WHAT HAPPENS AFTER CONCEDING: THREE-STATE EXTENDED NUCLEAR DETERRENCE

D.M. Vossenberg

BSc Political Science: International Relations and Organisations

Bachelor thesis on Nuclear Deterrence Theory

Tutor: dr. H. Pellikaan

Final version submitted on June 12 2017

8300 words

In contemporary international relations, states attempt both direct and extended nuclear deterrence. In order for game theory scholars to adequately analyse extended deterrence, an adequate model has to be constructed. This model is inquired via a comparative analysis of two cases in two game theoretical models of nuclear deterrence. The selected cases are based on scenarios that are likely to happen in the contemporary multipolar world. The two cases simulate the nominal nuclear deterrence game and another case with a nonviolent protégé. The first model is the Quackenbush's extended deterrence game. The second being an original three-state nuclear deterrence game. The analysis yields a new perspective on the role of allied nuclear powers in game theory on nuclear deterrence. The results conclude as that the ever-present option of a retaliation changes the strategy of the challenger. It adjusts whether and whom he should attack.

CONTENTS

LIST OF FIGURES

INTRODUCTION

In the international political spectrum there is the prevailing view that the amount of nuclear warheads in the world should be structurally reduced in order to reduce the chances of a nuclear conflict. There has been a global paradigm shift from the traditional necessity for nuclear weapons as a requirement for nuclear deterrence to the contemporary multipolar world in which there is a demand for the reduction of the amount of warheads (UN General Assembly, 2016). One of the significant factors in this is the greatly increased amount of nuclear powers. This horizontal expansion has led to a total of nine nuclear powers: the United States, Russia, China, the United Kingdom, France, Israel, India, Pakistan and North Korea (Van der Meer, 2016). More states than these nine have made an effort to build their own nuclear arsenal, yet have either failed or lost the ambition to do so. This reality that not every state has nuclear capabilities causes nuclear deterrence to be divided into two types (Morgan, 1983; Huth 1988).

These types are direct and extended deterrence. This causes nuclear deterrence to have two features. It deters a challenger from challenging the defender. The other utility is that it grants the ability to deter a challenger from challenging a third state, such as an ally. An evident example of the coexistence of the two has been the conflict between the United States and the Soviet Union during the Cold War. The United States had the objective to not only prevent itself from being struck, but also saw it as vital to deter any attacks on its NATO allies in Europe (Pellikaan, 2017). This example also shows that direct deterrence and extended deterrence are equal in its usage. Evidently, they are not in terms of their impact.

Several scholars have attempted to model the abovementioned events into models. Quackenbush came with the extended-deterrence game involving three actors in 2006. However, Quackenbush's extended deterrence game does not allow for a retaliation by a third state after the struck state has conceded (2006). The newly constructed 'original nuclear deterrence game' does allow for this reaction. The difference between these games is, thus, that the original game has integrated the possibility for a third state to retaliate, even if the protégé is not willing to do it itself. This, in turn, could lead to a better understanding of extended nuclear deterrence on a game theoretical level. The proposition is that Quackenbush's model does not allow for the retaliation of third safeguarding states if the defender immediately concedes.

Due to the great importance of nuclear deterrence for the sake of the entire world, it is vital that there is a grounded understanding of it within the fields of international relations and game theory. In order to perfect the model of Quackenbush it is important to expose its flaws (2006). The research question is therefore: 'How does the strategic interaction of the challenger change in nuclear deterrence when a retaliation by the defender is allowed after the concession of the protégé?' The aim of this thesis is to determine the difference in the outcomes of nuclear deterrence between two nuclear deterrence games.

The structure of the thesis is as follows. Firstly, the important contemporary literature is elaborated upon. The relevant subjects vary from the grand theories up to the usage of game theory as a method of analysis. After this, the concepts that are utilised are explained in detail. Moreover, the new concepts of nuclear collateral damage, the safeguarding state and the, original, three-state nuclear deterrence game. The methodology that is used for the analysis is explained after this. It is explained how backwards induction can determine the outcome of the nuclear deterrence games. Moreover, the case selection and analysed cases are elaborated upon in order to demonstrate their motives and preference orderings. Next in order is the analysis of the cases. The three aforementioned cases are exposed to the two nuclear deterrence games. Finally, the conclusion and discussion are drawn up.

DETERRENCE IN INTERNATIONAL RELATIONS

Waltz and Mearsheimer recognize in their works that states act within the structure of several concepts with the grand theory of neorealism (1979; 2001). These being self-defence, rationality, relative gains and the security dilemma. The premises behind these behaviour-explaining concepts is that states are inherently in constant conflict and act in a world of anarchy. There is no single supranational organ that rules over all. Oneof the key features of neorealism is the assumption that states act rationally. The concept of rationality is a vast component of game theory.

