Risk Manipulation and Economic Extortion: A Game Theoretic Approach to Diplomatic Relations between the United States and North Korea

By Melanie Hohlfeld
Economics 191AB
“Nuclear weapons have one, pre-eminent use in politics: to support threats. It is the threat to compel or punish – in short, coercion – which is the peacetime tool of diplomacy.”

For more than a decade the United States and the Democratic People’s Republic of Korea have contested the proliferation of nuclear weapons by North Korea in a game that has threatened to disrupt the security and stability of a region occupied by more than 90,000 U.S. troops and home to some of the world’s largest economies.

Using a game-theoretic lens to analyze the ambitions and interactions of each state, this paper seeks to explain the origins and strategies of nuclear brinkmanship between the United States and the North Korean regime under the Clinton and George W. Bush administrations. My analysis will examine the interactions between North Korea and the United States as a bargaining situation where the basic game is repeated over time.

The first two sections will illuminate basic elements of each player’s payoff structure. Section I will answer two basic questions: what has motivated North Korea’s dedication to nuclear weapons development; and why has the U.S. attempted to prevent said development. It will present the social, political, and economic factors that have influenced the DPRK’s nuclear pursuits and the U.S. interests in countering these moves under the Clinton and Bush administrations. In a similar form section II will identify the known nuclear capabilities of North Korea and illuminate the principal sources of uncertainty about its weapons inventory through a brief summary of its nuclear energy program.

The subsequent sections will develop and employ game theoretic techniques to analyze the nuclear brinkmanship in which North Korea and the United States have been engaged. The third section of the paper presents a theoretical discussion of bargaining and risk management in the context of international relations: how does game theory explain the diplomatic tools and strategies of the state? This will build the foundation for a formalization of the game in sections IV and V. Section IV will

---

examine the conflict of the Clinton Administration as a bargaining game of incomplete information that concluded with the formation of the Agreed Framework in 1994, a contract by which North Korea agreed to freeze nuclear weapons development in exchange for economic and aid concessions from the international community. In the first half of this discussion we will restrict our analysis to a simple two-stage game; this will illuminate the decision-making procedures of each player in an environment of uncertainty. The latter half of section IV will consider a more dynamic contract bargaining process by expanding the discussion to include multiple rounds of negotiation. This will illustrate a more fluid process by which players explore and increase the risk of mutual disaster to coerce his opponent’s behavior. Section V will explore the second game, the current crisis facing the second Bush administration, as a repeat of previous play. Having defined the parameters and payoffs of the game in section IV, section V will show how the flow of information across time can influence the actions and associated payoffs available to each player.
Section I: A Brief Narration of Historical, Political, and Economic Motivations

To understand how bilateral relations deteriorated into nuclear brinkmanship during the Clinton and Bush administrations, it is necessary to first understand the incentives of each state.

Understanding North Korea

While the entire Korean peninsula merited the sobriquet “the Hermit Kingdom” as early as the 19th century, in modern contexts this title most aptly applies to the Democratic People’s Republic of Korea (DPRK), also known as North Korea. North Korea has sustained a fifty year-long isolationist foreign policy informed in part by its communist party structure and in part by the ideology of juche, or self-reliance. This isolation has constrained its avenues for engagement with the international community, both economically and politically, and fostered that conditions that have motivated its nuclear weapons development.

The DPRK’s ambitions for a nuclear arsenal can be broadly defined as a function of two forces: the economic decay of the DPRK since the late 1980s, and an insecurity complex shared by the Kim Il-sung and Kim Jong-il regimes. The former motivates the DPRK to develop a nuclear weapons as economic tools. A nuclear arsenal may increase its bargaining power to gain economic and political concessions from other international actors or it may act as a revenue raiser for the DPRK by opening a black market for nuclear weaponry. The latter explains DPRK motivations as emanating from the ruling regime’s paranoia of takeover by international powers.

Nuclear weapons as economic tools. The North Korean economy has been in ruin since the late 1980s. Exhibiting many of the problems suffered by other socialist economies, the DPRK saw its economic deterioration accelerated with the collapse of the USSR as it could no longer rely on goods at subsidized prices from the Soviet Union. In January 1991, the Soviet Union withdrew nearly forty percent of the trade concessions previously offered to the DPRK, which contributed to an estimated twenty-five
percent decline in the DPRK’s foreign trade.\(^2\) While this decay in trade was partially offset by increases in trade with China and South Korea, the demise of the USSR marked the beginning of a severe economic downspiral for the North.

In more recent years North Korea’s economic decline has been exacerbated by attempts to reform. In July 2002 the DPRK loosened price controls over staple goods and promised increased wages to counter the expected price increases. As expected, prices have risen in response to these reforms,\(^3\) in some areas inflating the price of staple goods by as much as 400%,\(^4\) but unfortunately many of the promised salary increases have not been honored.

Parallel to, and a direct consequence of, its economic isolation and contraction has been a decade-long famine, which has affected most of the North Korean population. The signs of food shortages surfaced under the rule of Kim Il-Sung, Kim Jong-Il’s father, who instituted the “Let’s Eat Two Meals a Day” campaign. The problems of the father have continued to plague the son. As its economic problems have compounded, so have its food shortages. Since 1994 North Korea has lost perhaps as many as two million people, or 10 percent of its population, to famine\(^4\) and it is estimated that today more than 45 percent of children under the age of five are chronically malnourished.\(^5\) The famine serves as the most tangible evidence that the interests of the ruling regime may no longer align with the people of the DPRK. It has been shown that food relief offered by the U.N. World Food Program to ameliorate the suffering of the North Koreans has been siphoned from the peasant population to feed the army,\(^6\) a body which preserves regime power.

One of North Korea’s largest obstacles to economic recovery has been a chronic shortage of energy. The regime once maintained that this was the sole purpose of its nuclear reactors: an alternative

---

\(^2\) Official economic statistics are not published by the North Korean government but estimated by academics and government agencies from other countries.


\(^4\) Daalder, Ivo H. and James M. Lindsay, “Where are the Hawks on North Korea?” The American Prospect, Volume 14 Issue 1, February 1, 2003.


source of energy. The North Korean economy, built largely on heavy-industry products, has remained dependent on oil to its detriment. Prior to the collapse of the Soviet Union in 1991, the DPRK obtained oil at subsidized prices from the USSR. But with the fall of the USSR, and lacking its own domestic resources, North Korea has since been saddled with insufficient energy supplies. Feedback between the DPRK’s energy shortages and its economic decline has tied the North Koreans in a Gordian knot: the decay of its economy has hampered its ability to purchase necessary oil supplies, which in turn has accelerated the deterioration of its economy.

To combat the decline of its economy and concomitant energy crisis, the DPRK regime has found its nuclear program to be a strategic asset to gain economic concessions, energy aid, and reduced sanctions from the U.S. and its allies. For example, on June 16, 1998 Pyongyang publicly announced:

If the U.S. really wants to prevent our missile export, it should lift the economic embargo as early as possible and make a compensation for the losses to be caused by the discontinued missile export.8

Threats of nuclear war, as well as threats to sell its nuclear warheads, have proved successful in the extortion of additional food aid and heavy fuel oil from other nations beyond the food aid provided by the United Nations World Food Program.9 Thus the depression of the DPRK’s economy explains in part the regime’s apparent willingness to risk everything to build nuclear weapons, hoping to extract concessions from the outside world.

Self-Preservation Concerns of the Kim Il-sung, Kim Jong-Il Regimes. Fifty years of economic and political isolationism have left the DPRK a nation without allies. The paranoia that plagues the ruling class motivates its pursuit of nuclear weapons as a nuclear arsenal looms large as the ultimate equalizer

---

7 It has been quoted as 2/3 the world’s market price. See Noland, Marcus. Avoiding the Apocalypse: the Future of the Two Koreas, Washington, DC, Institute for International Economics, 2000.
among states, a tool to prevent potential rivals like the United States from toppling or challenging the legitimacy of its communist elite.

If the regime’s commitment to its own preservation is questioned, its importance can be seen first in the continued deterioration of the North Korean economy, and as a corollary, in the continued hardship of its people. Denny Roy of the National University of Singapore illuminates the regime’s interests no longer represent the national interests:

While liberalization would strengthen North Korea’s economy, it would also threaten the survival of the regime, something the Kim [regimes] value even more than national prosperity. Any economic liberalization would entail a freer flow of information, quickly destroying the myths upon which the regime’s legitimacy rests.10

Scholars have debated the merits of an ostensible opening of the DPRK through recent reform measures. Unilaterally, it has enacted a succession of economic and market reforms: increases in wages, floatation of good prices, and the establishment of a special economic zone in the northwestern town of Sinuiju, complete with its own legal codes and election processes.11 In terms of working bilaterally with South Korea, the DPRK agreed to re-establish road and rail links with its southern capitalist neighbor. Other examples of a DPRK reversal of past isolationism include summits with high-level officials of other states, sending representatives to the 2002 Asian Games, and de-mining portions of the Demilitarized Zone.12 Taken as a whole, North Korea has demonstrated by its various commitments a greater willingness to engage on an international scale in both political and economic spheres. However some more skeptical scholars cast these ostensible “openings” of North Korea as the mere “spasms of a dying regime”13 pressured by the forces of globalization. These critics argue that recent steps to reform constitute illusory promises of committed change. The underscore that all reform measures continue to minimize contact between DPRK civilians and foreigners. As it remains uncertain how willing or

12 Laney, James T. and Jason T. Shaplen, “How to Deal with North Korea” Foreign Affairs, March/April 2003, p17.
prepared is the DPRK regime to open its doors to the international community, experts have concluded that regime survival achieved through economic and political isolationism remains an important priority of the ruling elite.

Some scholars locate the insecurity of the DPRK regime as a direct consequence of continued international hostilities towards North Korea in the last 60 years. In particular, Dr. Leon Sigal of Columbia University identifies the source of North Korean paranoia and nuclear ambitions as continued aggression against the DPRK by the U.S. and its allies. He notes that “no country the target of more American nuclear threats than North Korea – at least seven since 1945.” Similarly the Nautilus Institute describes the North Korean procedural rule for engaging international relations as “respond-in-kind.” Game theorists label this rule “tit-for-tat.” From this perspective, DPRK incentives to obtain nuclear weapons are renewed with each U.S. move to disarm the regime. Evidence of U.S. “aggression” increasing the regime’s incentives to possess nuclear weapons was manifest in the DPRK’s first public response to the U.S. invasion of Iraq. North Korea explained that the war in Iraq had only strengthened its resolve to possess nuclear weapons: “The Iraqi war teaches a lesson that in order to prevent a war and defend the security of a country and the sovereignty of a nation it is necessary to have a powerful physical deterrent.”

Understanding U.S. interests

The United States has a long history of investment in the political and economic stability of Northeast Asia. Three interests in particular have motivated both the Clinton and Bush administrations to disarm North Korea: economic stability with regional trade partners; a history of bilateral security alliances with South Korea and Japan; and the forestallment of a nuclear arms race. In addition to these

---

interests, shared by both administrations, the Bush administration has also added another primary consideration: the protection of U.S. civilians from the threat of terrorism.

