

Missile Defense Within the Context of Extended
Deterrence: The Uncertain Security Commitments
in the Korean Peninsula

by

Dahyun Choi

Department of Political Science
Duke University

Date: _____

Approved:

Kyle Beardsley, Supervisor

Bahar Leventoglu

John Aldrich

Thesis submitted in partial fulfillment of the requirements for the degree of
Master of Arts in the Department of Political Science
in the Graduate School of Duke University
2018

ABSTRACT

Missile Defense Within the Context of Extended Deterrence:
The Uncertain Security Commitments in the Korean
Peninsula

by

Dahyun Choi

Department of Political Science
Duke University

Date: _____

Approved:

Kyle Beardsley, Supervisor

Bahar Leventoglu

John Aldrich

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Master of Arts in the Department of Political Science
in the Graduate School of Duke University
2018

Copyright © 2018 by Dahyun Choi
All rights reserved except the rights granted by the
Creative Commons Attribution-Noncommercial Licence

Abstract

Previous research examining the deterring effect of a missile defense has asserted that such a missile defense—damage-limiting capability —decreases the vulnerability of the United States(US) and enhances its security guarantee to allies. However, the puzzle of how to convince the potential challenger of the credibility of the threat inherent in an extended deterrence remains unsolved. In this paper, I present a game-theoretic model of the deployment of a missile defense which I use to demonstrate the gap between the resolve of the US as inferred from that deployment and the actual security commitment of the US to the region in which the missile defense is installed. Such a discrepancy weakens the credibility of the US security umbrella and consequentially lowers the likelihood of success of a policy based thereon which is intended to compel the behavior of an opponent. This finding suggests that a damage-limiting force actually undermines the credibility of the deterrence threat in an extended deterrence setting.

For my parents who have spent most of their lives for my happiness

Contents

Abstract	iv
List of Figures	vii
Acknowledgements	viii
1 Introduction	1
2 Literature Review	3
2.1 Does a missile defense result in successful compellence?	3
2.2 Missile defense as a signal to directly involved states	6
3 The Model	10
3.1 The structure of the model	10
3.2 Information structure and beliefs	14
4 Empirical Implications	16
4.1 Perceived level of resolve and the actual level of commitment	16
4.2 Possible outcomes	21
5 Conclusion	26
A Appendix	28
Bibliography	30

List of Figures

3.1	The bargaining game with missile defense	13
4.1	The expected utility of North Korea when resisting	17
4.2	The effect of missile defense on North Korea's belief	18
4.3	The likelihood of war depending on the US resolve	19
4.4	War cost incurred by the US and recognized level of the US resolve	22

Acknowledgements

I appreciate the guidance from Professors Beardsley, Leventoglu and Aldrich. Their help has enabled me to successfully develop this thesis. Especially, Professor Beardsley's feedback regarding my thesis and career advice have provided a milestone in my academic journey. I also would like to take this opportunity to express my gratitude to Professors Malesky and Rohde whose insight and help have led me to be a better person during my academic life at Duke University.

Last but not least, I also hope to thank my friends who have been my side for the past two years. Especially, special thanks to my dearest friend B.H.E and I sincerely pray for their happiness and success.

1

Introduction

Can the deployment of a missile defense by the US in a state without nuclear weapons successfully compel a would-be nuclear proliferator to relinquish its nuclear weapons? Or conversely, does such a deployment actually undermine the credibility of the threat to use force, weakening the US guarantee of security to the state lacking nuclear weapons in the face of a challenger's possible attack? In July 2016, the decision to deploy THAAD (Terminal High Altitude Area Defense) system in the Korean peninsula was made public¹, and resulted in the resumption of the heated discussion on predicting future outcomes based on the deployment of a missile defense. Although the purpose articulated by the US for this deployment was to coerce North Korea into giving up nuclear weapons by incapacitating its nuclear weapons through the use of the THAAD system, thereby reducing the incentive for North Korea to provoke further, the incessant missile tests and brinkmanship conducted by Pyongyang even after the THAAD deployment raised doubts, as to whether the decision made by South Korea and the US has been effective, in terms of coercion. Despite the abun-

¹ The US moved key components of the THAAD battery in Seongju, South Korea and transported the AN/TPY-2 radar and other equipment needed to set up the THAAD unit.

dant amount of literature on the study of nuclear coercion, the question of whether a missile defense can pose a credible threat in an extended deterrence, in which one state shares weapons with its partner to demonstrate the former's commitment to the latter, remains unclear. On a broader level, this puzzle can be generalized to the question of compellence success; whether extending security guarantees to a nonnuclear state can spur a challenger with nuclear capacity to take action such as relinquishing its nuclear weapons.

In this article, I construct a game-theoretic model to examine the effect of missile defense deployment on the success of compellence. I exploit the deployment of THAAD system in the Korean peninsula as the basis for my model and examine the debates concerning the function of a missile defense as a signal of commitment to directly involved players. This paper contributes new findings to the scholarship on nuclear deterrence, especially that on extended deterrence, and advances the literature on the effects of investment in defense on the consequences of extended deterrence. To date, pre-existing research in this area has treated the threat of punishment and defense as complementary tools in augmenting a states security umbrella. However, my finding suggests that defense exploiting a missile defense could undermine the credibility of the punishment threat and eventually fail to disarm a potential challenger. My finding in this regard represents a contribution to heated and high-profile debates concerning defensive postures including missile defense within the context of extended deterrence and aids in the understanding of misgivings about weak commitments under such situation.

First, I briefly discuss the literature on coercion which forms the basis of my model. In the second section, I provide details on the game, and in the third section, I discuss the implication of my model for missile defense under a policy of extended deterrence. Finally, I state my conclusions.

2

Literature Review

2.1 Does a missile defense result in successful compellence?

Compellence refers to a threat on the part of one entity, which is intended to compel a possible challenger into taking actions that satisfy the coercer's intention (Sechser and Fuhrmann 2017, Schelling 1980). To be more specific, the coercer uses a threat of force in order to induce a challenger to change its action or behavior to one that the challenger would not have taken, were it not for the threat of force. If compellence is successful, it brings about changes in the situation, such as relinquishment and policy change. Scholars have generally agreed that the success of compellence hinges on the coercer's ability to project its will, the war cost incurred by both sides, and the credibility of the threat. Based on the last of these factors, a challenger gauges the coercer's resolve and decides whether to change its behavior (Schelling 1980). In sum, compellence is the ability to intimidate a potential challenger into backing down by increasing the cost the challenger will incur by continuing the behavior and to alleviate the damage a coercer may suffer if the compellence fails.

Yet, despite the recent emphasis on compellence, most research, with some exceptions, has been conducted in isolation and lacks the sophisticated assessment of

credible threats based on specific weapons (e.g., Fearon 1992, Bueno de Mesquita 1981, Morrow and Zorick 1997). This shortcoming is particularly troubling because the absence of empirical analysis on this issue has generated two divergent insights in the theory of compellence related to whether a missile defense system can be a credible threat capable of coercing a challenger armed with nuclear weapons to relinquish those weapons. Camps of researchers have developed different beliefs concerning compellence theory based on these insights. One camp believes that a missile defense enables a deployer to look more resolute and increases the maximum risk of attack that a deployer is willing to endure. In the same vein, deploying a missile defense, according to this camp, reduces the expected political and military utility of the challenger's preemptive strikes and increases the challenger's war costs. Thus, a missile defense undermines an aggressor's confidence of winning a war against a deployer, and as a result, the challenger chooses to back down rather than risk a nuclear war (Hynek 2010).

