

**EXPLAINING NORTH KOREAN NUCLEAR BEHAVIOR:
NORTH KOREA'S SECOND NUCLEAR CRISIS
Andrew Roskos-Ewoldsen**

This paper attempts to explain North Korean nuclear behavior during the time period known as the second nuclear crisis. This crisis is characterized by erratic behavior by the North Korean government from 2002 to 2009, where the state swung wildly back and forth from cooperative, non-threatening nuclear behavior to aggressive, threatening nuclear behavior. This was significant because it had major security implications for Northeast Asia and the world with the possibility of nuclear violence, but also because this type of behavior is unprecedented. Normally, states pursuing nuclear proficiency are steadfast in their pursuit, but North Korea displayed inconsistent and oddly timed transitions in nuclear policy. Two models, neorealism and a synthesized theory called constructivist prospect theory, are used to explain these unusual transitions. Their findings help us in understanding this strange and critical period of history, and also reveal insights into the strengths and weaknesses of both approaches.

Andrew Roskos-Ewoldsen is a Master's candidate of International Cooperation at the Graduate School of International Studies, Seoul National University. He received his Bachelor's degree from Rhodes College in Memphis, Tennessee, U.S.A. His research interests are North Korean security, US-Korean relations, and North-South Korean relations.

In 2003, the Democratic People's Republic of Korea (the DPRK, or North Korea) withdrew from the Nuclear Non-Proliferation Treaty, reactivated its nuclear facilities, and conducted its first known test of a nuclear device, while making threats and warnings to the outside world. This marked the beginning of a second nuclear crisis in North Korea, and it stands out as a landmark event in the nuclear age. The stakes were high for the region and the world alike, but what makes the crisis so seminal is the type of nuclear behavior North Korea exhibited. The first nuclear crisis in North Korea (1992-1994) established an erratic precedent: the DPRK's desire for nuclear weapons, alongside a willingness to cooperate. The second crisis (2002-2009) represents a period of drastic changes in nuclear behavior, whose timing and dynamics appear irrational, inconsistent, and unexpected. In the past, states that have gone nuclear have pursued nuclear proficiency consistently, until attaining a nuclear stockpile, but the second crisis is a case in which a state did not follow this expected path. Instead, North Korea transitioned back and forth from passive nuclear behavior to aggressive nuclear behavior,

characterized by agreeing to nuclear agreements and then suddenly breaking them in favor of testing nuclear weapons or threat issuance. This seemingly irrational behavior even included the first known instance of a state acquiring nuclear weapons capabilities and then abandoning its nuclear programs (He and Feng, 2013). In short, the highly erratic behavior of North Korea both scared and puzzled bystanders, policymakers, and social scientists.

Academically, efforts to explain this high variance in North Korean nuclear behavior have fallen largely to the predominant school of thought in international relations: neorealism. Neorealism's focus on structural factors of the international system with a positivist analytical approach, and its general applicability, provide an easy-to-use framework for studying North Korea. It generates falsifiable hypotheses by focusing on the power distribution of rational states in the international system, and therefore has been a useful tool for studying typical state behavior. Yet, while this theoretical approach has contributed greatly to the overall study of international relations, neorealism has shortcomings that limit its to explain certain

phenomenon effectively. Specifically, for this paper, neorealism explanatory capacity might be limited if a unitary actor assumption can be made, i.e. the entire behavior of a state is a reflection of the thoughts and actions of a single leader. This paper contends that North Korea and the second nuclear crisis presents a unique case that cannot be effectively explained by neorealist assumptions, and introduces an alternative theory called constructivist prospect theory (CPT) that provides a more convincing and thorough explanation of North Korea's highly variant and radical behavior. In contrast to neorealism, in which state decisions are rational and resultant of structural pressures, CPT focuses on identity and the decision-making process of domestic political leaders.

Acknowledging North Korea as an atypical case of state behavior, as studied from the perspective of CPT rather than neorealism, could have serious implications for how the world understands and interacts with North Korea in the future.

In this paper these two models will be tested and compared in order to evaluate their explanatory powers on the case of North Korea's second nuclear crisis. First, the neorealist model will be introduced and assessed. Second, the CPT model will receive the same treatment. Finally, the significance of the findings and their implications for future research will be discussed.

Neorealism

Neorealism seeks to explain international events by taking the foundations of classical realism and adding an emphasis on the structural aspects of the international system. Kenneth Waltz, one of the pioneers of neorealism, argues that differences in the relative power of states under the condition of anarchy is the primary causal mechanism for international actions taken by states (see also Shimko, 1992). Within Waltz's strand of neorealism, states are all equal and identical in nature save for one important value: their aggregate power. Specifically, Waltz argues that the anarchic system creates differences in power which the system will try to balance as self-interested and rational states try to secure their own safety and power relative to other states. Consequently, differences in aggregate power that

arise between states forces states to address the imbalance. Waltz argued that states improved security primarily through military power and alliances, with the strongest security measure being nuclear weapons (Waltz, 2001).

Modifications of Waltz's assumptions have been proposed. Offensive realism asserts that the pressure of the international environment makes it impossible for states to know whether they have attained the necessary power to ensure survival, such that they continuously maximize their utilities and relative advantage over others (Mearsheimer, 2001). Other scholars have argued that the way in which states conceptualize themselves and others affects how they respond to international pressures (Jervis, 1988). In what is known as the security dilemma, Jervis contends that states will try to increase their own security by building up their military capabilities, thus neighboring states do the same based on this initial threat. The perception (or misperception) of military buildup by states can be key in explaining why conflict occurs or is avoided.