Protection of Economic Interests. Economic interests in the Northeast Asian region have been a critical motivation for the U.S. prevention of the DPRK’s acquisition of a nuclear arsenal. Many regional actors are valuable trading partners with the U.S. South Korea is the United States’ sixth-largest export market, representing $22.5 billion, $22.1 billion and $27.8 billion dollars in 2002, 2001, and 2000, respectively. In 1994 the U.S. exported $18 billion worth of goods to South Korea. Conversely, the U.S. is South Korea's largest export market, followed by the European Union and Japan: its exports to the U.S. were $19 billion in 1994, $40 billion in 2000 and $35 billion in 2002.17 Another key regional market is Japan. Japan's industrialized, free market economy is the second in size only to the United States. As the United States' third-largest trading partner, U.S. exports to Japan were $53 billion in 1994 and $51 billion in 2001; U.S. imports from Japan were more than $121 billion in 2002, an increase from $119 billion in 1994.18 China has also experienced rapid growth and its stature with the U.S. has increased commensurately. The influence of China in other international markets will only increase with its ascension to membership in the World Trade Organization. As relations between the U.S. and China have warmed since 1994, so has trade between the two nations. U.S. exports to China have more than doubled from $9 billion in 1994 to $22 billion in 2002 and U.S. imports of Chinese goods have more than tripled from only $38 billion in 1994 to $125 billion in 2002.19

A nuclear Kim Jong-II could threaten to cripple trade activity with these key actors through the extortion or potential destruction of regional economic centers.

---

18 U.S. Census Bureau, Foreign Trade Division, Data Dissemination Branch, Washington, D.C. 20233.
19 U.S. Census Bureau, Foreign Trade Division, Data Dissemination Branch, Washington, D.C. 20233.
Protection of Regional Allies and Prevention of Nuclear Arms Proliferation. America’s access to and influence within Northeast Asian markets has been fundamentally linked to its security alliances with South Korea and Japan. This year marks the 50th anniversary of the South Korea-United States Mutual Defense Treaty, signed October 1, 1953; parallel to its agreement with South Korea, the U.S. has maintained the U.S.-Japan Treaty of Mutual Cooperation and Security, signed January 19, 1960. The U.S. currently maintains approximately 90,000 troops in the region: 53,000 troops in Japan and approximately 37,000 service personnel in Korea in support of its commitments. Historical and geopolitical considerations have inextricably linked the two security alliances. As explained by Dr. Robyn Lim at Nanzan University in Japan: “The American presence in South Korea has been as much about protecting Japan as it was about South Korea.”

The emergence of a nuclear North Korea adds to an already formidable threat to both states. North Korea has developed the fifth largest standing army in the world, comprised of 1.1 million members, and represents the most militarized state in the world. About 80 percent of the DPRK’s active conventional weapons units are within 100km of the Demilitarized Zone. The North possesses 13,000 artillery pieces and could fire 400,000 shells, many combined with sarin and anthrax, in the first hour of an attack on the 21 million people residing in Seoul. A launch of the Taepodong-1 missile on August 31, 1998 demonstrated that Japan is equally vulnerable to an attack by the DPRK. From state-sanctioned terrorism, demonstrated by the bombing of a South Korean airplane before the 1988 Olympics in Seoul, to assassination attempts on South Korean presidents in 1968 and 1983, the bellicosity of North Korea has been well-documented.

---

The advent of a nuclear DPRK could trigger a regional nuclear arms race.\textsuperscript{24} The United States has pledged nuclear protection to South Korea and Japan to convince them that the acquisition of nuclear weapons is unnecessary. However if its northern neighbor possesses nuclear weapons, South Korea is more likely to arm itself with this technology. If both Koreas wield nuclear weaponry, Japan would be the only major regional power without and thus would also seek nuclear armaments as a defense. CIA Director George Tenet explains, “Additional countries may decide to seek nuclear weapons as it becomes clear their neighbors and regional rivals are already doing so. The ‘domino theory’ of the 21\textsuperscript{st} century may well be nuclear.”\textsuperscript{25} The greater the number of states that possess nuclear weapons, holding all other things equal, the greater the chance of nuclear war.

In addition to the prevention of a regional nuclear arms race, U.S. commitments to the nonproliferation of nuclear weapons in North Korea have supported its threats and promises in other regions. To illustrate, if the U.S. did not oppose North Korea’s weapons development, this could undermine the credibility of its commitments to prevent the proliferation of weapons of mass destruction in other regions, such as the Middle East.

Thus to forestall the proliferation of nuclear-weapons proliferation, to honor the security commitments to its regional allies, and to preserve regional stability, the U.S. has sought to prevent North Korea from acquiring nuclear-weapons capabilities.

\textit{Homeland Security.} The September 11 attack on the World Trade Center marked a shift in American foreign policy from a doctrine of measured response to one of pre-emptive defense. In a speech at West Point, New York on June 1, 2002, President Bush laid plain the concerns of his administration:

\begin{quote}
When the spread of … nuclear weapons, along with ballistic missile technology – when that occurs, even weak states and small groups could attain catastrophic power to strike
\end{quote}

\textsuperscript{24} Japan has already pledged to launch a pre-emptive conventional weapons strike against North Korea if it believes Korea is preparing a ballistic missile attack on it, according to Japan’s Defense Minister Shigeru Ishiba.

great nations. They want the capability to blackmail us, or to harm us, or to harm our friends – and we will oppose them with all our power.26

A nuclear North Korea would present on the one hand a direct threat to American soil: if the DPRK possesses eight or ten weapons, it has a much greater chance of delivering one to the U.S. successfully.

But a nuclear North Korea could also present a second, indirect threat to American security: it could sell nuclear weapons to the highest-bidder – Iran, al-Qaeda, or otherwise. As Colin Powell asked, “What [is North Korea] going to do with another two or three nuclear weapons when they’re starving, when they have no energy, when they have no economy that’s functioning?”27 The value of nuclear weapon exportation is plain, and North Korea has a history of ballistic missile exportation. CNN has reported that North Korea earned $560 million per year on missile exports, which is “equivalent to about 40 or 50 percent of its total exports.”28 It has also regularly bartered weapons including missiles, missile launchers, nuclear technology, tunneling technology, and warhead canisters. The buyers have often been to countries in the Middle East, including Syria, Pakistan, and Iran, and the deals have often been made in exchange for oil. In December 2002, Spanish warships working with American military and intelligence officials stopped a Korean freighter found to be transporting 15 Scud missiles bound for Yemen.29 It is in the interest of the United States to prevent the opening of a nuclear bazaar available to rogue states, conventional nations, or non-state terrorist groups.

27 Daalder, Ivo H. and James M. Lindsay, “Where are the Hawks on North Korea?” The American Prospect, Volume 14 Issue 1, February 1, 2003.
Section II: North Korea’s Nuclear History

A short exposition of North Korea’s nuclear program will illuminate the sources of uncertainty that have influenced the bilateral interactions of the U.S. and the DPRK.

Initial Steps: 1950s-1990

The pursuit of North Korea’s nuclear ambitions began as early as the 1950s. With assistance from nearby nuclear powers, Russia and China, North Korea was able to construct and mobilize a small two-megawatt nuclear reactor at Yongdong in 1967. However the DPRK did not initiate broader efforts to utilize nuclear technology until the late 1970s and early 1980s. With the completion of the Yongdong reactor in 1979 the DPRK broke ground on a five-megawatt nuclear reactor at Yongbyon and two gas-graphite reactors, one at Yongbyon and another at Taechon.

Suspicions of nuclear weapons development for military purposes have mushroomed since the inception of North Korea’s nuclear program. Rumors circulated that concurrent with the construction of the Yongbyon reactor in 1979 North Korea had also initiated building of underground weapons facilities within the geographic vicinity of this reactor. This suspicion of the DPRK’s objectives was not limited to the United States intelligence communities. The USSR, uncertain of North Korea’s true ambitions, required North Korea to accede to the Nuclear Nonproliferation Treaty before providing four light-water nuclear reactors (LWR).30

In 1986, the Yongbyon five-megawatt reactor became operational; international anxiety of nuclear weapons-development has since centered on this site. After only one year of operations, U.S. intelligence reported the DPRK was preparing the Yongbyon reactor for purposes other than energy generation, namely the reprocessing of spent reactor fuel into weapons-grade plutonium. A succession of shutdowns of the Yongbyon reactor in 1989, 1990, and 1991 only elevated the fears of U.S. intelligence communities. While the aim of the shutdowns remained uncertain, a primary reason to

---

30 The DPRK joined the NPT December 12, 1985. However as relations between North Korea and USSR deteriorated, the USSR never supplied North Korea with the four LWRs.
suspend operations would be to remove reactor fuel.\textsuperscript{31} It has remained unclear if the DPRK removed reactor fuel at these junctures. If the DPRK had seized the opportunity of the reactor shut-down to remove reactor fuel, it would be possible that it had reprocessed this plutonium to make a nuclear warhead before IAEA inspectors entered the DPRK to inspect its nuclear facilities in 1993. Thus, the DPRK could have in its possession one or two nuclear weapons. According to Marcus Noland, “This unanswered question is the kernel of uncertainty surrounding the North’s \textit{existing} nuclear weapons inventory.”\textsuperscript{32}

\textit{Nuclear Development during the Clinton Administration}

The Yongbyon reactor remained on-line until its operations were frozen under the terms of the Agreed Framework, formed in 1994. In the years leading up to the Agreed Framework, U.S. intelligence and IAEA inspection teams confirmed former suspicions that the DPRK had actively pursued the development of nuclear technology for military purposes. Despite the conclusion of the Joint Declaration on the Denuclearization of the Korean Peninsula\textsuperscript{33} between the two Koreas in January 1992, IAEA inspectors reported in May of that year that the DPRK had been erecting weapons-production facilities at the Yongbyon reactor. North Korea demonstrated in 1993 it had also developed nuclear-capable missile technology by firing a Nodong-1 missile into the Sea of Japan in 1993.

With these revelations North Korea stood in public violation of the NPT\textsuperscript{34} and the Korean Denuclearization Treaty. However the question of whether the DPRK regime the held in its possession

\begin{footnotesize}
\begin{itemize}
\item The fissile material of a nuclear bomb takes one of two forms: plutonium or enriched uranium. Plutonium is created in the core of a nuclear reactor through the process of nuclear fission. After extracting plutonium from spent fuel rods of a nuclear reactor, one can reprocess this plutonium into weapons-grade fissile material.
\item Noland, Marcus. \textit{Avoiding the Apocalypse: The Future of the Two Koreas}, Institute for International Economics, Washington DC, 2001, p 146.
\item The 1991 Joint Declaration on the Denuclearization of the Korean Peninsula treaty between North and South Korea called for inspections of nuclear reactor sites by a Joint Nuclear Control Commission to ensure that both states employed nuclear technology for peaceful energy production. Item three (3) of the treaty specifically condones the enrichment of uranium or the reprocessing of plutonium. For the full text of the treaty, see the Carnegie Endowment for International Peace at http://www.ceip.org.
\item Article II of the Nuclear Non-Proliferation Treaty: “Each non-nuclear-weapon State Party to the Treaty undertakes not to receive the transfer … of nuclear weapons or other nuclear explosive devices … not to manufacture or otherwise acquire nuclear weapons or other nuclear explosive devices … not to seek or receive any assistance in the manufacture of nuclear
\end{itemize}
\end{footnotesize}
reactor fuel to reprocess remained unanswered. Then in May 1994, North Korea removed 8,017 spent fuel rods from the five-megawatt reactor at Yongbyon. These rods could yield 25 to 30 kilograms of plutonium, which could be reprocessed to produce five to six nuclear warheads. This public display confirmed that North Korea had all the pieces it needed to build and deliver a nuclear bomb. Unable to accept a nuclear North Korea, the United States entered into bilateral negotiations that resulted in the Agreed Framework. North Korea’s responsibilities under this 1994 accord included the shutdown of operations at three of its reactors, continued membership to the NPT, the admission of IAEA inspectors, and the storage of the spent fuel rods removed in 1993. The U.S. and other nations were to reciprocate this good faith effort by the DPRK with, among other economic aid concessions, the construction of two LWRs to compensate for the termination of operations at the other reactors.

