Nuclear Proliferation, Inspections, and Ambiguity

Brett V. Benson
Vanderbilt University

Quan Wen
Vanderbilt University

May 2012

Abstract

This paper studies nuclear armament and disarmament strategies with and without a verification mechanism. We compare two models of nuclear development. The first model analyzes a government's development and disarmament decisions under "ambiguity," where the absence of external verification makes it possible for states to develop nuclear weapons secretly. The second is an "inspections" model, in which a government's arming decisions are verifiable. Comparative statics show that deterrence by doubt is Pareto optimal under limited conditions but ambiguity also leads to arming and conflict in other circumstances. In most states of the world, inspections are more likely to result in peace and non-proliferation. Additionally, there are not any extortion benefits of ambiguity that do not also exist with inspections. In fact, a counter proliferator is more willing to offer transfer payments with inspections, implying that governments might be willing to pay countries to join an inspections regime like the NPT.

Nuclear Proliferation, Inspections, and Ambiguity

Abstract

This paper studies nuclear armament and disarmament strategies with and without a verification mechanism. We compare two models of nuclear development. The first model analyzes a government's development and disarmament decisions under "ambiguity," where the absence of external verification makes it possible for states to develop nuclear weapons secretly. The second is an "inspections" model, in which a government's arming decisions are verifiable. Comparative statics show that deterrence by doubt is Pareto optimal under limited conditions but ambiguity also leads to arming and conflict in other circumstances. In most states of the world, inspections are more likely to result in peace and non-proliferation. Additionally, there are not any extortion benefits of ambiguity that do not also exist with inspections. In fact, a counter proliferator is more willing to offer transfer payments with inspections, implying that governments might be willing to pay countries to join an inspections regime like the NPT.

Introduction

Why do countries join the nuclear non-proliferation treaty (NPT) and subject themselves to weapons inspections? There are some advantages to remaining outside the NPT and leaving the status of one's nuclear weapons program ambiguous. One such advantage is that a government might achieve "deterrence by doubt,"[1] which is induced by uncertainty about whether a government possesses nuclear weapons. Another possible upside is "extortion by ambiguity," which is how we describe a government's effort to leverages uncertainty about its nuclear program to extract concessions from countries that would be willing to pay to prevent nuclear arming (see Benson and Wen 2011).

In spite of these two advantages of ambiguity, only four countries -- India, Israel, Pakistan, and North Korea -- are currently not signatories to the NPT. Since the NPT came into force in 1970, 190 countries have signed on to the inspections and safeguards agreement in the treaty. Overwhelmingly, countries have opted for inspections over ambiguity. Why would a government give up the deterrence and bargaining benefits of ambiguity and subject themselves to weapons inspections? Additionally, the five major powers in the NPT -- the US, UK, France, China, and Russia -- consistently encourage other countries to join the NPT. If ambiguity helps facilitate deterrence without weapons development, then why would a counter proliferator, who is primarily interested in preventing nuclear armament, push a nuclear aspirant to subject itself to weapons inspections if so doing might have the unintended effect of leading the aspirant to arm to gain deterrence?

[1] "Deterrence by doubt" is a phrase coined by General Hamdani, Iraqi Republican Guard commander, to describe Saddam Hussein's goal of being ambiguous about its weapons of mass destruction. For the reference, see Michael R. Gordon and Bernard E. Trainer, "Even as US Invaded, Hussein Saw Iraqi Unrest as Top Threat," *New York Times*, March 12, 2006.

While there may be benefits to using nuclear ambiguity to enhance one's security or bargaining payments, the fact that most countries are party to the NPT indicates that there may be advantages to inspections that outweigh the benefits of ambiguity for nuclear aspirants. And that counter proliferators in practice rarely encourage countries to remain ambiguous suggests the Pareto improving qualities of ambiguity may be rare. This paper analyzes the different incentives for nuclear aspirants and counter proliferators interacting under both "ambiguity," when counter proliferators are unable to observe whether they have armed or not, and "inspections," when counter proliferators can verify if an aspirant has armed.

In addition to examining the incentives of the aspirant and counter proliferators to determine whether we can reconcile the differences between the theory and observable trends in counter proliferations, we also investigate an equilibrium puzzle of ambiguity. When ambiguity generates deterrence by doubt and extortion by ambiguity, then it may benefit the aspirant to take advantage of these advantages and not pay the costs of arming. Given this incentive, it may not make sense for any other government to believe the aspirant has armed or will arm. If countries do not think the aspirant has an incentive to develop nuclear weapons, then the deterrence and bargaining benefits of ambiguity disappear. Thus, in our analysis, we examine whether deterrence by doubt and extortion by ambiguity are consistent with equilibrium when the counter proliferators may do nothing, attack, or offer a transfer to the aspirant. Further, we make welfare comparisons between a game of ambiguity and a game of inspections to determine whether the deterrence and bargaining benefits of ambiguity hold when players may make an ex ante choice to enter an inspections regime.

To facilitate these comparisons, we develop two models: an ambiguity game and an inspections game. Then we examine aspirants' decisions to arm in both games. Our approach builds on existing models of nuclear proliferation and counter proliferation. Sobel (1992) develops a model of strategic ambiguity in which countries are endowed with military capabilities and uncertainty about those capabilities can deter attacks. Feaver and Niou (1996) endogenize the decision to arm and evaluate the conditions under which a counter proliferator will attack or assist the aspirant to acquire weapons. While they consider a broader array of responses by the counter proliferation than just the standard attack option, they do not model ambiguity. Baliga and Sjostrom (2008) also endogenize military capabilities, but they also allow the aspirant to choose inspections, which leads to inferences about the conditions under which aspirants will prefer ambiguity over inspections and whether there is a deterrence benefit to ambiguity. In their account, ambiguity serves as a substitute for costly arming. Countries that might otherwise be interested in using force against a target may be deterred from doing so if there is a chance that the target possesses nuclear weapons. The benefit of deterrence reduces pressure on the target to arm and enables them to receive a cost savings by not arming. However, if countries are required to verify that they are not armed, they lose deterrence by doubt and will pay the costs of arming to regain deterrence.

In our approach, nuclear aspirants may pay to acquire weapons under both ambiguity and inspections. Additionally, we allow the counter proliferator to offer a transfer to an aspirant in exchange for verifiable abandonment of its weapons program. The standard modeling approach for achieving disarmament or deterring weaponization only considers whether a counter proliferator might attack to deter weaponization or to forcibly disarm a nuclear government (Sobel 1992, Feaver and Niou 2006, and Baliga and Sjostrom 2008). There is also a literature that examines when governments might use attacks to target nuclear programs (Fuhrmann and Kreps 2010, Goldstein 2006, Feaver 1997). Additionally, scholars have studied the general question of preventative strikes (Levy 2008 and 1987; Reiter 1995). And, many studies have analyzed Israel's decision to attack Iraq's Osirak facility in 1981 (Raas and Long 2007, Reiter 2005) as well as the US decision not to attack North Korea (Levy 2008, Oberdorfer 1997, Sigal 1997) and China (Burr and Richelson 2000, Chang 1988).

These studies provide important insights into the use of force as a counter proliferation tool. However, since they do not also examine the use of bribes and bargaining, we do not have many comparisons of the counter proliferation tools actually used in practice. An exception is Benson and Wen (2011), which adds a bargaining option to the standard options of doing nothing and attacking the aspirant. Their model, however, is only an "ambiguity" model; they do not consider an environment in which it is possible for a counter proliferator to verify the aspirant's arming decision. They make an argument for extortion by ambiguity. In their model, there are parameter values such that ambiguity can lead the counter proliferator to make an offer to the aspirant. The logic for this argument is that uncertainty about the aspirant's nuclear program raises the risk to the counter proliferator enough that it would be willing to pay the aspirant to verifiably abandon its weapons program, but the risk is not so high that it is cost efficient for the counter proliferator to attack the aspirant. However, unless we may compare this result to an inspections setting, it is not clear whether the concessions gained through bargaining under ambiguity are clearly better than what an aspirant may gain through inspections.

Our approach here builds on Benson and Wen (2011). We develop an inspections model so that we can make welfare comparisons between ambiguity and inspections. The key difference between these two models is that the counter proliferator faces two different informational problems in each game. In the inspections game, it is uncertain about how motivated a prospective nuclear power is to acquire nuclear weapons but it can observe whether the aspirant arms or not. How worried or threatened the counter proliferator is depends on whether the aspirant is actually motivated to acquire weapons, which is private information. The inability to distinguish a "motivated" nuclear aspirant from a "normal" type complicates counter proliferation strategies. In the ambiguity model, the counter proliferator is uncertain about the type of aspirant it faces, but it is also unable to observe the aspirant's arming behavior. The inability to observe the aspirant's arming behavior applies when the counter proliferator does not have a mechanism, such as weapons inspections, for verifying the aspirant's arming decisions.

In making comparisons across these two games, we show that deterrence by doubt is Pareto optimal under a limited set of conditions but that nuclear ambiguity also leads to arming and even conflict in other circumstances. Further, in most states of the world, inspections are much less likely to result in conflict and more likely to result in non-proliferation. Bargaining and proliferation is beginning to receive some attention in the empirical literature. We also show that under some conditions, providing security assurances to prospective proliferators can remove their motivation for developing nuclear weapons. There are types of nuclear aspirants who will only arm if they feel threatened. Thus, we provide some theoretical justification for the findings by Bleek and Lorber (this issue) that providing a security guarantee can reduce the chances that governments will develop nuclear weapons.

We also show that there are not any extortion benefits of ambiguity that do not exist with inspections. In fact, the counter proliferator may be more willing to make transfer payments to a government under inspections because the ability to verify arming decisions decreases the chances that a counter proliferator is wasting its payment. An implication of our finding is that the counter proliferator might be willing to pay countries to induce them to join an inspections regime like the NPT. This result is consistent with and may help explain the empirical findings in

Brown and Kaplow (this issue) that countries are attracted to Technical Cooperation (TC) assistance offered to countries in the grand bargain for joining the NPT.