However, even if a missile defense reduces the military utility of the challenger's preemptive strikes, it may not be the primary determinant which decides the success of compellence (Drell et al. 1984). The other camp posits that a missile defense may lead a potential challenger to believe that a deployer is attempting to deprive the challenger of a retaliatory capability. Since deploying a missile defense affects not only the challenger's incentive to strike first but also the reliability of the defender's retaliatory ability (Powell 2003), the military balance between these opposing states collapses and mutual deterrence fails. Consequentially, a possible challenger may become afraid that a defender may commit a preemptive strike and invest more of its resources to build up its military capacity. Snyder (2015) alludes that if compellence fails as a preventive measure, it may stimulate a potential challenger to impose more severe damage upon a defender. Other than this unintended effect, a missile defense also hampers the strategic stability since the uncertainty now present eliminates

incentives for both sides to relinquish offensive capabilities (e.g., Lennon 2002; Wirtz and Larsen 2001; Ivanov 2000; Quackenbush 2006). The increased possibility of accidental war cannot be overlooked either.

The primary danger associated with exploiting a missile defense under the aim of compellence is that a potential challenger would end up putting priority to counter-measure necessary to assure its safety and enhance its capability further, deploying more warheads or nuclear weapons that might be needed to equal or exceed the level of capability possessed by the defender. Under conditions of incomplete information, a challenger may be unsure whether a deployer will go to war or be satisfied with the status-quo ante, meaning that it is unsure of a deployer's level of willingness. In such cases, states may choose to gamble that risk-taking behavior because it may be more rewarding when the action that wages war can seem to be pareto-optimal even if the outcome caused by such action, war, is pareto-inferior.

The situation is rendered complex when a deployer places its missile defense in the territory of its partner with the avowed intention of providing a security guarantee. Using a credible threat to protect its partner is often more complicated than deterring a challenger from engaging in a direct attack (Pifer et al. 2010). How is a deployer able to make a credible threat of the use of force against a challenger, in response to the challengers attack on the deployers partner? It sounds more plausible that a deployer would use force against a challenger in response to a direct attack on the deployer than in response to an attack on the deployer's partner. Addressing this issue, in recent work, scholars have argued that, clear signals such as explicitly effective damage limitation capabilities are required to inform a challenger about the deployer's high level of resolve (Acton 2009). However, the defensive nature of a missile defense leads to uncertainty concerning whether a defender is willing to become involved in retaliatory attacks after further provocations. Even if the effectiveness of the missile defense represents the level of resolve a deployer has

(Powell 2003), it is likely that the messages conveyed by a missile defense may also denote a relatively low level of commitment on the part of the deployer towards its partner. Such possibility implies that a missile defense intended for compellence may not be able to provide a firm security guarantee to the partner.

2.2 Missile defense as a signal to directly involved states

The unclear message denoted by the deployer's action of placing the weapon on its partner's territory is one of the factors that cloud the possibility of compellence success. If sharing weaponry is interpreted as a sign of weak commitment, the partner may begin to fear being abandoned by the deployer, who is the partner's ally. (Snyder 1984). Since the ally tends to suggest tentative commitment in order to maximize its bargaining leverage regarding the issue and avoids a firm commitment in an attempt to escape from being entrapped, any weapon deployed by the ally raises the question of its reliability to its partner.

This logic implies that further research should be conducted on whether a missile defense system can ensure the ally's offer of a security guarantee under the circumstance of extended deterrence. Some scholars believe that a missile defense would enhance the ally's guarantee of extended security to a nonnuclear state, raising the bar for that state to become a nuclear weapon state (Gavin 2012). However, the level of commitment remains obscure since the missile defense still leaves vulnerable nonnuclear powers unprotected. In this paper, I therefore address the question by focusing on the uncertain nature of missile defense and consider how a missile defense transforms the dynamic of compellence.

To address this question, I construct a model of a bargaining game based on incomplete information about the willingness of the ally. Regardless of whether a challenger is a rogue regime, the main parameter in the model is the level of the ally's resolve as conveyed by its credible threat of the use of force. Since the essence

of compellence is the level of resolution, the level of reliability conveyed by deploying a missile defense is a crucial measure which I can use to anticipate the outcome of extended deterrence.

Moving forward, I show that deploying a missile defense can weaken the credibility of the deployer's threat, which may be counter to a defender's intention, and I reveal how a missile defense —damage-limiting capability—leads to the success or the failure of compellence. During the process of tracking these changes, I introduce a signaling mechanism to examine in more detail the twofold epistemic process regarding a missile defense. First, if the deployment of a missile defense can serve as a reliable signal demonstrating the credibility of a defender's actions representing the defender's high level of resolve, the determinants of compellence success would be irrelevant to the ally's real resolve. Next, if the missile defense deployment fails to send a credible signal, thus implying that the deployer's resolve may be low, misinterpretation would instead lead to an underestimation of the ally's actual resolve, leading to a conflict. Of course, states may have an incentive to misrepresent the other side's willingness to fight as bluff and bluster (Fearon 1992). I therefore assume that the real resolution of the ally to provide aid to the defender may or may not be different from the perceived resolve indicated by missile defense deployment under the extended deterrence because the security guarantee is not about its territory but about the partner's land and this uncertainty may cause a state to overestimate or underestimate the signal.

The model I develop here builds on two eminent works to address the uncertainty that persists in missile defense. Fearon (1997, 1994) suggests that a threat can be regarded as credible, when the act of sending it engenders the costs that the sender would be reluctant to incur if its actual level of resolve to carry out the threat is low. In other words, to make the threat more credible, the signal must be accompanied by associated costs or risks to dissuade the other side. A potential challenger infers

a defender's resolve based on the cost to the deployer of emplacing a missile defense. Thus, my model incorporates the cost of deploying a missile defense and defines it as the sunk-cost signal. The second work on which I base my model is Powell (2003) who argues that a missile defense is the signal which represents the willingness of states to intervene in a crisis. As the effectiveness of a missile defense increases, the willingness of a deployer to intervene or provide aid to her partner in crisis appears more likely; the low level of missile defense effectiveness suggests the ally is more reluctant, therefore less likely, to get involved in a conflict. Based upon Powell's model, the game I construct in this paper also contains elements of uncertainty by acknowledging the probability of misperception, intentional underestimation, or overestimation about the ally's resolve.

This argument provides a significant initial step, but it stops short of actually measuring the effectiveness of a missile defense since the usefulness of missile defense cannot be exactly measured until it proves its capability against a potential challenger's preemptive strike. As Snyder (1960) explicates, we are dealing here with factors which are highly subjective and uncertain, not subject to exact measurement, and not commensurate except in an intuitive way. Thus, I address such empirical limits by deducing from the costs the defender bears when the country chooses the deployment by its ally of a missile defense in its territory as a signal.

I enhance the flexibility of my model by encompassing the possibility that the ally's actual level of resolve is not fixed but varies in response to the inferred level of resolve by other states from the ally's damage-limiting capability. In other words, if the ally is perceived as highly resolute by the would-be nuclear weapon challenger, the ally does not necessarily need to project its power which corresponds to its level of resolve in order to compel a potential challenger. In Powell (2003)'s model, the ally's level of resolve remains fixed, and the missile defense only makes a defender look more resolute. Yet, as Snyder (2015) suggests, the resolve of the player is influenced

by individual and systemic variables. Here, one of the variables that establish the ally's actual resolve is a would-be challenger's perception of the ally's commitment, which the challenger inferred from the effectiveness of the missile defense deployed by the ally. In this model, the ally's actual resolve is veiled since its revealed conviction is the outcome of a strategic calculation which can maximize its payoff.