The Post-Agreed Framework World

Even with the permanent shutdown of operations at the Yongbyon facilities and the admission of IAEA inspectors into North Korea, suspicions of a nuclear weapons program were not abated by North Korea’s commitment to the Agreed Framework. Questions remained unanswered about the series of Yongbyon reactor shutdowns in the early 1990s: did the North Koreans extract undisclosed reactor fuel during these shutdowns? To complicate matters further, widespread public suspicion of an underground complex at Kumchongri surfaced in 1998.

In response to the increased scrutiny and pressure, North Korea threatened to end the reactor fuel storage process and on August 31, 1998 it launched a three-stage Taepodong-1 rocket over Japan. This raised fears that the DPRK already held a multi-stage rocket capable of delivering a conventional or nuclear weapons to the continental U.S.

weapons or other nuclear explosive devices.” The full text of the NPT can be accessed at the Monterey Institute Center for Nonproliferation Studies at http://cnsdl.miis.edu.

35 See the Nautilus Institute for Security and Sustainability at http://www.nautilus.org/.
These fears, among others, were confirmed first by U.S. intelligence and then by the Korean regime itself in 2002. On October 4th of that year, Assistant Secretary of State James Kelly confronted Korean elites with evidence of a covert nuclear weapons-development program. The DPRK responded with a public admission that it had been developing a new, highly-enriched uranium (HEU) nuclear weapons program. In the months following its admission of the secret HEU program, North Korea expelled all IAEA inspectors, withdrew from the NPT, restarted the Yongbyon nuclear reactor, and fired two surface-to-vessel, short-range Silkworm missiles into the Yellow Sea.

In addition to admitting to its HEU program, the DPRK’s foreign ministry official Ri Kwang Hyok reported the North Korea possessed an untested, three-stage missile capable of reaching the continental U.S. This intelligence was corroborated by Lowell Jacoby of the U.S. Defense Intelligence agency.

While the HEU put the DPRK in breach of the NPT, the Agreed Framework, and a number of other non-proliferation treaties, the Yongbyon reactor and plutonium reprocessing facilities presented the more immediate threat. If North Korea started to reprocess plutonium at Yongbyon the fissile material could come from one of two potential sources: the 8,017 spent fuel rods canned under the Agreed Framework or an uncertain supply of plutonium removed from the Yongbyon reactor during the 1989-1991 shutdowns. Most experts have asserted that the DPRK at the time of its acknowledgement of the HEU program most likely already had reprocessed some plutonium from the 1989-1991 shutdowns and thus probably possessed one or two nuclear weapons. On February 12, 2003, CIA director George

---

36 The technology for the HEU program was acquired from a Pakistani nuclear laboratory, the A.Q. Khan Research Institute, in exchange for North Korean missiles that can carry nuclear weapons. The Bush Administration imposed a two-year ban on any trade with the Research Institute in April 2003. See Sanger, David E. “U.S. Rebukes Pakistanis for Lab’s Aid to Pyongyang,” The New York Times, April 1, 2003.
37 December 31, 2002.
38 The DPRK Government Statement of the NPT withdrawal was released January 10, 2003.
40 Anti-ship missiles were launched February 24, 2003, on the eve of the inauguration of Roh Moo Hyun as South Korea’s new president, and March 10, 2003. Both failed to hit their targets.
43 See “Peace in Korea,” Foreign Affairs, March/April 2003; Noland (2001); Sigal (1998); The Nautilus Institute.
Tenet reported during a Senate Armed Services Committee hearing in that North Korea probably already has “one or two plutonium-based devices today [built from plutonium other than the 8,000 fuel rods placed in storage under the Agreed Framework].” Only three months later, Li Gun, chief of the DPRK’s negotiating team, claimed it had already processed the 8,000 fuel rods. While U.S. intelligence continued to debate the veracity of this claim, if the DPRK’s statement was true it could have possessed as many as eight nuclear weapons by June 2003.

45 “The chemical process of reprocessing spent fuel into plutonium lets off a distinct signature – a form of krypton – that can be detected by sensors used by American intelligence agencies for decades … so far there has been no evidence of that gas, officials say, or other evidence that reprocessing has begun.” Sanger, David E. and Howard W. French “North Korea Prompts U.S. to Investigate Nuclear Boast” The New York Times, May 1, 2003.
Section III: General Theory of International Relations and Brinkmanship

The focus of my analysis now turns to a discussion of international relations, and the special case of general diplomatic strategy: brinkmanship. This section attempts to situate elements of the US-DPRK game within a broader theoretical framework developed most notably by Daniel Ellsberg and Thomas C. Schelling.

Distinguishing International Relations and Brinkmanship Games

International relations, at their most basic interpretation, are bargaining games. Through diplomacy players negotiate new solutions, bargain new contracts. The basis of any negotiation game is a system of promises for contingent performance or behavior. Coercive devices facilitate the bargaining process by communicating to the structure of one’s incentives and are employed to induce other players to move in such a way that one’s own payoffs are maximized. In international relations, one player may threaten or commit to an action that could result in the destruction of realizable gains.

Brinkmanship is a special case of international diplomacy. It is the manipulation of risk to achieve new ends. In brinkmanship, solutions are negotiated using threats that raise the risk of war to both parties, the risk that the bargaining process will break down. Threats in brinkmanship do not propose the certainty of disaster but the possibility of it, and so players engaged in brinkmanship must assess which moves will be effective in arriving at an advantageous contract given the parameters of the game.

Through the paradigm of brinkmanship the relations between the DPRK and the U.S. come into greater focus. These two states have employed brinkmanship tactics to negotiate in a repeated bargaining game marked by a mixture of both conflict and mutual dependence. We assume when looking at diplomacy between North Korea and the United States that there exists a feasible contract

---

48 Schelling, The Strategy of Conflict, p 87. Schelling refers to these games as non-zero-sum, cooperative games marked by an “imperfect-correlation-of-preferences.”
space and mutual dependence among parties. If the former assumption did not hold, then no solution
could be found because the bargaining space the empty set; if the latter assumption failed the parties
would be better off acting independently.

The Convergence of Expectations on Obvious Contracts

In bargaining with each other, players use coercive devices to alter the game in such a way that a
contract is concluded in their favor. However it should also be noted, before a discussion of coercive
tactics, that some contracts may be more obvious choices than others, regardless of the coercive
stratagems employed. Schelling discusses at length the coordination of players’ expectations on certain
outcomes, especially in tacit negotiation situations, and proposes that outcomes built on simplicity,
precedent, or uniqueness are the solutions to which players will gravitate during the bargaining
process.49 This would suggest that the ultimate contract created between the U.S. and the DPRK may
center on the complete dismantling of any and all nuclear programs within North Korea. Why? Because
any solution that allows for some weapons or programs will be inherently unstable and most likely
challenged in the future. “Zero weapons” is a more obvious stopping point. Many scholars have argued
that this was the greatest benefit of the Agreed Framework: it only increased the conspicuousness of “no
nuclear weapons” as the ultimate answer to the DPRK-U.S. conflict.50 It may prove difficult for the

49 “The final outcome must be a point from which neither expects the other to retreat…If we then ask what it is that can bring
their expectations into convergence and bring the negotiation to a close, we might propose that it is the intrinsic magnetism of
particular outcomes, especially those that enjoy prominence uniqueness, simplicity, precedent, or some rationale that that
makes them qualitatively differentiable from the continuum of possible alternatives.” Schelling. The Strategy of Conflict,
p.70.

50 A brief commentary on the events following the Agreed Framework. The confrontation during the Clinton administration
which ended with the conclusion of the 1994 Agreed Framework set the tone for the remainder of the decade. Critics of
Clinton’s policy throughout the 1990s have dubbed the relations policies of appeasement or “checkbook diplomacy.”
Following the 1994 Agreement North Korea succeeded more than once in raising confrontations with the U.S. to win
concessions of economic aid and sanction reductions. For example, the DPRK’s August 1998 public threats to suspend the
fuel rod storage process and to increase its missile exports, underscored by the launch of a three-stage rocket over Japan into
the Pacific Ocean, won it an additional shipment of grain, valued at $120 million. Later in November 1998, DPRK threats to
abandon the Agreed Framework yielded a larger shipment of food from the U.S. This was followed in December 1998 by an
agreement forged under missile talks whereby North Korea agreed to suspend missile and missile exports for a period of
three years in return for $300 million from the U.S. And again in 1999 North Korea threatened a missile test launch of the
Taepodong-2 rocket, which eventually earned a reduction in sanctions by the United States.

However to suggest the Clinton administration was a passive, reactionary body is an oversimplification that ignores
the bargaining tools employed by the Clintonites to prevent war and secure the freezing of North Korea’s plutonium
DPRK to alter the game in such a way that this focal point becomes less attractive to involved parties. Another solution to which parties may gravitate is the polar opposite of the first: a normalization of relations with the DPRK as a nuclear state. In May 2003, the DPRK claimed during talks with the U.S. and China that it had already reprocessed the fuel required to produce nuclear weapons. The U.S. and its allies now face far greater costs to verifiably disarm an already nuclear North Korea. It may be that the solution derives from the simplest option: acquiesce to the status quo.

Defining Commitments and Threats

Coercive devices – classified as either a commitment or a threat – form the communication structure of brinkmanship games and are tools to limit the choices of the other player. Players employ these devices to change the payoff structures in such a way that some moves become prohibitively costly and thus, unattractive decisions. Both stratagems – the commitment and the threat – facilitate the bargaining process by communicating the intent, reservation prices, political factors, economic limitations, psychological thresholds, and good will of each player.

With each device the object is to “relinquish the initiative” to the other player: to force him to choose, and more to the point, to force him to choose to one’s own advantage. The difference between the two tools then lies in their relation to the moves of the other player. In game-theoretic terminology, the commitment translates into a “first move” tactic. In converts an indeterminate negotiating game into

---

facilities. The political environment of the region was vastly different to today’s landscape: following the Agreed Framework “relations between China and South Korea were [under-developed]… South Korea is now vastly more important to China, probably China’s third largest economic partner after the United States and Japan.” Unlike today, the U.S. was less likely to win support from other powers in the region. By ratcheting up the pressure on North Korea through threats of sanctions and war preparations the U.S. negotiated a solutions that circumvented war and facilitated a shutdown of nuclear reactors which otherwise would have produced enough plutonium to make up to 30 nuclear warheads.

The downside risks of war: US military commander in South Korea, General Gary Luck, estimated [war] would have killed a million people, including 80,000 to 100,000 Americans. (Foreign Affairs, p 23). South Korea’s GDP was 323 trillion won; today its GDP is roughly 544 trillion won. Had we gone to war this transformation, in combination with the 1997 Asian financial crisis, would have been improbable. China also experience rapid growth. Also, the US could not have counted on Russia or China to support its position towards North Korea. Also, had the Agreed Framework not been signed in 1994 and if we had not gone to war, the North’s plutonium based program would by today have produced enough plutonium for up to 30 nuclear weapons.