A Model of Ambiguous Nuclear Weapons Development

Suppose there are two governments, A, a more powerful *counter proliferator*, and B, a potential *aspirant*, that interact in a bargaining game. Government A is uncertain about how much government B values possessing nuclear weapons. Accordingly, government B has two possible types: motivated type Z or normal type N. A government is a motivated type if it is strongly motivated to acquire nuclear weapons for whatever reason. Government B has a prior probability p of being a motivated type and $1 - p$ of being a normal type. A motivated type derives a benefit from acquiring nuclear weapons while a normal type receives no benefit if it arms. Government A is worried or receives a threat cost w if a motivated type acquires nuclear weapons, but it is not threatened if a normal type arms.

We will first present the equilibrium conditions and players' equilibrium payoffs for both the ambiguity and inspections model. Since the ambiguity model can be found in Benson and Wen (2011), we summarize it here and present the full description of the model in a web appendix. Then we provide the full analysis of the inspections model. This approach will enable us to make comparisons for identical parameter values between the two models.

[Figure 1: Ambiguity model here]

In the ambiguity model, there is ambiguity about arming because A is unable to observe B's decision to arm. Figure 1 depicts the extensive form model. At the beginning of the game, B is not armed with nuclear weapons and nature determines its type. Government B learns its type, but A does not. Government B then decides whether to arm or not. The cost of arming is κ regardless of B's type. Unlike Baliga and Sjostrom (2008), we assume that B will successfully acquire the weapon if it chooses to arm. If B does not arm, it does not pay the development cost κ

and does not acquire nuclear weapons. If B is a motivated type, then it receives a benefit δ from developing and acquiring weapons, and A suffers a worry or threat cost w. On the other hand, if B is a normal type, then it does not benefit from arming. After B has made its arming decision, A decides how to respond. A does not know B's type and A is unable to observe whether B armed or not. The ability to observe B's arming decision is the key distinction between the ambiguity model and inspections model, which we will consider later. In the inspections model, government A can observe B's decision to arm. In the ambiguity model, A does not know if B has armed unless A attacks B and discovers in a war that B is armed or A makes an offer to B and the payment is accepted in exchange for verifiable disarmament.

In the ambiguity game, when it is A's turn to move, it chooses whether to attack B or to make an offer to B in exchange for giving up its weapons. If A attacks, the game ends in war. We assume that A will win a war against B, but attacking is costly. Government A pays c_A if it attacks an armed B and pays c_N if it attacks an unarmed B. Attacking an armed opponent is more costly, so it is reasonable to assume that $c_A > c_N > 0$. Government A's payoff for attacking an armed B is $1 - c_A$, and its payoff for attacking an unarmed B is $1 - c_N$. If B is attacked, possessing weapons is better than being unarmed. This benefit is captured by $\gamma > 0$, and B's cost of being attacked is $\beta > 0$. Assume that $\gamma > \kappa$, which implies that in war the benefits of having advanced weapons outweighs the costs of developing them. Thus, if B is attacked when armed, it receives $\gamma - \beta - \kappa$, otherwise receives $-\beta$ if B is attacked when unarmed.

Instead of attacking, A can offer $x \in [0, 1]$ to B in exchange for B giving up its weapons. Government B can either accept or reject such as offer. If B accepts, A receives a payoff of $1 - x$. Government B's payoff for accepting an offer depend on whether it is armed. It receives a payment of x if it is not armed and $x - \kappa$ if it is armed. Payoffs for rejecting also depend on B's type and whether B is armed or not. If it arms and rejects an offer, *B* keeps possession of its weapons and receives the value for having nuclear weapons. Type *Z*'s value possessing nuclear weapons if it rejects an offer is δ, but type *N* receives no benefit from arming and keeping nuclear weapon. On the other hand, if *B* does not arm and it rejects *A*'s offer, then it receives nothing regardless of its type. The higher the value of δ the more *Z* values having the nuclear weapons for whatever reason. We do not make distinctions about the reasons why *B* might wish to acquire nuclear weapons, only that the intensity of the motivation may differ by degree.[2]

Government *A* suffers a cost $w > 0$ if *Z* acquires nuclear weapons. This cost represents *A*'s level of worry or the degree of threat it experiences from an armed *B*. Government *A* might worry about the threat to itself and its allies posed by other nuclear powers, the possibility that weapons or weapons technology will be transmitted to terrorists, and the political implications from a rival possessing greater power and influence due to its nuclear weapons. Naturally, the possession of nuclear weapons by some governments is more threatening than others. For example, from the US perspective, Iran acquiring nuclear weapons today presents a greater threat or level of worry than did India's nuclear program in the 1990s. To account for such differences, we allow w to vary and inspect different values of w for which *A* might find different counter-proliferation strategies optimal. The worry cost accrues only if a motivated type *B* arms and keeps weapons. Thus, government *A*'s payoff if *B* arms and rejects an offer is $1 - w$, and *A* receives a payoff of 1 if *B* does not arm and rejects an offer.

[2] There is a sizable literature that explores governments' motivations for pursuing nuclear weapons. Since this question is not central to our analysis, we do not make distinctions along these lines in our models. For our purposes, it is sufficient to distinguish between two broad motivations for developing nuclear weapons: the value of possessing nuclear weapons versus the value of using weapons in a war. These dimensions are captured by the parameters δ and γ respectively. Our approach for assigning payoffs to government *B* follows Baliga and Sjostrom (2008). For studies documenting the many different motivations leaders have for developing nuclear weapons, see Sagan (1996/97), Solingen (2007), Singh and Way (2004), Hymans (2006), Jo and Gartzke (2007), Kroenig (2009), Fuhrman, 2009.

To facilitate the analysis, we make some assumptions about the parameters in the models. They apply to both the ambiguity and inspections games. First, we assume $c_A > \kappa > c_N > 0$. This assumption states that it is more costly for *A* to attack an armed than an unarmed *B*. It also means that a war with an aspirant who possesses nuclear weapons costs more than it does to develop weapons. To get some traction on the relationship between a powerful counter proliferator and today's problem of nuclear weapons development by relatively small countries, we assume that the cost of weapons development for *B* is greater than the costs to *A* of attacking an unarmed government *B*.

We also assume $\delta > \kappa > 0$. This assumption states that it is cost-effective for a type *Z* aspirant to develop weapons. If it does not pay for *Z* to arm, then *B* would never arm and *A* would never attack or offer anything to *B*.

Assume that *A* always feels some amount of threat $w > 0$ if type *Z* arms, and $\gamma > \kappa > 0$. This latter assumption states that *B*'s value for possessing and using technology acquired from arming is greater than the costs of developing weapons for both types of government *B*. This assumption guarantees that a government that expects to be attacked derives some value from having advanced weapons capability even if it does not derive intrinsic value from acquiring nuclear weapons when there is no threat it will be attacked.

We solve both models for perfect Bayesian equilibrium. In the ambiguity model, there are several equilibria in both pure and mixed strategies. We summarize the equilibria that are interesting for our analysis. There are two pure strategy equilibria, which we refer to as WAR and WMD, and three mixed strategy equilibria, which we call MIXED ATTACK, MIXED ATTACK/OFFER, and MIXED OFFER. These equilibria are depicted in Figure 2 in terms of δ and w.

[Figure 2: Ambiguity equilibria here]

In the WAR equilibrium, government A attacks and both types of government B arm when $c_A \leq \min\{pw, \delta\}$, *i.e.*,

$$w \geq \frac{c_A}{p} \; and \; \delta \geq c_A.$$

Since both types of government A behave the same way, there is only one possible outcome, which is war. In this equilibrium, payoffs are as follows: $U_A = 1 - c_A$ and $U_Z = U_N = \gamma - \kappa - \beta$.

In the WMD equilibrium, the motivated type, Z, arms while the normal type, N, does not, and A does not attack but does not offer anything when

$$w \leq \frac{\tilde{c}}{p} \; and \; \delta \geq pw,$$

where $\tilde{c} = pc_A + (1-p)c_N$. There are two possible outcomes: either the game ends with peaceful non-development of weapons or Z develops nuclear weapons peacefully. Players receive the following payoffs: $U_A = 1 - pw$, $U_Z = \delta - \kappa > 0$, and $U_N = 0$.

There is a MIXED ATTACK equilibrium if

$$\delta \geq pw \quad and \quad \frac{\tilde{c}}{p} \leq w \leq \frac{c_A}{p},$$

such that A attacks with probability $q_A \in [0,1]$, offers $x = 0$ with probability $q_0 = 1 - q_A$, and offers $x = \delta$ with probability $q_1 = 0$; type Z arms with certainty and type N arms with probability π_N and does not arm with probability $1 - \pi_N$ where

$$q_A = \frac{\kappa}{\gamma} \quad and \quad \pi_N = \frac{(c_A - pw)}{(1-p)(c_A - c_N)}.$$

The possible outcomes in this equilibrium include war against an armed B (where either type Z or type N arms), war against an unarmed B (where type N does not arm), peaceful arming (where either type Z or N arm and A does not attack), or peaceful non-arming (where type N does not arm and A does not attack). In this equilibrium, players' payoffs are as follows:

$$U_A = 1 - pw,\ U_Z = q_A(-\beta) + (1 - q_A)\delta, \text{ and } U_N = q_A(-\beta).$$

In the MIXED ATTACK/OFFER equilibrium, A attacks with probability $q_A \in [0, 1]$, offers $x = \delta$ with probability $q_1 = 1 - q_A$, and offers $x = 0$ with probability $q_0 = 0$; and type N arms with probability π_N and type Z arms with probability π_Z where

$$q_A = \frac{\kappa}{\gamma} \quad and \quad \Pi = p\pi_Z + (1 - p)\pi_N = \frac{\delta - c_N}{c_A - c_N}.$$

The equilibrium conditions under which this equilibrium hold are

$$\delta \leq c_A, \delta \leq pw, and\ \delta \geq c_N \text{ (assumed)},$$

The possible outcomes that might obtain under this equilibrium include war, military strikes against an unarmed government B, transfer payments made to an armed B, and payments to an unarmed B. Players' equilibrium payoffs are as follows:

$$U_A = 1 - \delta \text{ and } U_Z = U_N = q_A(-\beta) + (1 - q_A)\delta.$$

Finally, there is a MIXED OFFER equilibrium if

$$\delta \leq pw, \delta \leq \frac{c_N w}{w - c_A + c_N}, and\ \delta \geq \kappa\ (assumed),$$

such that A offers $x = \delta$ with probability $q_1 \in [0,1]$, offers $x = 0$ with probability $q_0 = 1 - q_1$, and attacks with probability $q_A = 0$; N does not arm with certainty and Z arms with probability π_Z and does not arm with probability $1 - \pi_Z$ where

$$q_1 = \frac{\delta - \kappa}{\delta} \;\; and \; \pi_Z = \frac{\delta}{pw} > 0.$$

In this mixed strategy equilibrium, war never occurs, since *A* never attacks. The possible outcomes include *A* making a bribe payment to disarm *B*, *A* making a payment to *B* even though it does not arm, *A* offering nothing and *B* peacefully arming, and *A* offering nothing and *B* not arming. Players' payoffs are $U_A = 1 - \delta$, and $U_Z = U_N = \delta - \kappa$. The web appendix fully describes these equilibria of the ambiguity game.

