROK can pursue the conflict to its higher damage thresholds and have a better chance of winning a settlement it sees as preferable.

If US damage thresholds were higher than those of the ROK, a nuclear ROK would have other advantages. It could, for example, veto any US invasion of the DPRK, and thus provide a credible guarantee of DPRK security once it and the DPRK had agreed to a settlement. Given their agreed settlement, the DPRK and the ROK could then collude to “defuse” whatever reasons US might have to continue the war beyond their desired settlement point. These possibilities can give the DPRK stronger incentives to make a deal with the ROK.

Finally, note that if the ROK were to insist on dissolving the Alliance and having real security independence, the US should make it very clear whether and how its subsequent relationship with ROK might imply some residual security guarantees in special circumstances. The US might, for example, want China to understand that if China chose to support a DPRK attack on the ROK, the US can be expected to intervene on behalf of the ROK. In the next section, we will examine the consequences of possible intervention by China.

CONCLUSIONS

In this section, we analyzed the consequences for war on the Korean Peninsula of three different security relationships between the US and ROK. These relationships can be viewed as three examples drawn from a continuum of possible relationships, starting at one end with the “perfect” alliance analyzed in Section 3, and ending with the last case considered in this section, of “strategic relations” between the US and ROK. Moving

along this continuum, each case featured less and less coordination of decision-making between the US and ROK, resulting from larger and larger differences in interests between the two.

The first case featured "sincere relations" between the two Allies, wherein minor differences in interests (in the form of disagreements about what the Alliance's damage thresholds should be) could be sincerely negotiated, so that the Alliance could present a unified front in the war. A suitable compromise could be reached, and adhered to in war, because small differences in preferred damage thresholds would seem less important than the long-term value of preserving the Alliance.

Nevertheless, these differences would mean that, for any given settlement, one ally's damage threshold would be lower, and so the DPRK would do better by concentrating its attacks on that ally. Since the DPRK may not know which Ally has the lower threshold, it faces an additional source of uncertainty to the several we discussed in Section 3 (e.g., uncertainty about the effectiveness of Alliance missile defenses). Thus, while the best strategies of the Allies are substantially the same as those in Section 3, the DPRK has yet another difficult choice to make: which ally to attack.

In the second case, we allowed for the possibility of "insincere relations" between the Allies. Here, substantial, uncertain differences in interests created incentives for misrepresentation on the part of the Allies. This in turn undermined the possibility of honest negotiations over Alliance damage thresholds.

This situation posed special dangers for the Allies. As the damage mounted in war, each ally would discover not only the true limits of the DPRK's willingness to tolerate damage, but also those of the other ally. If differences in interests are discovered in this way, the Alliance might only need to renegotiate the previously agreed damage thresholds. But if the revealed differences are large enough, the Alliance could fracture, with the Allies behaving more independently. One Ally might press on in the war, as the other sought a separate settlement with the DPRK or abandoned the conflict altogether. More insidiously, the need to reconsider agreed damage thresholds in the face of a nuclear war could seriously deteriorate the ability of all sides to control its severity.

If these differences were anticipated, the Allies might prefer to alter the terms of their relationship in peacetime. If it anticipated the possibility of abandonment, or the need to strengthen its bargaining power in negotiating Alliance strategy, the ROK would do best to acquire its own nuclear weapons. Of course, given its regional and global interests in non-proliferation, the US might then downplay differences in interests in

peacetime so as to assuage ROK concerns. But this would only contribute to the potential for a wartime fracture.

In the third case, we assumed that major differences in interests had led the ROK and US to dissolve their Alliance in peacetime, enabling each to pursue an independent strategy in the war. Without its own nuclear weapons, the ROK might be reduced to little more than an instrument for the other combatants. For the US, a reasonably intact ROK would make it much easier to invade and occupy the DPRK, and in the case of a more limited settlement would bear much of the burden the North's recovery. For the DPRK, the ROK is a potential invasion route for the US, but also the key to resuscitating the North's economy.

If the US did abandon the conflict, the ROK would be left unable to respond in kind to nuclear strikes, and so would find itself at the mercy of the DPRK. It therefore has good reason to acquire its own nuclear weapons. Though it might then have to bear more of the burden of defending itself, it would also be more able to ensure that its interests were respected in the conduct and outcome of the war.

This last case can be considered a trilateral war, in the sense that it features three actors, each making independent decisions. Of course, in a war over the future of the Korean Peninsula, there is at least one other state which might have a compelling interest in the outcome of the war, and therefore might intervene in it. In the next section, we consider a quadrilateral war, in which China is added to the set of potential combatants.

5. POTENTIAL INTERVENTION BY BOTH THE US AND CHINA

We turn now to the question of how both the US and China might intervene if the DPRK were to attempt to invade and defeat an independent nuclear-armed ROK. We will explore the intervention strategies the two major powers might choose in light of their concerns not to engage each other directly in a nuclear war, how the ROK and the DPRK might react to such interventions, and how anticipation of such interventions might affect the behavior of all four states even before the war starts.

We limit ourselves to the case where the US and China have opposed, partisan, strategic interests in the conflict. "Strategic" here is used in the sense of the last case in the previous section; it means that both the US and China are assumed to have no formal obligations to defend either the ROK or DPRK. By "partisan", we mean that each major power wishes to ensure that the war ends well for its favored Korea, and by "opposed", we mean that China favors the DPRK while the US favors the ROK.

Note that there are other possible configurations of the major powers' interests. For example, suppose that the DPRK alienates China and poses so clear a risk to the strategic balance in Asia that both China and the US judge that they cannot allow it to defeat the ROK. In this case, the major powers' interests would be strategic and partisan, but common rather than opposed. Alternatively, both powers could conclude that their concerns for the outcome of the war were dominated by the desire to eliminate nuclear weapons from the Peninsula, rather than favoring one combatant or the other. Here, the major powers' interests would be strategic and common, but non-partisan. These other cases are plausible, and may give different results than the ones discussed here, but we must defer them to future research.

We begin by exploring how an initial war between the DPRK and the ROK might evolve. But in contrast to our investigation at the end of the previous section, we will not assume that the US is so committed to ensuring an outcome it prefers that it enters the war as soon as the ROK clearly needs help. Rather, as we will explain below, the US and China, having no obligations to assist either of the two Koreas, would remain on the sidelines until the initial war has been largely settled.