51 “Skillful diplomacy in the absence of uncertainty consists in arranging things so that it is one’s opponent who is embarrassed by having the last clear chance to avert disaster by turning aside or abstaining from what he wanted to do.” Schelling, Arms and Influence, p. 101.
a two-move game where one player makes a commitment to a certain action by altering his existing payoffs and the other player makes a decision in response to this changed incentive structure. As Schelling explains, “The commitment is a strategic move, a move that induces the other player to choose in one’s favor. It constrains the other player’s choices by affecting his expectations.”\textsuperscript{52} The threat, by contrast, is a promise to react to an undesired move. The threat proposes costs to both players and, if effective, will change the payoff structure and expectations of the target so as to induce a move in the threatener’s favor. Thus the threat is a “second move” tactic that converts an indeterminate negotiating game into a two-move game; one player threatens to retaliate only if the other player fails to choose a desired action.

Both coercive devices carry the risk of failure. If they fail to influence the other player’s behavior, one or more parties – including the player who made the threat or commitment – may suffer the costs of who makes the commitment or threat will be penalized.

\textit{Functions of Coercive Devices in DPRK-U.S. Relations}

\textit{Deterrence.} One use of coercion is to prevent an opponent from moving in one’s disfavor. In the case of relations between the U.S. and the DPRK, one can find numerous devices aimed to deter conflict on the Korean peninsula. In particular, it has been argued that the 37,000 troops stationed in South Korea not to fight for Seoul but to serve as a deterrent threat to the DPRK. The U.S. presence does not significantly contribute to the defense of South Korea in the event of a short warning attack. What the 37,000 troops represent is not a retaliatory force but the “honor and prestige” of the United States. Were the DPRK to launch a short-warning attack against the South, the U.S. would have no choice but to defend the honor and memory of the U.S. troops killed in the attack. Thus the troops garrisoned in South Korea serve as a deterrent threat against the North by increasing the cost to the U.S. of non-

\textsuperscript{52} Schelling, \textit{Strategy and Conflict}. p 122.
response. The threat changes payoff structure of the U.S. and, as a corollary, constrains the choices of North Korea.

This also illustrates how the physical presence of troops acts as a more effective deterrent threat than words alone. During the Bush Administration the U.S. attempted to deter North Korea from the reprocessing the spent fuel rods frozen under the Agreed Framework by casting these moves as crossing “red line.” However verbal threats are often not sufficient to deter action because verbal commitments or threats are easily reversed, as evidenced by a DPRK screening device in May 2003. The DPRK seemed to first “test the water” with a “mistranslated” statement that claimed it had already reprocessed the spent fuel rods. The U.S. response, despite threats that reprocessing of fuel would bring grave consequences to the DPRK, was one of confusion and intelligence recalculation. While the DPRK retracted and corrected the erred translation, two weeks later the North Korean negotiating team renewed the claim that it had already reprocessed the 8,017 fuel rods into nuclear weapons. Again, the U.S. failed to respond with the automatic retaliation that would have followed if it had changed its incentive structure. Its verbal threats against reprocessing failed to deter the DPRK from building nuclear weaponry. A deterrent threat will fail if it is not credible, or in other words, if one’s opponent does not believe that alternatives have been surrendered:

Talk is not a substitute for moves. Moves can in some way alter the game, by incurring manifest, costs, risks, or a reduced range of subsequent choice; they have an information content. Talk can be cheap when moves are not.54

Compellence. While some coercive devices attempt to deter certain actions, other coercion tactics seek to compel the other player to move in one’s own favor. This may be to either preserve or alter the status quo. Thus the character of a compellent threat differs from the deterrent threat in that it “often requires the punishment be administered until the other acts rather than if he acts.”55 In diplomacy

53 It remains uncertain whether the DPRK’s claim that it had reprocessed the fuel rods is valid. Nonetheless, the point stands. Had the U.S. threat been credible, had North Korea believed the U.S. would automatically respond with military force, it would never have made the claim because a military conflict would have automatically resulted in retaliation for transgressing the “red line.”
between the U.S. and the DPRK the specter of economic sanctions or military build-up serve as compellent threats against the DPRK. Both seek to persuade the regime to abandon its nuclear ambitions and development.

Just as the U.S. employs compellent threats in the form of economic sanctions to induce the DPRK to move in its favor, the DPRK employs nuclear weapons serve as coercive devices to win economic concessions. Although it has also been argued that the DPRK’s nuclear program serves as a threat to deter attack from other nations, the North Korean weapons program has also evolved into a device of compellence, a strategic asset to compel other countries to grant economic concessions and reduced sanctions. Nuclear brinkmanship the regime compelled other nations to reduce sanctions in 1994; test launches of nuclear-capable missiles into the Sea of Japan won the DPRK further economic aid in 1998; and saber-rattling by the DPRK earned it millions from the South Korean economy in 2000.

Communicating Threats and Commitments

Communication and its Destruction. In order for coercion to be effective, a player must be able to communicate his threats and commitments to the opponent whose behavior he wishes to influence because it is through the communication of threats and commitments that players disseminate information about their expectations and incentives. If an opponent does not receive the threat or understand the commitment, he will not update his beliefs regarding the payoffs of each player and thus, the device will fail. Communication should not be mistaken for words alone: by withdrawing from the NPT, expelling IAEA inspectors, and conducting missile launches, the DPRK attempted to communicate to the second Bush Administration and regional actors information about its toughness and resolve to become a sovereign, nuclear state. If these moves had credibly signaled to the U.S. that the DPRK had no incentive to surrender and every intention to protect its sovereignty at any cost because this represented a better payoff than submitting to the U.S. will, it may have prevented the U.S. from pressuring the DPRK further for fear of a potential suicidal retaliation by the regime.
Because the necessity of communication is common knowledge, it can be equally advantageous for a player to work to destroy or confuse channels of communication. A party to which a threat or commitment cannot be communicated is no longer vulnerable to coercion. This is also inextricably linked to the process of identification: if a player cannot be identified as one who would respond to a threat, then they are also immune to coercion. As noted by Schelling, “If the person to be threatened can arrange before the threat is made to share the risk with others … he may become so visibly unsusceptible to the threat as to dissuade the threatener.”

In the game leading up to the Agreed Framework, the Clinton administration allowed the U.S. to be identified as a receiver of coercive devices by action as Seoul’s protector and attempting to engage the DPRK in negotiations. During the second Bush administration, diplomats have tried to undo the identification of the U.S. as the sole-receiver of the DPRK’s threats. The Bush administration rejected bilateral negotiations and cast the disarming of North Korea as a multilateral, regional issue.

Making Threats Credible

There is no limit to the number of threats a player can issue against an opponent; but any threats that fail to appear credible will fail to alter an opponent’s behavior. Credible coercion mandates that one’s opponent must believe a player has altered the incentive structure of the game. One must change an opponent’s expectations. More specifically, one’s opponent, after updating his beliefs, must estimate the probability that a threat will be carried out to be greater than some minimum “critical risk” threshold, or the point at which an action becomes unprofitable given the risk that a punishment or mutual disaster may occur. As explained by Ellsberg, the problem facing the coercer is to convince the other players that he is “too likely” to carry out the threat or commitment:

My problem as a blackmailer is to ensure – by actions that either change your payoffs, hence your critical risk, or that increase your expectation of punishment – that your estimate of the actual risk is greater than the critical risk.

---

Thus the critical risk is the “minimum credibility” any threat must carry to change the behavior or actions of the other players. Estimations of actual risk relative to the critical risk threshold are built on an opponent’s expectations of the coercer’s payoffs, signaled to him by threats and commitments.

How can a player, like the DPRK, ensure that he is too likely to carry out his threat, too willing to launch a preemptive strike against the U.S. or its allies? Two methods are available to the coercer: the building and staking of one’s reputation, and the elimination of attractive alternatives.

Building and staking reputations. “A potent means of commitment… is the pledge of one’s reputation.”

Reputation building is important in any bargaining situation. A player that may be perceived as a “tough” or “hard” negotiator is perceived to be less likely to concede, less likely to compromise, more ready to meet disagreement or shared disaster. Because hard players are less willing to compromise, the likelihood that they will be challenged is often diminished. Therefore building a reputation for toughness can enhance one’s credibility when one makes a threat or commitment to a particular action. The loss of a reputation, assuming he does not follow through with his threat, can represent a huge cost to the coercer. If one can demonstrate to his opponent that if he did not uphold his threat following a move in his disfavor that the costs suffered from loss of face would be unbearable, this can also raise the credibility of the threat. When demanding North Korea disarm, the U.S. must consider the costs that would be incurred by North Korea were it to acquiesce and their derivative incentives: is the cost to the regime of losing face with the international community so grave that it would prefer to launch a preemptive attack than suffer damage to its reputation?

Further, reputations are interdependent across time and across bargaining games. For example, the U.S. has been less willing to allow damage to its reputation for toughness to occur in its relations with North Korea because it has been simultaneously engaged in other bargaining games with other states and sub-national actors, such as Iraq and al-Qaeda. Similarly, if players appear weak in current

---

games, this may damage their reputations in future bargaining games. As explained further in *The Strategy of Conflict*:

‘If I conceded to you here, you would revise your estimate of me in our other negotiations; to protect my reputation with you I must stand firm.’ The second party is simultaneously the third party to whom one’s bargaining reputation can be pledged.\(^59\)

Of course reputation building can give rise to bluffing in situations where a player’s knowledge of his opponent’s willingness or resources is incomplete. Often in games of brinkmanship, players are unable to assess how willing or capable the other player is to carry out a threat. In game-theoretic terms, one would say that players are unable to distinguish between different types of players: a state could be engaged with a despot who will never surrender costs be damned, or it could be engaged with a government that would rather compromise but is seeking bargaining leverage in the immediate situation.\(^60\) As players move, they presumably learn more about what type of opponent they are facing; but when information across the players is incomplete, one type of player can pretend to be another type of player. For example weak player can bluff. He can act as a tough player would, and gain bargaining leverage against his opponent.\(^61\)

These bluffing or reputation building tactics can be seen in relations between the US-DPRK. Both players have employed threats early and often in the game to establish reputation. From the U.S. perspective, in this game of incomplete information, it cannot discern what type of North Korean regime it is asking to disarm. Because the U.S. lacks complete information about the incentives of the DPRK, it can be rational for a weak type North Korea to choose an aggressive strategy that will convince the U.S. it is dealing with an apparently irrational regime that would rather go to war than surrender to the demands of the impudent U.S. Some have argued that the belligerence of the DPRK in 1994 was a bluff. If so, the Agreed Framework illustrates the value of building a reputation for irrationality:


\(^{60}\) In game theory the disarticulation of types of players into one or more discrete categories is useful in solving games of incomplete information; this method is used in sections IV and V.

\(^{61}\) It should be noted that it is not always to a player’s advantage to pretend to be a strong player; in certain games weakness can offer its own advantages.
With the presumption of irrationality on its side … North Korean threats of suicidal war have elicited conciliatory responses from South Korea and the United States. These threats increased Pyongyang’s bargaining power in subsequent negotiations with the United States and South Korea.62

The Bush Administration, like the preceding administration, cannot know with certainty what type of North Korea it is challenging. As long as this uncertainty exists, it may be rational for the DPRK player – it its true type is weak or compromising – to establish a reputation for aggression; the sum benefits of “bluffing” may exceed the immediate costs and risks.

Elimination of Alternatives. A coercive device can also made more credible by the elimination of alternatives. A common analogy used to illustrate this concept is the burning of a bridge in battle. With no route to escape, one’s army is only given an incentive to fight harder. The enemy, understanding this, may be compelled to retreat. Any room for casuistry or subtle exits weakens the credibility of a threat. Any method by which one can make the option of reversing one’s threat so costly as to be eliminated from consideration altogether should be employed to ensure complete credibility in the eyes of the opposing player.