Inspections Model

We now consider an "inspections" model where government *B*'s decision to arm is observable to government *A*. Government *A* still does not does not observe *B*'s type, but *A* knows whether *B* has chose to arm. This implies that *B*'s decision to arm can be a signal about its type. Accordingly, there are two information sets for A's decision, one in which *B* is armed and *B* is not armed. Figure 3 depicts the extensive form of the inspections model.

[Figure 3: Inspections extensive form here]

There are two cases to consider. In one case, $\gamma - \beta > \kappa$, and in the other $\gamma - \beta < \kappa$. The first case denotes the situation in which *B*'s benefit of using a weapon in war minus its costs of fighting exceeds the costs of developing the weapon. From this perspective, an armed *B* derives a positive payoff from developing a weapon and fighting. In the second case, the benefits of fighting with a weapon do not justify the costs of developing and fighting.

We first consider the pure strategy equilibria. First note that if *B* is not armed, *A* will offer $x = 0$ and *B* will accept it. Hence, *B*'s continuation payoff by not arming itself is zero. Now, let us consider the two cases beginning with $\gamma - \beta > \kappa$.

1.1. Case 1: $\gamma - \beta > \kappa$

There are three pure strategy equilibria, which we identify as WAR, OFFER, and WMD. In the WAR equilibrium, government A attacks, and both types of government B arm. In the OFFER equilibrium, both types of B also arm but A offers them a payment of $x = \delta$ to give up their weapons. Finally, in the WMD equilibrium, A offers $x = 0$ while Z arms and N does not. This implies two possible outcomes depending on B's type. If B is a normal type N, then there is no arming and no offers or attacks. If B is a motivated type Z, then it arms, and A does nothing to prevent it from doing so. The formal conditions for these pure strategy equilibria are defined in the following proposition:

Proposition 1 *When* $\boldsymbol{\gamma} - \boldsymbol{\beta} > \boldsymbol{\kappa}$*, there are three pure strategy equilibria described as follows.*

1. (WAR) *Both types B arm and A attacks when* $c_A \leq \min\{pw, \delta\}$*, i.e.,*

$$w \geq \frac{c_A}{p} \text{ and } \delta \geq c_A. \quad (1)$$

2. (WMD) *The motivated type Z arms while the normal type N does not, and A does not attack but does not offer anything when*

$$w \leq c_A \text{ and } w \leq \delta, \quad (2)$$

3. (OFFER) *Both types B arm and A offers* $\delta \leq \min\{pw, c_A\}$*, i.e.*,

$$\delta \leq pw \text{ and } \delta \leq c_A. \quad (3)$$

[Figure 4: Case 1 equilibria here]

From these equilibria, which are presented in Figure 4, we note some similarities to the ambiguity game. The WAR equilibrium occurs under the same conditions in both games, and the payoffs for all players are the same. The logic for this equilibrium, therefore, is similar to the ambiguity game. The OFFER equilibrium in the inspections game occurs under the same conditions as the MIXED OFFER equilibrium in the ambiguity game. To see why A now makes an offer $x = \delta$ in pure strategies under inspections, consider its options given that both types of B arm in equilibrium. If A attacks, it would receive $1 - c_A$. Thus, as long as $\delta < c_A$, making the offer is better for A than attacking. Government A could also offer nothing. Since B arms, A would receive $1 - pw$, because A believes B is type Z with probability p at the information set where B arms. Therefore, government A will offer $x = \delta$ with certainty if $\delta < pw$. Now consider B's options given A offers $x = \delta$. Government B will choose not to arm in pure strategies when $\gamma - \kappa < 0$. But since $\gamma - \kappa > 0$ (by assumption), therefore, both types of B will arm when A attacks. The payoff for A is $U_A = 1 - \delta$, and the payoffs for B are $U_N = U_Z = \delta$.

No other pure strategy equilibrium in the inspections game obtains under the same conditions as the ambiguity game. The WMD equilibrium under inspections is different than the WMD equilibrium of the ambiguity game because the right boundary condition under inspections is $w \leq c_A$ instead of $w \leq \tilde{c}/p$ in the ambiguity game. The difference in this boundary condition results because A's ability to observe B's arming decision in the inspections game changes A's equilibrium payoffs. Under ambiguity, government A receives $1 - pw$ if it offers nothing when Z arms and N does not arm. Since A receives no information from B's arming decision, it receives $1 - \tilde{c}$ if A attacks. Under inspections, however, observing B's actions allows A to condition its strategy on B's action. Government A knows only Z will arm and only N will never arm in equilibrium. Since A prefers to offer $x = 0$ rather than attack an unarmed N, then either A offers $x = 0$ regardless of what B does, or it makes a conditional offer, attacking if it observes arming and offering $x = 0$ if it observes no arming. It will offer $x = 0$ no matter what it observes if and only if $w < c_A$.

Notice that the difference between the WMD equilibrium conditions under inspections versus ambiguity is that the right hand bound under inspections is less than the right hand bound under ambiguity. That is, $c_A < \frac{\tilde{c}}{p}$. Comparing inspections to ambiguity, this implies that, under inspections, *A*'s ability to fully resolve its uncertainty about *B*'s type as a result of *B*'s equilibrium strategy and observability of *B*'s actions leads *A* to be more willing to attack for lower values of *w*.

Now we will examine the mixed strategy equilibria in this game. There are two mixed strategy equilibria, which we call MIXED ATTACK and MIXED OFFER. In the MIXED ATTACK equilibrium, *A* mixes between attacking and offering $x = 0$, *Z* arms, and *N* mixes between arming and not arming. The possible outcomes are the same as the MIXED ATTACK equilibrium in the ambiguity game. In the MIXED OFFER equilibrium, *A* mixes between offering $x = 0$ and $x = \delta$, *Z* arms for sure, and *N* mixes between arming and not arming. The possible outcomes include *A* making a payment in exchange for disarmament, *A* making a payment when there is no arming, *Z* arming peacefully without any response from *A*, and no arming by anyone and no one receiving any payment. The formal conditions for the MIXED ATTACK equilibrium are as follows:

Proposition 2 *There is a mixed strategy equilibrium* (MIXED ATTACK) *if*

$$c_A \leq \delta \text{ and } c_A \leq w \leq \frac{c_A}{p}, \tag{4}$$

such that A attacks with probability $q_A \in [0, 1]$, *offers* $x = 0$ *with probability* $q_0 = 1 - q_A$, *and offers* $x = \delta$ *with probability* $q_1 = 0$; *Z arms with certainty and N arms with probability* π_N *and does not arm with probability* $1 - \pi_N$ *where*

$$q_A = \frac{\kappa}{\gamma - \beta} \quad and \quad \pi_N = \frac{p(w - c_A)}{(1-p)(c_A)}. \tag{5}$$

To see why this equilibrium obtains under these conditions, first notice that if A mixes between attacking and offering $x = 0$, Z will arm because $\gamma - \beta > \kappa$ and $\delta - \kappa > 0$. Government N will mix between arming and not arming if and only if

$$q_A(\gamma - \beta - \kappa) + (1 - q_A)(-\kappa) = 0,$$

which gives q_A in (13). For A to mix between attacking and offering $x = 0$, it must be that

$$\rho = \frac{c_A}{w},$$

where ρ is A's belief that B is a Z type given that it observes B arm. Thus, ρ is

$$\rho = \frac{p}{p + (1-p)\pi_N}.$$

Solving for the condition under which A will mix gives π_N in (12).

For A not to offer $x = \delta$, it must be the case that $\rho w = c_A \leq \delta$. Thus, $c_A \leq \delta$ defines the lower bound condition as seen in Figure 4. To establish the left and right boundaries, it must be that $\pi_N \in [0, 1]$. To ensure $\pi_N \geq 0$, it must be that $w \geq c_A$, and for $\pi_N \leq 1$, we need $pw \leq c_A$. Thus, we have the equilibrium conditions in (12). Players equilibrium payoffs are

$$U_A = 1 - pw, \; U_N = 0, \text{and } U_Z = \frac{(\gamma-\beta-\kappa)}{(\gamma-\beta)}\delta \in (0, \delta).$$

Now we define the conditions for the MIXED OFFER equilibrium.

