Credible Promise: Conflict and Security on the Korean Peninsula*

Jin Yeub Kim

December 2, 2019

Abstract

This note proposes a desideratum to resolve the North Korean conflict and to promote peace and security around the Korean peninsula. A political change in North Korea may be possible only with China's full support and cooperation with international sanctions against North Korea. The key to inducing China's cooperation is for the US alongside of South Korea to make a credible promise that respects Chinese security concerns in the northern part of the Korean peninsula in the event of political change.

JEL classification: C72, D82, F51.

Keywords: Credibility, promise, threat, conflict, security, cooperation.

---

*I am deeply grateful to Roger Myerson for sharing his insights on the topic of the Korean crisis and for encouraging me to write this note.

Kim: School of Economics, Yonsei University, 50 Yonsei-ro Seodaemun-gu, Seoul 03722, Republic of Korea. E-mail: shiningjin@gmail.com

1. Introduction

North Korea's provocative actions such as missile and nuclear tests have been an ongoing threat to international peace and security. A peaceful solution of the Korean crisis would involve some sort of moderation and political change in North Korea. The political change in North Korea may emerge in the form of destabilization or collapse of the authoritarian regime, complete denuclearization, an end-of-war declaration to the 1950-53 Korean War, or even a reunification of the peninsula under the government of South Korea.

The Hanoi summit between President Trump and Kim Jong-un (February 27-28, 2019) ended without a deal. The US wants North Korea to fully denuclearize in exchange for sanctions relief, while North Korea allegedly wants a more gradual rollback of its nuclear program. However, based on the nature of the North Korean regime and their past activities, there is no guarantee that North Korea is ever going to give up its nuclear ambitions. Given those conflicting demands, the North Korean regime is able to partly circumvent the economic sanctions by illegal trades with China's help. In fact, North Korea have been receiving shipping services from two Chinese shipping companies on which the Trump administration recently imposed new sanctions (Rappeport 2019).

If bringing about political change and peace in the Korean peninsula is the US' long-term foreign policy goal, then would its threats to extend trade sanctions against China be effective for achieving such goal? This short paper provides some insight into what must constitute an effective strategy for solving the North Korean crisis. I suggest that China's cooperation with sanctions against North Korea's nuclear program is the key factor toward moderation and political change in North Korea. China is North Korea's largest trading partner and has the most leverage on North Korea; so if China fully cooperates with the US, then the North Korean regime would feel great pressures to change its political course. While a threat to punish China by means of trade sanctions may motivate it to cooperate with the US, a promise of reward in the event of political change in North Korea is essential to elicit China's full cooperation.

China certainly would prefer that North Korea not have nuclear weapons, but its greatest fear is political change in North Korea which would expose China to a risk of foreign powers moving north to the China-Korea border. Eliminating such fear and respecting its vital security concerns in the region would bring China's interests into closer alignment with the US' interests, making China prefer full cooperation with the US against North Korea. Thus, a promise must include some form of recognition that northern Korea will be in China's sphere of influence and the US military forces will not move north in the event of political change in North Korea.

The credibility of a promise not to expand military forces over the entire Korean peninsula may be questionable, as the US already hosts military forces in South Korea. Further, such promise can be seen as a concession by the US. However, this paper argues that a credible promise could benefit the US interests by increasing China's willingness to cooperate and strengthening the US' ability to influence other countries in the world today.

Section 2 introduces a simple model, which is not intended to provide a complete analysis of the issue but merely to highlight the key point, and characterizes equilibria of the model. Section 3 examines the role of promise by comparing the equilibrium probability of cooperation in my model with that in a model without promise, and shows how the equilibrium probability of cooperation is affected by various factors in the model. The importance of making a credible promise is also illustrated, and several implications of my analysis for the North Korean conflict are discussed. Section 4 offers concluding comments.

2. A model

2.1. Setup

The model presented here is just one example of a conflict game that describes an international interaction between two countries, the U.S. (A, "she") and China (B, "he"), who have stakes involved in the conflict of North Korea (NK, not a player in the game).

Country A enjoys a benefit of $v > 0$ in the event of a political change in NK. This benefit can include monetary payoff, political gain, or any pecuniary or non-pecuniary value acquired from expanding A's sphere of influence on the Korean peninsula when NK changes its political course. Country B's full cooperation with sanctions against NK would increase the possibility of political change in NK. For simplicity of analysis, I assume that the probability of political change in NK is $\pi > 0$ if B cooperates, otherwise it is zero.