The failure of coercion

Ultimately, the goal of a threat is not to follow through with the punishment. If the threat is exacted, coercion has failed. A coercive device can fail for a number of reasons. As we have explained, if either party miscalculates the incentives or payoff structures of the other players, the threat will not appear credible and may consequently fail to move the other player’s expectations.

A threat may also fail in brinksmanship because neither state is in full control of all people and processes involved in the conflict: “faulty information, faulty communication, misunderstandings, misuse of authority, panic, or human or mechanical failure”63 could precipitate war where war is not intended. For example, in the DPRK-U.S. game played under the Bush administration, consider the

“mistranslation” of reprocessed plutonium. What if this had been a true error and not a screening device? What if the U.S. had been fully committed to punishing the DPRK if it transgressed the established “red line”? The DPRK may have stumbled into war when neither party desired this outcome. Because coercion carries these risks, it is not costless to make threats.
Section IV: Conflict during the Clinton Administration

To analyze relations between the DPRK and the U.S. identify two games: the confrontation under the Clinton administration and the crisis of the Bush administration. We treat these separate conflicts as the same game repeated over time. This section is a formal treatment of brinkmanship under the Clinton presidency, which ended with the conclusion of the Agreed Framework. After defining the players, parameters, and payoffs of a two-stage game, I analyze the process by which players develop strategies to optimize their payoffs in every outcome and through this identify stratagems that will never be optimal for a player to pursue, given the conditions of the game. I will then expand the game to include multiple rounds of negotiation. In this discussion I intend to give a more dynamic perspective of the bargaining process between the two states that draws further on the theory developed in section III.

Chronology of Moves in the Indeterminate, Determinate Bargaining Games

As circumstantial evidence of weapons facilities at Yongbyon mounted at the end of the 1980s, the U.S. and international community increasingly called for IAEA inspections of North Korean nuclear sites. While North Korea initially rejected these inspections on the grounds that U.S. tactical weapons were positioned within South Korea, eventually IAEA inspectors entered the DPRK in 1992. They confirmed that the North Koreans had been building weapons facilities and contended that they may have already extracted spent fuel rods from the Yongbyon reactor during previous reactor shutdowns. The U.N. cited North Korea in violation of international nonproliferation accords and considered punitive actions.

By challenging North Korea’s weapons development, the U.S. initiated phase I of the game with North Korea. In response to confrontation over its weapons program, North Korea, in

---

64 My analysis builds upon the work of Dixit and Skeath’s “Oldstar v. Nova” game and discussion of the Cuban Missile Crisis; Gates and Hume’s “Sanctions” game; Krep’s “Beer-Quiche” game; and Selten’s “Chain Store” game (1978).
March 1993, threatened to withdraw from the NPT. The U.S. countered this unprecedented threat by announcing the possibility that it might bring a sanctions resolution to the U.N. Security Council. North Korea, rather than acquiesce to U.S. demands to negotiate disarmament, test fired a potentially nuclear-capable missile, the Nodong-1, into the Sea of Japan. Alarm sounded throughout Washington, Tokyo, and Seoul. The U.S., fearing the future increase in plutonium production that could follow with the imminent completion of two additional graphite reactors, again threatened a sanctions resolution if the DPRK refused to come to the negotiation table. North Korea responded that “Sanctions are a declaration of war” and continued to refuse talks. These moves are then echoed and elaborated. Following a condemnation of the DPRK approved by the UN General Assembly (vote count 140-1, with China abstaining and the DPRK dissenting) a North Korean negotiator remarked that it could turn Seoul into “sea of fire.”

Then in May 1994, North Korea made its most decisive move: it removed spent fuel rods from the Yongbyon reactor. As noted by Marcus Noland: “The removal of the spent fueling rods in defiance of the international non-proliferation regime was instrumental in the US policy shift from preventative to coercive diplomacy.” The U.S. and IAEA immediately requested Security Council Action to which North Korea responded again, “Sanctions mean war, and there is no mercy in war.”

In early June the U.S. threatened war itself. President Clinton, working in conjunction with the Pentagon, set June 16 as the date to begin a war-buildup in the Asian theater. North Korea retaliates with a verbal warning that any military buildup in the region will trigger a pre-emptive strike against the U.S. by the DPRK. North Korea does not follow through on its threat.

On June 16, as Clinton consulted with Pentagon officials over plans for war with the DPRK, former president Carter and Kim Il-Sung struck a preliminary agreement that laid the foundation for the Agreed Framework. The preliminary agreement, as well as the Agreed Framework, called for a freeze of nuclear activities in return for promises of economic concessions, energy aid, and normalization of relations.
Phase I ends with the conclusion of the Agreed Framework.

The risk of mutual disaster was raised by each player’s subsequent move, the risk that the bargaining process would fail to find a feasible agreement and result in war. War, which will yield the worst payoff for each player, is the alternative to the negotiated agreement. The threats and counter-threats of North Korea and the United States can be broadly defined as signaling tactics in an indeterminate bargaining situation. The decision of the North Koreans to remove the fuel rods transformed the indeterminate dialogue into a determinate cooperative bargaining game.

Complications to the game: Incomplete Information

The interactions between the U.S. and North Korea are marked by incompleteness of information. When the U.S. applies pressure to North Korea to disarm through threats of military action or economic sanctions, it is unaware with what type of regime it is negotiating: is the DPRK a “hard” player in that it would rather initiate a war than acquiesce to the U.S. demands, or is the DPRK a “soft” regime, a player that would rather abandon its nuclear ambitions than risk war with the U.S.? The U.S. is uncertain of the DPRK’s payoff structure and therefore cannot ascertain with full knowledge the effect of the DPRK’s threats.

Game theory represents this incompleteness of information by disarticulating one player, in this case North Korea, into different types of players: one “hard” type player who is willing and able to engage in war to become a nuclear state, and one “soft” type player that is either not willing or not able to wage war to become a nuclear state. If North Korea is the “soft” type of player, meaning that it is either unwilling or incapable (or both) to follow through with a threat of preemptive nuclear action, it prefers to abandon its nuclear ambition to challenging the U.S. If North Korea is the “hard” type of player, meaning that it is either willing, capable, or willing and capable, to launch a preemptive nuclear strike against the U.S., it would prefer to challenge the United States to abandoning its nuclear weapons program.
Also added to the extensive form game is a move by Nature, who chooses at the outset of the game chooses which type of DPRK is negotiating with the U.S. Once Nature has chosen, North Korea learns its type.

The U.S. does not know which type of North Korea it faces in the bargaining game. Before the U.S. moves in response to North Korea, it must assign a probability distribution to assess the likelihood it is interacting with any one type of DPRK regime. The probability that Nature has chosen a “hard” DPRK is represented by $w$, a parameter that measures the probability that the DPRK is willing and capable to initiate nuclear war and would prefer this outcome to acquiescing to the U.S. As the game progresses, this parameter is updated using Bayes Theorem following each move by the DPRK; presumably with each move the DPRK reveals more about which type of Korea it is: hard or soft. Conversely, the probability that the U.S. faces a “soft” Korea is equal to $(1-w)$.

Here the model closes. The U.S. generates a probability distribution to measure the likelihood that North Korea is a hard regime and the estimation of $w$ by the U.S. is common knowledge among all players.

**Analysis of the Game**

As noted by James Friedman, $^{65}$ “it is possible to specify an equilibrium concept for some incomplete information games by making assumptions about the way that incompleteness enters the game.” The following treatment of the 1994 brinkmanship game builds on elements of Friedman’s discussion of non-cooperative games under incomplete information.

There is a set of $N$ players $= \{1, 2\}$, where 1 represents the DPRK and 2 represents the U.S. Each player knows his own payoff function for every outcome, but is uncertain as to the payoff functions of the other player. For player 1, there is a finite number of types selected from the set $V$ possible types such that player 1’s true type, $b_i \in V = \{1, 2\}$ for $i = 1$. Type 1 represents the hard-type player 1 (DPRK)

and type 2 represents the soft-type player 1. Therefore in this incomplete game the number of players that may play the bargaining game has expanded from two players, player 1 (DPRK) and player 2 (United States), to three potential players: soft-type DPRK, hard-type DPRK, and the U.S.

Each player selects an action from the available action set $A_i$ for $i = \{1, 2\}$ at stage one of the game. All available actions are common knowledge among players, thus the U.S. has complete information regarding the hard and soft DPRK’s feasible moves at each stage of the game and conversely, the DPRK knows all moves available to the United States at each stage of the game. Both types of DPRK draw from the same set of feasible actions. After nature chooses which type of player 1 will be active in the game, player 2 moves. The feasible action sets for each player in stage one are the following:

$$A_{1b, \text{stage } 1} = \{\text{Challenge U.S. with display, Challenge U.S. without display, Acquiesce}\}$$
$$A_{2, \text{stage } 1} = \{\text{Challenge, Acquiesce}\}$$

Whenever a player chooses to acquiesce, the game is over and it is assumed a contract is formed by which either 1) the DPRK is disarmed or 2) the U.S. normalizes relations with the DPRK as a nuclear power. When the DPRK challenges with display this is intended to mean that it rejects the U.S. proposal to disarm and makes some show of force to illustrate its resolve against the U.S.; examples of display by the DPRK include withdrawing from a peace treaty, conducting missile tests, or extracting or reprocessing spent fuel rods. This contrasts with the alternative of challenging without display. A challenge without display would constitute a verbal threat alone, such as threatening to “set Seoul afire.”

If neither the U.S. nor the DPRK acquiesces in stage one, the game progresses to stage two, again with player 2 moving before player 1. The feasible action sets for each player in stage two are the following:

$$A_{1b, \text{stage } 2} = \{\text{Initiate War, Acquiesce}\}$$
$$A_{2, \text{stage } 2} = \{\text{Military Response, Sanctions Response, Acquiesce}\}$$

---

66 This example can readily be extended to more than one player. The number of players $N = \{1, \ldots, n\}$, the types of players $V = \{1, \ldots, v\}$ and the action sets of each player $A_i$ for $i = \{1, \ldots, a\}$. It could also be extended to incorporate different action sets that would change depending on which type $v$ of player $i$ was chosen by nature to play: $A_{iv}$.

67 Later in this section we will complicate the feasible actions to include an additional challenge option to the DPRK that will allow for more than two stages of negotiation, which will progress until a party acquiesces or war is initiated.
Whenever war is initiated or a player acquiesces, the game is over. If a player acquiesces, it is consistently assumed a contract is formed by which allows for a nuclear DPRK or provides for the complete dismantling of the DPRK nuclear weapons program, as in stage one. If war is initiated no contract is formed but it is assumed the DPRK will be disarmed due to a lack of military and economic resources to successfully win a battle against the U.S. However, a breakdown in negotiations and initiation of nuclear conflict, even if it brings a disarmament of North Korea represents the worst outcome for both players – even the U.S. – as we will soon illustrate.

Case 1: Limiting the Game to a Two-Stage Negotiation Process

As previously noted, nature chooses a vector of player types $\beta \in V^n$ at the start of the determinate bargaining game. This vector indicates which two of the three players will play. Because the game is incomplete, player 2 does not know the type of its opponent, but each player has full knowledge of its own type once nature has moved. Vector $\beta$ will be called a state of the world. Here we have two possible states:

$$\beta = [11, 2:] , [12, 2:]$$

The first state of the world $\beta = [11, 2:]$ describes a hard-type DPRK playing the U.S. and occurs with probability $w$. The second state of the world $\beta = [12, 2:]$ describes the soft-type DPRK and occurs with probability $(1-w)$.