Proposition 3 *There is a mixed strategy equilibrium* (MIXED OFFER) *if*

$$\delta < w < \frac{\delta}{p} \quad and \quad \kappa < \delta < c_A, \tag{6}$$

such that A offers x = δ with probability $q_1 \in [0, 1]$*, offers x = 0 with probability* $q_0 = 1 - q_1$*, and attacks with probability* $q_A = 0$*; Z arms with certainty and N arms with probability* π_N *and does not arm with probability* $1 - \pi_N$ *where*

$$q_1 = \frac{\kappa}{\delta} \quad and \quad \pi_N = \frac{p(w - \delta)}{(1 - p)\delta}. \tag{7}$$

To see the justification for this equilibrium, consider *Z*'s best response if *A* mixes between offering $x = 0$ and $x = \delta$. In this case, type *Z* will arm for sure because $\delta - \kappa > 0$. For type *N* to mix, it must be that

$$q_1(\delta - \kappa) + (1 - q_1)(-\kappa) = 0.$$

Solving for q_1 gives *A*'s equilibrium mixed strategy in (14). For *A* to mix given π_N we need $\rho w = \delta$. Solving this expression for π_N, we have *N*'s equilibrium mixed strategy in (14).

To establish the boundary conditions, it must be that *A* does not have an incentive to arm. Thus, $\rho w = \delta < c_A$. Additionally, for A's mixed strategy $\pi_N > 0$, it must be that $w > \delta$, and for $\pi_N < 1$, we find that $\delta < pw$. Players' equilibrium payoffs are $U_A = 1 - pw$, $U_Z = \delta - \kappa$, and $U_N = 0$.

1.2. Case 2: $\gamma - \beta < \kappa$

Now we evaluate case 2, in which $\gamma - \beta < \kappa$. In this case, government *B*'s war value of weapons does not outweigh the costs of developing the weapon and paying the costs of fighting a war with *A*. We begin by evaluating the pure strategy equilibria. In case 2 of the inspections game, there are three pure strategy Nash equilibria: WMD, OFFER, and PEACE. The conditions and payoffs for the WMD and OFFER equilibria are the same as in case 1 of the inspections game. However, now there is the PEACE equilibrium in which neither type of *B* arms and *A* attacks. There is one outcome in this equilibrium, which is peaceful non-armament. Since *B* does not arm, A does not need to carry out its threat to attack. Thus, attacking is a credible threat that is never realized in equilibrium. The conditions for the PEACE pure strategy equilibrium in case 2 may be stated formally as follows:

Proposition 4 *When* $\boldsymbol{\gamma - \beta < \kappa}$, *there is a pure strategy equilibrium* (PEACE) *in which both types B do not arm and A attacks if B arms and A does not attack if B does not arm. The conditions for this equilibrium are*

$$c_A < w \ \ and \ \ c_A < \delta. \tag{8}$$

[Figure 5: Case 2 equilibria here]

To see the argument for the PEACE equilibrium, consider *B*'s best response if *A* attacks. If B arms, it receives $\gamma - \beta - \kappa$. Since $\gamma - \beta < \kappa$ in this case 2, then B's payoff from not arming is greater than its payoff from arming. Given that *B* does not arm, government *A* cannot do better than its equilibrium strategy. To ensure that *A*'s threat to attack is credible off the equilibrium path, the *A* must have an incentive to attack if *B* armed. For this to be the case, we must have $c_A \leq \min\{\delta,\ w\}$, which gives the equilibrium conditions in (16). Players' equilibrium payoffs are $U_A = 1$, $U_N = U_Z = 0$.

There is one mixed-strategy equilibrium in case 2, which is MIXED OFFER. The strategies, conditions, and payoffs for this equilibrium are identical to case 1 of the inspections game and are defined in Proposition 7.

Deterrence by Doubt

We are now prepared to address the questions of whether deterrence by doubt is consistent with equilibrium in our approach and whether there are circumstances under which all players prefer it to inspections. Our analysis shows that deterrence by doubt is rare relative to other possible results, ambiguity can also be dangerous, and inspections is effective at reducing conflict and achieving non-proliferation for more states of the world. To demonstrate these conclusions, we show that three results follow from our analysis. First, deterrence by doubt occurs under some restrictive parameter values. Second, there are no other parameter values such that deterrence by doubt holds. Third, there are many more states of the world in which inspections leads to peace and/or non-proliferation including some parameter values in which there is both arming and conflict in the ambiguity game.

To make an argument for deterrence by doubt, three conditions need to be met. First, there must be an equilibrium in the ambiguity game such that at least one type of aspirant will not arm and the counter proliferator will not attack. Second, for the corresponding parameter values in the inspections game, the same type(s) of aspirant will arm and there is, at least, some risk that the counter proliferator will attack. Third, for these parameter values, ambiguity is Pareto superior to inspections.

We find that there are parameter values that meet these conditions. Let us consider $w \in [c_A, \tilde{c}/p]\ and\ \delta \geq c_A$, which corresponds to Region 1 in Figure 6. The relevant equilibrium comparisons in this region are the pure WMD under ambiguity and the MIXED ATTACK equilibrium in case 1 of the inspections game where $\gamma - \beta > \kappa$. If $\gamma - \beta < \kappa$, then the comparison is between pure WMD under ambiguity and pure PEACE under inspections. In the pure WMD equilibrium, government *N* does not arm but *Z* does arm. Comparing B's behavior under inspections when $\gamma - \beta > \kappa$, we see that *N* mixes between arming and not arming and *Z* arms with certainty. Turning to *A*'s equilibrium behavior, under ambiguity *A* will not attack, but under inspections it will mix between attacking and doing nothing. Thus, under ambiguity *N* will not arm and will not be attacked, but if there are inspections it might arm and there is a risk it will be attacked.

The intuition for this difference in equilibrium behavior is the effect that verification of B's behavior has on A's incentives. Under ambiguity, A is uncertain both about B's type and B's behavior. Being uncertain about both induces some caution in A's behavior, because it does not want to attack a normal type. Additionally, since it cannot observe B's actions, it cannot gain any leverage about B's type from its actions. Under inspections, however, B's action signals some information about its type. Government A knows that only N will choose not to arm. This increases A's confidence that B is a motivated type if it arms. When A observes an arming decision under inspections, it is more likely to attack, even though it is not certain an armed B is a normal type. The increased risk of getting attacked leads type N to arm with positive probability. Thus, for the parameter values in region 1, ambiguity deters A while inspections increases the likelihood both that A will attack and that N will arm itself to defend against the risk it will be attacked.

Now we can compare players' payoffs under ambiguity and inspections to determine whether ambiguity is Pareto superior. In both equilibria in the two different games, N receives the same payoffs. Government A's payoffs are also the same in both games as long as $\gamma - \beta > \kappa$, but, if $\gamma - \beta < \kappa$, then A receives a payoff of 1 for inspections versus $1 - \delta$ under ambiguity. Therefore, in region 1, A will only find ambiguity acceptable if $\gamma - \beta > \kappa$.

Since both A and N are indifferent between ambiguity and inspections when $\gamma - \beta > \kappa$, the motivated type Z determines the Pareto ranking. Government Z has the following payoffs from the two games

$$U_Z^A = \delta - \kappa \quad and \quad U_Z^I = \frac{(\gamma-\beta-\kappa)}{(\gamma-\beta)}\delta,$$

where U_Z^A is motivated type Z's utility under ambiguity and U_Z^I is Z's utility under inspections. Note that $U_Z^A > U_Z^I$ if and only if $\delta > \gamma - \beta$. Therefore, we conclude that deterrence by doubt obtains when $w \in [c_A,\ \tilde{c}/p], \delta \geq \max\{c_A, \gamma - \beta\}$, and $\gamma - \beta > \kappa$.

These conditions are relatively restrictive. For them to hold, A's worry can be only so high. Additionally, Z's payoffs to using its weapon in a war must outweigh its costs of developing the weapon, but the payoff to Z for possessing a weapon must be greater than the value to Z of using the weapon in war. These conditions likely rule out many actual cases. For example, for most countries in the world, developing a non-conventional weapon to use in war or to keep is not cost efficient. There are likely only a few exceptions, and India may be an illustrative case. India did not join the NPT, and prior to its 1998 nuclear test there was a great deal of ambiguity about the status of its nuclear program. From India's perspective, there was positive value both to possessing a weapon and using it in a conflict if necessary. It faced threats from both China and Pakistan, both of whom were nuclear powers at the time of the 1998 nuclear test. Thus, there were clearly military benefits to acquiring advanced weapons should a conflict break out between India and its two powerful neighbors. Additionally, by the 1990s, the costs to India of developing a nuclear weapon were likely relatively low, since, according to US intelligence estimates, it had accumulated a large amount of fissile material over many years and could easily weaponize that material in a short amount of time (Perkovich 1999, 340). In spite of the wartime value of weapons for India, its strongest motivation for developing a nuclear weapon was its perceived value of possessing a nuclear weapon to achieve the same international status as other nuclear states (Perkovich 1999). Hence, from India's perspective, it was likely the case that $\delta > \gamma - \beta > \kappa$.

The US valuation of India's weapons program also likely approximates the parameter values in region 1. India possessing a weapon was not nearly as costly to the US as some other countries, such as North Korea, whom they pressured not to arm and to join the NPT a great deal more than they did India at the time. The US hoped India would not conduct the test, but it did not make an attempt to coerce India's government and then accepted the existence of India's nuclear program after the nuclear test. Notably, the US did not pressure India to make the details of its weapons program clear and has, in fact, generally favored India's leaving the status of its program ambiguous (Perkovich 1999, 345 and 438).

[Insert figure 5: Comparing ambiguity to inspections, here]

Does deterrence by doubt occur under any other conditions? To demonstrate that it does not, we need only to show that there does not exist any other range of parameter values such that, under ambiguity, at least one type of *B* does not arm and *A* does not attack, and under inspections *B* arms and *A* attacks with positive probability. To see this, notice that with the exception of region 1 in Figure 6, whenever there is at least one type of *B* that will not arm under ambiguity, either that type of *B* is not more likely to arm or *A* is not more likely to attack under inspections. Consider that under ambiguity, there are only two equilibria in which there is some type of *B* that does not arm: pure WMD and MIXED OFFER. In every other equilibrium, all types of *B* arm with positive probability. For the parameter values that correspond to pure WMD under ambiguity, the relevant equilibrium with which to draw a comparison under inspections is also the pure WMD. The equilibrium strategies for all players are identical in both WMD equilibria. Thus, deterrence by doubt does not obtain for those parameter values.