Neorealism focuses on systemic factors and power relations for its explanation of North Korea's second nuclear crisis. Since North Korea's inception it has been present in a very tumultuous region with a delicate balance of power. The effects and after-effects of the Cold War positioned North Korea to be at odds with the regional powers South Korea, Japan, and the United States, while being allied with China and, temporarily, the Soviet Union. With the collapse of the Soviet Union, North Korea was faced with competing powers that threatened its security. Using the neorealist model one can advance the following hypotheses about North Korea's expected behavior: (1) North Korea would pursue aggressive nuclear behavior when it is the most efficient option for state security; (2) North Korea would likely transition to passive nuclear behavior when an alliance offers stronger state security than aggressive nuclear policy. The self-interested nature of all states make it rational for North Korea to pursue nuclear weapons to ensure its own survival while gaining a relative advantage over rival powers. Not pursuing aggressive nuclear behavior would be irrational unless security and

survival were attained through other means, such as an alliance or passive nuclear behavior.

Neorealism and the Second Nuclear Crisis

North Korea was consolidated as a communist state in 1948. Kim Il Sung attempted to unify the peninsula during the ensuing Korean War. This war solidified North Korea's allies and enemies as it received support from China and the Soviet Union, while fighting against South Korea and the United States. In a stalemate, North Korea signed an armistice agreement in 1953 that technically kept the state at war with South Korea and the United States, but both sides agreed to end military hostilities. Kim Il Sung had established himself as a militaristic dictator with total control over all aspects of North Korean affairs, while also clarifying the power distribution of Northeast Asia during the Cold War.

Kim Il Sung relied heavily on the Soviet Union and China throughout the Cold War, while having little positive contact with South Korea, United States, or Japan until the Soviet Union's collapse. North Korea's self-reliant domestic policy called *juche* was intended to preserve state security, but the state's dependence on its powerful communist neighbors made substantial contributions to its national security through economic and military support (He, 2013).

Beginning in the late 1980s, North Korea suffered a major downturn when Mikhail Gorbachev reevaluated the Soviet Union's imbalanced aid to North Korea as untenable (Beal, 2005).

Furthermore, the Soviet Union briefly engaged with South Korea diplomatically, a betrayal of North Korea's strongest political ally. Likewise, China established diplomatic relations with South Korea in 1992. With the collapse of the Soviet Union, virtually all the economic subsidies and aid on which North Korea was so reliant came to an end. China, while maintaining its defense of and aid to North Korea, also began cooperation with the United States and South Korea as it dealt with its own domestic issues. This pivoting of allies to North Korea's enemies only heightened security concerns for the state.

Going Nuclear

In the post-Cold War era, North Korea chose two simultaneous approaches for increasing its security in the wake of the Soviet Union's collapse: the formation of alliances and the pursuit of nuclear weapons. North Korea attempted the former by making several breakthroughs with South Korea, including the signing of several bilateral agreements that promised denuclearization and cooperation on the peninsula. North Korea also began engaging with the international community by signing the Non-Proliferation Treaty (NPT), joining the United Nations in 1991, and signing the nuclear safeguard agreement with the International Atomic Energy Agency (IAEA). Contrary to the values of these agreements, Kim Il Sung also started nuclear weapons production in secret. North Korea had several active nuclear facilities and reactors that were intended for plutonium extraction and development (Pinkston, 2009). According to the CIA, North Korea had produced enough plutonium for at least one nuclear weapon by the time the IAEA started their inspection in 1993 (He, 2013). The ensuing inspection found evidence of three separate instances of plutonium reprocessing, while the United States also offered satellite photos of nuclear reactors, resulting in the IAEA declaring that North Korea had violated the non-nuclear agreement.

This period is described as the first nuclear crisis. The discovery of North Korea's pursuit of nuclear weapons resulted in the international community calling for increased measures to halt North Korea's nuclear programs. The United Nations Security Council passed a resolution that requested North Korean facilities be inspected under the NPT and IAEA agreements, but North Korea refused and threatened withdrawal from the NPT. Tensions escalated as the United States threatened economic and military sanctions over North Korea's non-cooperation. North Korea responded with war threats against South Korea and the U.S. (McAllister, 1994). North Korea continued its aggressive nuclear behavior in the face of international pressure before an agreement was finally reached in 1994 in the Agreed Framework between the DPRK and the United States. North Korea agreed to freeze its nuclear program in

exchange for light-water reactors and 500,000 tons of heavy oil annually, along with security assurances (He, 2013). In addition, North Korea reentered bilateral talks with South Korea. Kim Il Sung died a few months before the signing of the Agreed Framework, but Kim Jong Il, his successor, continued his father's legacy of a passive nuclear policy until the early 2000s.

The Case

Neorealism focuses on the balance of power in the region and the rational options available to North Korea in the face of these systemic pressures. The structural pressures of the anarchic system during this time period are quite clear: North Korea was surrounded by more powerful states that threatened its national security.