INITIAL WAR BETWEEN THE DPRK AND AN INDEPENDENT, NUCLEAR-ARMED ROK

For purposes of this investigation, we will assume initially that neither of the initial combatants expect intervention by the US and China, and that none occurs before this conflict has been concluded. Given this assumption, we will see how such a conflict might evolve and be settled, and thus the conditions that might then motivate and confront interventions by the US and China.

How then might a potential war between these two small nuclear-armed states differ from the previous cases of war on the Korean Peninsula that we have examined? To illustrate some of the more important potential features of the initial war, we first contrast it with one of the examples we have already examined: war between the DPRK and the “perfect” US/ROK alliance.

A first difference is that the more global concerns of the US would not be a factor in determining the damage level that might be reached in this war. Neither the DPRK nor the ROK would need to preserve its reputation as a reliable ally. Neither seems likely to be significantly concerned about setting the right global precedent for conflict between nuclear-armed states. Both seem likely to regard their nuclear disarmament as a much less important issue than the nature of any political agreement they might reach to end the war.

A second significant contrast is that common interests such as the ethnic identity of the two Koreas and their shared history and culture would be more salient. And both Koreas would be concerned about their reputation with each other, and determined to save face and avoid appearing subservient. Of course, their reputations with one another would matter only if the initial war did not lead to the permanent end of one of them.

Perhaps the most important contrast is that this conflict solely between the DPRK and the ROK, both nuclear-armed, would risk the end of the combatants; i.e., both would see the stakes as existential. At most a few tens of nuclear weapons detonated on the more important cities, industrial facilities, and key military bases of either state seems likely to cause more than enough damage to render either nation incapable of continuing the conflict. And absent some major technical advances to which these states would have access plus some major investments, it seems likely that neither would be able to build defenses or preemption capabilities good enough to protect it from such nuclear destruction.

Of course, the US could not be defeated in this way by the DPRK, and in this sense, the DPRK cannot totally defeat a US/ROK Alliance. The US and the Alliance would be defeated if they were to suffer more than their maximum damage tolerances for resisting defeat, and substantially more damage were possible. But that defeat would be by choice, at least by the US, not as a result of losing the physical capability to carry on. Whether a strong state has the *will* to endure an exchange of nuclear damage long enough to defeat a weak state is a different question.

Our example war between the ROK and the DPRK pits two weak states against one another. And it is possible they could both prove willing and able to destroy or exhaust each others' physical capability to continue. Both would then need foreign help, and might call first on former patrons: China for the DPRK and the US for the ROK.

Furthermore, while mutual exhaustion is a possibility, it is also possible that one side in the initial conflict might prove to have substantially higher damage thresholds than the other and thus be able to force its opponent to accept a most disadvantageous settlement if not total defeat. We'll look at such a possibility in some detail because it presents a plausible path to intervention by both the US and China. It is easy to imagine that total defeat of one of the original combatants could prove unacceptable to one of the major powers and that the other major power would also intervene with the goal of ensuring that any new settlement takes its interests into account.

Consider then the example conflict shown in Figure 5.1. Note that in this example, all the DPRK damage thresholds exceed all the ROK damage thresholds. Thus, as the damage mounts for both sides the ROK offers every settlement but the DPRK refuses them all and the ROK surrenders. Note that the figure shows the ROK begins the war short of the damage capability needed to force the DPRK to accept defeat or regime change with immunity for the leadership. It does start with sufficient damage capability to drive the DPRK to accept the long-term confederation or disarmament and development settlements, but does not have the will to accept the damage that the DPRK would impose before accepting either of those. Of course, short of espionage, the ROK would not know the DPRK damage thresholds.

Note also that the damage level at which this initial war would end determines how much of each side's initial capability to impose damage would have been expended. In this example, the DPRK would expend nearly 80% of its damage capability in defeating the ROK, while the ROK would be willing to accept defeat having expended 70% of its initial damage capability.

This outcome poses a number of important questions. What dangers and opportunities might China and the US see in intervening in this conflict to change its outcome and how might they go about intervening? How might the DPRK and the ROK react to such interventions? How might DPRK and ROK change their strategies for the original war if they were to anticipate the possibility of such interventions?

BEST STRATEGIES FOR INTERVENTIONS BY THE US AND CHINA

Recall that we have assumed that the US and China would have opposed, partisan strategic interests in the outcome of this example conflict – the US favors the ROK and China the DPRK. As we have assumed that neither of the major powers is obligated in any way to come to the aid of either the ROK or the DPRK, the major powers would determine their strategies without regard for the interests of their former allies.

The most important strategic interest for both China and the US would be to minimize the damage they impose and suffer in seeking a tolerable settlement of the original war between the two Koreas. For both, the greatest risk of damage would arise if their interventions somehow led to exchanges of nuclear strikes against each other. A smaller risk would come from the possibility that their interventions would result in exchanges of nuclear strikes between each of the major powers and the Korean state it opposes, and even exchanges with both Koreas—if both so resent the interventions that they join forces to resist.⁹ Note that the first risk is existential for China, and perhaps also for the US, and perhaps very unlikely. The second risk is not existential for the major powers but might be more likely.

⁹ It is interesting from a theoretical point of view to consider the possibility that two weak states, initially antagonists, might band together to resist the intervention of an outside power. But we repeat our earlier view that the idea of the ROK joining with the DPRK to launch nuclear strikes on the US is beyond the realm of plausibility, at least in the foreseeable future.

