

Uncertainty in Foreign Policy Making: A Bayesian Game Analysis of Korea

HeeMin Kim and Jun Y. Choi

Department of Political Science
Florida State University
Tallahassee, FL 32306-2230, U.S.A.
hkim@garnet.acns.fsu.edu
jyc4405@garnet.acns.fsu.edu

Abstract

Bayesian games are used to analyze situations where at least one player is uncertain about the other's preferences. For the past decade or so, the Bayesian models have been rigorously applied to various aspects of international relations involving uncertainty. These models have contributed to our understanding of international relations by uncovering complicated strategic interactions through deductive reasoning and by generating many empirically testable hypotheses. These models have contributed to our understanding of international relations by uncovering complicated strategic interactions through deductive reasoning and by generating many empirically testable hypotheses. Apart from these efforts for general theory developments, however, scholars rarely applied Bayesian models to analyze real-world international events, although many of them involve situations where one or more players are uncertain about the other's preferences. Some of these events are interesting and important in their own right, given their potential impact on regional security and the amount of attention paid by scholars and politicians alike. Therefore, these cases warrant independent studies utilizing Bayesian models. In this paper we study one such situation, the changing relationships North Korea has with South Korea and the United States. We demonstrate that since the late 90's, both South Korean and U.S. policy makers have attempted to formulate their new North Korea policies while they were uncertain about the true intentions of the North Korean leadership. By developing two Bayesian models of the Korean situation, we make some very interesting predictions and policy recommendations. We also discuss another interesting aspect of North-South Korea-United States relations: the North Korean leadership is simultaneously playing similar, but separate, games with South Korea and the United States. This makes the North Korean choice of strategies more difficult since, if the North sends a certain signal (reveals its intentions) in one game, its competitor in the other will try to use that information as well.

An earlier draft of this paper was presented at the annual meeting of the American Political Science Association, 2001. We thank Patrick James, Jacek Kugler, Chae-Jin Lee, Will Moore, Tong W. Park, and Nancy Powers for their valuable comments and Ashley Leeds for providing us with an overview of the Bayesian game applications in the field of International Relations.

Introduction

Bayesian (incomplete information) games are used to analyze situations where at least one player is uncertain about the other's preferences. For the past decade or so, the Bayesian models have been rigorously applied to various aspects of international relations involving uncertainty. They include international conflict (e.g., Morrow 1989, Fearon 1994, 1995, Kim and Bueno de Mesquita 1995, Powell 1996, Bueno de Mesquita, Morrow, and Zorick 1997, Wagner 2000, and Werner 2000), alliance formation (e.g., Morrow 1994, Smith 1995, 1998), deterrence (e.g., Nalebuff 1991), domestic constraints on foreign policy (e.g., Iida 1993, Fearon 1994, 1997, Schultz 1998) as well as reputation building in the world political economy (e.g., Alt, Calvert, and Humes 1988).

These models have contributed to our understanding of international relations by uncovering complicated strategic interactions through deductive reasoning and by generating many empirically testable hypotheses. Apart from these efforts for general theory developments, however, scholars rarely applied Bayesian models to analyze real-world international events, although many of them involve situations where one or more players are uncertain about the other's preferences. Some of these events are interesting and important in their own right, given their potential impact on regional security and the amount of attention paid by scholars and politicians alike. Therefore, these cases warrant independent studies utilizing Bayesian models.

One such situation would be the changing relationships North Korea has with South Korea and the United States. As we will demonstrate in the following section, both South Korean and U.S. policy makers have attempted to formulate their new North Korea policies

while they were uncertain about the true intentions of the North Korean leadership, especially since the late 90's. One potential payoff of developing Bayesian games of a real world situation such as that on the Korean peninsula is that we can make predictions and policy recommendations specific to the situation and observe empirically how the players act as the situation unfolds. Another aspect of North-South Korea-United States relations, which makes them interesting, is that the North Korean leadership is simultaneously playing similar, but separate, games with South Korea and the United States. If the North sends a certain signal (reveals its intentions) in one game, its competitor in the other will try to use that information as well. This necessarily constrains the North Korean choice of strategies, as we will demonstrate below.

In the first section of this paper, we discuss the changing political situations in the Korean peninsula. By doing so, we also establish that the situation is amenable to Bayesian game analysis. In the two following sections, we present the Bayesian games between North Korea and South Korea, and North Korea and the United States, respectively. After solving for the equilibria of these games, we discuss the insights and substantive implications these models provide.

The Changing Political Situations in the Korean Peninsula

In June 2000, the North Korean leader, Kim Jong-il accepted the South Korean President, Kim Dae-jung's call for a North-South summit. The first-ever summit of the two Koreas' leaders was held in the North Korean capital of Pyongyang, followed by the visits of separated families and the flow of South Korean capital across the border. This rapprochement was

culminated at the Sidney Olympiads when the athletes from the two Koreas walked together under one flag in the opening ceremony. This was a dramatic and quite unexpected turn of events, since the North Korean regime (and its leader Kim Jong-il) had been known for its aggressiveness, inflexibility, and unwillingness to compromise with the outside world. These events also prompted the Clinton administration to re-assess its North-Korea policy, leading to the U.S. Secretary of State, Madeleine Albright's visit to Pyongyang (*Hankuk Ilbo* December 14, 2000, June 9, 2001).