The theory of deterrence has been extensively investigated in the past (Pellikaan, 2017). It is widely researched in international relations theory (Lebow and Stein, 1995; Huth, 1988; Stein 1987). The aforementioned articles are all focussed on conventional deterrence with two actors: a challenger and a defender. In the past, scholars included only the challenger and the defender in their models. Werner follows the nominal route as she encompasses only the two actors in her model (2000). This results in a model that puts an emphasis on the available choices of the target. However, by ignoring the environment in which the defender exists she neglects the role of the defender in the strategic calculation of the challenger.

Kilgour and Zagare include the "the role of the pawn in extended deterrence" (1994). However, the scholars do not include this as an actual actor in their model. This evidently leads to a different outcome than when the pawn or protégé is included. The addition of the third actor in the game is the vast difference between all the aforementioned models and the models of Quackenbush and the newly presented model in this research.

Quackenbush himself includes three actors in his model: the challenger, defender and protégé. According to Quackenbush his model provides three benefits. Firstly, it grants a formal model of extended deterrence that is in line with the rationale behind extended deterrence. Next to this, it complies with the argument that extended and direct deterrence are administered simultaneously by the actors (Snyder, 1961). On a third note, it includes all three actors. Unlike Kilgour and Zagare's model it allows for the analysis of whom the challenger should strike first. This issue has seldom been addressed in literature before.

Powell has researched the concern whom the challenger should strike in a model with aligned decisions (1999). He concludes that the challenger attacks the larger state. The third state will then 'bandwagon' with the challenger. This splits the alliance. Powel uses the premise that every strike turns into a full-blown war. Realistically, states can concede or compel the challenger to back down after a first strike has been performed. The model that is introduced and Quackenbush's model allow for the struck actor to concede.

Bueno de Mesquite researched which allied actor the challenger should strike upon first (1981).

He came to the conclusion that the challenger should attack the strongest ally. This would be the defender. One of his premises, however, is that the fighting will take place on the territory of the attacked. This premise, however, is too restrictive for a nuclear deterrence model. There would be no fighting at all, as a nuclear strike should demobilize the army of the struck state. Moreover, it would be useless for the third actor to send its reinforcements to the struck territory, as there are no opposing forces present.

Zagare and Kilgour have used game theoretical analyses by use of their extended-deterrence game to give insights into the British policymaking in the countdown to World War 1 (2006). The extended-deterrence game is closely related to the three-state nuclear deterrence game utilised in this research. Moreover, Wu has revised Huth's similar model of deterrence in order to improve its prediction rate of successful deterrence (1990; 1988). Game theory has thus proven to be of relevance especially the field of deterrence theory.

The new model also allows the third state to retaliate in the case of the capitulation of the struck state. This is also the foundation of this thesis. It is devoted to the construction of a game theoretical model that allows for the retaliation by the nuclear state, the defender, to retaliate even after the protégé has conceded.

NUCLEAR DETERRENCE THEORY

What is the sole reason that you do not go over the speed limit on a completely empty road? It is the deterrence of the possibility of a fine. This occurs between states in a similar manner. Instead of a fine, an unwanted action 'D' by state A can be retaliated by state B. Originally, there were two state actors involved in deterrence. The first state, state A, is the challenger. It has the intention to attack state B, the defender, with a first strike. The challenger has to have the intention to gain utility of the attack. This can range from resourceful territory to prestige for the politicians. The optimal operation would yield the objective with as little costs as possible, so presumably a quick and effective strike. The defender on the other hand has the exact opposite objective. It wants to decrease the likelihood of an attack by the challenger. It would be a strategy to make the challenger struggle as much as possible in his objective. A long-lasting and costly war could be a turning point in the cost-benefit analysis of the challenger.

The roles of the actors do not change when there is a defender with nuclear capabilities. The challenger yet has to have an incentive to do a first strike on the defender that will result in more gains than losses. It is important to note, however, that it is not always the case that the challenger will use their nuclear weapons as a first strike. The notion of deterrence is fully about the possibility of a retaliation or strike-back from the defender. The difference now is that a nuclear-capable defender has the capability to annihilate entire cities of the challenger. This will deter the challenger even more, as the costs of war have increased.

In the scenario that the challenger would be the only state to have nuclear capabilities, its costs for attacking the defender would be minimal. This would cause the tables to turn and let the challenger favour a war. Moreover, it is possible for both parties to have acquired nuclear weapons. This gives the defender an incentive to create a credible strike-back capability.

The credible strike-back accounts for the likelihood that the challenger deems the defender to retaliate if it were to be struck by nuclear arms first (Morgan, 2003). Without a strike-back credibility, there is no nuclear deterrence possible, as the challenger simply does not believe that it will be struck back with a nuclear weapon. Therefore, it is important for the defender to create a status that shows that it is not afraid to use its arsenal. This way, the status quo can be maintained. As the first strikes and strikes-back have become catastrophically destructive, states actors are forced to determine their set of overriding preferences (Morgan, 2003). This set of limits can serve as the starting point for a rational decision method.