Nuclear War between DPRK and Independent ROK

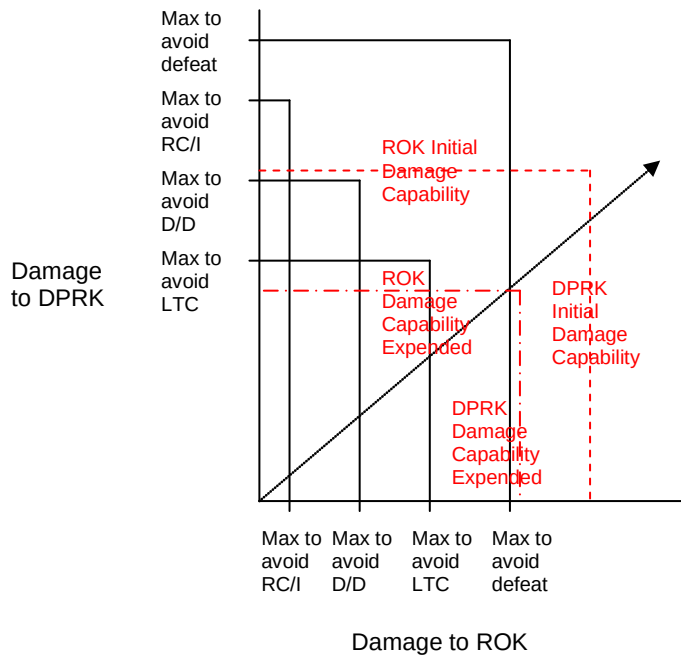


Figure 5.1

These major power concerns and the purely strategic views we have assumed they would have of the war suggest that any intervention to change the outcome of the initial war should occur as late as possible. At the end of a war, especially one that had led to a lopsided settlement or even total defeat for one of the combatants, the two sides could have reached a very high damage threshold, meaning that both would have suffered tremendous damage.

This damage would imply corresponding reductions in their readiness to take more damage, at least until they had substantially recovered from their war. The combatants would also have had to expend however many of their nuclear missiles were required to reach the damage levels at which the initial war had been settled. Thus, the capabilities of both original combatants to impose damage and to resist accepting new settlements as a result of external interventions could have been greatly reduced.

Another advantage for the major powers of waiting is that the resulting outcome of the initial war might prove to be a tolerable settlement. This would eliminate any need for intervention and allow both powers to avoid the inherent risks.

Of course, if intervention were required neither would want to wait much longer. If the loser of the original war had been totally defeated, waiting could give the winner the time needed to gain control of the loser's remaining nuclear forces. Thus, at the end of our example initial war the US would not want the 30% of the ROK nuclear force that was not used to become available for use by the DPRK. China might feel the same if it saw that with both the major powers intervening, a final outcome of the war tolerable to both would be virtually assured. It might then reason that any further use of nuclear weapons by the DPRK would be wasted suffering and, given the high likelihood of US retaliation, dangerous.

Clearly, securing or destroying left-over ROK nuclear weapons would be a very high priority for the US. Note that the DPRK might try to get the ROK to transfer these weapons to it quickly by offering some sweetening of the settlement ending the original war. In the event that the ROK were to see intervention by the US as even more objectionable than the settlement that had been forced on it by the DPRK, it might go along with such a transfer.

It seems more likely that the ROK would become aware of the impending intervention by the US and China and realize that a new settlement more to its liking might be reasonably assured. In this event the ROK would surely not agree to transfer its remaining nuclear weapons to the DPRK. And if the DPRK were close to seizing these weapons, the ROK might even enable their transfer to or destruction by the US. For obvious reasons, the US should not wait too long to make its intentions to intervene known to the ROK.

Figure 5.2 depicts the situation facing the US as our example initial conflict between the ROK and the DPRK comes to a close. The damage thresholds for the US are the same as those shown in the figure in Section 4. Those for the DPRK have been reduced from the ones in Figure 5.1 to account for the damage the DPRK would have already suffered. The figure shows the amount of DPRK damage capability remaining at the end of the initial war, as well as the sum of the damage capabilities remaining to it and the ROK.

Note that if the DPRK can somehow gain control of the remaining ROK nuclear forces it would then be able to impose enough damage to cause the US at least to accept disarmament and development, though of course it would not be sure of that. Clearly, the disposition of those forces is of great importance to all involved in the continuing conflict.

Assuming that US intelligence had been able to establish a reasonably tight upper bound on the size of the original combatants' nuclear forces, and that the evolution of the initial war had been observed in some detail, the US would have a reasonably tight bound on the remaining DPRK nuclear capabilities. It would likely be impressed at how tough the DPRK had turned out to be in neither offering nor accepting any settlement short of total defeat for the ROK. It would not know the DPRK damage thresholds but would expect them now to be lower, and perhaps much lower, than they had been originally.

Note that the DPRK damage tolerances for resisting intervention by the US and China might not be simply the differences between its original damage thresholds and the amount of damage it had suffered in the original war. Perhaps the DPRK would be somewhat more motivated to resist offering or accepting the same settlements when the adversary is the US than when it is the ROK. But substantial differences seem unlikely as the fundamental nature of the war outcome seems likely to count for much more than who imposes it.

Finally, note also that the US would find itself in an even more advantageous position to intervene if it has substantial missile defenses and preemption capabilities that would further limit DPRK capabilities.

Potential Subsequent War between US and DPRK

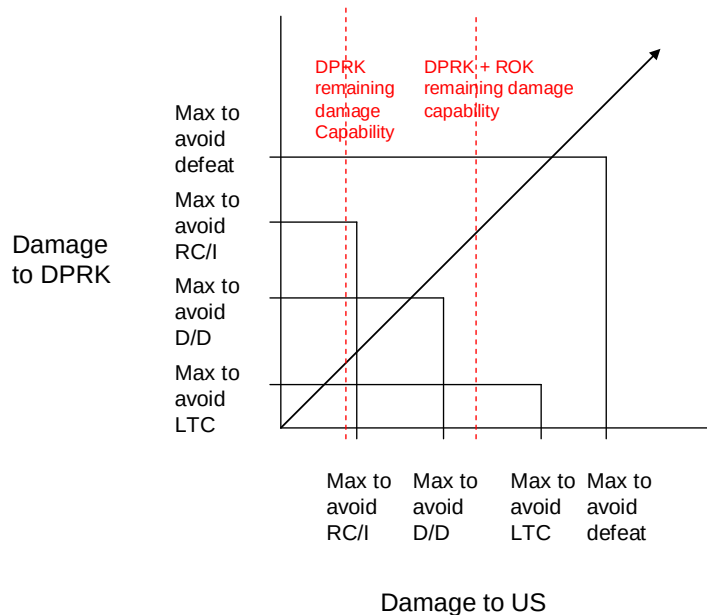


Chart 5.2

China should have a generally similar picture of the dangers and opportunities of its intervention. Given its partisan interests, it would want to ensure that the DPRK is not totally defeated as a result of a US intervention. Given the dangers of direct combat with the US, it might propose that it and the US intervene in the DPRK and the ROK, respectively, to ensure that both these weak states survive. Once these states had been pacified, China and the US would oversee the disarming and possible unification of the two Koreas under mutually acceptable political arrangements.