Figure one illustrates the game, taking into consideration the two possible states of the world. The model assumes that both the DPRK and the U.S. are rational, payoff-maximizing players and also assumes that the payoffs earned by each player at every outcome are common knowledge. Under these assumptions, each player will look ahead in the game to develop strategies that will optimize their payoff at any particular node, in any possible contingency. Before we discuss the payoffs and risks associated with every outcome for each player, we want to consider if there exist any actions that, while feasible, will never be optimal for a player to choose.
In figure one below the dotted arcs in show the information sets faced by the U.S.; in other words, the U.S. cannot distinguish between the two nodes as it does not know which type of North Korea Nature has chosen to play. However it can observe whether the DPRK displays when it challenges or not.

The decision tree of figure one above shows the payoffs won by each player at every outcome. Whenever the U.S. secures an outcome that requires the DPRK to disarm, it earns a payoff of $b$; conversely, whenever the DPRK secures a contract that allows for the DPRK to keep its nuclear weapons, it earns a payoff of $b$ for itself, regardless of its type. However from the payoffs to the DPRK bolded in the figure, one can see that different types of DPRK players incur different costs for displaying. For example, a hard-type DPRK can choose the action “challenge and display” without
incurring a cost to display because it is both willing and capable to fulfill its threat, whereas a soft-type DPRK incurs greater costs for displaying when in challenges, even when the U.S. acquiesces. The soft-type DPRK must assume these additional costs because it risks more by imitating the hard-type DPRK: it is either not willing or not able to uphold a threat of war, unlike a type 1 DPRK.

Comparing these payoffs provides some insight into which actions will never be optimal for any type of DPRK. To explain further, let us take each information set in turn:

1. \{NK hard/NK displays, NK soft/NK displays\}. Here the U.S. has observed some form of display by North Korea. It is common knowledge among players that a soft-type DPRK will incur a cost for imitating the strong type North Korea and that a hard-type DPRK displays without incurring any cost.

2. \{NK hard/NK no display, NK soft/ NK no display\}. Here the DPRK rejects the U.S. proposal to disarm but does not make any display to show its commitment to become a nuclear state. Although the U.S. does not know which type of DPRK has been chosen by Nature, the payoffs at each outcome are, as previously stated, common knowledge. Therefore the U.S. knows that a hard-type North Korea does not suffer any additional costs to display. Because display will enhance the credibility of its threat, the hard-type DPRK will do so automatically. The hard-type DPRK will never choose to challenge without display. Thus the U.S. can assume that anytime the DPRK chooses an action “Challenge without display” that it is facing a soft-type North Korea. The soft-type DPRK, knowing that this move will reveal its type to the U.S. automatically, will instead choose to acquiesce.

Hence an examination of the second information set indicates it will never be optimal for any type of DPRK to challenge without display and allows for a simplification of the model.

Figure two shows the revised model and the payoffs to each player at each outcome.
**DPRK payoffs.** We assume that whenever an outcome allows for North Korea to keep its nuclear weapons program, it earns some positive number $b$ as a payoff. The only question is, what costs must the DPRK assume to achieve an outcome with it as a nuclear power? North Korea’s best outcome occurs when the U.S. does not challenge it and the prevailing status quo is the condition where North Korea is a nuclear power. This generates a payoff to the DPRK of $b$, whether the DPRK is the hard or soft-type regime. The next “least-cost” avenue for the DPRK to become a nuclear state is when the U.S. challenges it but later acquiesces to a North Korean challenge. If the DPRK is the hard-type regime, it can challenge and display without incurring any additional costs if the U.S. acquiesces. Therefore if the U.S. acquiesces, the hard-type DPRK earns a payoff of $b$ for counter-challenging. However if the DPRK is the soft-type regime, when it displays and imitates the type 1 DPRK, it also incurs a cost equal to $c$.
Thus, if the weak DRPK challenges with display and the U.S. acquiesces, it will earn a payoff equal to $b-c$, which is still greater than zero, but less than $b$. If North Korea chooses not to challenge and simply acquiesces, it receives a payoff of zero.

If the U.S. does not acquiesce to a DPRK counter-challenge, it is assumed that the DPRK cannot become a nuclear state and thus cannot earn $b$. This is based on the assumption that the U.S., with greater technological and economic resources will ultimately force an end to the regime, with a verifiable dismantling of the DPRK’s weapons program or both. We also assume, as we have stated before, that a soft-type DPRK will not respond to either a military or sanctions threat with a preemptive strike against Seoul or the U.S. because it is either not willing, not capable, or both, to initiate conflict with the U.S. So if the U.S. threatens a military build-up or sanctions resolution, the soft-type DPRK will acquiesce with a probability equal to 1.00 and suffers some cost $-2c$: the cost of a failed bluff. The hard-type DPRK by contrast, is willing and capable to launch a short-warning attack against either the U.S. or Seoul and does so with a probability $q$. Whenever the hard-type DPRK responds to a military or sanctions threat by the U.S. with the initiation of conflict, it earns some negative payoff $d$, where $d > c$. The value of $d$ represents an outcome of complete disagreement, an irreparable breakdown in negotiations. It is the worst outcome for both players. While the DPRK may gain some satisfaction from standing up to its rival the United States, it will ultimately result in the demise of the regime in either case and so represents the worst payoff to the DPRK. The expected payoff from the “war” response is $dq$. If the hard-type DPRK acquiesces to either a military or sanctions threat by the U.S. it suffers a negative payoff of $c$: again, the cost of a failed bluff.

_U.S. payoffs._ Just as we assume that North Korea will earn some positive payoff $b$ whenever the final contract allows for it to be a nuclear state, we assume that whenever the final solution dictates a dismantling of the DPRK’s weapons programs the U.S. earns some positive payoff, which we also represent by $b$. Thus from the U.S. perspective, the best outcome occurs when it initially challenges the DPRK and the regime subsequently acquiesces by abandoning its nuclear ambitions. This earns the U.S.
a payoff of \( b \). If the DPRK does not immediately acquiesce, but acquiesces after additional threats by the U.S. in the form of either a military build-up or a sanctions resolution, the benefits to the U.S. will be diminished in accordance with the costs it must incur to eventually bring the DPRK to disarm. We assume that the cost of a sanctions resolution is less than that of a military threat. This is because a sanctions resolution by the U.N. Security Council is enforced multilaterally, whereas the cost of a military threat will most likely be born out by the U.S. alone. Further, a military threat risks the loss of U.S. lives and reconstruction of the country after a potential regime-change. Thus the payoff to the U.S. of securing a non-nuclear North Korea by way of a sanctions resolution threat is some positive number equal to \( b-c \), whereas the payoff to the U.S. of securing a non-nuclear North Korea by way of a military action threat is some positive number equal to \( b-2c \). Whenever the U.S. fails to negotiate a contract that disarms the DPRK and war is initiated between the two states, it earns a negative payoff of \( d \), where \( d > 2c \). If at any point it acquiesces, it faces future risks in the form of a possible nuclear strike by the DPRK, the exportation of weapons to non-state actors or rogue states, a nuclear arms race in the region. The payoff to acquiescing at any point is represented by \(-c\).

**U.S. Decision-Making.** Both the DPRK and the U.S., as rational, profit-maximizing players, use the available information to look ahead and solve recursively to develop a strategy that dictates the optimal decision in any contingency. The U.S., before it challenges the DPRK, must determine how it will respond if North Korea counter-challenges the U.S., considering that if its opponent is the hard-type DPRK it will initiate war with probability \( q \). If the DPRK responds with a counter-challenge the U.S., the U.S. has three available actions: 1) a military threat, 2) a sanctions response, and 3) acquiescing to the DPRK. Using all available information it examines its the expected payoff of each action and then calculates the expected utility (EU) of any one action. We will take each action in turn. Before we do, we will make a substitution. The worst payoff resulting from war can be any number greater than \( 2c \), so we will substitute an arbitrary \( 3c \) for \( d \) to show more clearly the inter-relationships among the U.S. decisions.
and payoffs. Let us first look at the expected utility (EU) of acquiescing to the DPRK. The U.S. earns the same payoff with certainty, regardless of the type of the DPRK it faces, so this provides a benchmark to make other more risky choices:

\[
EU (aquiesce) = (1-w) (-c) + (w) (-c)
\]

\[
EU (aquiesce) = -c
\]

Now consider the expected utility of a military threat:

\[
EU (military threat) = (1-w) (b-2c) + (w) (q) (-3c) + (w) (1-q) (b-2c)
\]

\[
EU (military threat) = b - 2c - cwq - bwq
\]

When the EU (military threat) is greater than \(-c\), the payoff it earns with certainty if it chooses to acquiesce, it will consider a military threat as a tool to coerce the DPRK to disarm. If it is less than the payoff that it earns by acquiescing, and it cannot earn a higher payoff by any other action, then it will acquiesce. The precise condition for the EU (military threat) to be greater than \(-c\) is satisfied when \(w < (b - c)/ [(q) (c + b)]\). What parameters affect this threshold? Taking the partial derivative of \(w\) with respect to \(c\), one can see that as \(c\) increases, or as the cost to the U.S. of implementing its threat increases, the U.S. requires the probability that the DPRK is the hard-type regime to be smaller. Similarly, the decision to use the threat of military action is also influenced by the probability \(q\) that a hard-type DPRK will initiate war. As \(q\) decreases, holding all other variables constant, the critical risk threshold for \(w\) increases. The U.S. only willing to make a military threat against the DPRK when its estimation of \(w\) is below this risk threshold. Compare these results with the expected utility of posing a sanctions response:

\[
EU (sanctions threat) = (1-w) (b-c) + (w) (q) (-3c) + (w) (1-q) (b-c)
\]

\[
EU (sanctions threat) = b - c - 2cwq - bwq
\]

When the EU (sanctions threat) is greater than \(-c\), it will also consider this as a tool to coerce the DPRK to disarm. If it is less than the payoff that it earns by acquiescing, then it will acquiesce. The specific condition for the EU (military threat) > \(-c\) is satisfied when \(w < (b)/ [(q) (2c + b)]\). Like before, as the

---

68 For both the DPRK and the U.S. the cost coefficients have been represented by numbers like one or two to clarify the discussion. We could just as easily have substituted some non-negative variable such as “x” in place of “2” which we could allow to vary with the threats issued. We will do this in case two when we discuss a game with multiple negotiation stages.
cost of succeeding in this procedure increases, the threshold decreases, and as the probability of the hard-type DPRK initiating war increases, the threshold falls.

These two thresholds are the answer to Ellsberg’s question of “how likely” must a threat be to deter his opponent: these conditions constitute the critical risk thresholds of the U.S. It is the aim of the DPRK in negotiations to convince the U.S. through the communication of a credible threat when it “challenges” the U.S. that it is “too likely” to be the hard-type DPRK, that it is “too likely” to initiate war. When the U.S. estimates $w$ and estimates the DPRK is “too likely” the risks and costs of these actions become unpalatable and the U.S. would choose to acquiesce.

**DPRK payoffs.** In the simple two-stage bargaining model, if the DPRK challenges the U.S., the U.S. will respond with a sanctions response unless the expected payoff to disarming the DPRK drops below the certain payoff of $-c$, in which case the U.S. would prefer to acquiesce to the regime as a nuclear state. This leads my analysis to another question: when will the DPRK challenge? The answer to this question depends on the size of the cost to display, $c$, which depends on its type. For example if the cost to a soft-type DPRK of imitating the hard-type DPRK is too high, it will not be optimal for the type 2 DPRK to challenge. Specifically, when $2c$ is greater than $b$, it becomes optimal for the DPRK to acquiesce immediately and take a payoff of zero. The condition that $2c$ must be greater than $b$ is necessary and sufficient to produce a separation of types.