Where the MIXED OFFER equilibrium obtains in the ambiguity game, there is a pure OFFER equilibrium under inspections. This range of parameter values corresponds to Region 5 in Figure 6. In the MIXED OFFER equilibrium, type *N* does not arm, but it does arm in the OFFER equilibrium under inspections. However, the probability *A* attacks is 0 in both equilibria. In fact, in the OFFER equilibrium under inspections, *A* makes a guaranteed offer of $x = \delta$. Therefore, we conclude that there is no other range of parameters for which deterrence by doubt, as we have defined it, exists.

We have shown that deterrence by doubt only applies to a limited set of circumstances. Now we will examine whether inspections game is Pareto optimal in more states of the world than ambiguity and whether it may lead to peace and/or non-proliferation under more circumstances. To facilitate the welfare comparisons, we will briefly inspect each region in Figure 6. First, we already determined that ambiguity is Pareto superior in region 1 if $\delta \geq \max\{c_A, \gamma - \beta\}$ and inspections is Pareto superior if $c_A < \delta < \gamma - \beta$. In region 1, if $\gamma - \beta < \kappa$, then there is no Pareto ranking between the two games, because *A* always has higher payoff from PEACE under inspections and *Z* has higher payoff under Ambiguity. For these parameter values, conflict and arming is more likely under inspections. In region 2, however, ambiguity is never Pareto preferred, but inspections is Pareto superior either when if $\delta \in [c_A, \gamma - \beta]$ or when $\gamma - \beta < \kappa$ and $\delta \in [c_A, \ ((\kappa\beta)/(\gamma - \kappa))]$. Since the comparison in this region is between MIXED ATTACK in both games, both ambiguity and inspections might result in arming and war. In region 3, inspections are strictly preferred by all players when $\gamma - \beta < \kappa$. Otherwise, there is no Pareto ranking. Note that when $\gamma - \beta < \kappa$, there is no war and no arming under inspections. This stands in contrast to ambiguity where both types of *B* arm and *A* attacks. When parameter values correspond to region 4, ambiguity is never Pareto superior, but inspections are preferred by both types of *B*, and *A*'s payoffs are the same in both games when $\delta \in [\gamma - \beta, \ c_A]$.

In region 4, we compare the MIXED ATTACK equilibrium under Ambiguity and the MIXED OFFER equilibrium under Inspections. In both cases, Z arms and N mixes between arming and not arming. But, under ambiguity A mixes between attacking and not attacking while under inspections it mixes between offering $x = \delta$ and $x = 0$. Conflict is more likely to occur under ambiguity, and there is no chance of war under inspections. In region 5, ambiguity is Pareto superior if $\delta < \gamma - \beta$; otherwise, ambiguity and inspections yield the same payoff for both players. War does not occur in either game, but ambiguity leads to a positive probability of arming only for type Z and, under inspections, both types of government B arm. Finally, in region 6, inspections are preferred to ambiguity when

$$\max\left\{\frac{c_N w}{(w - c_A + c_N)},\ \gamma - \beta\right\} \leq \delta \leq \min\{pw,\ \gamma - \beta\}.$$

But, ambiguity is preferred if

$$\frac{c_N w}{(w - c_A + c_N)} \leq \delta \leq \min\{pw,\ \gamma - \beta\}.$$

Notably, in this region, inspections leads to a higher probability of arming, but ambiguity leads to a higher probability of war.

Gathering these points together, we can draw some inferences about whether inspections are more welfare improving for a greater range of parameter values as well as whether conflict and the proliferation of weapons is more consistent with ambiguity or inspections. First note that ambiguity and inspections may both be Pareto superior in regions 1 and 6 depending on the value of $\gamma - \beta$. Region 5 is the only state of the world in which ambiguity may Pareto dominate while inspections does not under any condition. By contrast, there are several regions for which inspections are Pareto preferred but ambiguity is not. These areas include regions 2, 3, and 4. Thus, there are more circumstances under which inspections make all players better off than ambiguity.

Furthermore, overall ambiguity is more likely to result in conflict than inspections. Ambiguity only reduces the likelihood of war in region 1. For all other regions, either there is less violence under inspections (region 3 if $\gamma - \beta < \kappa$, region 4, and region 6) or there is no difference between inspections and ambiguity (region 2, region 3 if $\gamma - \beta > \kappa$, and region 5).

Proliferation, where B develops and keeps nuclear weapons, is also more prevalent under ambiguity. Proliferation is less likely with inspections in region 3 if $\gamma - \beta < \kappa$, since under inspections neither type of B arms but under ambiguity both types arm. This is an interesting case because it represents the set of parameters for which A is most worried about proliferation (w is highest) and Z is most motivated to develop and possess a weapon (δ is highest). Yet, under inspections, not only will neither type of B arm, but A does not have to pay to prevent arming. The reason is that being attacked is costly for B and A's threat to attack in this region is credible, since inspections makes it possible to verify violations.

Proliferation is also less likely under inspections than ambiguity in regions 5 and 6. Under inspections, both types arm but are offered a payment of $x = \delta$ to disarm in both regions 5 and 6. But, under ambiguity in region 5, Z mixes between arming and not arming while A mixes between offering $x = 0$ and $x = \delta$. There is, therefore, some probability that Z will arm and A will offer $x = 0$, leaving a weapon in Z's possession. The logic is similar in region 6: under ambiguity both types mix arming and not arming while A mixes between attacking and offering $x = \delta$. Thus, there is a chance that either type of B will end up with a weapon.

We have seen that proliferation is more likely under ambiguity than inspections in region 3 if $\gamma - \beta < \kappa$, region 5, and region 6. There is only one region, region 1, where proliferation is more likely with inspections. This is our deterrence by doubt region. For the remaining parameter values (region 2, region 3 if $\gamma - \beta > \kappa$, and region 4), proliferation may occur under both inspections and ambiguity. In region 2, the comparison is between the MIXED ATTACK equilibria in both games. In region 3 when $\gamma - \beta > \kappa$, inspections and ambiguity both result in WAR. And, in region 4, *Z*'s equilibrium strategy in both games is to arm for sure while *N* mixes between arming and not arming. Since *A* mixes between doing nothing and some other strategy (attacking under ambiguity and offering $x = \delta$ under inspections), there is some probability in both games that a weapon gets developed and remains with *B*.

To summarize the main points of this section, the comparisons between ambiguity and inspections show that ambiguity is preferred by all players in fewer states of the world, and it leads to more conflict and weapons in more circumstances. We find that there are parameters such that ambiguity is Pareto preferred, but there are more states of the world in which all players prefer inspections. There are only two states of the world where ambiguity is Pareto superior, one of which achieves deterrence by doubt. While there are, therefore, some parameter values that give rise to deterrence by doubt, there are also parameter values such that ambiguity can actually be dangerous, leading to more conflict. By contrast, inspections results in peace and non-proliferation for many more circumstances, and may even be the most peaceful and inexpensive mechanism for countering proliferation in the toughest cases -- when there is a risk that the most worrisome countries are highly motivated to possess nuclear weapons.

Extortion by Ambiguity

Extortion by ambiguity is the idea that a potential aspirant may be able to extract concessions from countries who are willing to pay to prevent the proliferation of nuclear weapons. The logic for this argument is that uncertainty about an aspirant's nuclear program raises the risk to the counter proliferator enough that it would be willing to pay the aspirant to verifiably abandon its weapons program, but the risk is not so high that it is cost efficient for the counter proliferator to attack the aspirant (Benson and Wen 2011). There are three conditions for extortion by ambiguity to obtain. First, there is at least one type of government B that does not or is unlikely to arm under ambiguity. Second, A is more likely to offer a positive transfer to B under ambiguity than under inspections. Third, the type of B that benefits from the transfer prefers ambiguity to inspections. In this section we show that even though there are some extortion benefits to ambiguity, inspections yields a payment to B with a higher probability under every circumstance in which B might expect to receive a transfer under ambiguity. Additionally, inspections results in payments under a wider range of parameter values. These counter-intuitive results suggest that being subject to inspections can better facilitate bargaining over nuclear weapons.

To see that inspections results in a higher probability that B will receive a transfer, we will inspect the regions in the parameter space in which transfers will be made under ambiguity. In the ambiguity game, transfers are only made in the MIXED OFFER and MIXED ATTACK/OFFER equilibria, which correspond to regions 5 and 6 in Figure 6. In region 5, A mixes between offering $x = \delta$ and $x = 0$, and in region 6 A mixes between offering $x = \delta$ and attacking. Thus, in both regions B receives a transfer of $x = \delta$ with some probability. Now compare this with the corresponding parameter values in the inspections game. The appropriate comparison is the pure OFFER equilibrium in the inspections game, which covers all of regions 5 and 6. In this equilibrium, both types of government B arm, and A makes a guaranteed transfer $x = \delta$ to both types to get them to disarm. While the transfer is made in the ambiguity game with some probability, it is guaranteed when there are inspections. Hence, the argument for extortion by ambiguity cannot be supported by our approach, because there is no type of B that is more likely to receive a transfer under ambiguity.

Additionally, there is separate set of parameter values for which transfers will be made under inspections but not ambiguity. Under inspections, there is a MIXED OFFER equilibrium, in which A mixes between offering $x = \delta$ and $x = 0$, Z arms, and N mixes between arming and not arming. For these same parameters under ambiguity, no offer is ever made. Therefore, not only is there no condition under which ambiguity is more likely to result in a transfer, there are more states of the world where inspections result in transfers being made in exchange for disarmament.