There are two explanations for the North's change of mind. First, the North Korean leaders have come to accept that the only way out of poverty and isolation is to open up its border to the South and the United States, accepting the inevitability of an open society and the potential for unification under the South's initiative. The second possibility is that this is an attempt to take advantage of outside aid and lower tensions on the Korean Peninsula temporarily, until the North recovers from its economic problems (see for example, an editorial in *The New York Times*, October 15, 2000).

Obviously the South's and the United States' best courses of action differ depending upon the real intentions of the North. Without complete knowledge about the North's true intentions, what would be the South's and the United States' best strategy in their interactions with the North? What would be the most likely outcome of these interactions, given the uncertainty about the North's intentions? On the North's part, it needs to be consistent about the portrayal of its intentions since revealing its true intentions to either the South or the U.S. may reveal the same to the other. These are important questions for students of Korean studies as well as those of international relations. Given the uncertainty involved, however, most explanations given so far have been based on conjectures.

Bayesian game theory offers tools capable of analyzing a situation like this where at least one player is uncertain about the other's true preferences. In the following sections, we build Bayesian game-theoretic models in which the South's and U.S. policy makers are uncertain about the North's true intentions. Once the North sends a signal of accommodation as it did in 2000, the leaders of the South and the United States must determine how far they want to go in terms of accommodating back.¹

Model 1: the North - South Korea Interaction

In this section, we analyze the interaction between North Korea and South Korea, which is depicted by the extensive form in Figure 1. There is no uncertainty about South Korea's preferences, as its leaders have consistently suggested reconciliation between the two Koreas. There are two possible types of North Korean leadership, a sincere one and a deceitful one. The sincere type truly pushes for opening-up of its border to the South (and the United States). The deceptive type attempts to take advantage of outside aid and lower tensions on the Korean Peninsula temporarily, until the North recovers from its economic ills.²

(Figure 1 about here)

At the time of the presidential election in South Korea in 1997, then candidate Kim Dae-jung pledged to attempt to improve relations with the North. Ever since his election, President

¹ For previous studies applying game-theoretic models to inter-Korean relations, see Ahn (1995) and Kim (1995), among others.

² Another way to interpret the uncertainty about the North Korean preference is that, even if Kim Jong-il wants to push for a true change in North Korea, he may not be able to sustain it without potential loss of power given the disagreements among the top leaders of the communist party and the military about the method and the pace of the reform (see *Hankuk Ilbo* December 15, 2000).

Kim has consistently maintained that his administration is open to the idea of a North-South summit and the subsequent reconciliation between the two. This is commonly known as the “Sunshine” policy. So, our game in this section begins with the North Korean response to the Kim administration’s offer of the summit already on the table.

In this game, Nature moves first and determines the North’s type. Then the North can *either* accept or accommodate South Korean President Kim Dae-jung’s offer to engage *or* not. If the North does not accommodate, then Status Quo (SQ) continues. If it chooses to accommodate, then the South gets a chance to determine whether to accommodate back fully or not. If the South somehow withdraws from its commitment to engage, then the North’s accommodation is rejected by the South (REJ_S). If the South fully reciprocates and the North stops accommodating at some point in the future ($\sim A'$), then the outcome is exploitation by the North (EXP_N). If the North fully accommodates back (A'), then the outcome is reconciliation (RECON), leading toward long-term cooperation and potential unification in a peaceful manner.

As we described above, this extensive form succinctly summarizes the recent relationship between the two Koreas. For us to solve for equilibrium in this game, we first need to determine the players’ preference structure. There are some assumptions about their preferences that we can make without controversy. For example, South Korea’s most preferred outcome probably is RECON, while its least preferred outcome would be EXP_N. REJ_S would be in between these two outcomes. By definition, the sincere North Korea prefers RECON to EXP_N, while the deceitful North Korea prefers EXP_N to RECON. Regardless of its type, North Korea’s least preferred outcome would be REJ_S. The information above leads to the following inequality conditions:

$$\text{SK: } U(\text{RECON}) > U(\text{REJ}_S) > U(\text{EXP}_N)$$

$$NK_{\text{SINCERE}} : U(\text{RECON}) > U(\text{EXP}_N) > U(\text{REJ}_S)$$

$$NK_{\text{DECEITFUL}} : U(\text{EXP}_N) > U(\text{RECON}) > U(\text{REJ}_S)$$

This leaves the location of Status Quo in relation to other possible outcomes as the only source of potential debate. For South Korea, SQ will most likely rank just ahead of REJ_S because the Kim Dae-jung government does not really gain anything by suggesting a summit and reconciliation, as it did earlier, and then backing away from it after the North accepts the offer (REJ_S). For a sincere North, SQ is probably slightly above or below EXP_N while it is somewhere near RECON for a deceitful North. As we will demonstrate below, the location of SQ in the preference ordering of North Korea does not affect our conclusion as long as it is not the most or the least preferred outcome. With the inequality conditions above, we are now ready to solve for the equilibrium conditions. Due to its technical nature, we state the details of the equilibrium solution in Appendix A.