Let us take the opposite view: what if the cost to display is low such that $2c < b$? If $c$ is low enough, it could still be profitable for a soft-type DPRK to imitate a hard-type regime. This could increase the expected payoff to the soft-type DPRK such that it is optimal to bluff the U.S. Thus a necessary condition for a “pooling equilibrium” to emerge would be that $2c$ is less than $b$. This is a necessary condition, but not sufficient to produce a pooling of types. In this case we also require that the probability of a hard-type DPRK, $w$, to be sufficiently large. The probability of the hard-type DPRK must exceed the higher of the two U.S. critical risk thresholds developed above. This makes it sufficiently probable that the U.S. will meet a hard-type DPRK such that it would never be optimal to
randomize its threats against the DPRK. If \( c \) is sufficiently small and \( w \) is sufficiently large, one would expect a pooling of types to develop where the soft-type DPRK could increase its payoffs by pretending to be the type 1 regime.

*Equilibria of the game.* Therefore in a game of incomplete information, it is common that multiple equilibria exist. In the case of the DPRK and the U.S., we have already identified two: a separating equilibrium and a pooling equilibrium. There also exists a third possibility: a semi-separating equilibrium. This emerges when \( c \) is sufficiently low, such that it is not optimal for the weak-type DRPK to always choose to immediately acquiesce to a U.S. challenge to disarm, and when the probability of meeting a hard-type DPRK is also sufficiently low, such that it becomes optimal for the U.S. to randomly challenge an opponent.

In my analysis of two-stage negotiations I have highlighted the process by which the U.S. decides to employ a threat in a game of incomplete information. I have also shown how a weak-type DPRK, given certain conditions, can exploit the uncertainty in the game to increase its payoffs by imitating a stronger type of player; from my discussion of the real-life game that concluded with the Agreed Framework, one can see that there is strong evidence that the DPRK did precisely this in the conflict with the Clinton administration. The DPRK exploited the elements of uncertainty to achieve a higher payoff in the short run.

*Case 2: Allowing for more than two stages of negotiations*

With a better understanding now of how each player operates in a game of incomplete information, we can consider a game that allows for multiple bargaining rounds, which will serve to capture the more dynamic character of brinkmanship in the indeterminate game but as we will explain does not alter the decision criteria by which players maximize their payoffs.
We start by assuming that multiple stages of negotiations can only occur if the game progresses into stage 2. We also assume that both DPRK types have more response actions available than simply “initiate war” or “acquiesce.” Now the DPRK may choose from the following feasible action set:

\[ A_{1b, \text{stage } t} = \{ \text{Initiate War, Counter-Challenge (with display), Acquiesce} \} \]

The U.S. chooses from the same feasible action set:

\[ A_{2, \text{stage } t} = \{ \text{Military Response, Sanctions Response, Acquiesce} \} \]

We still assume that whenever either party initiates war or acquiesces to its opponent that the game is over. With the addition of the counter challenge action, the DPRK now has the option of rejecting the contract without ending negotiations.

As section III discussed in theoretical terms and as evidenced by the real-life events of 1994, each party graduated its threats over time, starting out with smaller threats to explore the tolerance and limitations of its opponent and increasing its threats over time. In order to incorporate more fluidity between bargaining stages, we now also assume that the military response and sanctions threats available to the U.S. and the counter-challenge threats of the DPRK to be continuous and updated following the conclusion of any negotiation round that does not result in a contract or war between the two states. Thus the military response threat made by the U.S. can vary from “Patriot” war exercises with South Korea to a build-up of armaments and bombers within the region; the sanctions response can be as small as limited sanctions imposed only by the U.S. against certain industries or could be increased to a multilateral embargo against the DPRK. The “counter-challenge (with display)” action available to the DPRK could range from a missile test launch to the more risky extraction of fuel rods. Whenever the DPRK chooses to counter-challenge, second stage negotiations begin and each player can choose to make an increased threat against its opponent. With each threat a player’s aim is to coerce the other player to acquiesce to its contract. However as a player increases the potency of its threats, it also incurs greater costs. Figure three illustrates a game that allows for more than two stages of negotiation:
As one can see from the decision tree, the cost coefficients have changed in this model. In the previous model numbers like “2” or “1” were substituted for the general coefficients of “x,” “y,” and “z” to simplify the discussion of how a player chooses strategies to maximize its payoffs in the two-stage negotiation game. But this is not necessary. What is important is the relationship of the payoffs of one outcome to the next. Here we define x, y, and z each to be some number greater than or equal to 1.00 that will scale the costs in accordance with the new threats made against the other player. As threats are made, the values of x, y, and z vary to increase with the increase the cost of making more risky threats. This allows a player to credibly coerce his opponent, as discussed in section III. As a player issues a new threat he renews his opportunity to communicate that the alternative action of acquiescing is so costly as to be eliminated from consideration; here either player can do this credibly by increasing the value of z.

With each subsequent stage of negotiations, each player can threaten the other with an action that will
lower his payoffs and can enhance the credibility of its threat by making an alternate choice more costly. The outcome of war still represents the worst outcome to both players, so the coefficient $d$ will always introduce the greatest cost to a player. As larger, more risky threats are made against one’s opponent, one’s payoffs will converge towards the worst possible outcome. The two-stage negotiation game can be interpreted as the final round of negotiations in the game that allows for multiple stages of negotiation; the player that credibly communicates a threat or commitment will transform this indeterminate bargaining game into a determinate bargaining situation, which will conclude with a contract in his favor or a disagreement outcome.

This discussion is provided to illustrate the manner by which players can increase their threats over time but concludes that because a player ultimately must compare the expected utility of any one move to another that allowing for an infinite number of negotiation stages does not materially change the interactions of players in the game.
Section IV: The Crisis in the Bush Administration

The Bush Administration reversed President Clinton’s policy of engagement. Returning to a containment policy more reminiscent of the Cold War Era, the hard-line approach of the Bush administration changed tenor of the game established by President Clinton. As we will show in this section, the tactics of the Bush administration derive from the outcome of the bargaining game concluded under the Clinton administration. As relations between the two states are essentially the same game repeated in time we will not go into depth, as we did in section IV, about how each player makes a decision at each node. This section is used to highlight how the flow of information across time affects the costs incurred by players and to illustrate how players update information based on previous bargaining outcomes. We will also restrict our discussion to the two-stage negotiation game, having concluded in the previous section that allowing for additional bargaining rounds does not materially change the decision criteria or dynamics of the player’s interactions.

Brinkmanship under the Bush Administration

Suspicions of continued weapons development never fully dissipated, even with the conclusion of the Agreed Framework and the concurrent admission of IAEA inspectors. South Korean intelligence and North Korean defectors, in addition to U.S. sources, indicated that the DPRK had been developing a second, covert highly-enriched uranium (HEU) weapons program.69

While President George W. Bush put the DPRK regime “on notice” with his first State of the Union address in January of 2002 by categorizing the DPRK as a constituent of the so-called “axis of evil,”70 it was not until later that year that the second phase of the game began. On October 16, 2002, U.S. Assistant Secretary of State James Kelly confronted North Korean officials with evidence of the

---

69 Today experts believe this program was started as early as 1995. See Sanger, David E. “Bush Takes No-Budge Stand on Talks with North Korea,” The New York Times, April 17, 2003.
70 The State of the Union Address, January 29, 2002, named three countries in the axis of evil: North Korea, Iraq, and Iran. Subsuming these countries under a single umbrella of “rogue states” or dissident international actors increased the interdependency of U.S. commitments. Its credibility in making threats or commitments to any one of these countries was thereafter closely correlated to its willingness to uphold its commitments to the other members of the axis of evil.
HEU program while visiting Pyongyang. With this challenge the U.S. initiated a second bargaining
game with the DPRK.

North Korea officials, ostensibly nonplussed by the accusation, admitted to the clandestine HEU
weapons program and immediately offered to end its HEU program in exchange for a nonaggression
agreement, recognition of its sovereignty, an agreement to not obscure its economic development.71

Contrary to the moves of his predecessor, Bush refused to open dialogue. Holding the DPRK in
material breach of the Agreed Framework, the U.S., the European Union, South Korea and Japan cut off
the shipment of 500,000 tons heavy oil fuel the following month72 to which the DPRK responded with a
series of actions to signal to the U.S. that it was willing and capable to risk nuclear conflict rather than
submit to the will of the U.S.73

In February of 2003, while building up forces in the Middle East for war in Iraq, Defense
Secretary Donald Rumsfeld also announced the U.S. would be sending 24 B-52 and B-1 bombers74 and
additional military equipment to Guam75 and was considering a drastic reconfiguration of U.S. forces in
South Korea that would possible reduce and relocate U.S. garrisons stationed within the range of a short-
warning strike by the DPRK.76 Concurrently with this threat of military build-up in the region, the Bush
administration also announced that it was planning a sanctions resolution that would prevent any
weapons exportation from the DPRK among other restrictions.77 The sanctions resolution was not
approved by the U.N. Security Council due to reservations by regional actors, most notably China.

Former Ambassador to China and South Korea James R. Lilley explained that “the Chinese remained
deeply ambivalent about squeezing North Korea, fearful that it will collapse and send millions of

---

71 A statement by the DPRK’s Foreign Ministry Spokesman, October 25, 2002.
72 November 11, 2003.
73 On December 12, 2002 the DPRK announced it would reactivation the Yongbyon nuclear reactor, the nuclear plant frozen
under the 1994 Agreement. Then, on the heels of the start-up of the Yongbyon plant the DPRK followed through on threats made to the Clinton Administration by withdrawing from the Nuclear Non-Proliferation Treaty73 and expelling all IAEA inspectors from the DPRK’s borders.
74 “Each B-1 bomber can carry up to two dozen one-ton, satellite guided bombs. The payload of the giant B-52 is 70,000
pounds of bombs and missiles. See Sanger, David E. and Thom Shanker, “U.S. Sending 2 Dozen Bombers in Easy Range of
75 February 3, 2003.
impoverished refugees into Northeastern China.” But he also noted, “The Chinese were concerned about North Korea’s nuclear programs and had been willing at least twice in the past decade to reduce food and fuel shipments...in apparent efforts to pressure North Korea to freeze its weapons programs.”  

North Korea continued to challenge the U.S. and the U.S. continued to pressure the DPRK. On February 27, the U.S. concluded that the Yongbyon reactor had been restarted, as threatened, but did not believe that the reprocessing of fuel rods had been started at the same time. The U.S. warned regional powers that if the DPRK were to start fuel reprocessing, it could produce a nuclear bomb per month for the next five or six months.

While North Korea launched another anti-ship missile some concern over the DPRK’s technology waned: “North Korea’s talk may prove scarier than its technology. In the last two weeks, NK tested land-to-sea missiles that failed to hit their targets.” Furthermore, the war in Iraq seemed to temper the saber-rattling of the DPRK, as did quiet protestations from China. In April 2003, North Korea announced a reconsideration of the forum in which it would negotiate the status of its nuclear program; it would no longer demand a “[bilateral] dialogue format.” This was followed a week later with a reconvening of nuclear-arms talks between North and South Korea, after more than a month of cancelled summits between the two states.