While it may be that A is more willing to make transfers under more circumstances under inspections, it remains to be seen whether B ever prefers inspections to ambiguity for the purpose of receiving transfer payments. To see this, we compare B's payoffs in the two games for the parameter values for which A will make payments with positive probability in inspections. The relevant parameters are regions 4, 5, and 6 in Figure 6. In the previous section, we explained that there are conditions such that inspections Pareto dominate in regions 4 and 6. Hence, in region 4, both types of B prefer inspections when $\delta \in [\gamma - \beta, c_A]$. Otherwise, N prefers inspections but Z prefers ambiguity. In region 6, inspections is preferred by both types if $\gamma - \beta < \delta$. Otherwise, ambiguity is preferred by both types. In region 5, ambiguity and inspections yield the same payoff for both types if $\gamma - \beta < \delta$. Otherwise, ambiguity is preferred by both. The determining factor in each case is the value to B of using the weapon in war. When this value is lower than the value of possessing a weapon, then both types of B prefer receiving the transfer payment under inspections rather than being under ambiguity where there is a chance of getting attacked (regions 4 and 6) or receiving the transfer with a lower probability (regions 5 and 6).

The parameter values for which B prefers inspections in regions 4, 5, and 6 likely describe the conditions of many, if not most, the governments in the world. For most governments, the value of possessing weapons δ is likely relatively low, and the value of using the weapon in a war is even lower so that $\gamma - \beta < \delta$. For these countries, joining the inspections regime can pay, since both types of B can receive transfer payments for joining. Why would A offer guaranteed benefits when there are inspections but not when there is ambiguity? In the inspections model, A offers guaranteed benefits because its expected costs of fighting are greater than the amount it is required to pay to induce B to disarm. Under ambiguity, however, its expected costs of fighting and paying the inducement are the same, and so A is indifferent between attacking and offering δ. Hence, in regions 4 and 5, A will only make transfers some of the time under ambiguity, and it will make them all of the time under inspections. Either way, A will receive the same payoffs. However, there may also be reasons why A might actually prefer inspections. Although we do not formally solve for equilibria under risk aversion here, a risk averse government A might prefer inspections to ambiguity for some high values of w and low values of δ. Intuitively, risk averse counter proliferators will prefer to gain the ability to verify the weaponization decisions of potentially dangerous aspirants in exchange for making a guaranteed transfer payment rather than lose the ability to monitor arming decisions and then randomize between making the payment and attacking.

Are inspections actually associated with transfer payments in practice? Part of the bargain of the NPT is that member states promise not to develop nuclear weapons and subject themselves to weapons inspections in exchange for gaining access to civilian nuclear technology and training (Brown and Kaplow, this issue). The International Atomic Energy Agency (IAEA) was created specifically for the purpose of providing peaceful nuclear energy while monitoring countries' programs to identify whether nuclear development gets channeled into weaponization efforts. The transfer payments made under the IAEA are not trivial. According to Brown and

Kaplow, "In 2010 alone, the Department of Technical Cooperation disbursed $114 million for 890 active projects in such areas as nuclear safety, human health, food and agriculture, radioisotope production, and the nuclear fuel cycle." Thus, while offers are seldom made to governments outside the NPT, governments can receive significant benefits by joining the NPT and submitting to weapons inspections.

In summary, our analysis does not support the argument for extortion by ambiguity. While transfers payments are made under some conditions in the ambiguity game, they are more likely to be made under a greater range of circumstances when there are inspections. Additionally, when countries do not place a high value on using weapons for fighting wars, then countries will be willing to submit themselves to inspections in exchange for receiving transfer payments. This result is consistent with what we observe in practice in the NPT. When countries join the NPT and accept the IAEA safeguards agreement, they gain access to valuable civilian technology and training as well as other related benefits.

Conclusion

We began by asking why countries join the NPT. Our analysis shows that there are both counter-proliferation and peace-enhancing benefits to weapons inspections that make all countries better off. When normal type countries are unable to verify that they are not arming, there might be a risk that they will be attacked by some counter proliferator. In this case, the normal type government, who gains no intrinsic benefit from possessing a weapon, has an incentive to arm just to help it in war. This government benefits from inspections, because verifying that it is not armed can eliminate the risk of it being attacked and enable it to avoid the cost of developing a weapon. There is also the additional benefit that inspections can result in transfer payments being made from counter proliferators to aspirants.

Ambiguity affords countries few benefits by comparison. There are three reasons why countries may favor ambiguity. First is deterrence by doubt. When a government values having a weapon in case there is war and is even more highly motivated to keep nuclear weapons, then ambiguity may be welfare enhancing and normal type states might achieve deterrence by doubt. However, the conditions for deterrence by doubt are limited to these restrictive conditions and may explain why only few cases fit the logic. The second situation in which ambiguity may be better than inspections is when the war benefits of a weapon outweigh the value of acquiring and possessing a weapon. In this case, the transfer payment from the counter proliferator is insufficient to compensate a government for giving up its war asset, and it will prefer to arm in secret to reduce the risk of being attacked. The third reason why a government might prefer ambiguity is that it may be a motivated type that places high value on possessing a weapon and ambiguity may reduce the chance it will be attacked. In this case, the aspirant wants to possess a weapon, but arming might be dangerous. Since counter proliferators prefer not to attack normal types, and the ability to verify non-armament increases the counter proliferators' confidence that arming states are motivated types, therefore, the motivated aspirant might resist inspections to lower the risk that it will be attacked.

We have shown that deterrence by doubt is rare and extortion by ambiguity is not consistent with our argument. How, then, do we explain Iraq, which is cited as an example of deterrence by doubt, and North Korea, which is the standard example of extortion by ambiguity? The anecdotal evidence indicates that Saddam Hussein's evasiveness prior to the Iraq War was largely motivated by his desire to deter his enemies, both domestic and foreign. Michael Schrage claimed that prior to the war in Iraq, Saddam Hussein's ambiguity "kept the West at Bay while keeping Hussein's neighbors and his people in line."[3] Saddam Hussein and Tariq Aziz, the Iraqi deputy prime minister under Saddam Hussein, both confirmed Hussein's motive to create the perception that the government had weapons of mass destruction to deter Iranian aggression and a Shiite uprising.[4] David Kay, the US chief weapons inspector, claimed that Hussein pursued "creative ambiguity" to "enjoy the benefits of having . . . weapons without having to pay the costs."[5] So well-guarded was the secret that Iraq did not possess weapons that Hussein's top military leaders did not know until just months before the war that Iraq did not have weapons of mass destruction (Gordon and Trainer 2006). Iraq is also a good illustration of a case in which there were clear deterrence motives behind the strategy of ambiguity.

Yet, Iraq is an impure case of deterrence by doubt because it was a member of the NPT, and the US pushed aggressively for it to be compliant with weapons inspections. Rather than both aspirant and counter proliferator mutually preferring ambiguity, as is the case with deterrence by doubt, both appeared to favor inspections. It appears cooperating with weapons inspectors while maintaining some ambiguity was part of Hussein's plan. This implies two possible avenues of ambiguity: refusing inspections so nothing can be observed and allowing inspections while creating uncertainty about actual actions. In our modeling approach, the first form of ambiguity is a structural feature of the ambiguity game, such that the counter proliferator is unable to observe the aspirants arming decision. The second form of ambiguity is possible even when there are inspections but the aspirant chooses a mixed strategy and thus arms with some probability. That Hussein elected to pursue the second option suggests that there may be

[3] Michael Schrage, "No Weapons, No Matter. We Called Suddam's Bluff," *Washington Post*, May 11, 2003, B2.
[4] See Gordon and Trainer (2006); Woods, Lacey, and Williamson (2006); and Myre (2006).
[5] David Rennie, "Saddam was bluffing, says top US arms hunter," *The Telegraph*, January 30, 2004, URL: http://www.telegraph.co.uk/news/worldnews/northamerica/usa/1453068/Saddam-was-bluffing-says-top-US-arms-hunter.html

circumstances under which a government might join the NPT and still create uncertainty in mixed strategies.

Our analysis, therefore, may provide some insight into the Iraq case. In region 4, inspections is preferred by both the aspirant and counter proliferator if the war fighting value of non-conventional weapons is greater than the value of possessing them. With its history of conflict with Iran and domestic threats from the Shiites and Kurds, it is reasonable to believe that Iraq believed their value for using weapons in the case of a war was high relative to the intrinsic value of developing weapons for the purpose of having them. Iraq, then, would have preferred inspections, and its equilibrium strategy would have been to mix between arming and not arming. The alternative would have been to refuse inspections and still mix between arming and not arming. However, it is plausible that Hussein believed the US was more likely to attack if it did not participate in weapons inspections. In fact, Hussein may actually have believed that being in the NPT and participating in inspections decreased the probability that the US would attack (Gordon and Trainer 2006). The US decision to attack is more difficult to understand with the logic of our analysis. One possibility is that the US believed that Hussein's efforts to obfuscate information about its weapons program sufficiently hindered inspectors' ability to verify that Iraq was not arming that it was functionally equivalent to being outside the NPT and being fully ambiguous. In this case, the US would mix between attacking and not attacking, and the probability of attacking is positive in Iraq's costs of developing weapons, which were high. This illustration and the logic of the argument points to the possibility that countries might find deterrence benefits from ambiguity by being inside an inspections regime but nevertheless creating uncertainty about the status of their weapons program. The main advantage to this approach compared to being ambiguous by not joining the NPT is that it reduces the chance of being attacked. Iran may now be following this course.