Country B resists any possibility of political change in NK out of fear that foreign power will expand into NK. So to induce B's cooperation, A must provide some assurance that B's security interests would be respected in the event of NK's political change. For example, A can make a promise that in such event NK will remain in B's sphere of influence. I assume that A promises to give B a benefit of size $\theta > 0$ in the event of political change if B cooperates. This parameter θ captures any monetary, political, or security-related gain or value that B gets from the northern part of Korea being in B's sphere of influence.[1]

However, A may not act as she promised and cannot be enforced to do so after B has cooperated and NK's political change has occurred. Let p represent the probability that A will keep her promise. Alternatively, p can be interpreted as the fraction of the benefit for which B will be awarded (possibly greater than one). Country A knows her true p, but B believes that p is drawn from a non-degenerate distribution $F(\cdot)$ over the interval $[0, M]$ with density $f(\cdot)$. I refer to p as A's type.

[1] I do not require that $\theta < v$.

The game begins with A's decision either to maintain the status quo or to "promise-and-threat" by issuing a statement to B of the following sort:

> "If you (B) do not cooperate with international sanctions against NK, then our government (A) will extend trade sanctions against you. But if you fully support and cooperate, then not only we will not extend sanctions on you but also we will keep American forces out of North Korea in case of any moderation or political change in NK."

This statement consists of both a short-term threat of an immediate punishment for B's non-cooperation and a long-term promise of a future reward for B's cooperation.

If A chooses to maintain the status quo (followed by no political change in NK), I shall assume without loss of generality that both A and B get their utility payoffs normalized exogenously at zero. If A issues the promise-and-threat statement, then B must decide whether to cooperate or not without knowing A's probability of keeping the promise, p. If B ignores the statement and does not cooperate, then A will immediately impose trade sanctions on B, in which case B incurs damage of size $d > 0$ while A receives a (small) benefit of size $l > 0$.[2] If B cooperates, then A receives an expected payoff of $\pi(v - p\theta)$ and B receives an expected payoff of $\pi p\theta - c$.[3] The parameter $c > 0$ captures B's costs that are incurred from losing his diplomatic relations with NK consequent upon its cooperation with A (and being unsupportive of NK) regardless of whether there is a political change. In this game, all of the parameters except p are common and public knowledge. The game tree illustrated in Figure 1 describes the sequence of events with the expected payoffs specified.

[2]One might argue that imposing trade sanctions against B may be at least as harmful to A as to B, so that $l < 0$. I briefly provide the basis for my assumption of $l > 0$ as follows. For the sake of simplicity, the game assumes that the probability of political change in NK is zero when A issues the statement and B does not cooperate. In such case, however, there is still a possibility that NK will face a change of its political course. Let such probability be $\pi' < \pi$. Then A's expected payoff can be characterized by $\pi' v' - t$ where v' is the benefit that A enjoys in the event of political change in NK and t is the loss from a trade war with B. Even so, $\pi' v' - t$ can be positive given that v' is very large. For example, A might gain a lot from expanding its power and sphere of influence into NK without having to "concede" part of the power to B upon B's non-cooperation. Then I can represent $\pi' v' - t > 0$ by the parameter $l > 0$.

[3]It is assumed that A gets zero payoff when there is no political change in NK after B has cooperated.

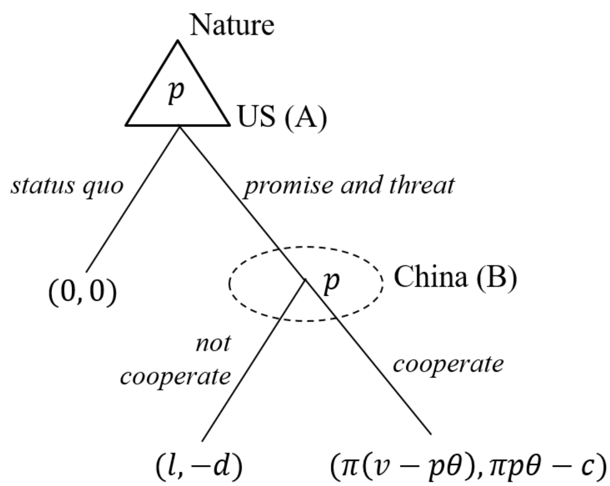

Figure 1: The game tree

2.2. Equilibrium

Country B's strategy is represented by δ that specifies his probability of cooperating if A makes a promise-and-threat. Country A of type p compares a zero payoff from maintaining the status quo with the expected payoff from making promise-and-threat, $\delta\pi(v-p\theta)+(1-\delta)l$. Country A's equilibrium strategy will have a cut-off property that if type p makes promise-and-threat, then any type $p' < p$ will also promise-and-threat. So A's best response against B's strategy $\delta > 0$ is to maintain the status quo if $p > \tilde{p}$ and to promise-and-threat if $p \leq \tilde{p}$, where $\tilde{p}$ is the cutoff type who is just indifferent between maintaining the status quo or making promise-and-threat when B is expected to cooperate with probability $\delta > 0$.[4] The indifference condition is