3

The Model

3.1 The structure of the model

The most intriguing case of a test of the theory of extended deterrence regarding missile defense is the deployment of the THAAD system in the Korean peninsula. The South Korean government finalized its decision in coordination with the US to deploy THAAD last year. Both parties have publicly announced their intention to stop North Korea from pursuing nuclear weapons, and the THAAD deployment may be the prelude in their effort to compel North Korea to change its behavior. By analyzing my model of the interaction between South Korea and North Korea, I attempt to estimate the credibility of the commitment of the US to extended deterrence and the effectiveness of a missile defense under the purpose of compellence.

My model describes a bargaining process between South Korea and North Korea, the former in an asymmetric alliance with the US, in which both the US and South Korea intend to compel North Korea to give up nuclear weapons by deploying a missile defense. And North Korea—the challenger or the threatening power—determines whether to give up its nuclear weapons, given its estimation of

the levels of resolve the US possesses, which is denoted by the effectiveness of missile defense in my model, a proxy of North Korea's perception of the US commitment to its guarantee of the security of South Korea. Also, missile defense reduces the war cost which South Korea and the US might otherwise expect to incur since it fundamentally reduces the predicted utility of offensive nuclear weapons.

From North Korea's perspective, the missile defense is a signal of the sturdiness of the alliance between South Korea and the US. In theory, an aggressor considers the credibility of its target's alliances before resisting and is not likely to back down if the reliability of signals sent by the target and its ally is low since the aggressor assumes that the ally is reluctant to intervene in disputes (Gartner and Siverson 1996). Furthermore, since missile defense's value materializes only when the challenger chooses a preemptive strike given its nature as a defense mechanism, it does not contribute to military capability of South Korea. Thus, South Korea still expects the US to project its power to advance their mutual interests in changing the status quo (Morrow 1991). This explains why North Korea takes the willingness of the US seriously, and why it is important whether the deployment of a missile defense successfully convinces a potential challenger of the ally's commitment to extended deterrence. Based on these assumptions, we can anticipate the success or failure of indirect compellence.

I make two assumptions about missile defense in the model. First, I assume that the effectiveness of a missile defense corresponds to the perceived level of resolve possessed by a deployer. Powell (2003) articulates that deploying the missile defense makes a state more resolute and the country becomes seemingly more determined as the effectiveness of missile defense increases. I base my first assumption on this argument. Second, I assume that deploying a missile defense does not remove the uncertainty a potential challenger has about the defender's level of resolve. The concept of missile defense programs is still fraught with considerable uncertainty

related to their technical feasibility, practicality, and cost. Thus, even though a challenger can infer the level of resolve the defender possesses from the estimated effectiveness of the missile defense, it cannot obtain complete information about the ally's payoff on which the ally bases its decision whether to intervene.

The sequence of events is as follows. First, nature determines the effectiveness of the missile defense denoted as β . Next, South Korea decides whether to deploy the missile defense, and North Korea determines whether to relinquish its nuclear weapons based on the information it has about the missile defense. Finally, South Korea decides whether to back down or stay firm. As I mentioned earlier, the deployment of the missile defense is a proxy which reveals the willingness of the US to become involved in this dispute, and North Korea guesses the willingness of the US based on its expectation of β . As shown in Figure 1, there are four possible outcomes. When South Korea does not take any action in response to North Korea missile tests, the result is the continuation of the status quo, and if North Korea gives up its nuclear weapons, that indicates the success of compellence. Two possible scenarios of the failure of compellence are as follows: 1) when North Korea resists, and South Korea stays firm and 2) when North Korea resists, and South Korea acquiesces to North Korea's nuclear ownership. Of course, there is no certainty that South Korea's uncompromising attitudes towards North Korean nuclear crisis will necessarily result in war. However, if the pattern in which North Korea continues its nuclear tests and South Korea keeps pushing North Korea to refrain from such provocation is repeated, the probability of conflict increases intuitively; for brevity, this outcome is simply described as the war in my model.

This extended compellence game with missile defense is illustrated in Figure 1. Fundamentally these two players are bargaining over a pie which follows a uniform distribution from 0 to 1. Based on Fearon (1997), missile defense belongs to sunk-cost signals that are costly for the state in the first place and may change the state's

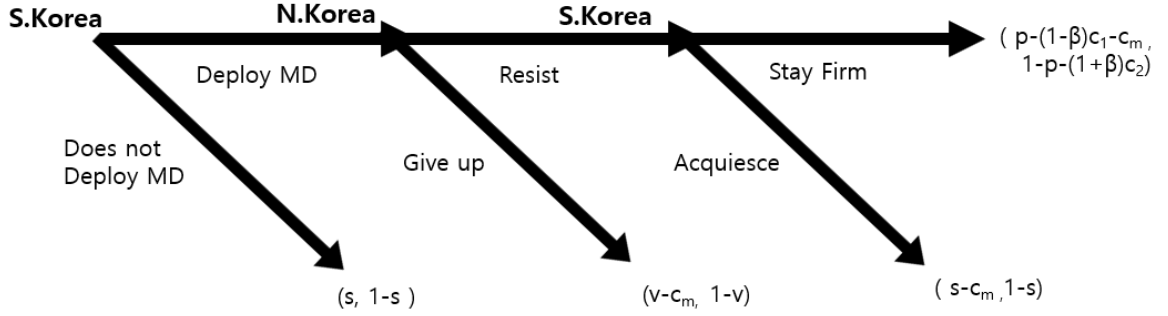


FIGURE 3.1: The bargaining game with missile defense

expected value for fighting versus backing down in a challenge. Thus, South Korea incurs the cost of c_m for missile defense deployment and gets a portion of the pie equal to $v - c_m$ if nuclear compellence is successful. If North Korea resists, South Korea determines whether to stay firm, thus receiving $p - (1 - \beta)c_1 - c_m$ or back down, thus receiving $s - c_m$. p denotes South Korea's chance of winning the war, and c_1 refers to its war cost. North Korea gets a portion of the pie equal to $1 - v$ if South Korea chooses to acquiesce. If South Korea stays firm, North Korea receives $1 - p - (1 + \beta)c_2$, where c_2 is North Korea's war cost. If North Korea chooses to resist and South Korea decides to back down, the former receives $1 - s$, which is the same as it receives when South Korea does not deploy a missile defense. Even though North Korea gains benefits in the negotiations after being acknowledged by South Korea as a nuclear weapon state, the cost is also accompanied by the coexistence with the missile defense which incapacitates the preemptive strike capability. Thus, I presume that the payoff North Korea obtains when South Korea backs down is $1 - s$ in this model. β is the perceived effectiveness of the missile defense and decreases the war cost of South Korea and amplifies the war cost incurred by North Korea.