The equilibrium solution of this game provides several insights about the North-South Korean interaction. First, the North's behavior would be the same regardless of its type (Observation 1 in Appendix A). That is, whether the real type of the North's leadership is sincere or deceitful, it will consistently be accommodating or consistently unaccommodating. This will make it difficult for the South to collect more information about the North's real intentions. Second, the South must continue to engage the North only if it is fairly certain that the North is truly interested in reforming its system for the long-term (Observation 2). Otherwise, the South should pull away from its initial accommodation even if it results in the loss of face on the part of the Kim Dae-jung government both in and outside of the Korean Peninsula. Finally, the North Korean leadership, regardless of its real type, would try to

convince the South that it is sincere (Observation 3). As a matter of fact, it had been the case for the past few years until the change in parties in power in the United States in 2001. Why the advent of the Bush administration affects the North's behavior toward the South will be discussed in the following sections.

Model 2: the United States – North Korea Interaction

Next we build a Bayesian model in which U.S. policy makers are uncertain about the North's true intentions as in Model 1 above. This model is more complicated than the North-South interaction, though, as there is reason to believe that North Korean policy makers are also uncertain about whether the United States truly wants eventual reconciliation with the North. Some have suggested that the new Bush administration can use "a threat to world peace," such as North Korea, to sell its new military concept of "Missile Defense" to its allies (see for example, *Los Angeles Times* March 8, 2001).

(Figure 2 about here)

This situation is depicted in extensive form in Figure 2. As before there are two possible types of the North Korean leadership, picked by the nature, a sincere one and a deceitful one. Further there are two possible types of the U.S. leadership, which we will call hawkish and dovish, for lack of better terms. On its first move, the North can voluntarily accommodate or not. Due to the inter-connectedness of international relations situations, we can assume that this decision occurs when the North responds to the South's initial offer of a dialogue in the first game above. If they choose to accommodate, then the United States gets a chance to determine whether to accommodate back or not. If the North accommodates and the United States

reciprocates, then the outcome is reconciliation initiated by the North ($RECON_{NK}$). If the United States does not reciprocate, then the North's accommodation is rejected by the United States (REJ_{US}). If the North does not initiate accommodation, then the United States has a choice between initiating accommodation or not. If the United States chooses not to, the outcome is Status Quo. If the United States chooses to initiate accommodation, then the North gets a chance to determine to accommodate back or not (A' or $\sim A'$). If the North fully accommodates back, then the outcome is reconciliation initiated by the United States ($RECON_{US}$). If the North does not, the U.S.'s offer is rejected by the North (REJ_{NK}).

For us to analyze this game, we first need to determine the players' preference structure. There are several assumptions about their preferences that we can make without much controversy. By definition, the sincere North Korea prefers the $RECON$ outcomes to REJ_{NK} , while the deceitful North's preference over these outcomes would be reversed. Regardless of its type, the worst possible outcome for North Korea would be REJ_{US} . By definition, the hawkish United States prefers REJ_{US} or SQ to the $RECON$ outcomes, while the situation is reversed for the dovish United States.

This leaves the location of Status Quo in relation to other possible outcomes to be determined. For the sincere North Korea, which pushes for reconciliation, Status Quo should be placed in between the $RECON$ outcomes and REJ_{US} . For the deceitful North Korea, which does not want long-term conciliation with the United States, Status Quo must be better than the $RECON$ outcomes. For the hawkish United States, which does not want long-term reconciliation with North Korea, Status Quo must be preferred to the $RECON$ outcomes. For the dovish United States, Status Quo must be placed in between the $RECON$ outcomes and REJ_{NK} . The information in this paragraph does not give us the precise location of Status Quo in the

preference ordering of all possible types of players. It does, however, along with the information above, give us enough information to be able to analyze the U.S.-North Korea relations game. The information we have presented so far can be summarized by the following inequality conditions, where a comma indicates that we do not have enough information to differentiate the two:

$$NK_{\text{SINCERE}} : \text{RECON}_{\text{US}} > \text{RECON}_{\text{NK}} > \text{SQ}, \text{REJ}_{\text{NK}} > \text{REJ}_{\text{US}}$$

$$NK_{\text{DECEITFUL}} : \text{REJ}_{\text{NK}}, \text{SQ} > \text{RECON}_{\text{US}} > \text{RECON}_{\text{NK}} > \text{REJ}_{\text{US}}$$

$$US_{\text{HAWKISH}} : \text{REJ}_{\text{US}}, \text{SQ} > \text{RECON}_{\text{NK}} > \text{RECON}_{\text{US}} > \text{REJ}_{\text{NK}}$$

$$US_{\text{DOVISH}} : \text{RECON}_{\text{NK}} > \text{RECON}_{\text{US}} > \text{SQ}, \text{REJ}_{\text{US}}, > \text{REJ}_{\text{NK}}$$

With these inequality conditions, we are now ready to analyze the game. Due to its technical nature, we state the details of our analysis in Appendix B. As was the case with the North-South Korea Bayesian game above, the equilibrium solution of the U.S.-North Korea Bayesian game provide us with several interesting insights.