Can we understand North Korea's behavior, given that it withdrew from the NPT and offers to induce it to disarm were ineffective? If it was similar to Iraq in that the war fighting value of non-conventional weapons is greater than the value of possessing them, then making a transfer payment in the correct amount would likely be an effective counter-proliferation approach. Yet, payments have been unsuccessful, and North Korea, unlike Iraq, developed weapons. The US twice offered to transfer assistance to North Korea to gain verifiable disarmament, first in the 1994 Agreed Framework and then again during Six-Party Talks from 2003-2007. North Korea left the NPT in 1993 and then agreed to the 1994 Agreed Framework with the United States. North Korea rejoined the NPT and froze its weapons program in exchange for security guarantees, two light-water nuclear reactors, and shipments of oil to North Korea. The deal fell apart in 2002 when North Korea withdrew from the NPT again. Six Party Talks took up negotiations over North Korea's weapons program, and several deals were proposed. However, talks were discontinued in 2009, a deal was never reached, and North Korea tested a nuclear device and has announced plans to test another.

One possible explanation is that North Korea is actually a motivated type, and its value of acquiring a weapon is greater than its value of using it. It is unlikely the case that North Korea places high war fighting value on a nuclear weapon, but it is reasonable that it wants to possess nuclear weapons for status and other reasons. Victor Cha (2009), former director of Asian affairs on the US National Security Council from 2004-2007 and deputy head of the U.S. delegation to the Six-Party Talks during that time, claims that experience has taught American negotiators that North Korea really wants to be a nuclear state like India, and while the level of its weapons program is negotiable, its status as a nuclear power is not. If North Korea differs from Iraq in this respect, then the logic of our argument suggests that North Korea would prefer not to be party to weapons inspections and would find it advantageous to be ambiguous about its weapons program to reduce the risk that it will be attacked by the US. Ironically, if the logic of our argument is correct, offering transfer payments and security assurances may have been more successful with Iraq, and being more aggressive may have been a more effective strategy to prevent North Korea from arming.

In this paper, we have developed a model of nuclear weapons development and our examples are mostly related to nuclear proliferation. In the Iraq illustration, however, one of the concerns discussed prior to the Iraq War was the possibility that Saddam Hussein might be developing chemical or biological weapons. The threat of proliferation of non-conventional weaponry other than nuclear weapons raises the question about the applicability of our model to bargaining over other kinds of weapons systems. Horowitz and Narang (this issue) explore the relationship between nuclear weapons and chemical/biological weapons and show that under many circumstances they are substitutes. To the extent that these non-conventional weapons systems do, in fact, exhibit similar properties, our results might also be relevant to a broader range of proliferation issues beyond nuclear weapons.

A final point ties our study about bargaining over proliferation to existing empirical studies and future work in this area. In our model, a nuclear aspirant either arms or does not arm, and then the counter-proliferator decides whether to attack or to make an offer to induce armed aspirants to give up their weapons. We regard this as a positive first step in a theoretical approach to studying bargaining over nuclear weapons. There are, however, other factors related to bargaining over governments' nuclear weapons programs that should be taken up in future theoretical work. One area, in particular, that stands out is the aspirant's arming decision. An aspirant might adjust the type and amount of its weaponization to its expectation of how the counter-proliferator will respond. This topic is addressed empirically by some of the studies included in this issue. Kroenig (this issue) analyzes proliferation decisions as a function of other governments' nuclear arsenal size, suggesting there may not be a link between governments' decisions to reduce arms. Sechser and Fuhrmann (this issue) examine proliferators' decisions to forward-deploy nuclear weapons and Gartzke, Kaplow and Mehta (this issue) investigate the proliferators' choice of force posture including decisions regarding missile ranges and types of deployment systems. The model we have developed here can be extended in many ways to further examine the incentives that shape governments' decisions to develop different types and amounts of nuclear weapons.

References

Baliga, Sandeep and Tomas Sjostrom. 2008. "Strategic Ambiguity and Arms Proliferation." *Journal of Political Economy* 116: 1023-1057.

Benson, Brett V. and Quan Wen. 2011. "A Bargaining Model of Nuclear Weapons: Development and Disarmament," in *Causes and Consequences of Nuclear Proliferation*. Edited by Robert Rauchhaus, Matthew Kroenig, and Erik Gartzke. New York: Routledge Press.

Brown, Robert L. and Jeffrey M. Kaplow. 2012. "Talking Peace, Making Weapons: IAEA Technical Cooperation and Nuclear Proliferation" (this issue).

Burr, William and Jeffrey T. Richelson. 2000/2001. "Whether to 'Strangle the Baby in the Cradle'." *International Security* 25(3): 54-99.

Cha, Victor D. 2009. "What Do They Really Want?: Obama's North Korea Conundrum." *The Washington Quarterly* 32(4): 119-138.

Chang, Gordon. 1988. "JFK, China, and the Bomb." *The Journal of American History* 74(4): 1287-1310.

Fuhrmann, Matthew. 2009. "Spreading Temptation: Proliferation and Peaceful Nuclear Cooperation Agreements." *International Security* 34 (1): 7-41.

Hymans, Jacques. 2006. *The Psychology of Nuclear Proliferation: Identity, Emotions, and Foreign Policy*. Cambridge: Cambridge University Press.

Feaver, Peter. 1997. "Neo-optimists and the Enduring Problems of Nuclear Proliferation." *Security Studies* 6 (4): 93-125.

Feaver, Peter, and Emerson Niou. 1996. "Managing Nuclear Proliferation: Condemn, Strike, or Assist?" *International Studies Quarterly* 40 (3): 209-234.

Fuhrmann, Matthew and Sarah E. Kreps. 2010. "Targeting Nuclear Programs in War and Peace: A Quantitative Empirical Analysis, 1941-2000." *Journal of Conflict Resolution* 54(6):831-859.

Goldstein, Lyle. 2006. *Preventive Attack and Weapons of Mass Destruction: A Comparative Historical Analysis*. Stanford: Stanford University Press.

Jo, Dong-Joon, and Erik Gartzke. 2007. "Determinants of Nuclear Weapons Proliferation." *Journal of Conflict Resolution* 51:1-28.

Kroenig, Matthew. 2009. "Exporting the Bomb: Why States Provide Sensitive Nuclear Assistance." *American Political Science Review* 103 (1): 113-33.

Levy, Jack. 2008. “Preventive War and Democratic Politics.” *International Studies Quarterly* 52 (1):1-24.

Levy, Jack. 1987. “Declining Power and the Preventive Motivation for War.” *World Politics* 40(1): 82-107.

Myre, Gregory. 2006. “In a Slip, Israel’s Leader Seems to Confirm its Nuclear Arsenal.” *New York Times* (December), A5.

Oberdorfer, Don. 1997. *The Two Koreas: The Impact on Global Proliferation*. Berkely and Los Angeles, CA: University of California Press.

Perkovich, George. 1999. *India’s Nuclear Bomb*: *The Impact on Global Proliferation*. Berkeley and Los Angeles, CA: University of California Press.

Raas, Whitney and Austin Long. 2007. “Osirak Redux? Assessing Israeli Capabilities to Destroy Iranian Nuclear Facilities.” *International Security* 31(4): 7-33.

Reiter, Dan. 2005. Preventive Attacks Against Nuclear Programs and the ‘Success’ at Osiraq.” *Nonproliferation Review* 12 (2): 355-71.

Reiter, Dan. 1995. “Exploding the Powder Keg Myth: Preemptive Wars Almost Never Happen.” *International Security* 20 (2): 5-34.

Sagan, Scott. 1996/1997. “Why Do States Build Nuclear Weapons? Three Models in Search of a Bomb.” *International Security* 31(4): 7-33.

Saunders, Phillip C. 2009. “Confronting Ambiguity: How to Handle North Korea’s Nuclear Program.”

Sigal, Leon V. 1997. *Disarming Strangers: Nuclear Diplomacy with North Korea*, Princeton University Press.

Singh, Sonali, and Christopher Way. 2004. “The Correlates of Nuclear Proliferation: A Quantitative Test.” *Journal of Conflict Resolution* 48:859-85.

Sobel, Joel. 1992. “How (and When) to Communicate with Enemies,” in *Equilibrium and Dynamics*. Edited by Mukul Majumdar. London: Macmillan Press.

Solingen, Etel. 2007. *Nuclear Logics: Contrasting Paths in East Asia and the Middle East*. Princeton, NJ: Princeton University Press.

Woods, Kevin, James Lacey, and Williamson Murray. 2006. “Saddam’s Delusions: The View from the Inside.” *Foreign Affairs* 85 (May/June): 2-26.

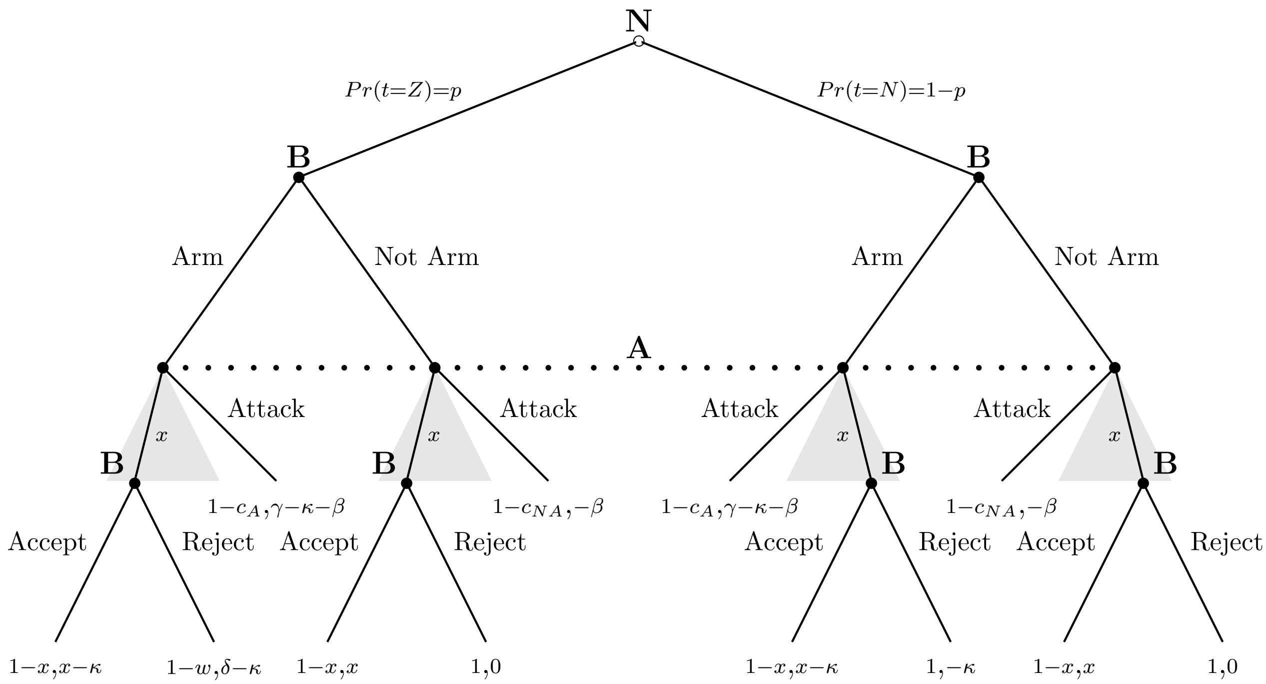

Figure 1: Ambiguity game (Benson and Wen 2011).