The US also has a utility function over the issue space, which favors South Korea. As I mentioned previously, β , perceived effectiveness as a proxy, represents two

Koreas' perception of the US's degree of resolve to intervene. The US also incurs the cost, c_a , if the conflict occurs. When North Korea resists and South Korea stays firm, the US receives $p - (1 - \hat{\beta})c_a - c_m$ when it decides to intervene because the US also has a mutual interest with South Korea in compelling North Korea and shares the cost of deploying the missile defense. In this case, c_a is the US's war cost, and c_m is the same amount of expenditure incurred by South Korea for deploying THAAD. However, if the US is not willing to intervene in the event of a crisis, its payoff is defined as $s - c_m$. In this model, the US explicitly reveals its true resolve only when North Korea retains its nuclear missiles, and South Korea stays firm. Both players do not know whether the US is actually willing to push North Korea harder to give up its nuclear weapons if the game ends at other nodes. The genuine willingness of the US to provide aid to South Korea in response to a North Korean threat is revealed only when conflict occurs between South Korea and North Korea and it is denoted as $\hat{\beta}$.

3.2 Information structure and beliefs

The game involves incomplete information about the US; North Korea is uncertain about the US level of commitment and estimates the commitment of the US based on the perceived effectiveness of the missile defense. Since the uncertainty in the model concerns the willingness of the US to engage in the war, I assume that nature randomly selects β . All parties' payoffs follow a uniform distribution and $\beta \in [0, 1]$. South Korea prefers issue resolution over the status quo, and therefore, v is greater than s . The likelihood of winning the war, which is denoted by p , is also larger than s .

Based on backward induction, I consider the last node of the game first. South Korea must decide whether to stay firm or back down at the bottom of the nodes after North Korea chooses to resist. South Korea chooses to stay firm when the payoff from

conflict is greater than the payoff from backing down. Thus, when $\beta > \frac{p-s+c_1}{c_1}$, South Korea incurs the high cost of war and chooses to back down; when this condition is not met, South Korea decides to stay firm. I denote this point as β^* . Consequently, North Korea shapes its prior belief about the US willingness and updates it based on the effectiveness of the missile defense. ϵ encompasses the probability that North Korea may underestimate, overestimate or ignore the effectiveness of the missile defense and $1 - \epsilon$, otherwise. I assume $1 - \epsilon$ is greater than ϵ .

Based on its utility function, the US also decides whether to intervene considering its expected war cost from intervening. I do not assume the actual value of β is always the same as the actual willingness of the US to become involved in the possible disputes since the perceived effectiveness of the missile defense based on the information from the missile defense might be correct or incorrect. The decision made by the US—intervene or not intervene—can be observed only when the game reaches the node where North Korea resists and South Korea stays firm.

Empirical Implications

4.1 Perceived level of resolve and the actual level of commitment

Since the goal in this paper is to explore whether the extended deterrence by missile defense is effective in terms of coercing a would-be nuclear weapon state to relinquish its nuclear weapons, I focus on comparing the equilibria of the game proposed above within the range of β . Thereby, I can observe how outcomes change when a potential challenger takes the offer of a security guarantee seriously. Capturing this phenomenon does not require a new model, and I can set the parameter β equal to zero if North Korea does not take the US willingness seriously, while keeping all other assumptions as given in the game. Thus, the core result that interests us is the effect on the game when β shifts from zero to a number closer to one. The formal solution to the model is given in the Appendix, and I present the main results here.

Figure 2 indicates how the expected utility of North Korea's resistance is transformed by the effectiveness of the missile defense, which is equal to the perceived resolve of the US to intervene.¹ Here, to minimize the expected payoff of North

¹ In Figure 3, the parameter values are $\epsilon = 0.4$, $p = 0.6$, $c_1 = 0.8$, $c_2 = 0.8$, $s = 0.2$.

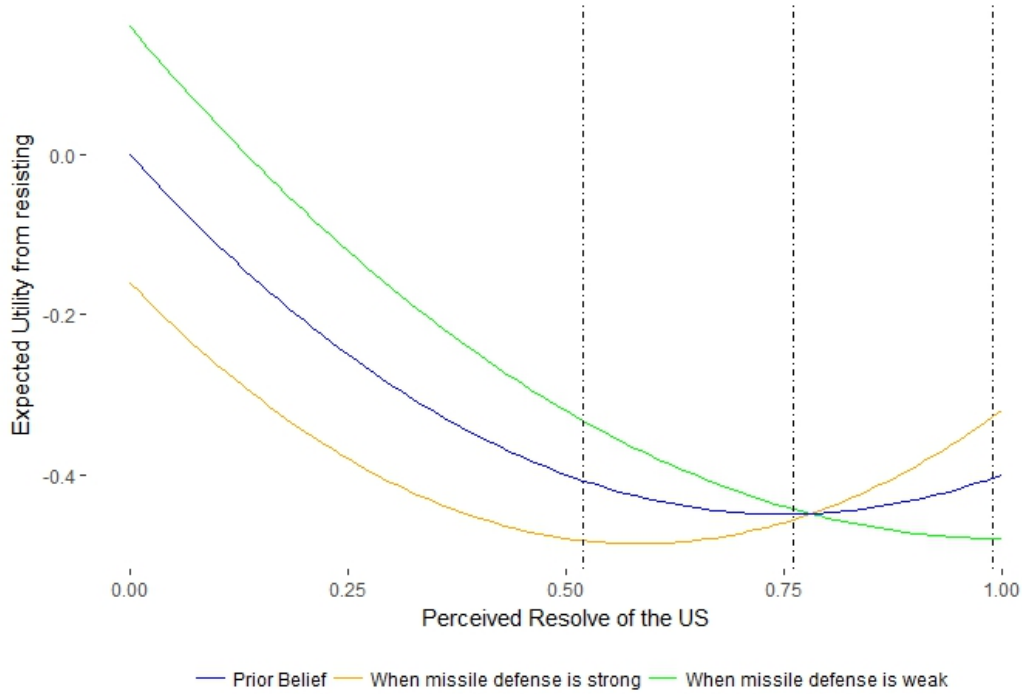


FIGURE 4.1: The expected utility of North Korea when resisting

Korea when it perceives that $\beta > \beta^*$,² the level of the US resolve should be around 0.6. Without the missile defense deployment, the resolve of the US should be around 0.52. If North Korea perceives that $\beta < \beta^*$, the willingness of the US to intervene should be around the maximum level necessary to reduce North Korea's incentive for resisting. Given that the perceived resolution of the US is constant, we can observe that the expected utility when the missile defense is strong outweighs that of North Korea when missile defense is regarded as weak. This is because other factors influence the potential payoffs of North Korea, which implies that the missile defense does not reduce the adversary's expected utility if the level of the US resolution is greater than a certain threshold.

The facts I derive from Figure 2 can be generalized to Figure 3, which suggests insights on how the missile defense serves as a mechanism of generating a credible

² β^* is defined in the appendix

threat perception to a potential challenger when its effectiveness is high. Without the missile defense, North Korea would relinquish its nuclear weapons given that the commitment of the US is higher than $R(US)$. In other words, $R(US)$ is the threshold of the US willingness above which North Korea can be coerced to give up nuclear weapons. If the missile defense is highly efficient, the level of actual resolve required to make North Korea back down is $R_H(US)$, which is lower than $R(US)$. However, if the missile defense is not effective, the US must be highly resolute to coerce North Korea, by becoming actively involved in the Korean peninsula.