While the DPRK allowed China to attend talks, the regime refused any representatives from South Korea and Japan, thus allowing diplomats from only China, North Korea, and the United States to

79 The DPRK threatened to withdraw from the 1953 Korean War armistice if the U.S. followed through with a military buildup in the region and on the eve of a diplomatic tour of the Pacific region by Colin Powell, February 20, a Korean MiG 19 entered South Korean airspace. The following week, in the days before the inauguration of Roh Moo-Hyun the DPRK (February 24, 2003) launched a short-range anti-ship Silkworm missile into the Sea of Japan. See Eckert, Paul. “N. Korea Fires Missile Ahead of Roh Inauguration” Reuters News on Yahoo!, February 25, 2003.
80 On March 2nd North Korean MiGs intercept a U.S. RC-135 “Cobra Ball” surveillance plane that specializes in photography of foreign ballistic missile tests at long range and attempt to force its landing; the U.S. contended that it was in international airspace.
83 In late March, Chinese ministers to North Korea warned the DPRK that “continued provocations could harm its relationship with China.” China also halted oil supplies to the DPRK for several days; experts estimate that the DPRK currently relies on China for “as much as 80 percent” of its oil and food supplies. For more information see French, Howard W. “North Korea’s Reaction on Iraq Subdued So Far” The New York Times, April 2, 2003.
negotiate. To many, it seemed the path was paved for progress. But then on April 18 North Korea seemed to test the water with a statement by the DPRK’s official KCNA news agency that it was “successfully reprocessing” the 8,000 that had been relegated to storage under the Agreed Framework.\textsuperscript{85} The DPRK quickly retracted the “mistranslated” statement. But it had gained valuable information about the U.S., as Park Jong Chul, senior research fellow at the Korea Institute of National Unification explained, “We might think North Korea intentionally made a mistake to see how hard-line is the U.S. attitude.”\textsuperscript{86}

Negotiations continued as planned. Ra Jong Il, South Korea’s national security advisor, dismissed the KNCA statement as an “effort to gain leverage ahead of the talks.”\textsuperscript{87} However optimism that negotiations would end the brinkmanship between the U.S. and the DPRK in China were again eclipsed by a DPRK weapons declaration. During negotiations North Korea announced it had already reprocessed the fuel rods and coupled with its announcement with a threat to export the weapons or to use them against the U.S. Li Gun, chief of the North Korean delegation, indicated that it would adopt a respond-in-kind procedural rule whereby the outcome would be determined by U.S. actions.\textsuperscript{88} Before the negotiations in Beijing were brought to a close, the DRPK offered to dismantle its nuclear program, suspend ballistic missile sales, and halt missile exports in exchange for security assurances from the U.S., a softening of demonstrated hostility, recognition of the DPRK as a sovereign state, energy assistance in the form of oil shipments and the continued construction of two light-water nuclear reactors, food aid, and economic benefits.\textsuperscript{89} The DPRK reiterated, too, that sanctions would result in war.

While the DPRK had crossed a “red line,” the U.S. contended it was unable to conclude with certainty that the DPRK declaration was true. As the Bush administration shifted its focus towards

\textsuperscript{88} The declaration was made on April 24, 2003. See Schuman, Michael and Donald Macintyre, “Joining the Club” \textit{Time Asia}, May 5, 2003.
prohibiting the sale of nuclear armaments and sent various U.S. congressmen to talk with Pyongyang 
officials, the DPRK followed up its nuclear weapons declaration by withdrawing from The Joint 
Declaration on the Denuclearization of the Korean Peninsula.90

The Effects of Previous Play on the Bush Administration Game

The actions and strategies of the first game under the Clinton administration are the same under 
the second game during the second Bush presidency. Specifically, we assume that there is a set of \( N \) 
players = \( \{1, 2\} \), where 1 represents the DPRK and 2 represents the U.S. For player 1, there is a finite 
number of types selected from the set \( V \) possible types such that player 1’s true type, \( b_i \in V = \{1, 2\} \) for 
\( i = 1 \). Type 1 represents the hard-type player 1 (DPRK) and type 2 represents the soft-type player 1. 
Thus, there exist two possible states of the world:

\[ \beta = [11, 2\cdot], [12, 2\cdot] \]

Where the first state of the world \( \beta = [11, 2\cdot] \) describes a hard-type DPRK playing the U.S. and occurs 
with probability \( w \). The second state of the world \( \beta = [12, 2\cdot] \) describes the soft-type DPRK and occurs 
with probability \( (1-w) \). However, as this analysis will show the U.S. estimates of \( w \) have been updated 
based on information learned from outcomes of earlier games.

Each player selects an action from the available action set \( A_i \) for \( i = \{1, 2\} \) at stage one of the 
game, which is common knowledge among players.

\[ A_{1b, stage 1} = \{\text{Challenge U.S. with display, Challenge U.S. without display, Acquiesce}\} \]
\[ A_{2\cdot, stage 1} = \{\text{Challenge, Acquiesce}\} \]

Whenever a player chooses to acquiesce, the game is over and it is assumed a contract is formed by 
which either 1) the DPRK is disarmed, which happens when player 1 acquiesces, or 2) the U.S. 
normalizes relations with the DPRK as a nuclear power, which happens when player 2 acquiesces. If

---

neither the U.S. nor the DPRK acquiesces in stage one, the game progresses to stage two, again with player 2 moving before player 1. The feasible action sets for each player in stage two are the following:

\[
A_{1b, \text{stage 2}} = \{ \text{Initiate War, Acquiesce} \}
\]
\[
A_{2, \text{stage 2}} = \{ \text{Military Response, Sanctions Response, Acquiesce} \}
\]

Whenever war is initiated or a player acquiesces, the game is over. Figure 4 illustrates the two-stage bargaining game, complete with payoffs, played during the second Bush administration.

The game is a repeat of previous play, but it should be underscored that some parameters of the game have changed because the players have learned something of each other from earlier games.

One element of the game that has changed is the cost to display or make a threat. The cost of display \( c \) is history-dependent. What display tactics were effective in 1994 must be superceded in 2003.
to carry the same credibility. This raises the cost of \( c \), which may make it sub-optimal for the weak-type DPRK to challenge. As we noted in section IV, the precise condition that makes challenging unprofitable for the weak type DPRK is when \( 2c > b \). When this condition is satisfied, the soft-type DPRK earns a negative payoff from imitating the hard-type DPRK and thus will choose to acquiesce instead; a separating equilibrium will emerge. However if \( c \) does not increase to satisfy this condition and the probability that \( w \) is sufficiently large, the game will produce again a pooling equilibrium. In this case, as we explained in the previous section, it will still be optimal for a soft-type DPRK to imitate a hard-type regime because it can earn a higher payoff by bluffing.

As we noted above, the emergence of a revealing or pooling equilibrium depends not only on the cost to challenge and display, which has most likely increased, but it also depends on the U.S. assessment of \( w \), the probability with which nature chooses a hard-type DPRK. Using information observed from earlier games the U.S. will update its prior estimate of \( w \). In 1994, the DPRK challenged the U.S. with display but ultimately acquiesced to U.S. demands to disarm. Intuitively, one might think that this would automatically indicate the probability of the DPRK being a hard-type regime is zero (\( w = 0.00 \)). But the Bush administration continues to face uncertainty in the repeated game. Why? The source of the uncertainty emanates from the parameter \( q \), the probability with which the hard-type DPRK follows through with the threat of war in response to military or sanction actions. It emanates from the very definition of brinkmanship: posing the risk of disaster, but not the certainty of it.

In the first game, \( q \) was some number greater than zero and less than one, such that the U.S. knew the expected payoffs associated with a threat of war. The U.S. uses the information of the first game to generate a posterior probability of seeing a hard-type regime according to Bayes’ Theorem:

\[
\begin{align*}
w' &= \text{pr} (\beta = 11, 2 - | \text{DPRK didn’t go to war in game I}) \\
&= \frac{\text{pr} (\text{DPRK didn’t go to war} | \beta = 11, 2) \cdot \text{pr} (\beta = 11, 2)}{\text{pr} (\text{DPRK doesn’t go to war})} \\
&= \frac{[(1-q) (w)] / [(1-w) (1) + (w) (1-q))]}{[1-w + w-wq]} \\
&= [(1-q) (w)] / [1-wq]
\end{align*}
\]
Given this relationship of the prior estimation of $w$, the new estimation of the DPRK’s type by U.S. has been degraded. To make the discussion more clear, let’s assume that the U.S. assessment of $w$ in game one was $w = 0.50$. Using Bayes Theorem, $w' = [(1-q) (0.5)] / [1-0.5q]$ we can plot the relationship between the updated assessment of a hard-type DPRK and the probability of initiating war:

![Updated probability of w vs Values of q](image)

As we can see, the original estimation of $w$, the probability that the DPRK would have followed through with its threat to initiate war is the upper limit on the updated estimation of the DPRK as a hard-type player. When the probability $q$ that the DPRK would actually fight was very low in game one, the estimation of $w$ by the U.S. is relatively unchanged. However as $q$ increases, the knowledge that the DPRK ultimately acquiesced despite this high value of $q$ dramatically lowers the U.S.’s estimation of the DPRK as a hard-type player in the second game. This is consistent with intuition. Thus as play is repeated over time, if the DPRK continues to acquiesce to the U.S., the U.S. assessment of $w$ will commensurately diminish.

As a corollary, the benefits from bluffing decrease. From game one we know that the U.S. will use a military or sanctions threat to coerce the DPRK to acquiesce when its estimate of $w$ falls below the following critical risk thresholds:

- Military Threat is used when $w < (b - c)/ [(q) (c + b)]$
- Sanctions Threat is used when $w < (b)/ [(q) (2c + b)]$
As $w$ diminishes, the likelihood that the U.S. will acquiesce diminishes as well. North Korea must work harder to cause the U.S. to estimate the actual risk of $w$ to be below its critical thresholds above, which in turn raises the cost of $c$.

Our analysis has shown that a pooling equilibrium cannot persist as the number of games is increased over time. Again, because every game will be marked by incomplete information, multiple equilibria will exist. However as our the very parameters that allow for bluffing in earlier games by soft-type DPRK players eventually serve to reveal its type to the U.S. as the benefits from imitating hard-type players is degraded by the flow of information across time.
Conclusion

This analysis has of nuclear brinkmanship between the United States and the DPRK during the Clinton and Bush administration has treated the conflicts between the two states as the same basic game is repeated through time. It has sought to identify the incentives of each state for engaging the other in negotiations and to draw out the sources of uncertainty in the game. We have characterized the game as one of incomplete information where players look ahead in the game to solve recursively in such a way that they maximize their expected utilities at any given outcome. This necessarily called for the elimination of certain strategies as sub-optimal, in particular the action by the DPRK to challenge without display. Because the games are marked by incomplete information, multiple equilibria may emerge depending on the parameters of the game: a pooling of types equilibrium, a separation of types equilibrium and a semi-separating equilibrium. Over time the ability of a soft-type player to bluff its opponent by imitating a hard-type player is diminished by the flow of information across time.
Notes


Daalder, Ivo H. and James M. Lindsay, “Where are the Hawks on North Korea?” *The American Prospect*, Volume 14 Issue 1, February 1, 2003.


Laney, James T. and Jason T. Shaplen, “How to Deal with North Korea” Foreign Affairs, March/April 2003, p17.
“North Korea’s Creative Revenue Raisers” CNN.com, April 28, 2003. For more information on DPRK illegal exports.
The State of the Union Address, January 29, 2002.
U.S. Census Bureau, Foreign Trade Division, Data Dissemination Branch, Washington, D.C. 20233