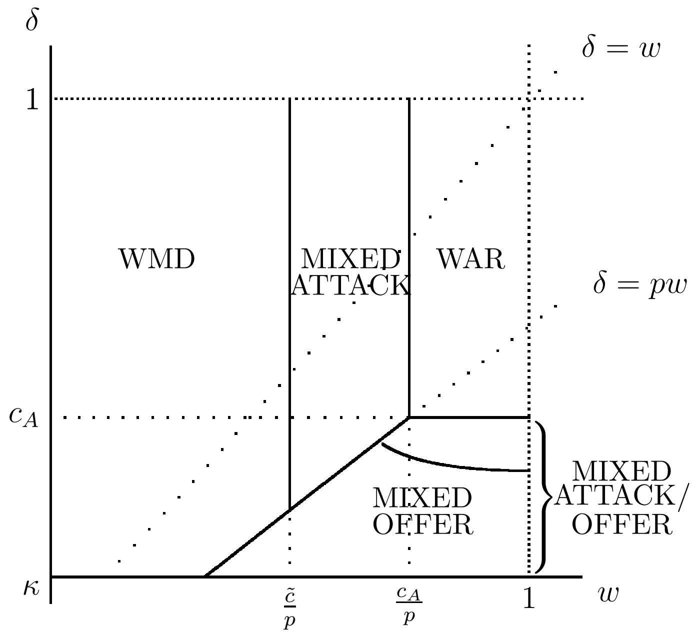

Figure 2: Ambiguity equilibria.

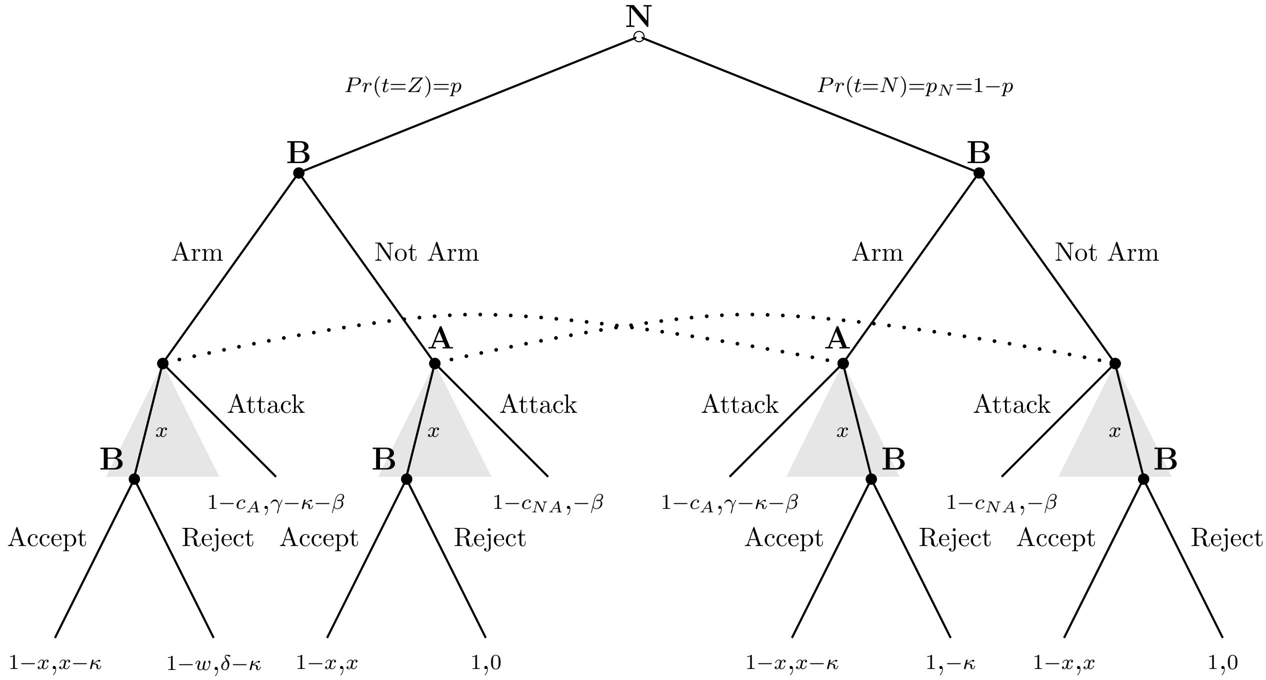

Figure 3: Inspections game.

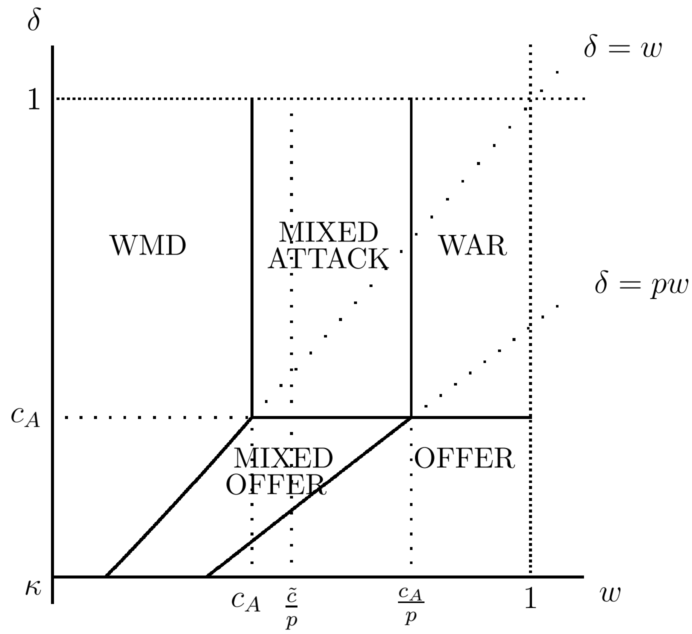

Figure 4: Inspections equilibria when $\gamma - \beta > \kappa$.

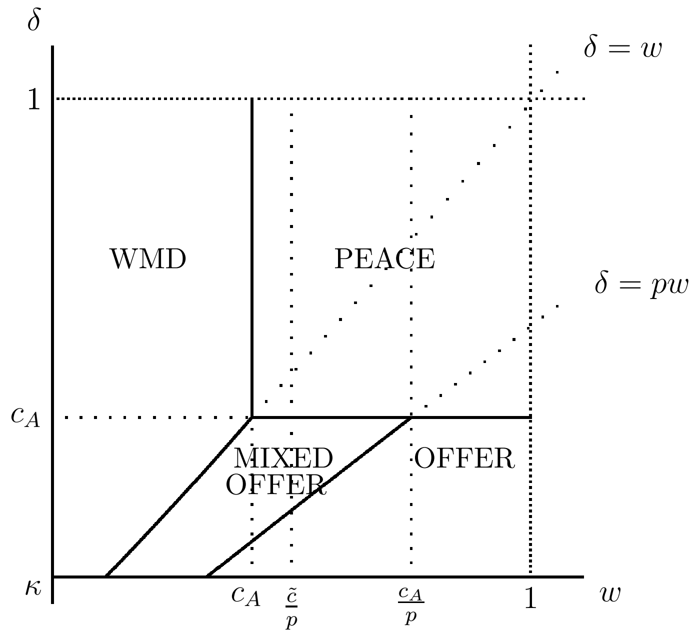

Figure 5: Inspections equilibria when $\gamma - \beta < \kappa$.

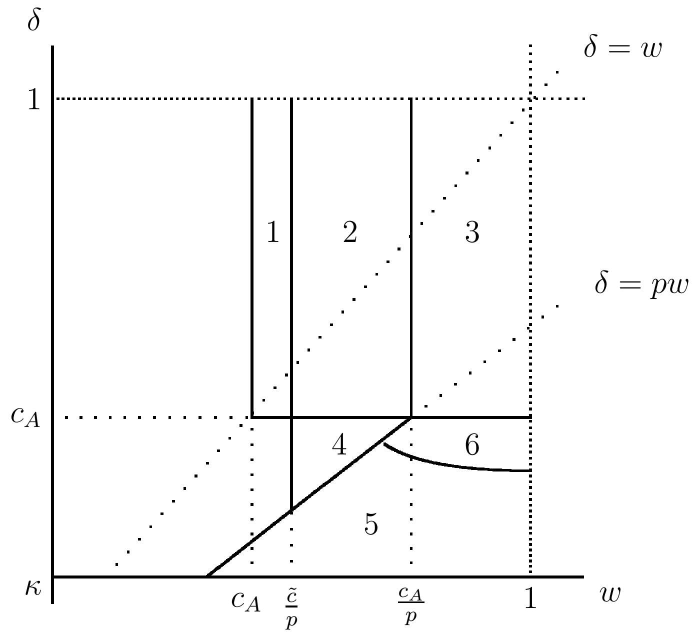

Figure 6: Comparing ambiguity to inspections.