In terms of military strength, North Korea during the beginning of the second nuclear crisis was inferior to South Korea and the United States. Records show that North Korea was at a disadvantage in terms of the quality of the army, military expenditure, weaponry, technology, and resources during the first decade of the 21st century (Cha and Kang, 2004; Goo and Kim, 2009; Goo and Lee, 2014; Pinkston, 2009). North Korea has a larger standing army than South Korea or the United States, at a size exceeding one million, but most reports agree that South Korea's standing army of around 600,000 and the ROK-US alliance forces enjoy superior training, equipment, and logistical support (Cha and Kang, 2004). In terms of military expenditure, North Korea was outspent by South Korea in terms of absolute expenditures (by around \$800 million) and percentage of public expenditure (by around 20-30%). The United States consistently dwarfs these absolute expenditures, but this can be misleading because not all of the US defense spending goes towards Northeast Asia. North Korea also has a disadvantage in weapons, technology, and resources. The United States had superior nuclear capabilities and, while officially denying the existence of nuclear weapons in South Korea, reports have confirmed the nuclear sufficiency of reactors in Seoul during the early 2000s (Pinkston, 2009). South Korea has a greater and higher quality stockpile of land-to-air weapons, and naval

vessels and aircraft. During the crisis, North Korea was presented with a clear need to respond to the unequal distribution of power in the region. Furthermore, during the second crisis, North Korea had one key defense alliance with China. However, the China-North Korea alliance does not come close to the military power of the combined forces of the ROK-US alliance (Goo and Lee, 2014). One reason for this was China's focus on economic reforms (He, 2013). This alliance disparity in terms of combined military power meant the distribution of power in the region was against North Korea.

In accord with the first hypothesis, North Korea began pursuing aggressive nuclear behavior. In 2002 North Korea had one rational option for self-preservation in the face of mounting external pressure: to pursue an aggressive nuclear policy in order to deter military actions against it. North Korea removed itself from the NPT and reactivated its nuclear facilities in 2003, before publicly admitting to possessing nuclear weapons in 2005 and finally testing a nuclear weapon in 2006. This behavior is consistent with the prediction that North Korea was trying to maximize its security in the face of difficult systemic pressures from the anarchic system. North Korea made its nuclear activity public so that rival states would know of its nuclear capabilities and be less inclined to attack. There was then a transition from aggressive back to passive behavior in 2007, only a year after completing a successful detonation of its first nuclear weapon (Pinkston, 2009). After attaining an operational nuclear deterrent, North Korea agreed to relinquish it by freezing all nuclear activity and dismantling its weapons as part of a Six-Party Talks agreement. Since there are only two ways to increase security under anarchy – military strengthening or alliance formation (Waltz, 1979) – the only rational explanation for this in a neorealist approach would be identifying passive nuclear behavior as a better security option. This points to the hypothesis that North Korea pursued a passive nuclear policy when an alliance could increase its security more than aggressive nuclear policy. The alliance offered by the Six-Party Talks must have offered stronger

security incentives than a fully functional nuclear weapon.

This proves to be an empirically weak argument. The Six-Party Talks were not very different than the Agreed Framework in which denuclearization was to be exchanged for normalized bilateral relations, security assurances, and the offering of economic, energy, and humanitarian aid in the form of oil and light-water reactors (Pinkston, 2009). While these agreements could be beneficial, the increase in security would not be as assured or at the level of a fully functional nuclear deterrent. The distribution of power remained grossly uneven and the historical tension between the agreeing states did not disappear. A security assurance by South Korea or the United States was not as guaranteed or as safe as a nuclear deterrent. Even more oddly, North Korea completed a transition back to aggressive nuclear behavior with the second successful nuclear weapons test in 2009. This inconsistent behavior raises questions about the applicability of the assumptions of the neorealist model. If North Korea's nuclear behavior was going to change, there must have been a structural change that encouraged a different, rational approach to nuclear policy. Yet there is no evidence of a significant change in the structure of the anarchic system during this period. These findings could also mean that North Korea does not choose behavior rationally, or that its preferences are not always based around state power, survival, and security. Critically, the neorealist model has difficulties explaining the timing of North Korean behavior.

Constructivism/Prospect Theory

Structural, top-down theories such as neorealism ironically rely on state agency. Neorealists contend that due to the anarchy of the international environment and differences in power, states will be forced to pursue power and military advancement for their own survival. The causality is actually not in the system-level variable alone, however, because this assumption relies on the idea that states are predatory, self-help-oriented actors who will rationally choose to pursue relative advantages over other states. The nature of the state is arguably as responsible for causing this

behavior as the anarchy of the system. This implicit usage of state agency belies the true importance of a micro-, bottom-up approach. Constructivist prospect theory (CPT) is a synthesized theoretical approach that combines the constructivist emphasis on preference formation with prospect theory's decision-making under uncertainty to create a comprehensive and systematic model for explaining state behavior. Under this model, state decisions can be explained by making a unitary actor assumption, analyzing the social identities of state leaders, identifying their socially constructed preferences, and then evaluating their available choices. Constructivism openly addresses the important role of state agency, claiming that state preferences can be different and unique according to social and historical context. Prospect theory explains how decisions are made based on these social identities, and how identity affects evaluations of choices and outcomes. CPT ultimately offers a blueprint for what decisions states are likely to take. When considered separately, constructivism and prospect theory have substantial limitations in explaining state behavior, but when combined form a compelling model.