Given the circumstances at the end of our example initial war – a DPRK thinking it had won complete control of the Korean Peninsula and a ROK so depleted as to be of limited assistance—the US might be inclined to accept China’s proposal. If the US were to refuse to cooperate, China might become even more convinced that it could not allow total defeat of the DPRK and the imposition of a single government for the Peninsula that would likely be shaped to serve a US savior’s interests. Thus, China might intervene despite US opposition.

It is also possible that China might decide to warn the US that it cannot accept the total defeat of the DPRK, to describe the kinds of settlements it might find acceptable, and then to hold off on its intervention until it became clear that the US would not accommodate its interests. China would have no obligations to assist the US in driving the DPRK to accept a new settlement between the Koreas that both the major powers could tolerate. Why not try to make the US bear all the burdens and risks of getting to such a settlement? If the US were to go too far, China could intervene then.

Note that if the ROK had defeated the DPRK in the initial war, the shoe would be on the other foot. In this case, a US with no formal obligations to the ROK might then hang back and let China do the work.

In making an intervention to oppose US efforts to impose total defeat or any revised settlement that China could not accept, China would want to minimize the prospects of direct engagement and especially nuclear war with the US, and the US would be similarly wary of direct conflict with China. Thus, we might expect China and the US to concentrate their military actions on the ROK and the DPRK and to limit any exchanges of nuclear weapons subsequent to their intervention to the original, weakened combatants.

Of course, the DPRK and the ROK should at least anticipate that any war that starts between the two of them could lead to intervention by the US, China, or both. To the extent they do so, they should alter their strategies to take these possibilities into account.

POTENTIAL EFFECTS OF DPRK AND ROK ANTICIPATION OF INTERVENTIONS

We continue to assume that China and the US would have the same opposed partisan interests examined above. Suppose for the moment that the DPRK and the ROK assume that the major powers would both intervene successfully if either the DPRK or the ROK were to suffer total defeat or even be forced to accept a particularly disadvantageous settlement. In this case, neither state should be especially worried that a war between them could lead to total and permanent defeat. But neither could expect to win big either.

Nonetheless, a desperate DPRK might still see an advantage in attacking the ROK. It could believe that the US and China might not intervene if the settlement

reached by the Koreans did not jeopardize either's existence. Thus, the disarmament and development and long term confederation settlements assumed possible in our example conflicts might be tolerable to both superpowers. And, a desperate DPRK might find them an improvement over its prewar circumstances.

A desperate DPRK might also win some sweetening of these agreements from the ROK if it were willing to test the limits of what kind of settlement the US and China might find tolerable, or if the ROK wanted a hedge against the possibility that the expected intervention by the US might not materialize.

Suppose that the desperate DPRK judges that US intervention might not be successful in undoing its defeat of the ROK or rolling back a settlement of great advantage to the DPRK. In other words, it thinks it might be possible that its damage thresholds and remaining capabilities to impose damage on an intervening US might be sufficient to drive the US above its damage thresholds for accepting the outcome that the DPRK had initially won.

To have the best chance of realizing this possibility, the DPRK should try to emerge from the initial war with as much remaining capability to impose and accept damage as possible. This suggests that it should set its damage thresholds for the initial war lower. And lower DPRK damage thresholds should lead it to accept a less advantageous settlement of the initial war. Thus, even the potential for US intervention should have useful stabilizing effects.

Note that the same kinds of arguments would operate on the ROK if it were to try for a settlement of the initial war that China would not tolerate. It would either have to limit its ambitions in that war, or arrange to have enough capability to impose and accept damage in a subsequent war to exceed China's damage threshold for accepting the original settlement won by the ROK.

Note also that the ROK and the DPRK cannot simply build more missiles and expect that alone to make a victory over their opponent stick. They must also emerge from the initial war with damage thresholds that exceed those the major powers would have for resisting acceptance of their victory. Of course, damage acceptance capability is not nearly so easy to build as nuclear-armed missiles.

Finally, let us suppose for the moment that the DPRK is not desperate, but opportunistic. In this case we would expect its decision to attempt to defeat the ROK by force to be calculated. And it would have to calculate whether it is ready to "pay twice" for the defeat of the ROK, first the missile and damage prices the ROK would extract

before accepting its defeat, and then the missile and damage prices a possible US intervener would extract before it also accepted the defeat of the ROK. Those calculations would be problematic given that the DPRK would be substantially uncertain of both the ROK and US damage thresholds.

CONCLUSIONS

In this section, we analyzed the consequences for war on the Korean Peninsula of the possibility of intervention by both the US and China. We dealt only with the case where the US favored the ROK and China favored the DPRK, but neither had any formal obligation to defend its favorite. We also assumed that the major powers would only consider using force directly against the opposing Korean state, and not against each other.

We identified two ways in which the major powers could be drawn in to the conflict. First, if the ROK and DPRK fought each other to the point of exhaustion, then each might call upon its former patron for assistance. Second, and more likely, if either major power anticipated that its favored state faced imminent defeat, or the imposition of a particularly lopsided settlement, it might intervene to coerce a more tolerable outcome.

Either way, each major power would face strong incentives to intervene as late as possible in the conflict. Obviously, a major power would not want to intervene if the outcome of the war would have been acceptable anyway, and will need to wait to judge whether this is likely or not. Additionally, the longer the DPRK and ROK are allowed to fight, the less capacity they will have to contest an eventual intervention, both in terms of remaining nuclear weapons to be used and also remaining willingness to suffer damage in pursuit of a particular outcome. Thus, even if a major power anticipated the need to intervene, the risks of doing so would decline as the war continues. Finally, if a major power's favorite state did so well in the initial war as to motivate intervention by the opposed major power, the first power still does best to wait. There is little point in intervening so long as the original combatant still has the ability and will to contest the intervention.

Still, the major powers should not wait too long. In particular, if the initial war ends, the victor may be able to assume control of its vanquished opponent's remaining nuclear forces, thus raising the danger to a major power of intervening. A major power that anticipated the need to intervene should take steps to ensure that its favored state's remaining nuclear forces cannot be captured in this way.