$$\tilde{p} = \left[v + l(1-\delta)/\pi\delta\right]/\theta. \tag{1}$$

Because $\tilde{p} > 0$ so that there is always some positive probability that A makes promise-and-threat, B uses Bayes' rule in forming his posterior beliefs at his decision node. That is, B's posterior expectation of p is given by $\mathbb{E}[p|p \leq \tilde{p}] = \int_0^{\tilde{p}} xf(x)/F(\tilde{p})dx$. Accordingly, B compares his payoff from not cooperating, $-d$, with the posterior expected value of co-

[4] I assume that the indifferent type makes promise-and-threat.

operating, $\pi\mathbb{E}[p|p \leq \tilde{p}]\theta - c$. This expected value of cooperation increases when more As promise-and-threat (higher $\tilde{p}$).

Notice that if $-c \geq -d$, then $\pi\mathbb{E}[p|p \leq \tilde{p}]\theta - c > -d$ regardless of B's posterior expectation of p. In such case, B's best response to any cut-off strategy of A is to cooperate with certainty, $\delta = 1$. Then A's best response is to promise-and-threat if $p \leq v/\theta$ and to maintain the status quo otherwise. These best responses uniquely define the equilibrium when $-c \geq -d$. I rule out such case for two reasons. First, if the size of B's damage inflicted by trade sanctions that A imposes on B is large, then it is also likely that such a "trade war" would be at least as harmful to A as to B. Then the assumption that A gets a positive payoff l upon B's non-cooperation is unjustified. Second, a more practical reason is that if $-c \geq -d$, cooperation is a dominant strategy: B's cooperation is not only profitable given the prior distribution of p but also preferred even against A with zero probability of keeping promise. This implies that A's promise of a positive θ was not needed to induce B's cooperation, so the case of $c \leq d$ represents a very uninteresting game. The following assumption eliminates such case.

Assumption 1. The amount of B's costs associated with losing relations with NK upon cooperating with A is larger than the size of B's damage inflicted by trade sanctions upon not cooperating with A. That is, $-c < -d$.

Under Assumption 1, there are two cases to consider.

Condition 1. $\pi\mathbb{E}[p]\theta - c \geq -d$.

Condition 2. $\pi\mathbb{E}[p]\theta - c < -d$.

Condition 1 requires that B's expected value of cooperating is greater than that of not cooperating given the prior distribution about p. Then there exists a unique value of $\tilde{p}$, denoted by p^*, that solves

$$\pi\mathbb{E}[p|p \leq p^*]\theta - c = -d. \tag{2}$$

That is, p^* is the cutoff value such that promise-and-threat by the types below p^* makes B indifferent between cooperating and not. Then B's best response at promise-and-threat by

As of type $p \leq \tilde{p}$ is $\delta = 1$ if $\tilde{p} > p^*$ (the value of cooperation is greater than $-d$); $\delta = 0$ if $\tilde{p} < p^*$ (the value of cooperation is less than $-d$); and any $\delta \in [0, 1]$ if $\tilde{p} = p^*$ (the value of cooperation is $-d$).

Under Condition 2, cooperating is not profitable for B if all types of A promise-and-threat in which case B's posterior distribution about p is identical to his prior distribution. If less As promise-and-threat, then B lowers his posterior expectation of p, so cooperating becomes even less profitable. So B's best response to any cut-off strategy of A is $\delta = 0$. Note that under Condition 1, cooperating is profitable if all As promise-and-threat, so it rules out the cases where B never cooperates upon observing promise-and-threat regardless of his posterior expectations.

The following proposition describes the equilibrium strategies in this game.

Proposition 1. *Suppose that Assumption 1 is satisfied.*

1. *Under Condition 1, there is a unique sequential equilibrium in which As of type $p > \tilde{p}$ maintain the status quo while As of type $p \leq \tilde{p}$ promise-and-threat, and B cooperates with probability $\delta > 0$ such that:*

   (i) if $v - p^*\theta > 0$, then $\tilde{p} = v/\theta$ and $\delta = 1$;

   (ii) if $v - p^*\theta \leq 0$, then $\tilde{p} = p^*$ and $\delta = l/[l - \pi(v - p^*\theta)] \in (0, 1]$.