Constructivism is a critical theory that focuses on the formation of preferences and identity for actors. Constructivists make the claim that international life is socially constructed rather than determined by the inherent nature or unavoidable character of states or the international environment. Alexander Wendt (1992) argues that the structural realities that neorealism and neoliberal institutionalism take as givens are formed by shared ideas and social practice, and the identities and preferences of key actors are determined by these shared ideas. Social interaction determines collective identity and preferences, so states and their interests in power, as assumed in neorealist explanations, are actually socially manufactured and subject to change in line with social practice. Where neorealism merely claims states are a certain way exogenously, constructivism explains why states are the way they are. The causal mechanism of state behavior, according to constructivism, is the state's preferences and identity. Studying how these are

formed through social interaction and practice reveals how and why states behave the way they do. It also provides an explanation for systemic factors as well, with social practice and shared ideas giving value to structural ideas such as anarchy, institutions, and regimes. Therefore, constructivists acknowledge the influence structure can have on state behavior, signifying a possible coexistence with structural models like neorealism in some instances.

Since states, as well as their preferences and interests, are socially constructed, the rationality assumption made by neorealism can no longer be assumed. States may not always act rationally according to cost-benefit calculations, and perhaps will not choose decisions that maximize security, survival, and material power. A state has its identity constructed, not automatically determined. Jacques Hymans (2008) gives an account of how state identity formation occurs. National Identity Conception (NIC) refers to the basic sense of what a nation stands for and how it compares to other nations. This is a partially subjective measure, as it takes into account emotional and psychological factors, and this is why states might seem to behave “irrationally” under the objective, material measures of neorealism. NIC establishes national preferences and interests.

While constructivism convincingly provides answers on how state identity and preferences are formed, it does not explain how states go through their decision-making process. The origins of state behavior can be explained, but not their timing and dynamics. Constructivism convincingly argues against all states being rational and existing within an anarchic system, with identical and fixed preferences, since both rationality and anarchy rely on assumptions about agency that are not adequately addressed. Yet, it fails to explain how, if a state is not necessarily rational, it evaluates and chooses options and consequences.

Daniel Kahneman and Amos Tversky developed prospect theory as a model for decision-making. Originally created at the individual level as an alternative to expected-utility theory, otherwise known as the rational actor model, Kahneman and Tversky (1979) developed prospect theory as a

model for explaining how actors evaluate and choose options around them.

Prospect theory assumes that actors’ decision-making and preferences revolve around a reference point that actors use to value assets and choices. This is in direct contrast to the expected-utility rational actor model, in which the costs and benefits are weighed completely objectively.

Based on the idea that the reference point is central, prospect theory provides a general process of choice making. Firstly, in an editing phase an actor identifies a reference point through a process called framing, using available information like expectations, aspirations, social norms, and social comparisons. The actor also assigns value and probability of outcomes to the available options around this reference point. Secondly, in the evaluation phase, the values of possible outcomes are combined with their weighted or compared probabilities, and the actor uses the product to make decisions. In other words, actors do not have a consistent and transitive set of preferences and use less than optimal calculations to maximize these preferences. Consequently, behavior becomes non-linear or possibly irrational. Instead of making completely objective decisions, actors select choices according to probable gains and losses of actions based on a reference point.

Another key assumption in prospect theory is that actors care more about changes to assets than net asset levels. Essentially, actors are more concerned with gains and losses that are given meaning relative to a reference point than the total or net worth of an option. This claim introduces a subset of assumptions about gains and losses. Firstly, actors treat gains and losses differently: they tend to overvalue losses compared to equal gains. This affects actors’ behavior in what is called *loss aversion*. For instance, an actor will probably not engage in an activity that has an equal chance of success or failure since the actor tends to overweigh the failure.

The *endowment effect* states that actors overvalue assets they already possess as opposed to comparable assets they do not have. This implies that losing an asset is worse than acquiring an asset of comparable value. An actor might work or pay more to keep an asset than they ever would to

attain that asset if it did not belong to them. Another assumption called *status quo bias* states that actors tend to treat the costs of moving away from the reference point as losses and the benefits of moving away as gains. Coupled with the tendency for actors to overvalue losses, it can be seen that they tend to remain at the status quo. The last assumption in regards to gains and losses is called *risk orientation*, which simply states that actors are risk-averse in the domain of gains and risk-acceptant when it comes to losses. This explains actors' choices: if the possible outcome is a gain then the actor will probably not take much risk to achieve it, but if there is a possible loss then the actor will be far more inclined to take risks in order to avoid it.