First, the hawkish United States will never initiate accommodation and will always reject the initial accommodation by North Korea (Observation 4 in Appendix B). On the other hand, the dovish United States will always reciprocate the accommodation initiated by North Korea. It will also initiate accommodation in the absence of North Korea's initial accommodation, but only if it feels highly certain that North Korea is sincere (Observation 5). Third, the deceitful North Korea never initiates accommodation and will always reject the initial accommodation by the United States (Observation 6).

Observations 4-6 lead to the conclusion that long-term reconciliation between North Korea and the United States is possible if and only if North Korea is sincere about political reform and the United States is dovish toward North Korea (Observation 7). This sounds rather intuitive at first, but what we are predicting in fact is that, out of eight possible RECON outcomes in Figure 2, only two are attainable in the interaction between North Korea and the United States. Finally, the combination of the deceitful North Korea and the hawkish United States always leads to Status Quo, since neither side is willing to risk being exploited by the other (Observation 8).

Recent U.S. Policy Toward North Korea

Shortly after the inauguration of the Bush administration in the United States, the new Secretary of State, Colin Powell, announced that the U.S. government would fundamentally re-examine its North Korea policy. During the summit with President Kim Dae-jung in March 2001, U.S. President George W. Bush made it clear that the United States still considered North Korea a rogue nation and a threat to the international community; that it was too early to resume the U.S.-North Korea talks; and that any future dialogue must be based on the “complete verification” of the 1994 Geneva agreement.³ In its annual report on terrorism in May, the U.S.

³ Under the Geneva agreement, North Korea will freeze its graphite-moderated reactor and related facilities and will eventually dismantle them. The International Atomic Energy Agency (IAEA) will be allowed to monitor this freeze, and North Korea will provide full cooperation to the IAEA for this purpose. In return, the U.S. will undertake arrangements for the provision to North Korea of light-water reactor power plants with a total generating capacity of approximately 2,000 MW by a target date of 2003. In addition, the U.S. will make arrangements to provide alternative energy in the form of heavy oil for heating and electricity production to offset the energy foregone due to the freeze of North Korea’s graphite-moderated reactors and related facilities. See further details of the agreement at www.kedo.org/Agreements.

State Department continued to include North Korea as one of the states sponsoring terrorism.⁴ Condoleezza Rice, the national security advisor for the Bush administration, stated in a White House function that Kim Jong-il was not trustworthy (*Hankuk Ilbo*, various dates, 2001).

From this information, it appears that the Bush administration's perception about the North Korean leadership is very different from that of the Kim Dae-jung government and the Clinton administration. Using the terms presented in previous sections, the Bush administration seems to place much higher probability on the deceitful North Korean type than its predecessor and its counterpart in South Korea. However, our finding in the previous section (and Observation 4 in Appendix B) indicates that the difference in the choice of actions by the Clinton administration and the Bush administration may indeed stem from the difference in their own type (preferences) and have little to do with their perceptions of North Korea.

In June 2001, President Bush signaled a change in his administration's policy toward North Korea by announcing that the United States will resume bilateral talks with North Korea. His proposed agenda for future talks included the North Korean nuclear development, missile development, and its conventional weapons deployment. President Bush further implied potential economic aid and political cooperation with the North, contingent upon the North's accommodation with the aforementioned agendas (*Hankuk Ilbo* June 8, 2001). Secretary Colin Powell, during his visit to Seoul in July, stated that the United States was ready to talk to North Korea without preconditions. He further called on President Vladimir Putin of Russia to persuade Kim Jong-il to abandon weapons of mass destruction and resume the dialogue for peace with both South Korea and the United States⁵ (*International Herald Tribune* July 29, 2001).

⁴ Other countries that "made" this list were Cuba, Iran, Iraq, Libya, Sudan, and Syria.

⁵ Secretary Powell also attempted to convince South Korean President Kim Dae-jung as well as Unification Minister Lim Dong-won and Foreign Minister Han Seung-soo of Washington's desire

We need to note that North Korea did not do anything to change the U.S.' perception of North Korea between January and June of 2001. Nevertheless, the Bush administration's North Korea policy began to change. The U.S. media interpreted this change of policy as a sign that the moderate State Department got an upper hand on the issue of North Korea over the more hawkish White House and Defense Department (see for example *Washington Post* June 7, 2001). If this was indeed the case, then the change of policy was not caused by the change in U.S. policy makers' perception of North Korea but by the change in the U.S.' own type from hawkish to dovish (or less hawkish)! This situation fits exactly with what our model is telling us (Observations 4 and 5).