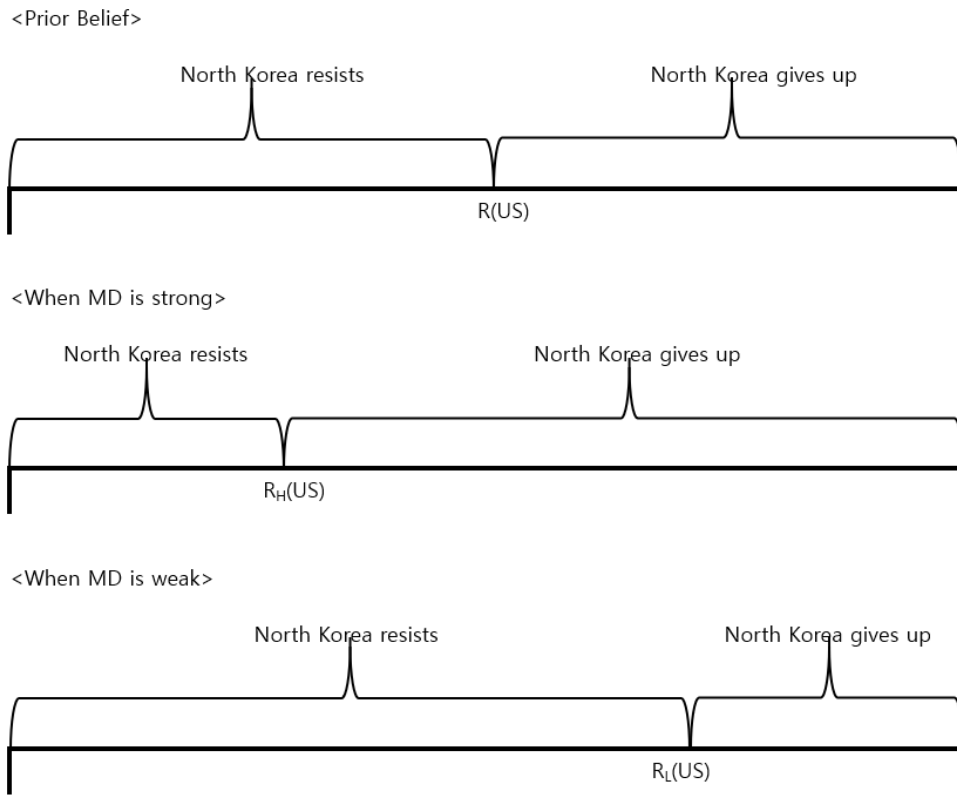


FIGURE 4.2: The effect of missile defense on North Korea's belief

The essential point is that once the missile defense is acknowledged as highly potent, the actual level of the US willingness does not necessarily need to be as high as that needed to deter North Korea without a missile defense. This implies that a

missile defense can serve as a mechanism by which the US or South Korea coerce North Korea to relinquish its nuclear weapons when the actual resolve of the US is low. Moving forward, if the US notices that North Korea considers the missile defense as a robust defense system, which discounts its anticipated utility from war, the US actually does not have any reason to devote effort equivalent to North Korea's expectations, and the US security umbrella will shrink. Such possibility induces two Koreas to be suspicious of their previous perception of the level of US resolution. On the other hand, if the missile defense turns out to be ineffectual, coercing North Korea requires a higher level of US resolve than the country is thought to have. If $R_L(US)$ is too high, North Korea may be the undeterrable enemy who cannot be coerced without the actual use of force. In this scenario, the only possible way to compel North Korea is the war, the decision for which involves both the degree of US commitment and South Korea's perception of the US's willingness.

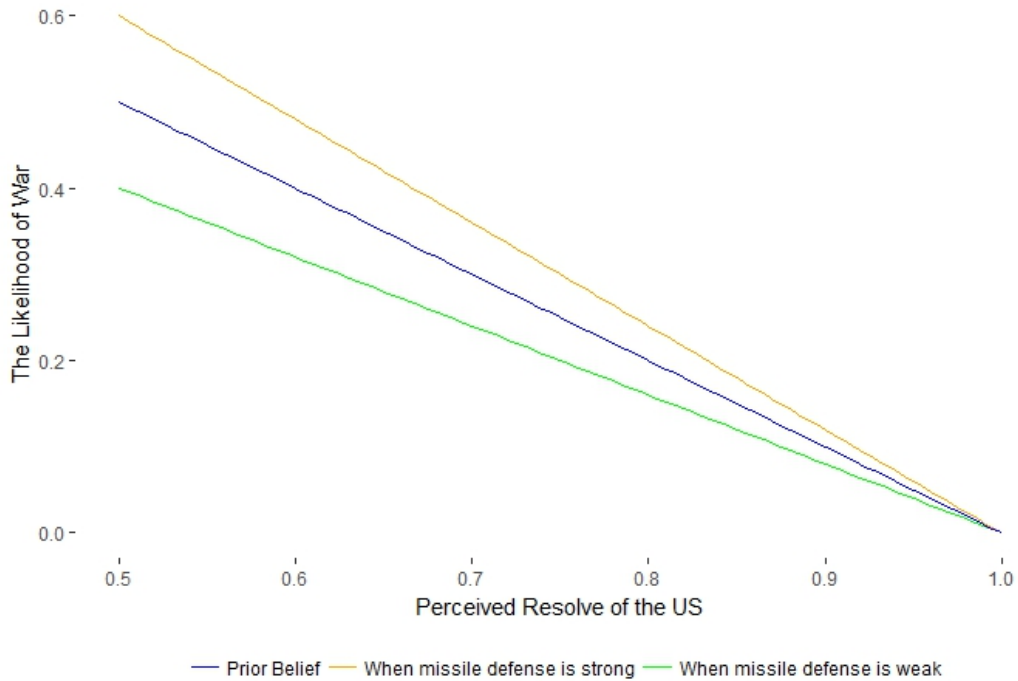


FIGURE 4.3: The likelihood of war depending on the US resolve

Figure 4 shows the likelihood of war depending on the US's resolve as perceived by South Korea and North Korea.³ As shown in the graph, the possibility of war decreases as the perceived US willingness increases to close to one. We can presuppose that this is because North Korea is hesitant to resist when the US seems to have a firm conviction regarding denuclearization in the Korean peninsula. Under the circumstance where this parameter assumes a value similar to that in footnote 2, South Korea invariably chooses to acquiesce when the perceived β is less than 0.5. This explains why the probability of war is zero when β is below 0.5 no matter what North Korea decides to do and implies that war is possible if and only if when $\beta > 0.5$ and North Korea resists.

Overall, the chance of war decreases as the US becomes increasingly interested in North Korean missile tests since North Korea will prefer to relinquish its nuclear weapons; if the resolve of the US gets closer to the maximum value of β , the possibility of war approaches zero, which suggests that an obviously resolute US can coerce North Korea to give up nuclear weapons before South Korea makes its decision. Intriguingly, given that the US resolve is constant, the probability of war is higher when the missile defense is durable rather than when it is fragile. This is because South Korea is willing to run higher risks when the missile defense is strong, escalating to war while dragging in the US. This finding is consistent with Snyder (1984)'s argument about entrapment, in which the ally is being dragged into a conflict against its actual interests. Since South Korea deduces that the US commitment is stronger when the missile defense is highly capable, there is no incentive for it not to stay firm.

³ The parameters are the same as before.

4.2 Possible outcomes

Originally, THAAD was introduced as one of the measures to prevent North Korea from engaging in possible provocations. Yet, simultaneously it can still be interpreted as a passive security guarantee against a would-be nuclear weapon state. The critical issues of THAAD deployment have been whether the US can credibly threaten North Korea by assuring its security commitment to South Korea, and what roles a missile defense plays in extended deterrence. In an attempt to address these concerns, the model shows that the missile defense actually undermines the credibility of the extended security guarantee of the US to South Korea, a nonnuclear power.