Finally, prospect theory assumes that framing is absolutely crucial to choice making. Framing is the identification of a reference point, which is the basis for all decision-making, since it determines what constitutes a gain or a loss. Most instances of framing depend on the setting or situation, and usually the status quo, or the pre-determined situation, is made to be the reference point. A change in frame is just as important as establishing the original frame, which can easily create new editing and evaluation phases. It is critical to take into account dynamic situations in which the reference point and frame have a high likelihood of changing, and an actor's behavior might be altered substantially with each changing situation. Prospect theory has received substantial support at the individual level, but challenges remain in applying it to international relations or other fields. Methodologically, it is hard for prospect theory to be applied because of the micro-level nature of its focus (Boettcher, 2004; Levy, 1992). International events are dynamic in nature, making the analysis of options, values, and frames much more difficult. There is no systematic formula for applying prospect theory to higher levels of analysis, so this has been done only on a case-by-case basis. Prospect theory has been used in combination with other theories such as neorealist power transition theory to fill in theoretical shortcomings and explain North Korean foreign policy (Cha, 2003; He & Feng, 2013). Prospect theory has been used in a standalone fashion as well: Haas (2001) has

used prospect theory to explain the Cuban missile crisis in comparison to expected-utility theory and Farnham (1995) has used it to analyze President Roosevelt's handling of the Munich crisis. International relations applications of prospect theory are almost always used in cases in which a unitary actor assumption can be made because the predictions of the theory are more straightforward with the assumption of a single reference point. Constructivism, and NIC in particular, can provide a framework for establishing a reference point at the international level. NIC describes the aspirations, social norms, and social expectations needed to identify a reference point. Once a reference point is established, prospect theory can explain and predict likely behavior. Consequently, CPT offers unique insights on state agency and decision-making.

When applied to the case of North Korea's second nuclear crisis, constructivist prospect theory proposes three hypotheses: (1) North Korea pursued aggressive nuclear behavior as the status quo. This hypothesis illustrates that North Korea's leadership identity, as an oppositional nationalist, established having nuclear weapons capability as the reference point or status quo. Therefore, aggressive nuclear behavior was the status quo and will be framed in a domain of gains. Due to this reference point, passive nuclear behavior involves the possibility of losing these nuclear weapons and consequently will be framed in a domain of losses. North Korea should be unlikely to move away from aggressive behavior because of the status quo bias. (2) North Korea only pursued passive nuclear behavior when the reference point changed, hence a new editing phase occurred and passive nuclear behavior was perceived as having no risk. This hypothesis indicates North Korea perceived a choice to limit their nuclear program as having no risks in damaging nuclear assets while also reaping the benefits of foreign aid and resources. This implies that it must be a very specific passive choice with strong benefits, such as making agreements that offer substantial concessions, with little to no perceived risk, such as the dismantling of nuclear facilities or stringent plans for nuclear downgrading. The available passive choice must allow Kim to leave his nuclear assets intact or have

the ability to escape any threat to his nuclear assets by abandoning the agreement. (3) North Korea would return to status quo aggressive nuclear behavior when the risk of continued passive nuclear behavior increased and was framed in a domain of losses. Kim should operate under the reference point that he was operating under earlier that was established by his oppositional nationalist identity. In this context, when the risk of damaging his nuclear assets increased, he would protect them and return to preference fulfilling, status quo behavior.

The Second Nuclear Crisis: Constructivist Prospect Theory (CPT)

The neorealist approach has contributed greatly in how state behavior is interpreted, but North Korea's preferences and choices may not be as rational as neorealism assumes. To explain the second crisis, we need to analyze the agency of the North Korean state. CPT's agency-centered approach provides a different perspective on state behavior.

Kim Jong Il

CPT operates under the assumption that the state and its actions reflect the identity and preferences of its leader, therefore one must operate as if Kim Jong Il and the state of North Korea are one and the same. North Korea has effectively been a dictatorship since its inception. This continued through the first nuclear crisis up till the death of Kim Il Sung in 1994. Kim Jong Il came to power having been groomed for the position by his father and quickly consolidated his control of North Korea (He, 2013). Based on the precedent his father set, Kim enjoyed absolute control over the state and its functions (He, 2013). It is important to note that Kim Jong Il maintained a highly centralized regime, conferring validity to the assumption that Kim Jong Il effectively was the state of North Korea.

The next step in testing the CPT is analyzing the identity of Kim Jong Il and, by extension, North Korea. Kim Jong Il's National Identity Conception (NIC) aids in understanding what factors were involved in the formation of his nuclear preferences. An NIC is a state leader's "basic

sense of what the nation naturally stands for and of how high it naturally stands in comparison to others in the international arena" (Hymans, 2008). Kim Jong Il can be characterized an oppositional nationalist NIC (Hymans, 2008). Oppositional nationalists form their identity in contrast to other key actors, and believe their nation's core values and interests should reflect this as well. This is combined with the belief that their nation should proudly deal with other nations. The *oppositional* component of their identity gives rise to security-seeking measures as a way of increasing security and the emotionally stabilizing the leader. The *national* component of the identity is associated with pride, and this gives rise to the seeking of a symbol for the state's unlimited power and potential. This drives oppositional nationalists to greatly desire nuclear weapons (Hymans, 2008). Kim is clearly an oppositional nationalist based on Hymans's research (2008) which uses historical events, measures of foreign influence, and content analysis of Kim's rhetoric to ascertain Kim's identity. The experience of the Korean War helped fortify a fear of other states and the belief that North Korea could overcome any obstacle, and his father and other military leaders passed down this attitude to Kim Jong Il while he was groomed for leadership. The continuance of Korean-first governmental policy that emphasized the state and military over all other things is an example of both a fear- and pride-driven identity in which security and military strength were valued. From a foreign influence standpoint, North Korea went to great lengths under Kim to impress the idea that North Korea is alone against the world. Kim revised history to dismiss almost all cases of Korea being reliant on other countries, while glorifying the history of the peninsula as a dominant kingdom (Hymans, 2008). Critically, Hymans (2008) content analyzed Kim Jong Il's speeches and written addresses to identify to which states he was in opposition and how he conceived his own nation. Hymans demonstrated that Kim saw North Korea as ideologically against the entire world, with only South Korea as a misguided, exploited ally. States perceived as imperialist were key comparison others and the United States and Japan represent the most mentioned states for this

category. In sum, North Korea sees itself, with South Korea, as one country and this Korea stands against an undifferentiated foreign world. This confirmation of Kim's oppositional nationalist identity is coupled with the finding that Kim Jong Il greatly sought nuclear weapons because they provide a security measure for the fearful component in Kim and also a trophy of power and respect for Kim's pride (Hymans, 2008).