The Simultaneous Nature of the Two Games

As we mentioned above, North Korea is playing the two separate games simultaneously. The dilemma it faces is that it is not flexible in its choice of strategies (and thus, its attempt to portray itself as sincere) since an opponent in one game can observe the North Korean choice of actions in the other game.⁶ For example, it is in the best interest of the North Korean leadership to portray itself as a true reformer in its game with South Korea (Observation 3). The same is not necessarily the case in the U.S.-North Korean interaction (Observation 6). Nevertheless, due

to cooperate closely with South Korea on relations with the North. In an effort at allaying Seoul's concerns that the Bush administration is not sympathetic to Mr.Kim's "Sunshine" policy of rapprochement with North Korea, Mr.Powell emphasized "President Bush's strong support for engagement with the North."

⁶ From the modeling point of view, an ideal way to analyze the North-South Korean interaction and the U.S.-North Korean interaction together would be to treat these two interactions as nested games in a single model (see Tsebelis 1990). Given the technically complicated nature of these interactions (as in two Figures and two Appendices), however, we chose not to incorporate them into a single model in this paper, leaving it as a future research topic.

to the simultaneous and observable nature of the two games, the North Korean leadership cannot portray itself to be two conflicting types in two separate games. This means that, if the relationship strains, and the negotiations stall in one game, the same is likely to happen in the other. This is exactly what happened in early 2001. That is, when the U.S.-North Korean dialogue stalled, North Korea also stopped responding to the South Korean gesture for continuous dialogue. The North Korean leadership further toughened its stance in August. After a summit in Moscow, Kim Jong-il and the Russian President Putin re-iterated their old argument that the withdrawal of the U.S. forces from South Korea was critical for peace and security in the Korean Peninsula and Northeast Asia.⁷ (*Joongang Ilbo* August 6, 2002).

The simultaneous nature of the two games analyzed in this paper and the restraint on the North Korean leadership in their choice of strategies explains why the President Bush's announcement to engage North Korea again received more public and positive response from South Korea than the North (*Hankuk Ilbo* June 8, 2001). The South Korean leadership understood that once the negotiation resumed between the United States and North Korea, the same thing would happen between North Korea and South Korea.

Conclusion

In this paper we studied the changing relationships North Korea has with South Korea and the United States. As we demonstrated above, both South Korean and U.S. policy makers

⁷ The two leaders further announced, "the North Korean missile development was for a peaceful purpose and did not pose any threat to those countries honoring North Korea's sovereignty." They also emphasized that the 1972 Anti-Ballistic Missile Treaty was the basis of global strategic stability. Obviously Mr. Putin utilized the summit with Kim Jong-il to criticize the Bush administration's concept of Missile Defense.

have attempted to formulate their new North Korea policies while they were uncertain about the true intentions of the North Korean leadership, especially since the late 90's. By applying Bayesian games to the Korean situation, we could make some very interesting observations.

In the interaction between North Korea and South Korea, for example, the North will be consistently accommodating or consistently unaccommodating regardless of whether it truly wants change or not. This will make it difficult for the South to collect more information about the real intentions of the North Korean leadership. We recommend that the South must continue to engage the North only if it strongly believes that the North is truly interested in reforming its system for the long-term. Otherwise, the South should pull away from its initial accommodation even if it results in the loss of face on the part of the Kim Dae-jung government. Finally, we expect that the North Korean leadership, regardless of its real intentions, will try to convince the South that it is sincere. We will continue to see the North Korean attempt to gesture its willingness to reform.

In the interaction between North Korea and the United States, we expect that the hawkish United States will never initiate accommodation and will always reject the initial accommodation by North Korea. On the other hand, the dovish United States will always reciprocate the accommodation initiated by North Korea. It will also initiate accommodation in the absence of North Korea's initial accommodation, but only when it feels highly certain that North Korea truly desires change. The deceitful North Korea never initiates accommodation and will always reject the initial accommodation by the United States. Based on these observations, we predict that the long-term reconciliation between North Korea and the United States is possible if, and only if, North Korea is sincere about reform and the United States is dovish toward North Korea.

Further, the combination of the deceitful North Korea and the hawkish United States always leads to Status Quo, since neither side is willing to risk being exploited by the other.

As we discussed above, the new Bush administration in the United States toughened its policy toward North Korea, citing that it perceived the latter as a threat to world piece. By all indications, the Bush administration seemed to place much higher probability on the deceitful North Korean type than its predecessor and its counterpart in South Korea. Since June of 2001, however, the U.S. government has softened its policy toward North Korea by announcing its willingness to engage North Korea without preconditions. We believe that this change in U.S. policy toward North Korea was caused by the change in the U.S.'s own type as our model indicates and as some media coverage reported, rather than the change in the Bush administration's perception of the North Korean regime.

Another aspect of the North-South Korea-United States relations, which makes them interesting, is that the North Korean leadership with private information about its real intentions is simultaneously playing two separate games, one with South Korea and the other with the United States. The dilemma it faces is that it is not flexible in its choice of strategies (and thus, its attempt to portray itself as sincere) since an opponent in one game can observe the North Korean choice of actions in the other game. For example, it is in the best interest of the North Korean leadership to portray itself as a true reformer in its game with South Korea (Observation 3). The same is not necessary the case in the U.S.-North Korean interaction (Observation 6). Nevertheless, due to the simultaneous and observable nature of the two games, the North Korean leadership cannot portray itself to be two conflicting types in two separate games. This means that, if the relationship strains, and the negotiations stall in one game, the same is likely to

happen in the other, as it happened in 2001 when the U.S.-North Korean stalemate caused the North-South Korean relationship to strain as well.