My model comprehensively captures the gap between the actual resolve and the perceived resolve of the US. In Figure 5, the x-axis is the war cost incurred by the US, and the y-axis indicates the actual willingness of the US.⁴ The black line is the minimum threshold of the willingness of the US given its war costs. Thus, the area above the black line represents the possible range of the US willingness given its incurrence of an equivalent war cost. This is consistent with my intuition that the US should have a higher willingness to intervene in the conflict as its war cost increases.

The yellow line is the $R_H(US)$, the minimum threshold of perceived resolution necessary to compel North Korea when the missile defense is highly effective. I used the same notation for Figure 3. Similarly, the green line is the critical value above which North Korea can be coerced when the missile defense is highly weak. Particularly noteworthy, the area below the black line and above the yellow line indicates the scenario where the US is not interested in coercing North Korea but compellence is possible because North Korea overestimates the US willingness based on the effectiveness of THAAD. This raises the possibility of bluffing, where the

⁴ The parameters are the same as before

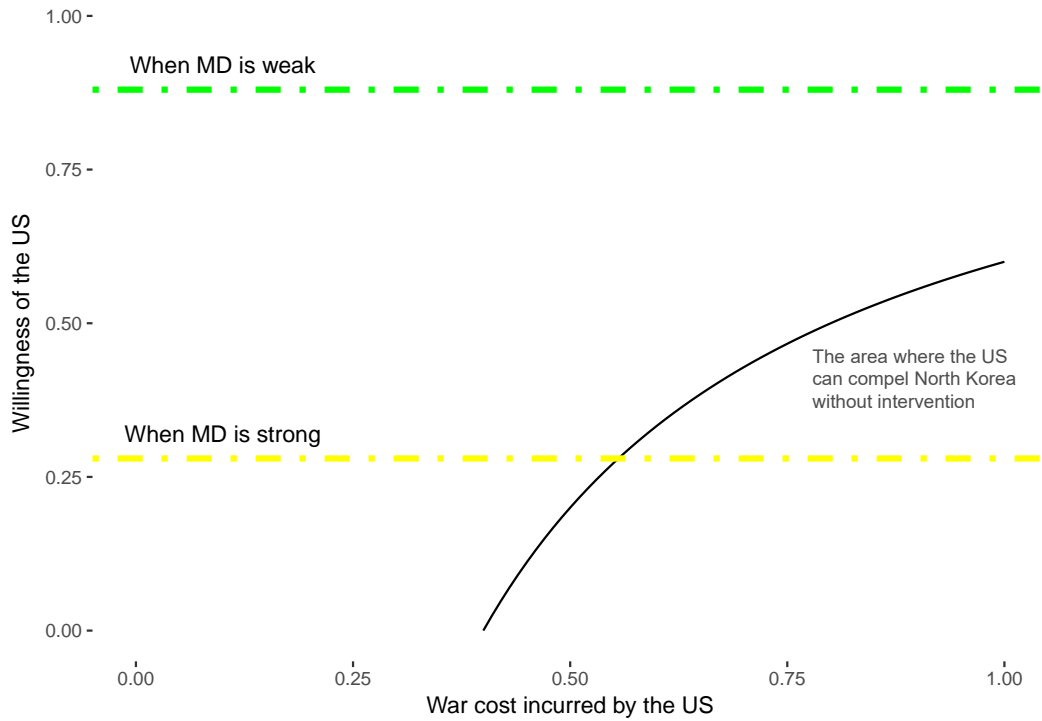


FIGURE 4.4: War cost incurred by the US and recognized level of the US resolve

US is not eager to get involved in disputes but has marginal interests in preventing North Korean provocations. This situation is equivalent to one of the equilibria in the model at which South Korea stays firm in response to North Korea’s resistance since South Korea expects the US to assist, but the US lacks the will to become involved in disputes.

The credibility of the US resolve is even more severely damaged when the missile defense engenders the impression that the US is less willing to provide aid to South Korea. Such a perception would draw North Korea to retain its nuclear weapons more than it would without the missile defense. Furthermore, the US should be highly resolute in efforts to persuade North Korea to relinquish its nuclear armament; the same level of resolve that could coerce North Korea when the missile defense is regarded as strong is insufficient to achieve a satisfactory result. This suggests that deploying the missile defense would rather result in an adverse effect on the

denuclearization efforts, which is counter to the interests of both the US and Korea.

Such an implications of deployment of THAAD in the Korean peninsula is reminiscent of the historical debates during the Cold War over missile defense and US security commitments to its ally. For an extended deterrence to be credible, the US should have both maintained power superiority and promised a nuclear response to a challenger's possible attack (Gavin 2014). However, the introduction of damage-limiting capability produces another issue of reliability for US allies; though a missile defense system protects the US, it still leaves Europe vulnerable and increases European doubt about the credibility of the US willingness to risk the US's destruction in order to protect them. This historical fact empirically supports the argument I draw from the model that damage-limiting capability rather hurts the credibility of allies and deploying missile defense is actually an ineffective deterrence strategy against a would-be challenger.

Similar to the aforementioned historical case, the fear of being entrapped or abandoned has complicated the situation in the Korean peninsula, increasing doubts about the actual US commitment after the announcement of THAAD deployment. With the calls for nuclear armament in the South growing, researchers in conservative think tanks and right-wing lawmakers in South Korea resumed the discussion on the Souths independent nuclear armament against North Koreas nuclear provocations. Cheong Seong-chang, a researcher at Sejong Institute which is one of the prestigious think tanks in South Korea, argued that the closer North Korea gets to the ownership of nuclear weapons, the more likely the South Koreans fear of being abandoned from the US to increase while the US would be more reluctant to get involved in full-scale war in the Korean Peninsula. Won Yoo-chul, a former representative of Saenuri Party, insisted that only nuclear weapons could be the successful deterrence strategy in response to nuclear weapons and many conservative politicians have shown support to such claim. These testimonies and North Koreas continuous provocations are

consistent with the argument that a would-be challenger becomes more unlikely to be deterred and the partners confidence about the US security guarantee cannot be maintained when the US seemed to be less willing to intervene.

After winning the 19th presidential election, Moon Jae-in has stated he would work closely with the US on addressing the security threats posed by North Korea, simultaneously emphasizing a strong message on national defense, rather than reversing THAAD deployment. In the same context, Moong Jung-in, President Moons special advisor for foreign affairs and security defined the elements of South Koreas security in two terms: 1) conventional deterrence and 2) American provision of nuclear umbrella. Despite the uncertainty over the US security guarantee denoted by THAAD deployment, he denied the deterring value of THAAD but admitted only its defensive advantage, implying that THAAD has not been intended to deter North Korea at the initial step. One other anonymous interviewee, a professor from Joint Forces Military University in South Korea insisted that THAAD may be an effective deterrence mechanism against North Korea, but the actual US willingness to coerce North Korea remains blurring. These understandings of missile defense in the Korean peninsula explicitly indicate that defensive posture —missile defense—fails at assuring the US security guarantee to its ally.