The anticipation of major power intervention by the ROK and DPRK should moderate the initial war. In choosing their strategies for the initial war, the original combatants must account for the possibility that whatever outcome is reached could be subject to revision by the major powers. The more the initial outcome favors one side over the other, the more likely is major power intervention to restore some balance. This undermines the incentives to pursue a lopsided settlement (or even total victory) in the first place. Either a state would see its victory taken away by the major powers, or it would have to "pay twice" to achieve victory: once to its original opponent, and again to the opposing major power. Thus, the DPRK and ROK would have less reason to pursue a decisive victory, and less reason to worry about the possibility of total defeat.

Even if one of the original combatants persisted in the desire for a decisive outcome, it would have to lower its damage thresholds for that outcome. This is because it would need to preserve its capacity to suffer damage (and also to inflict it) in pursuit of its goal in order to ensure that it could contest a major power intervention. A state cannot escape this tradeoff simply by building more nuclear weapons. Though this would raise the state's capacity for inflicting damage, it would not raise its willingness to suffer damage.

It is important to note that this moderating effect of the possibility of major power intervention will occur *only if intervention is sufficiently credible*. If the ROK and DPRK see the major powers as unwilling to risk nuclear strikes in order to alter the outcome of the initial war, then their original strategies for the war will not change. This has important implications for the stability of a more proliferated world, to which we turn in the next section.

6. GENERALIZING TO A MORE PROLIFERATED WORLD

This paper began with grand questions of proliferation, posed by Kenneth Waltz and others. Would a world in which more states possess nuclear weapons be more or less stable than the present one? Would this world be more or less favorable to US interests? How would the possibilities for alliances and third-party interventions affect these answers? This section discusses some of our work's implications for these questions, as well as the approach used to derive these implications and the difficulties inherent to the endeavor.

Our analyses of the case studies give us some grounds upon which to speculate about the nature of a more proliferated world. Sections 3, 4, and 5 described the effects of alternative *institutional structures* on the conduct and outcome of a potentially nuclear war on the Korean Peninsula. Here "institutional structures" simply refers to the various types of alliances we considered, as well as the cases where there was no formal alliance and where third-party intervention was possible. One can think of these institutional structures as alternative ways in which the involved states could organize themselves to pursue their interests, jointly or individually.

Our approach to studying a more-proliferated world as a whole is to consider some global or regional analogues to the institutional structures discussed in previous sections. For example, in Section 4 we discussed a structure in which the US had no formal alliance with the ROK, but sought to intervene in the conflict in order to ensure that the outcome would be tolerable to the US. Given the global scope of US interests, and particularly its interests in discouraging further proliferation and nuclear wars, we can imagine a world in which the US might intervene in *any* nuclear war, in order to ensure that nuclear aggression met with limited or no success. In this analogue, the US would be taking direct, global responsibility for the deterrence and defeat of nuclear aggression—it would be a "global sheriff."

Of course, there are many such analogues to the structures we have analyzed, and the world could be organized according to some mixture of these analogues; but we will limit ourselves to discussing just three. We chose these three because each seems quite different from the others, and because they struck us as particularly informative about the

difficulties and dangers of stabilizing a more proliferated world. However, we make no claims as to the likelihood that any of these particular structures are implemented.

The three structures are as follows. First, we consider the example given above, in which the US acts on its own, if necessary, to curtail nuclear aggression around the world. Second, we consider a world in which all or most of the major powers form a “global posse”, in which the responsibility for deterring and particularly defeating nuclear aggression is shared. Finally, we consider a world in which the smaller powers are left to fend for themselves and band together to provide for their collective security against nuclear aggression.

It is important to be clear about the limitations of this exercise. There are many steps involved in generalizing results from the specific cases analyzed in previous sections to the global scenarios discussed here. Along the way, many moving parts must be dealt with. To make the analysis tractable, we ignore some of these complications; in doing so, we also limit the scope of the conclusions that can be drawn. In the end, we can give only partial answers to only some of the questions posed above. Our view is that the exercise is nonetheless justified by the importance of these questions and by the potential to surmount these difficulties in future research.

We will concentrate on determining the stability of each of the three world structures, relative to the others. “Stability” here is taken to mean that states almost never initiate nuclear warfare; and if they do, they do so despite a very low expectation of achieving a profitable outcome. This is clearly a very narrow definition of stability. It ignores the possible dangers of arms races, the rise and fall of major powers, and many other well-known aspects of stability as it is traditionally construed. It also ignores some more novel aspects of broader stability, such as the potential for proliferation to non-state actors.

The virtue of this narrow definition is that it allows us to focus on the aspects of stability about which our work has the most to say. The principal limitation it imposes is that our assessment of relative stability is valid *only if other aspects of stability are held constant*. To be more explicit about this point, we must first review in a very general way what was done in previous sections.

One can think of Section 2 as building a device that, given some initial inputs, can calculate the evolution and outcome of a potentially nuclear conflict. These initial inputs include the set of countries potentially involved in the war; the capabilities of these countries, in terms of nuclear inventories, defenses, and preemptive strike; their interests,

in terms of the values they place on various outcomes; the information they have about the other states' capabilities, interests, and information; and the institutions that govern their interactions, such as alliances. The model developed in Section 2 tells us how these factors affect what happens in the war. Sections 3, 4, and 5 describe in detail how alternative institutions (e.g., different kinds of alliances) affect the war, while also varying the interests of the involved countries.

Now, the device built in Section 2 and used in later sections assumes that a conflict has already begun (because of a desperate DPRK). But it can also tell us how likely the *occurrence* of conflict is. In particular, it tells us how the countries that might be involved in a conflict would ascertain the likely outcome of a war. All other things equal, a state that expected a more desirable outcome would be more likely to go to war in the first place. Thus the device allows us to predict how likely it is that a conflict will occur in a particular situation. "Situation" here just means the set of inputs described above.

The world is characterized by a distribution of these situations. That is, in any given scenario, there are various places around the world where potentially nuclear conflicts could occur. The potential for conflict and the likely conduct and outcome of a conflict in each situation is predicted by our model as a function of the countries involved and their capabilities, interests, and information. So, given a distribution of these situations around the world, and all their associated inputs, the model could predict the expected frequency of nuclear conflict.

Of course, for any given scenario (including the current world), we do not actually know what values to use for all these inputs. At a minimum, we cannot specify in detail the interests of all the countries that might become involved in a nuclear conflict, and certainly not the information that is available to them about potential opponents' interests and capabilities. So, we cannot make claims like: "given Scenario X, we expect nuclear war to occur about once every ten years."