References

- Ahn, Byeonggil. 1995. "Domestic Uncertainty and Coordination between North and South Korea." In *Rationality and Politics in the Korean Peninsula*, eds. HeeMin Kim and Woosang Kim. East Lansing: Michigan State University, 119-142.
- Alt, James E., Randall L. Calvert, and Brian D. Humes. 1988. "Reputation and Hegemonic Stability: A Game-Theoretic Analysis." *American Political Science Review* 82: 445-466.
- Bueno de Mesquita, Bruce, James D. Morrow, and Ethan R. Zorick. 1997. "Capabilities, Perception, and Escalation." *American Political Science Review* 91: 15-27.
- Fearon, James D. 1994. "Signaling versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model." *Journal of Conflict Resolution* 38: 236-269.
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88: 577-592.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49: 379-414.
- Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands Versus Sinking Costs." *Journal of Conflict Resolution* 41: 68-90.
- Hankuk Ilbo* [The Korea Daily News], various dates, 2000, 2001.
- Iida, Keisuke. 1993. "When and How Do Domestic Constraints Matter?" *Journal of Conflict Resolution* 37: 403-426.
- International Herald Tribune*, July 29, 2001.
- Joongang Ilbo*, July 30, August 6, 2001.
- Kim, Woosang. 1995. "South Korea's Foreign Policy Strategies toward Main Actors in the Northeast Asia." In *Rationality and Politics in the Korean Peninsula*, eds. HeeMin Kim and Woosang Kim. East Lansing: Michigan State University, 101-118.
- Kim, Woosang and Bruce Bueno de Mesquita. 1995. "How Perceptions Influence the Risk of War." *International Studies Quarterly* 39: 51-65.
- Los Angeles Times*, March 8, 2001.
- Morrow, James D. 1989. "Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining." *American Journal of Political Science* 33: 941-972.

Morrow, James D. 1994. "Alliances, Credibility, and Peacetime Costs." *Journal of Conflict Resolution* 38: 270-297.

Nalebuff, Barry. 1991. "Rational Deterrence in an Imperfect World." *World Politics* 43: 313-335.

New York Times, October 15, 2000.

Powell, Robert. 1996. "Uncertainty, Shifting Power, and Appeasement." *American Political Science Review* 90: 749-764.

Schultz, Kenneth A. 1998. "Domestic Opposition and Signaling in International Crises." *American Political Science Review* 92: 829-844.

Smith, Alastair. 1995. "Alliance Formation and War." *International Studies Quarterly* 39: 405-425.

Smith, Alastair. 1998. "Extended Deterrence and Alliance Formation." *International Interactions* 24: 315-343.

Tsebelis, George. 1990. *Nested Games*. Berkeley: University of California Press.

Wagner, R. Harrison. 2000. "Bargaining and War." *American Journal of Political Science* 44: 469-484.

Washington Post, June 7, 2001.

Werner, Suzanne. 2000. "Deterring Intervention: The Stakes of War and Third-Party Involvement." *American Journal of Political Science* 44: 720-732.

Appendix A

We solve for the equilibrium of the North-South Korea Bayesian game. The most commonly used solution concept for dynamic Bayesian games is Perfect Bayesian Equilibrium (PBE), which we adopt here. First of all, it is easy to see that North Korea has a dominant action on its last move. The sincere type will always choose A' while the deceitful type will choose $\sim A'$. Now we narrow our attention down to North Korea's first move.

[1] AA (pooling on A)

$$U_{SK}(A) = p' U_{SK}(RECON) + (1-p') U_{SK}(EXP_N)$$

$$U_{SK}(\sim A) = U_{SK}(REJ_S)$$

If SK's best response to NK's choice of AA is $\sim A$, then NK has an incentive to switch to $\sim A$. Therefore the PBE does not include SK's choice of $\sim A$. If SK's best response to NK's choice of AA is A, then both types of NK get the best possible outcomes and do not have incentive to switch. This is true when $U_{SK}(A) \geq U_{SK}(\sim A)$. This inequality condition is satisfied when $p' \geq [U_{SK}(REJ_S) - U_{SK}(EXP_N)] / [U_{SK}(RECON) - U_{SK}(EXP_N)]$.

Therefore, a PBE is: $\{(A/A', A/\sim A'), A, p' \geq [U_{SK}(REJ_S) - U_{SK}(EXP_N)] / [U_{SK}(RECON) - U_{SK}(EXP_N)]\}$.

[2] A \sim A (separating with the sincere type choosing A)

$p' = 1$. SK's best response is A, and it receives $U_{SK}(RECON)$ or $U_{SK}(SQ)$ depending on NK's type. A \sim A is the equilibrium strategy for NK only if SK's best response to NK's choice of A is $\sim A$. This is incompatible with $p' = 1$. So, there is no PBE under this scenario.