Notwithstanding the testimonies, THAAD obviously has resulted in many discrepancies between Washington and Seoul, such as suspended rollout and balancing burden-sharing in addition to the equivocal US security guarantee. Resting on secured observations, the mechanism under which defense measure including missile defense operates cannot bolster the credibility of the punishment threat to disarm a would-be nuclear state, North Korea. This aggressor has already been in possession of nuclear warheads, delivery capability, and conducted the sixth nuclear test. Until this point, what THAAD has yielded in the Korean Peninsula has been the increased qualms about the US security commitment and much noise in the surrounding area,

obscuring the future of denuclearizing a would-be challenger. One prominent researcher at a think-tank in Washington D.C. said that the US should apply lessons from the BMD crisis in Europe to the Korean peninsula. As many politicians say, nature abhors a vacuum and so does power; when two parties alliance breaks down, there comes a time when a potential challenger seamlessly enters and fills that vacuum.

One of the things that may obscure the dynamics of missile defense is the influence of neighboring states, especially China. Although the relation between North Korea and China is not as amiable as before, China has been the traditional ally of North Korea, and it is undeniable that the country still possesses the capability of indirectly transforming the behavior of North Korea. China may affect North Korea's strategic calculation of its expected payoff from staying firm and the level of satisfaction from status quo through arbitration or mediation. Including other variables that may influence the payoffs of players in the game would be a meaningful extension of this research in the future.

Conclusion

This game-theoretic analysis shows how a missile defense intended to expand a security guarantee indirectly affects the behavior of states directly involved. Based on their estimation of US willingness, South Korea and North Korea decide whether to push the other side harder or to yield. My analysis of the bargaining game among three players including the US and the two Koreas suggests that a missile defense is unlikely to be a credible threat which can be used to coerce North Korea. Even when the missile defense gives North Korea the impression that the US is highly resolute, whether the US has precisely the same or higher level of resolve which is sufficient to engage in a real fight remains unclear. Furthermore, when the missile defense is viewed as weak, its presence rather increases the threshold above which the compellence is successful, thus increasing the probability of compellence failure.

Consequently, even without the disputes over its uncertain effectiveness, the missile defense cannot function as the primary mechanism of a credible threat. The inconsistency between the recognized resolution and the actual US commitments leads a would-be nuclear challenger to be suspicious of the actual level of the US resolve. This finding supports the skepticism of the promulgation of the damage-limiting

capabilities as a compellence mechanism. However, my model develops the idea further in two important ways. First, it shows that defense measure or second-strike counter-force capability is unable to credibly threaten a would-be nuclear challenger in the extended-deterrence setting. Second, I have demonstrated that, counter to conventional wisdom which presumes a strong alliance after deploying a missile defense system, such a missile defense can actually increase uncertainty about the level of the US security commitment. My finding suggests a reevaluation of the theoretical and practical underpinning of conventional discourse on defense measures including missile defense and extended deterrence.

Appendix A

Appendix

At the bottom of the nodes, the US makes a decision whether to intervene in case conflict occurs between two koreas. Based on its payoff function, the US will get involved in disputes when $1 - \frac{p-s}{c_a}$ is smaller than $\hat{\beta}$. Here, c_a is the private information, which is not known to North Korea and South Korea.

Depending on wheheter the US has decided to intervene, South Korea also considers its action. If South Korea receives a signal that $\beta > 1 - \frac{p-s}{c_1}$, it would stay firm. From now on, we denote $1 - \frac{p-s}{c_1}$ as β^* . The prior belief of North Korea about the US resolve is $1 - \frac{c_1-p+s}{c_1}$, and we represent it as $B(US)$. North Korea's belief about the US resolve is updated after receiving the signal from nature. If the signal is about the high effectiveness of missile defense, which means that the US is highly willing to intervene, from Bayes rule, $B_H(US) = \frac{(1-\epsilon)*B(US)}{(1-\epsilon)*B(US)+\epsilon(1-B(US))}$. After receiving the signal that the missile defense is not effective, North Korea's belief is $B_L(US) = \frac{\epsilon B(US)}{\epsilon B(US)+(1-\epsilon)(1-B(US))}$. The solution concept I employ is perfect Bayesian equilibrium.

Based on the prior belief, the expected utility of North Korea from resisting can be

defined as $B(US)(\frac{1-\beta}{1-\beta^*})(1-p-(1+\beta)c_2-c_m) + (1-B(US))(\frac{\beta^*-\beta}{\beta^*})(1-s)$. To figure out the global minimum, we can set the derivative equal to zero, $-(1-B(US))\frac{1-s}{\beta^*} + \frac{B(US)}{1-\beta^*}(-1+p+2c_2\beta) = 0$ and solve for β , which gives $\beta^0 = \frac{1-B(US)}{2c_2}\frac{1-s}{\beta^*}\frac{1-\beta^*}{B(US)} + \frac{1-p}{2c_2}$. If a missile defense is strongly effective, the payoff from resisting is, $B_H(US)(\frac{1-\beta}{1-\beta^*})(1-p-(1+\beta)c_2) + (1-B_H(US))(\frac{\beta^*-\beta}{\beta^*})(1-s)$. Likewise, if we take derivative and set it equal to zero, $-(1-B_H(US))\frac{1-s}{\beta^*} + \frac{B_H(US)}{1-\beta^*}(-1+p+2c_2\beta)$, and solve for β , $\beta^H = \frac{1-B_H(US)}{2c_2}\frac{1-s}{\beta^*}\frac{1-\beta^*}{B_H(US)} + \frac{1-p}{2c_2}$. Now, consider the case where a missile defense is not effective. Here the payoff from resisting is $B_L(US)(\frac{1-\beta}{1-\beta^*})(1-p-(1+\beta)c_2) + (1-B_L(US))(\frac{\beta^*-\beta}{\beta^*})(1-s)$. Once again, we differentiate and set equal to zero, $-(1-B_L(US))\frac{1-s}{\beta^*} + \frac{B_L(US)}{1-\beta^*}(-1+p+2c_2\beta) = 0$, and solving for x produces $\beta^L = \frac{(1-B_L(US))}{2c_2}\frac{(1-s)}{\beta^*} * \frac{(1-\beta^*)}{B_L(US)} + \frac{1-p}{2c_2}$. We already know that $B_H(US) > B(US) > B_L(US)$, thus it is easily inferred that $\beta^L > \beta^0 > \beta^H$, which implies that the level of resolve that minimizes North Korea's payoff from resisting gets low if the missile defense is effective. Also, based on the expected utility function, we can derive the likelihood of war.

If $EU_{NorthKorea}(Resisting|MD)$ is smaller than 1-s, North Korea would prefer to resist rather than back down. We can easily figure out the threshold $R(US)$ by solving the equation for β . Likewise, we can figure out the value of $R_H(US)$ and $R_L(US)$ according to North Korea's belief, $B_H(US)$ and $B_L(US)$.