Establishing a Reference Point

Kim's pride- and fear-based identity, along with his desire for nuclear weapons, allow us to establish a reference point of decision-making. Kim aspired to have nuclear weapons based on the history of nuclear development begun by his father, knew how to pursue nuclear weapons from his father's example, and actively compared North Korea to other regional states with nuclear weapons like the United States. The reference point for Kim is therefore quite simply the acquisition of nuclear weapons through aggressive nuclear policy. This reference point is the status quo for nuclear behavior for Kim as it validates his NIC, and is used to give value to choices that revolve around this reference point.

Transition 1: Passive to Aggressive (2002-2003)

From 1994 to 2002 North Korea operated under a passive nuclear policy. It was consistent with its private nuclear activity, non-threatening rhetoric, and cooperation with institutional arrangements. Prospect theory predicts that actors are risk-averse and will not take actions that could potentially damage their assets. This leads to the second hypothesis's prediction of passive behavior occurring when moving away from aggressive nuclear behavior is framed in a domain of gains (e.g. foreign aid) with no perceived risks. The Agreed Framework was conceived in a way that ensured no risk to North Korea's nuclear assets by cooperating. This institution lacked any mechanism for self-enforcement, leaving an easy way out for North Korea to defect. The institutional problems in the Agreed Framework ensured that North Korea could dictate its cooperation in the agreement, actively stalling or avoiding nuclear inspections while never actually

harming its nuclear facilities (He, 2013). The nuclear process was merely on hold, not suffering losses, and North Korea received substantial economic gains from this choice.

The perception of this passive nuclear behavior choice changed as the United States began to apply more pressure to North Korea on their nuclear disarmament. Entering the 21st century, the United States began making more demands of North Korea to speed up its denuclearization and the United States failed to deliver aid which had been agreed upon (Kang, 2004). The United States was dealing with domestic shifts as it transitioned from the Clinton to the Bush administration. The Bush administration took a much harder stance on foreign policy, particularly in the wake of the terrorist attack of September 11th, 2001. The United States gradually became less interested in cooperation with a state that had a history of non-compliance (Pinkston, 2009). In 2001, the United States requested improved inspection procedures and changes in North Korean military before giving aid (Kang, 2004). Later that year a senior administrative official stated: "we need to see some progress in all areas. We don't feel any urgency to provide goodies for them" (Kang, 2004). In 2002, Secretary of State Colin Powell demanded an additional condition to the Agreed Framework that stipulated a reduction in the North Korea's missile programs as well as an increase in denuclearization transparency (Kang, 2004). In what was the most famous indication of the United States' incentive to defect, President Bush announced that North Korea was among the nations considered the "axis of evil" in his State of the Union Address in 2002 (Kang, 2004). To Kim this led to a new evaluation phase in which the costs of cooperation increased substantially, and the probability of having his nuclear assets damaged became higher. North Korea was unsure of whether meeting the conditions of the Agreed Framework would increase its national security. The political history of warfare and tension fostered a deep mistrust of the United States (Kang, 2004). In addition, North Korea was unsure of whether the agreed benefits would be delivered to them if they complied. The United States could cease aid at any moment and even impose

sanctions or use military force. Over time the agreement had too many problems. The United States did not deliver on many of its conditions while being late on others, and these uncertainties became significant. North Korea had to make a choice upon reaching a turning point with the United States in 2002, when they were openly accused of violating the agreement. North Korea could continue passive nuclear behavior and have its nuclear assets damaged, or choose the status quo of aggressive nuclear policy framed in a domain of gains. It has been demonstrated that actors are risk-averse and overvalue losses as opposed to gains. Consistent with this and the third hypothesis, Kim decided to switch from passive to aggressive behavior and resumed the status quo. North Korea subsequently withdrew from the NPT, stopped all cooperative measures relevant to its nuclear activity, and publically revealed its nuclear efforts by reopening their nuclear facilities capped by a successful nuclear detonation test in 2006.