What we *can* do is to look at how the expected frequency of nuclear conflict changes between scenarios. To do so, we assume that the distribution of situations and all their associated inputs are the same in two scenarios, except for a change in one input (say, the institutional structure of the world). Then the model allows us to predict how this one change can be expected to *directly* affect the frequency of nuclear conflict, holding all the other inputs constant between the two scenarios. So we can make claims

such as: “we expect that nuclear conflict will be more frequent in Scenario X than in Scenario Y.”

This is what was meant earlier when we said that “our assessment of relative stability is valid *only if other aspects of stability are held constant*.” In what follows, we assess the relative stability of the three scenarios while assuming that the *only* differences among them are in the institutional structure. This ignores the possibility that different institutional structures could lead states to make different long-term choices about the capabilities they wish to possess, the way they view their interests, and so on.

So, for instance, in a world in which most states expect the US to intervene to curtail nuclear aggression, most states might elect to have at most a minimum set of nuclear capabilities. But in a world where small states band together to fend for themselves, they might opt for a fuller set of capabilities. The differences in desired capabilities between the two scenarios might alter the potential for arms races and thereby affect stability.

For the sake of tractability, we ignore these indirect effects of changes in the institutional structure. We analyze only the *direct* effect of institutional structure, and only the effect on our narrowly-defined stability. Having established the method we use to generalize earlier results, and its limitations, we turn to the three scenarios of interest.

In our first scenario, the US attempts to maintain stability, on its own when required, acting as a “global sheriff.” It must threaten—and, if necessary, carry out—interventions against any state that attempted to use its nuclear weapons for aggression. To the extent that the US can credibly do so, it can deter any nuclear aggression by ensuring that any gain from such aggression would be quickly reversed or minimized, at substantial additional cost to the “law-breaker”.

However, it is far from clear that the US would actually be willing to take on this role. In particular, it is difficult to imagine that the US would provide this guarantee to threatened states that were not friends or allies. As the last case of Section 4 makes clear, in order to substantially affect the outcome of a nuclear war, a third-party intervener may be required to employ nuclear strikes. In so doing, it must accept the risk of suffering nuclear retaliation in return, as one or more of the original parties to the war fights to preserve the original outcome of the war. Though stabilizing a more proliferated world might be very important to the US, would it be important enough to justify the risk of suffering nuclear strikes, especially when the beneficiary of the US guarantee is not even a US ally or friend?

Near-perfect combinations of missile defense and preemptive strike capabilities would mitigate this risk. A combination so exquisitely effective that the US could expect to suffer not even a single homeland strike might suffice to sustain the will of the US to act as the global sheriff. But even a very good set of capabilities, promising to allow only one or a few weapons to strike the US, might still leave the price of being sheriff so high that the US would not be able to afford it.

If the US had a near-perfect set of capabilities, aggressors would search for ways to subvert these capabilities. An aggressor might resort to covert or unconventional means of using his weapons, increasing the risk of proliferation to non-state actors. He might also seek to take a less well-defended friend or ally of the US as a hostage, imposing damage on this hostage with the hope of causing the US to back down. The results of Section 4 suggest that US friends or allies who anticipate this possibility might be led to separate from the US and acquire their own nuclear weapons in order to lessen their risk. A near perfectly-defended US also runs the risk of being perceived as too zealous in the execution of its duties, causing potential adversaries to form defensive alliances like those to be discussed in our third scenario.

Consider instead the second scenario, in which the US acts in concert with other major powers by forming a “global posse”. Clearly the US would prefer to share the potential costs of stabilization with other powerful states if at all possible. A credible, joint commitment by some or all of the major powers to putting down nuclear aggression would be an especially powerful deterrent. Together, these powers would have an enormous capacity to inflict damage on aggressors and possibly substantial defensive capabilities to limit their risk. Most potential aggressors would be incapable of putting more than a very small part of the interests of these powers at risk.

However, establishing the credibility of this commitment still poses problems. Our analysis of the “strong alliance” case in Section 4 suggests that making and sticking with joint decisions about when to intervene, and how far to push a given intervention, requires sincerity in the relations among the posse members. This in turn necessitates a strong commonality in the powers’ perceived interests. But of course the major powers may sometimes find their interests over a given conflict to be opposed, as in the discussion in Section 5 where the US and China favor different sides of the original conflict. The major powers might also differ in the effectiveness of their defensive capabilities, causing those with inferior defenses to anticipate suffering most of the strikes inflicted on the posse during an intervention.

The appearance of a concert of major powers acting to police the smaller states might also generate a backlash. Presumably, the posse would demand that states behave in accordance with rules set by the powers themselves, opening themselves to accusations of imperialism. This risks polarizing the world between large and small states and causing small states to seek their own weapons and band together to avoid “domination”.

Finally, the smaller states might band together to provide for their own security, either in the absence of major power guarantees, or to avoid the possible liabilities of relying on these guarantees. It is these states that would have the most to fear from nuclear aggressors, particularly those located in their regions. Getting their own nuclear weapons might eliminate the possibility of their being conquered by an aggressor or treated as a pawn by powerful interveners, but it might not always suffice to deter limited nuclear aggression.

The case analyzed in Section 4 where the US and ROK have strategic relations indicates an additional step small states could take to assure their security. That case indicated that an aggressor might have to “pay twice” for a given outcome: once to its original target and again to a third-party intervener. Anticipating this, the aggressor would be led to choose lower damage thresholds for the original conflict and might thereby be more easily rebuffed or deterred by the original target. Analogously, small states might band together to act as third-party interveners in case one was attacked.

This mechanism would work as follows. Suppose that a few small, nuclear-armed, status quo states came to perceive a growing threat from a potential aggressor in their midst. Suppose also that at some point the aggressor attacked one of these states and nuclear strikes were exchanged. During this war, the aggressor would have used up part of his nuclear forces and also his willingness to continue fighting. The other, unmolested states would then have incentives to take advantage of this opportunity, striking the weakened aggressor in order to lessen the future threat it might pose if it were to emerge victorious from the original war.