[3] \sim AA (separating with the sincere type choosing $\sim A$)

$p' = 0$. In response to SK's best response, ($\sim A$), NK has incentive to change its current strategy. Therefore, there is no PBE under this scenario.

[4] $\sim A\sim A$ (pooling on $\sim A$)

SK's information set is off the equilibrium path. For $\sim A\sim A$ to be part of the equilibrium profile, SK's best response to NK's deviation must be $\sim A$ (that is, $U_{SK}(A) \leq U_{SK}(\sim A)$). This is the case when $p' \leq [U_{SK}(REJ_S) - U_{SK}(EXP_N)] / [U_{SK}(RECON) - U_{SK}(EXP_N)]$.

Therefore, a PBE is: $\{(\sim A/A', \sim A/\sim A'), \sim A, p' \leq [U_{SK}(REJ_S) - U_{SK}(EXP_N)] / [U_{SK}(RECON) - U_{SK}(EXP_N)]\}$.

The details of the equilibrium solution above lead to the following observations:

Observation 1: All PBEs in the North-South-Korea game are pooling equilibria.

From [1] and [4] above.

Observation 2: SK should pursue A only when p' is high.

From [1] above.

Observation 3: The PBE under the first pooling equilibrium is more efficient for NK than the one under the second. This means that both types of NK will try to convince SK that p is high.

From [1] and [4] above.

Appendix B

We solve for the equilibrium of the U.S.-North Korea Bayesian game here. Again we use the notion of Perfect Bayesian Equilibrium (PBE). First of all, it is easy to see that North Korea has a dominant action on its last move. The sincere type will always choose A' while the deceitful type will choose $\sim A'$. Now we narrow our attention down to North Korea's first move.

[1] AA (pooling on A)

When NK is pooling on A, the US equilibrium strategy is $\sim AA \sim AA$. Since at least the deceitful NK always has an incentive to switch, there is no PBE under this scenario.

[2] A \sim A (separating with the sincere type choosing A)

When the US observes NK's choice of A, then $p' = 1$ and the US's best response is $\sim AA$. When the US observes NK's choice of $\sim A$, then $p' = 0$ and the US's best response is $\sim A \sim A$. The deceitful NK does not have incentive to switch. For sincere NK, $U_{NK}(A) = qU_{NK}(REJ_{US}) + (1-q)U_{NK}(RECON_{NK})$ and $U_{NK}(\sim A) = qU_{NK}(SQ) + (1-q)U_{NK}(RECON_{US})$. For sincere NK not to have incentive to switch from its equilibrium strategy, $U_{NK}(A) \geq U_{NK}(\sim A)$ must hold. It is true when $q \geq [U_{NK}(RECON_{US}) - U_{NK}(RECON_{NK})] / [U_{NK}(RECON_{US}) + U_{NK}(REJ_{US}) - U_{NK}(RECON_{NK}) - U_{NK}(SQ)]$.

Therefore a PBE is: $\{(A/A' \sim A/\sim A'), (\sim AA \sim A \sim A), q \geq [U_{NK}(RECON_{US}) - U_{NK}(RECON_{NK})] / [U_{NK}(RECON_{US}) + U_{NK}(REJ_{US}) - U_{NK}(RECON_{NK}) - U_{NK}(SQ)]\}$.

[3] $\sim AA$ (separating with the sincere type choosing $\sim A$)

When the US observes NK's choice of A, then $p' = 0$ and the US's best response is $\sim AA$. When the US observes NK's choice of $\sim A$, then $p' = 1$ and the US's best response is $\sim A \sim A$. Since at least the deceitful NK always has an incentive to switch its action, there is no PBE under this scenario.

[4] $\sim A \sim A$ (pooling on $\sim A$)

The hawkish US's best response is $\sim A$. For dovish US, its best response is A if $p' \geq [U_{US}(SQ) - U_{US}(REJ_{NK})] / [U_{US}(RECON_{US}) - U_{US}(REJ_{NK})]$. Otherwise, it is $\sim A$. The deceitful NK does not have an incentive to switch. If the dovish US's equilibrium action is A , then the sincere NK does not have an incentive to switch, either.

Therefore a PBE is $\{(\sim A/A' \sim A/\sim A'), (\sim AA \sim AA), p' \geq [U_{US}(SQ) - U_{US}(REJ_{NK})] / [U_{US}(RECON_{US}) - U_{US}(REJ_{NK})], q\}$.

If, on the other hand, the dovish US responds with $\sim A$ ($p' < [U_{US}(SQ) - U_{US}(REJ_{NK})] / [U_{US}(RECON_{US}) - U_{US}(REJ_{NK})]$), the sincere NK has incentive to switch to A if $U_{NK}(A) \geq U_{NK}(\sim A)$. To prevent this, $U_{NK}(\sim A) \geq U_{NK}(A)$ must hold for the sincere NK. This is true when $q \geq [U_{NK}(RECON_{NK}) - U_{NK}(SQ)] / [U_{NK}(RECON_{NK}) - U_{NK}(REJ_{US})]$.