Transition 2: Aggressive to Passive (2003-2007)

Kim had two choices regarding his nuclear behavior early in this period: return to the negotiation table and pursue passive nuclear behavior, or pursue aggressive nuclear behavior. The former choice was framed in a domain of losses, as it would cost North Korea to return to the negotiation table and risk giving up their nuclear assets. Meanwhile, pursuing aggressive, status quo behavior satisfied the pride and fear-based identity Kim is associated with, and this choice was framed in a domain of gains. This exemplifies the *status quo bias*, where moving away from the status quo is treated as a loss and staying at the status quo is treated as a gain. Since actors are loss-averse and risk-averse, actors tend to stay with the status quo. This makes aggressive nuclear behavior the standard behavior during the second nuclear crisis, consistent with hypothesis 1. As the second hypothesis suggests, North Korea would only pursue passive nuclear behavior when it would be framed within a domain of gains. This indicates that when North Korea entered passive nuclear behavior by agreeing to the Six-Party Talks in 2007, there was a change in the frame and

a new editing and evaluation phase occurred. This arrangement was signed officially in 2005, but did not take effect until 2007, and had the proscribed goal of denuclearizing North Korea in exchange for negative security assurances from the United States, promises of economic assistance, and an improvement in relations with Tokyo and Washington (Pinkston, 2009). The arrangement consists of three phases: (1) a shutdown or “freeze” of all nuclear programs and facilities, (2) disablement, and (3) dismantlement. Each phase had procedures and deadlines for completion, and upon the completion of a phase talks between the six member states would resume in order to renegotiate the agreement if necessary (Pinkston, 2009).

The way in which the cooperative choice was perceived by Kim changed based on the nature of the institution that was presented. This can explain why North Korea suddenly stopped its aggressive behavior soon after testing its first nuclear weapon successfully, a landmark for North Korean security and something neorealism has trouble explaining. This appears to be irrational, but Kim saw the choice through the lens of his reference point. Better institutional benefits, multiple exit opportunities in the proposed phases of the negotiations, and the inclusion of allies China and Russia in the Six-Party Talks situated the agreement in a domain of gains with no risk. Kim would not have to damage his nuclear assets, merely freeze them, while experiencing gains of economic, political, and humanitarian aid. This change in frame clearly changed the way Kim evaluated the probabilities and risks of cooperation, causing Kim to sign the Six-Party Talks and freeze his nuclear program in 2007.

Transition 3: Passive to Aggressive (2007-2009)

This period of the crisis is similar to the first transition, where Kim experienced another change in frame that altered how he valued his choices and assets. North Korea exhibited passive nuclear behavior by completing the first denuclearization phase while receiving its promised benefits (Pinkston, 2009). As the first phase of the Six-Party Talks came to an end, the second phase of disablement meant the potential for real damage to

Kim's nuclear assets. Kim's willing acceptance of risk to avoid damages to his assets is an example of the *endowment effect*. Kim valued protecting his already acquired nuclear assets much more than gaining the institutional benefits of the second phase of the Six-Party Talks, which included the promise of a valuable light-water reactor (Pinkston, 2009). Kim perceived passive nuclear choice to be risky, and status quo aggressive behavior to not be risky. Consequently, North Korea suddenly withdrew from the Six-Party Talks, stopped cooperative measures heading into the second phase, and resumed the public pursuit of nuclear weapons which culminated in the second successful nuclear weapons test in 2009 (He, 2013). Kim's oppositional nationalist identity still judged aggressive nuclear behavior to be the status quo, in which non-cooperative behavior was framed as a gain and passive behavior as a loss. CPT allows for the explanation of the timing of the transitions, as well as the variance in behavior, much more effectively than the structural realism approach by focusing on preference formation and decision-making processes. By addressing the agency of the North Korean nuclear behavior, patterns become explainable when considered relative to a reference point.

Conclusion

Neorealism has indicated that the power structure of states is important, and that specifically North Korea's geopolitical history has shown its tendency to take its neighbors' power into consideration when acting. While this approach demonstrates the role of security in North Korea's nuclear decision-making, it is weak in its ability to explain the variance and timing of North Korea's decisions. Furthermore, this model is not fully able to explain the events of the crisis without acknowledging the role of agency and preferences. The constructivist prospect theory model addresses these issues by studying the origins of North Korea's preferences under Kim Jong Il. By looking at Kim's identity, it has been shown why North Korea had different preferences than the ones expected as per the neorealist model. Based on this understanding of Kim's identity and preferences, how Kim perceived North Korea's

nuclear assets, choices, and consequences during the crisis were surveyed. Knowing Kim's reference point provides a convincing explanation for why Kim chose such erratic, variant, and sometimes irrational behavior at specific instances during the crisis. Overall, the hypotheses of CPT were supported.

This has major implications for research in the future. From a theoretical standpoint, it is important to keep fleshing out the ideas, assumptions, and testability of all approaches. The neorealist model needs to consider the possibility of varying preferences in state decision-making to enhance its explanatory power. CPT should also be developed further as a tool for studying state behavior, particularly in cases where a unitary actor assumption can be made. The study of identity should be more prominent and more effectively operationalized so it can be generalizable to more cases. Likewise, the operationalization of prospect theory at the international level should be developed more thoroughly. A general framework and procedure for measuring reference points, changes in frames, and evaluating choices will be vital for making this theory more applicable.

North Korea is an important state because of the problems it creates for regional security. Its nuclear behavior, in particular, has been hard to predict or explain. CPT may provide a rich source of explanation and prediction for future behavior. For instance, by knowing how Kim Jong Un is likely to form his preferences and evaluate choices, one can reasonably predict how he will respond to future institutional agreements. North Korea's back and forth shifting from aggressive behavior to passive behavior from 2002 to 2009 represents the research puzzle of this paper. The evidence suggests that Kim Jong Il's identity, preference formation, and choice-making process were mainly responsible for this variance. CPT most convincingly explains North Korea's nuclear behavior during the second nuclear crisis, and is a strong candidate for further research and development.