Of course, the aggressor would retaliate with at least commensurate strikes, so the degree to which each of the defensive opportunists took advantage of the situation would depend on the danger they perceived to themselves from the aggressor in the future. The higher the probability and severity of a future conflict with the aggressor perceived by the opportunist, the more damage it should be willing to suffer now to take advantage of the aggressor’s weakness. Put another way, the more dangerous a potential aggressor is

perceived to be in his region, the more he will have to worry about the possibility of having to pay twice, three times, or more for a favorable outcome in a war.

In peacetime, the status quo states would have strong incentives to advertise and aggrandize their fear of a potential aggressor and their willingness to take advantage of such opportunities. Though the potential aggressor might well suspect the other states of bluffing, calling their bluff would be extremely expensive. This is because the unmolested states would not intervene until the aggressor appeared to be making substantial gains against one of them. Thus the aggressor can only test their threats by suffering substantial damage and thereby undermining its ability to deter their opportunism.

This mechanism looks to be a very strong force for stability. In most cases, we would expect status quo states to outnumber violent revisionists by a substantial margin; the higher this ratio, the stronger is the deterrent effect. Notice also that, counter to intuition, this arrangement does not incite the status quo states themselves to engage in revisionism. The opportunists are able to contain the damage they suffer only because they do not initiate the conflict. So, a band of states cannot expect to use this mechanism for offensive gains because doing so would require that one of them must “stick his neck out first” and risk suffering much greater damage than during an opportunistic intervention.

CONCLUSIONS

This brief investigation of three alternative institutional structures for a more proliferated world exposes some interesting hypotheses. First, it would seem as though the US, and indeed the world, might not be able to get by simply by perpetuating the institutional structure left behind by the Cold War. Traditional US allies might remain assured by a US guarantee of extended nuclear deterrence, but what of all the other states that might reasonably come to fear nuclear aggression in a more proliferated world? In theory, the US could extend guarantees to these states as well, but the discussion above suggests that in practice these guarantees are unlikely to be credible and the US is unlikely to be willing to offer them as widely as would be necessary, short of truly exquisite defensive and preemptive capabilities. Absent these, the “global sheriff” might not be, and more importantly, might not be *seen* to be, reliable.

Second, the notion of a concert of major powers, acting together to preserve stability, may also be deeply flawed, and not just because of a possible lack of credibility analogous to that of the global sheriff. Though a “global posse” might well be consistent

with the charter of the United Nations Security Council, it might be just as prone to inefficacy because of differences in the interests of the major powers. If these differences are large enough, such a structure might even be a *source* of instability, as major powers intervened in opposition to one another in other conflicts!

Third, the absence or curtailed presence of these hierarchical structures doesn't imply nuclear anarchy. Kenneth Waltz argued that with respect to "the spread of nuclear weapons, more may be better," because the stabilizing effects of nuclear weapons on the relationship between the superpowers would be replicated among other nations. We have argued in this paper and in previous work that, given the uncertainties inherent to states' assessments of each other's interests, the potential for nuclear conflict in spite of its costs cannot be ignored. But even if Waltz's view was too rosy, the assay here of small states' ability to band together and deter aggression suggests that a more proliferated world need not be a nuclear free-for-all. If small states organized themselves in the way we have suggested, such a world could be relatively stable.

Though the limitations of the method we have used here to generalize our earlier results prevent us from stating any firm conclusions, we believe that these limitations can and should be overcome in subsequent studies.

7. GENERAL OBSERVATIONS AND CONCLUSIONS

We set out to better “understand conflicts in a more proliferated world.” Our study has only scratched the surface of this complex topic. Our approach to it has been to employ an idealized model of nuclear war to examine a series of hypothesized conflicts in which a desperate DPRK gambles that it can win control of the Korean Peninsula, or at least a tolerable settlement, by force.

The hypothesized conflicts were structured to provide insights into nuclear war with more than two states involved. We began by using our model to examine the nature of a pure bilateral war between the DPRK and a "perfect" ROK/US Alliance, where the Allies act as one. We then examined three case studies that moved toward trilateral conflict, in that the US and the ROK were treated as increasingly independent decision-makers. In the last case study, we considered a quadrilateral war that began as a war solely between the DPRK and ROK, but was subject to actual or anticipated interventions by the US and China. Finally, we generalized these results to provide a partial characterization of the overall stability of a more proliferated world.

Rather than recapitulate the specific results of previous sections, this concluding section will instead offer some general observations that are motivated by the earlier results.¹⁰

We begin with the observation that an objectively correct theory of nuclear warfare, if it exists, may be irrelevant. The problems of nuclear war have been subject to intense study in the US for over sixty years, and yet there is no consensus among researchers. Even if an agreement were reached, and US policy was set in accordance with the “correct” theory, there is no guarantee that other nuclear powers would arrive at the same theory. The best policy for the US must therefore depend upon the—potentially subjective—theories of the other nuclear powers.

Of course, the US is naturally inclined towards a theory, based on escalation-by-risk as that theory best suits a large, technologically advanced, militarily strong state. But we argued in Section 2 that a relatively weak state such as the DPRK could not be

10 For summaries of the specific results, we refer the reader to the concluding sub-sections of Sections 3, 4, 5, and 6.

expected to conceive of nuclear war in this way, especially if desperate. The theory of escalation-by-damage, or *limited* nuclear warfare, elaborated there is better suited to states like the DPRK, and it severely undermines the advantages of the US. Ignoring this possibility exposes the US to great danger, which only will grow if proliferation continues.

Even if the US and its opponent share the same conception, a limited nuclear war poses novel risks. During the Cold War, the US gained some experience of decision-making under the escalation-by-risk model, in the Cuban Missile Crisis and others. But neither the US, nor almost any other state, has any experience of suffering nuclear strikes. Making calm, careful decisions under such conditions would be exceedingly demanding for the US (and its opponent's) leadership. These leaders might be overwhelmed by panic in such a war; they might balk at the terrible damage that they might need to inflict to protect US interests; and they might lash out disproportionately in response to suffering nuclear strikes. These difficulties will only be compounded if decisions must be made jointly, with allies.

Obviously any war in which nuclear strikes *could* happen is extremely dangerous. In a war where nuclear weapons *are* used, however, states face the additional risk that they end up suffering—and possibly inflicting—more damage than was necessary to reach a tolerable outcome. *Disciplining the conduct of such a war will be of paramount importance.* There are conceivable steps to be taken, in peacetime and in war, to reduce the chance of an undisciplined war, but we leave the analysis of these steps for further work.