Therefore another PBE is $\{(\sim A/A' \sim A/\sim A'), (\sim A \sim A \sim A \sim A), p' < [U_{US}(SQ) - U_{US}(REJ_{NK})] / [U_{US}(RECON_{US}) - U_{US}(REJ_{NK})], q \geq [U_{NK}(RECON_{NK}) - U_{NK}(SQ)] / [U_{NK}(RECON_{NK}) - U_{NK}(REJ_{US})]\}$.

The equilibrium solution above leads to the following observations:

Observation 4: The hawkish U.S.'s equilibrium strategy does not depend on NK's type.

Simple backward induction shows that the hawkish US's optimal strategy is always $\sim A$ regardless of NK's type.

Observation 5: The dovish U.S. will always reciprocate the accommodation initiated by NK. Whether the dovish U.S. will initiate accommodation in the absence of NK's initial accommodation will depend on the value of p' .

Observation 6: The deceitful NK never initiates A .

This comes from the three PBEs under [2] and [4] above.

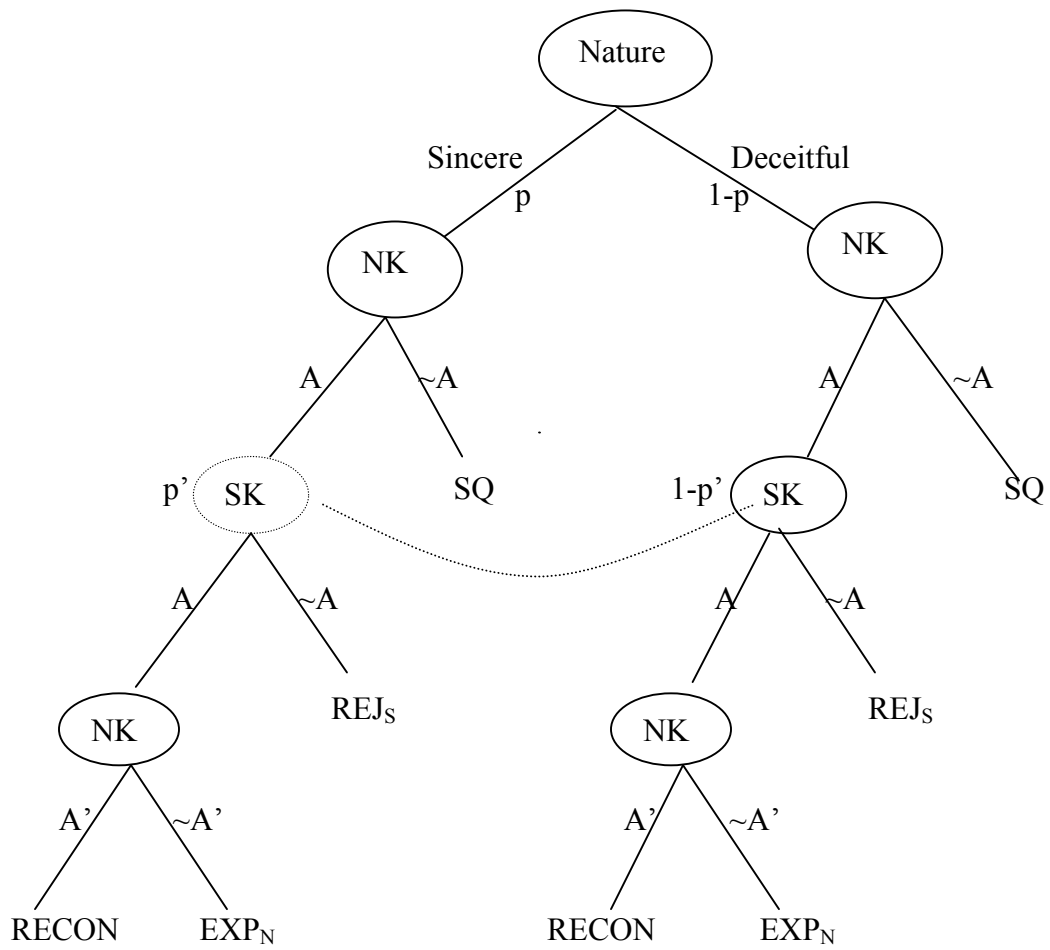
Observation 7: The long-term reconciliation is possible iff NK is sincere and the US is dovish.

Among the three PBEs above, only the combination of a sincere NK and a dovish US leads to RECON outcomes.

Observation 8: The combination of a deceitful NK and a hawkish US always leads to SQ.

From the three PBEs above.

Figure 1. The North-South Korea Bayesian Game



SK: $U(\text{RECON}) > U(\text{REJ}_S) > U(\text{EXP}_N)$

NK_{SINCERE}: $U(\text{RECON}) > U(\text{EXP}_N) > U(\text{REJ}_S)$

NK_{DECEITFUL}: $U(\text{EXP}_N) > U(\text{RECON}) > U(\text{REJ}_S)$

Figure 2. The U.S.-North Korea Bayesian Game