2. *Under Condition 2, there is a unique sequential equilibrium in which all As promise-and-threat and B chooses not to cooperate.*

Proof. Under Condition 1: (i) Given definition of $\tilde{p}$ in (1), $\tilde{p} \geq v/\theta > p^*$ when $v - p^*\theta > 0$. Because $\tilde{p} > 0$, when A makes promise-and-threat, B uses Bayes' theorem to compute his posteriors on A's type p given the priors. Country B's expected value from cooperation with beliefs concentrated on $[0, \tilde{p}]$ upon promise is then $\pi\mathbb{E}[p|p \leq \tilde{p}]\theta - c > \pi\mathbb{E}[p|p \leq p^*]\theta - c = -d$ where the equality follows from (2). So B's best response strategy must be $\delta = 1$. Against B's strategy $\delta = 1$, A's best response is to make promise-and-threat if $p \leq v/\theta$ and to maintain the status quo if $p > v/\theta$. This in turn justifies B's optimal strategy to be $\delta = 1$. These strategies of A and B are uniquely defined, so constitute the only sequential equilibrium. (ii) When $v - p^*\theta \leq 0$, whether $\tilde{p} \gtreqless p^*$ depends on δ. First suppose that $\delta = 0$. Expecting B to not cooperate with certainty, all As would make promise-and-threat, earning $l > 0$ instead of zero by maintaining the status quo. Upon promise-and-threat by all As, B learns nothing so that his posterior expectation of p equals his priors. By Condition 1, B will prefer cooperating to not cooperating, so $\delta = 1$, which is a contradiction. Now suppose that $\delta \in (0, 1)$, which requires B to be indifferent between cooperating and not. By backward induction, A's cutoff type must be $\tilde{p} = p^*$. For this to be A's best response strategy, δ is determined by $p^* = [v + l(1 - \delta)/\pi\delta]/\theta$, or equivalently, $\delta = l/[l - \pi(v - p^*\theta)]$, which is uniquely defined. Therefore, such δ and $\tilde{p} = p^*$ constitute a sequential equilibrium. Lastly, suppose that $\delta = 1$. If $v - p^*\theta < 0$, then $\tilde{p} = v/\theta < p^*$ so that $\delta = 0$ because B's expected value of cooperation is less than $-d$, which leads to a contradiction. So $\delta = 1$ is only possible when $v - p^*\theta = 0$, which is a special case of $\delta \in (0, 1)$ where $\delta = l/l = 1$ and $\tilde{p} = p^*$. Thus in case (ii), the sequential equilibrium is unique. Under Condition 2: After promise-and-threat by all As, B's posterior equals his prior, and so B chooses not to cooperate, which in turn justifies A of any type making promise-and-threat. This is the only sequential equilibrium. □

3. Discussions

In this section, I first study the role of making a promise by comparing the likelihood of B's cooperation in equilibrium of my model with that of a model with no promise but only a threat. I then identify how the likelihood of B's cooperation is shaped by various factors in the model using the equilibrium characterization. This analysis is applied to discuss several implications for understanding how the North Korean conflict can be affected by interactions between the US and China, as well as for the importance of credible commitment of a long-term promise and a short-term threat.

3.1. The role of promise

I begin by looking at the probability of B's cooperation in equilibrium. Formally, the equilibrium probability of cooperation is given by:

$$Pr(cooperation) \equiv F(\tilde{p})\delta. \tag{3}$$

When Condition 1 holds, B ex ante prefers cooperating to not cooperating at the start of the game. In this case, the equilibrium probability of cooperation is

$$Pr(cooperation) = \begin{cases} F(v/\theta) & \text{if } v - p^*\theta > 0, \\ F(p^*)\,[l/[l - \pi(v - p^*\theta)]] & \text{if } v - p^*\theta \leq 0. \end{cases}$$

When Condition 2 holds, B finds cooperating ex ante unprofitable than not cooperating so that B never cooperates in the unique equilibrium, in which case the equilibrium probability of cooperation is zero. This observation leads to the following corollary.

Corollary 1. *For B's cooperation to occur with positive probability in equilibrium, the level of damage inflicted on B by A's trade sanctions must be high enough, i.e.,* $d \geq c - \pi\mathbb{E}[p]\theta$*.*

Corollary 2. *In the model with no promise but only a threat of punishment, the probability of cooperation is zero in the unique equilibrium where all As promise-and-threat and B never cooperates.*

3.2. Comparative statics

[To be added]

3.3. Credible commitment

[To be added]

4. Concluding remarks

[To be added]

References

Rappeport, Alan. 2019. Administration Imposes North Korea Sanctions Amid Stalled Talks. *The New York Times*, March 21. https://www.nytimes.com/2019/03/21/us/politics/trump-north-korea-china-sanctions.html.