Bibliography

- Allison, Graham T., and Philip Zelikow. 1999. *Essence of Decision: Explaining the Cuban Missile Crisis*. Vol. 2. New York: Longman.
- Axelrod, Robert, and Robert O. Keohane. 1985. "Achieving cooperation under anarchy: Strategies and institutions." *World Politics* 38.01: 226-254.
- Boettcher, William A. 2004. "The prospects for prospect theory: An empirical evaluation of international relations applications of framing and loss aversion." *Political Psychology* 25.3: 331-362.
- Cha, Victor D. 2000. "Abandonment, entrapment, and neoclassical realism in Asia: the United States, Japan, and Korea." *International Studies Quarterly* 44.2: 261-291.
- Cha, Victor D., and David Chan-oong Kang. 2003. *Nuclear North Korea: A Debate on Engagement Strategies*. Columbia University Press.
- Cha, Victor D., & David Chan-oong Kang. 2004. "The debate over North Korea." *Political Science Quarterly* 119.2: 229-254.
- Checkel, Jeffrey T. 1998. "The constructive turn in international relations theory." *World Politics* 50.02: 324-348.
- Finnemore, Martha, and Kathryn Sikkink. 1998. "International norm dynamics and political change." *International Organization* 52.04: 887-917.
- Garst, Daniel. 1989. "Thucydides and neorealism." *International Studies Quarterly* 3-27.
- Goo, Y., & Kim, S. 2009. "A study on the military alliance of South Korea-United States with the existence of threat from North Korea: A public good demand approach." *Empirical Economics* 36.3: 597-610.
- Goo, Y., & Lee, S. 2014. "Military alliances and reality of regional integration: Japan, South Korea, the US vs. China, North Korea." *Journal of Economic Integration* 29.2: 329-342.
- Haas, Mark L. 2001. "Prospect Theory and the Cuban missile crisis." *International Studies Quarterly* 45.2: 241-270.
- Haggard, Stephan, and Beth A. Simmons. "Theories of international regimes." *International Organization* 41.03 (1987): 491-517.
- He, Kai, and Huiyun Feng. 2013. *Prospect Theory and Foreign Policy Analysis in the Asia Pacific: Rational Leaders and Risky Behavior*. Routledge.
- Hymans, Jacques. 2008. "Assessing North Korean nuclear intentions and capacities: A new approach." *Journal of East Asian Studies* 8.2: 259-292.
- Jervis, Robert. 1978. "Cooperation under the security dilemma." *World Politics* 30.02:167-214.
- Jervis, Robert. 1998. *System Effects: Complexity in Political and Social Life*. Princeton University Press.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica: Journal of the Econometric Society* 263-291.
- Kahneman, Daniel, and Amos Tversky, eds. 2000. *Choices, Values, and Frames*. Cambridge University Press.
- Keohane, Robert O. 2005. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton University Press.
- Keohane, Robert O., and Joseph S. Nye. 1977. *Power and Interdependence: World Politics in Transition*. 2nd ed. Boston: Little, Brown.
- Koremenos, B., Lipson, C., and Snidal, D. 2001. "The rational design of international

institutions.” *International Organization*, 55.4: 761-799.

Krasner, Stephen D., ed. 1983. *International Regimes*. Cornell University Press.

Krasner, Stephen D. 1982 “Structural causes and regime consequences: regimes as intervening variables.” *International Organization* 36.02: 185-205.

Levy, Jack S. 1992. “Prospect theory and international relations: Theoretical applications and analytical problems.” *Political Psychology* 283-310.

Levy, Jack S. 1997. “Prospect theory, rational choice, and international relations.” *International Studies Quarterly* 41.1: 87-112.

Martin, Lisa L. 1993. *Coercive Cooperation: Explaining Multilateral Economic Sanctions*. Princeton University Press.

McDermott, Rose. 2001. *Risk-taking in International Politics: Prospect Theory in American Foreign Policy*. University of Michigan Press.

Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics*. WW Norton & Company.

Morganthau, Hans J. *Politics Among Nations: The Struggle for Power and Peace*. AA Knopf, 1960.

Moravcsik, Andrew. 1993. “Preferences and power in the European Community: a liberal intergovernmentalist approach.” *JCMS: Journal of Common Market Studies* 31.4: 473-524.

Moravcsik, Andrew. 1997. “Taking preferences seriously: A liberal theory of international politics.” *International Organization* 51.04: 513-553.

O’Neill, Barry. 2001. “Risk aversion in international relations theory.” *International Studies Quarterly* 45.4: 617-640.

Ruggie, John G. 1982. “International regimes, transactions, and change: embedded liberalism in the postwar economic order.” *International Organization* 36.02: 379-415.

Shimko, Keith I. 1992. “Realism, neorealism, and American liberalism.” *The Review of Politics* 54.02: 281-301.

Walt, Stephen M. 1987. *The Origins of Alliance*. Cornell University Press.

Waltz, Kenneth N. 2001. *Man, the State, and War: A Theoretical Analysis*. Columbia University Press.

Waltz, Kenneth N. 2010. *Theory of International Politics*. Waveland Press.

Wendt, Alexander. 1992. “Anarchy is what states make of it: the social construction of power politics.” *International Organization* 46.02: 391-425.

Wendt, Alexander. 1999. *Social Theory of International Politics*. Cambridge University Press.