Bibliography

- Acton, J. M. (2009), "Extended Deterrence and Communicating Resolve," Tech. rep., Center for Contemporary Conflict.
- Altfeld, M. F. (1983), "Arms races And escalation? A comment on Wallace," *International Studies Quarterly*, 27, 225–231.
- Altfeld, M. F. (1984), "The decision to ally: A theory and test," *Western Political Quarterly*, 37, 523–544.
- Barnaby, F. and Boserup, A. (1969), *Implications of anti-ballistic missile systems*, vol. 2, NewYork: Humanities Press.
- Blank, S. (2008), "Strategic rivalry in the Asia–Pacific theater: a new nuclear arms race?" *The Korean Journal of Defense Analysis*, 20, 27–46.
- Brodie, B. (1976), "On the objectives of arms control," *International Security*, pp. 17–36.
- Bueno De Mesquita, B. (1981), "Risk, power distributions, and the likelihood of war," *International Studies Quarterly*, 25, 541–568.
- Cimbala, S. J. (1989), "Strategic impasse offense, defense and deterrence theory and practice," .
- Danilovic, V. (2001), "The sources of threat credibility in extended deterrence," *Journal of Conflict Resolution*, 45, 341–369.
- Diehl, P. F. (1983), "Arms races and escalation: a closer look," *Journal of Peace Research*, 20, 205–212.
- Drell, S. D., Farley, P. J., and Holloway, D. (1984), "Preserving the ABM Treaty: A Critique of the Reagan Strategic Defense Initiative," *International Security*, 9, 51–91.
- Fearon, J. D. (1992), "Threats to use force: the role of costly signals in international crises," Ph.D. thesis, Ph. D. diss., University of California, Berkeley.

- Fearon, J. D. (1994), “Domestic political audiences and the escalation of international disputes,” *American Political Science Review*, 88, 577–592.
- Fearon, J. D. (1995), “Rationalist explanations for war,” *International organization*, 49, 379–414.
- Fearon, J. D. (1997), “Signaling foreign policy interests: Tying hands versus sinking costs,” *Journal of Conflict Resolution*, 41, 68–90.
- Fortna, V. P. (2003), “Scraps of paper? Agreements and the durability of peace,” *International Organization*, 57, 337–372.
- Gartner, S. S. and Siverson, R. M. (1996), “War expansion and war outcome,” *Journal of Conflict Resolution*, 40, 4–15.
- Gavin, F. J. (2012), *Nuclear statecraft: history and strategy in America’s atomic age*, Cornell University Press.
- Glaser, C. L. and Fetter, S. (2001a), “The International Political Costs of National Missile Defense,” Tech. rep., DTIC Document.
- Glaser, C. L. and Fetter, S. (2001b), “National missile defense and the future of US nuclear weapons policy,” *International Security*, 26, 40–92.
- Glaser, C. L. and Kaufmann, C. (1998), “What is the offense-defense balance and how can we measure it?” *International security*, 22, 44–82.
- Grieco, J. M. (1988), “Anarchy and the limits of cooperation: a realist critique of the newest liberal institutionalism,” *International organization*, 42, 485–507.
- Haines, J. R. (2016), “Foreseeable, foreseen, ignored: Is Iran advancing its missile program at home while offshoring its nuclear program to North Korea?” *FPRI, E-Notes*, <http://www.fpri.org/articles/2016/01/foreseeable-foreseenignored-iran-advancing-its-missile-program-home-while-offshoring-its-nuclear-program-north-korea>.
- Huth, P. and Russett, B. (1988), “Deterrence failure and crisis escalation,” *International Studies Quarterly*, 32, 29–45.
- Huth, P. K. (1988a), “Extended deterrence and the outbreak of war,” *American Political Science Review*, 82, 423–443.
- Huth, P. K. (1988b), “Extended Deterrence and the Prevention of war,” *New Haven, Conn.*
- Hynek, N. (2010), “Missile Defence Discourses and Practices in Relevant Modalities of 21st-Century Deterrence,” *Security Dialogue*, 41, 435–459.

- Jervis, R. (1979), “Deterrence theory revisited,” *World Politics*, 31, 289–324.
- Kahn, H. and Wiener, A. J. (1967), “The next thirty-three years: a framework for speculation,” *Daedalus*, pp. 705–732.
- Kaufmann, W. W. (1989), “The requirements of deterrence,” in *US Nuclear Strategy*, pp. 168–187, Springer.
- Kupperman, R. H. and Smith, H. A. (1972), “Strategies of mutual deterrence,” *Science*, 176, 18–23.
- Larsen, J. A. and Wirtz, J. J. (2001), “US Missile Defenses: Three Scenarios and their International Consequences,” Tech. rep., DTIC Document.
- Lennon, A. T. (2002), *Contemporary nuclear debates: missile defense, arms control, and arms races in the twenty-first century*, MIT Press.
- Minhas, S., Hoff, P. D., and Ward, M. D. (2016), “Inferential Approaches for Network Analyses: AMEN for Latent Factor Models,” *arXiv preprint arXiv:1611.00460*.
- Mitchell, G. R. (2000), *Strategic deception: Rhetoric, science, and politics in missile defense advocacy*, Michigan State University Press.
- Morgan, P. M. (1983), *Deterrence: A conceptual analysis*, vol. 40, Sage Publications.
- Morrow, J. D. (1986), “A spatial model of international conflict,” *American Political Science Review*, 80, 1131–1150.
- Morrow, J. D. (1991), “Alliances and asymmetry: An alternative to the capability aggregation model of alliances,” *American Journal of Political Science*, pp. 904–933.
- Pifer, S., Bush, R., Felbab-Brown, V., Indyk, M., and Pollack, K. (2010), “US nuclear and extended deterrence: considerations and challenges,” .
- Podvig, P. (2017), “Did Star Wars Help End the Cold War? Soviet Response to the SDI Program,” *Science & Global Security*, 25, 3–27.
- Powell, R. (1990), *Nuclear deterrence theory: The search for credibility*, Cambridge University Press.
- Powell, R. (2003), “Nuclear deterrence theory, nuclear proliferation, and national missile defense,” *International Security*, 27, 86–118.
- Quackenbush, S. L. (2006), “National missile defense and deterrence,” *Political Research Quarterly*, 59, 533–541.
- Schelling, T. C. (1966), “Arms and influence,” *New Haven: Yale*.

- Schelling, T. C. (1980), *The strategy of conflict*, Harvard university press.
- Sechser, T. S. (2010), “Goliath’s Curse: Coercive Threats and Asymmetric Power,” *International Organization*, 64, 627–660.
- Sechser, T. S. and Fuhrmann, M. (2017), *Nuclear Weapons and Coercive Diplomacy*, Cambridge University Press.
- Shultz, G. P. and Goodby, J. E. (2015), *The War that Must Never be Fought: Dilemmas of Nuclear Deterrence*, Hoover Press.
- Signorino, C. S. and Tarar, A. (2006), “A unified theory and test of extended immediate deterrence,” *American Journal of Political Science*, 50, 586–605.
- Šimalčík, M. et al. (2016), “Ballistic Missile Defense and its Effect on Sino-Japanese Relations: A New Arms Race?” *Politologický časopis-Czech Journal of Political Science*, 23, 235–250.
- Snyder, G. H. (1960), “Deterrence and power,” *Journal of Conflict Resolution*, 4, 163–178.
- Snyder, G. H. (1984), “The security dilemma in alliance politics,” *World politics*, 36, 461–495.
- Snyder, G. H. (2015), *Deterrence and defense*, Princeton University Press.
- Taylor, A. (2017), “Why China is so mad about THAAD, a missile defense system aimed at deterring North Korea,” *The Washington Post*.
- Wallace, M. D. (1979), “Arms races and escalation: Some new evidence,” *Journal of Conflict Resolution*, 23, 3–16.
- Weede, E. (1980), “Arms Races and Escalation SOME PERSISTING DOUBTS (RESPONSE TO WALLACE’S ARTICLE, JCR, MARCH 1979),” *The Journal of Conflict Resolution (pre-1986)*, 24, 285.
- Wildavsky, A. (1965), “Practical consequences of the theoretical study of defense policy,” .