Despite the apparent chaos that a limited nuclear war might bring, one element of order survived in every case we considered. Regardless of how many states were party to a conflict, and of the relations between them, they could be partitioned at any given time into only two groups, with violence exchanged between groups but not within them. That is, there were always just two sides to the war. This is in keeping with the historical record of non-nuclear wars, of course, but it was conceivable to us that nuclear wars might be different. Perhaps the strategic independence that comes with nuclear arms might enable states to gain from forming a third side. We found no evidence to support this possibility.

Nevertheless, the order that bilateralism imposes in our case studies is quite limited. The composition of a side may change, as states within it exit the conflict or switch to the other side. New states may enter on one side or the other. The anticipation

of these possibilities adds new complexity to the relationships between states and the decisions they must make, and this complexity grows with the number of states engaged in the conflict. In particular, the best strategies for the combatants depend very sensitively on the number of independent states involved.

The best strategies are also sensitive to the relationships among states. States with quite similar interests in a potential conflict are best served by forming strong alliances and dealing with one another sincerely. It is only necessary that one of the allies possess nuclear weapons, as the others can rely upon it to defend their interests. But if two states have large differences in interests, then they are best served by remaining (or becoming) independent of one another, *but only so long as both have nuclear weapons*. Large differences in interest vitiate the rationale for alliance and thereby encourage proliferation, which in turn can lead to the dissolution of an existing alliance.

A particularly dangerous situation is when an existing alliance masks serious differences in interests. In the presence of such differences, relations between the allies are unlikely to be sincere, and so may blur each ally's perception of the other's interests. Revelation of the true differences may then occur *during* a war, and may severely undermine the allies' coordination, if not end it altogether. A fracture such as this would jeopardize the discipline with which the war ought to be conducted and so would represent a danger to *all* the combatants.

The possibility of insincere alliances exposes a difficult tradeoff for the US. As a provider of extended nuclear deterrence, now and in the foreseeable future, the US is vulnerable to the possibility of a dangerous wartime epiphany if it does not sever relations with allies with whom it has serious differences over the desired outcome of a conflict. Anticipating this, each such ally would just as soon not face potential abandonment—certainly not without nuclear weapons of its own. To stem this source of proliferation, the US must persuade the ally that the differences in interests between them are not so great. In so doing, the US increases its exposure to wartime fractures. Thus the US must choose a balance between its peacetime interest in non-proliferation and its wartime interest in maintaining only stable alliances.

If proliferation continues, the possibilities for nuclear wars will increase, and this balance must inevitably shift toward wartime concerns. In the limit, the US would do best to maintain only its very closest alliances. In the absence of exceptional defenses or preemptive capabilities, its nuclear guarantees will only remain credible with its closest allies. Moreover, we argued in Section 6 that the US cannot do better by extending

nuclear guarantees in concert with other major powers, because the likely differences in interests among these powers would undermine their credibility and potentially generate conflicts between them. In a more proliferated world, *extended nuclear deterrence may not suffice to suppress either proliferation or nuclear war.*

Compared to conventional warfare, or even massive nuclear exchange, limited nuclear warfare is much more a contest of resolve than of technology or size. This is precisely why it is so damaging to the influence of the US and other major powers. But it also acts as a great equalizer among the smaller states. In limited nuclear war, strategic depth is not solely about territory or economy, but about the will to suffer and inflict terrible damage. States that seem weak by conventional measures can band together, pooling their resolve, to resist a much stronger aggressor. These coalitions can create stability, not by promising to protect their members directly, but by threatening to take advantage of an aggressor that expended some of its own resolve in attacking and fighting a member.

These ideas give us modest cause for hope. A more proliferated world might still be stabilized: not by global guarantees of protection, but by local threats of opportunism. It is perhaps a less noble system, and certainly one in which US influence will be greatly reduced, compared to the one enjoyed by the West and its allies today. But if proliferation continues, it may have to do.

This study has been a revelation to us. It has shown us the depth of the complexity of a more proliferated world. It has strengthened our belief in the importance of studying such a world, despite the many analytical and empirical challenges. And it has reinforced our conviction that such study can be profitable.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YY) December 2008		2. REPORT TYPE Final		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Understanding Conflicts in a More Proliferated World				5a. CONTRACT NO. DASW01-04-C-0003	
				5b. GRANT NO.	
				5c. PROGRAM ELEMENT NO(S).	
6. AUTHOR(S) Andrew J. Coe, Victor A. Utgoff				5d. PROJECT NO.	
				5e. TASK NO. BB-6-2926	
				5f. WORK UNIT NO.	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Institute for Defense Analyses 4850 Mark Center Drive Alexandria, VA 22311-1882				8. PERFORMING ORGANIZATION REPORT NO. IDA Paper P-4426	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Net Assessment Pentagon Washington, DC 20301				10. SPONSOR'S / MONITOR'S ACRONYM(S) OSD(NA)	
				11. SPONSOR'S / MONITOR'S REPORT NO(S).	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT A new model of nuclear conflict is employed to explore the potential nature of a more proliferated world. The model postulates that deterring and winning a war with another nuclear armed state requires not just nuclear forces but also the capacity to absorb retaliations and escalations until the adversary is no longer willing to take more damage either to win the interest at stake or to defend it. The model is used to explore a series of case studies of nuclear war between two, a few, and finally an arbitrarily large number of nuclear armed states. We see that coalitions of small nuclear states can defeat a much larger nuclear state because they sum not only their nuclear arsenals, but also their capabilities to absorb damage and keep fighting. The incentives of small states to form such coalitions and honor their "obligations" appear strong. The potential formation of such coalitions should have a substantial stabilizing on a more highly proliferated world.					
15. SUBJECT TERMS alliances, case studies, China, coalitions, commensurate damage, defensive coalitions, desperate DPRK, escalation, highly proliferated world, Korean Peninsula, looser security relationships, conflict between more than two states, nuclear aggressor, nuclear conflict, perfect ROK/US Alliance, proliferated world, potential war outcomes, quadrilateral war, retaliation, strategy, trilateral war, US					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NO. OF PAGES 78	19a. NAME OF RESPONSIBLE PERSON
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (Include Area Code)



The Institute for Defense Analyses is a non-profit corporation that administers three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