RATIONALITY

Rationality is a fundamental part of the school of rational choice theory and is a defining part of the decision-making process of an actor with limited resources that uses a cost-benefit analysis (Morgan, 2009). This actor can be a person, state or anything that has to make a decision. All these actors have a selection of choices to pick from.

Arrow found that in order to be deemed rational an actor has to suffice to two axioms, Axioms I and Axiom II (1995). In other words, if the decision of an actor does not fulfil both requirements, the actor is irrational and has not maximized its utility. Axiom I is formulated by Arrow as follows:

(1) Axiom I: For all x and y, either x R y or y R x

This formal statement exerts on the rational choice theory that all actors must be able to formulate a preference or indifference in every pair of two options (Arrow, 1995). This means that there must a relationship in terms of preference between all possible options. For example, a customer in a café must be able to distinguish his preferences between coffee, tea and water. There has to be a relationship between the tea and coffee, tea and water, and coffee and water. Thus, all options have to be interconnected.

The second requirement, Axiom II, prescribes that all preference orderings must be transitive. Transitivity means that a logical preference relationship between the options is achieved. It is presented formally by Arrow in the following manner:

(2) Axiom II: For all x, y and z, x R y and y R z implies x R z

Accordingly to the previous example, if someone prefers tea over coffee, and coffee over water, he or she must prefer tea over water. This forms the transitive relationship T > C > W. Axiom II defines it as irrational to have the relationship W > C > T if the abovementioned pairwise relations are present. This premise of rationality means that actors are effectively free in their motives and the defining of utilities. The two axioms do not define a certain base from which all humans or states make their decisions.

The reason that Arrow's definition of rationality is chosen over other theories is that it does not predefines the motives of states. This is the thin theory within rational choice theory. An option could be to have the assumption that states to define their utility on a purely economic basis. However, this would reject the fact that states have addressed non-economically viable policies in the past, such as humanitarian law.

QUACKENBUSH'S EXTENDED DETERRENCE GAME

In his paper *Not Only Whether but Whom: Three-Party Extended Deterrence* Quackenbush elaborates on his extended deterrence model (2006). This model is seen in the appendix. The model is based on the asymmetrical power relationship between a defender and a protégé. The defender can be seen is presumed to be a great military power in its region or alliance. The protégé, on the other hand, is its inferior on that aspect. Therefore, the defender might act as an immediate ally in the case of an attack. The defender is comparable to the role of the United States in the Cold War. The Netherlands would, in this case, project the role of the protégé.

The other state in this game is defined as the challenger (Quackenbush, 2006). In the model of Quackenbush, the challenger has the possibility to either attack the defender, attack the protégé or to maintain the status quo. These are the options at Quackenbush's first decision node. It depends fully on the decision of the challenger made at this point that sets the possible reactions by the defender and protégé.

The challenger can choose to execute a first strike on either the defender or the protégé (Quackenbush, 2006). This moves the game up to the second decision node. If the challenger chooses to attack the protégé, the protégé has the two options. The first option is to concede. Conceding in this model means that the protégé surrenders to the challenger. The other option is to fight back to the challenger. If the challenger would strike the defender first instead, the defender has the same two options. It can either concede or fight back. If either actor chooses to fight back, the game will continue to the third decision node.

On the third decision node the actor that has not yet made a decision now comes into play. If the challenger would have struck the protégé first, it is the turn of the defender. However, if the defender has previously been attacked, the turn is to the protégé. Both the protégé and the defender have the option to join the attacked actor or to take on a passive role. If either state chooses to intervene, the result will be a multilateral war between the challenger, the protégé and the defender. The passive role means that the actor does not intervene. If the third actor chooses to intervene, the game advances to the fourth and last decision node.

PREMISES FOR NUCLEAR DETERRENCE

In this research there is the aim to determine whether Quackenbush's model is usable in the contemporary world. Therefore it is necessary to set certain additional preferences for this model. These premises include that the implementation of nuclear strikes and nuclear strike back possibilities. Moreover, the last decision node is exempted from the model. This results in three premises.

Premise 1

Instead of conventional interstate warfare the nuclear model uses nuclear strikes. This means that at least the challenger and defender are nuclear powers. It reality it can occur that the protégé has nuclear capabilities as well. However, this is the case in this analysis. As analysing conventional warfare is not the aim of this research non-nuclear warfare by the challenger and defender is excluded from the deterrence model. This means that the first strike by the challenger on an actor will be a nuclear strike. The struck actor will then have two options. It can either retaliate or concede.

Premise 2

The protégé state actor has two options. The first option is to concede and allow its unilateral destruction. The other option would be to execute a nuclear strike as a means of retaliation. In Quackenbush's model this results in the bilateral nuclear war instead of a conventional bilateral war.

Premise 3

The third premise is that the last decision node is removed from the game. If this node is included in the method of backwards induction becomes impossible to use. It is unclear whether the challenger would prefer multilateral war (MW) over challenger defeated (CD). Therefore it is not possible to use backwards induction on from the fourth node as a starting point.

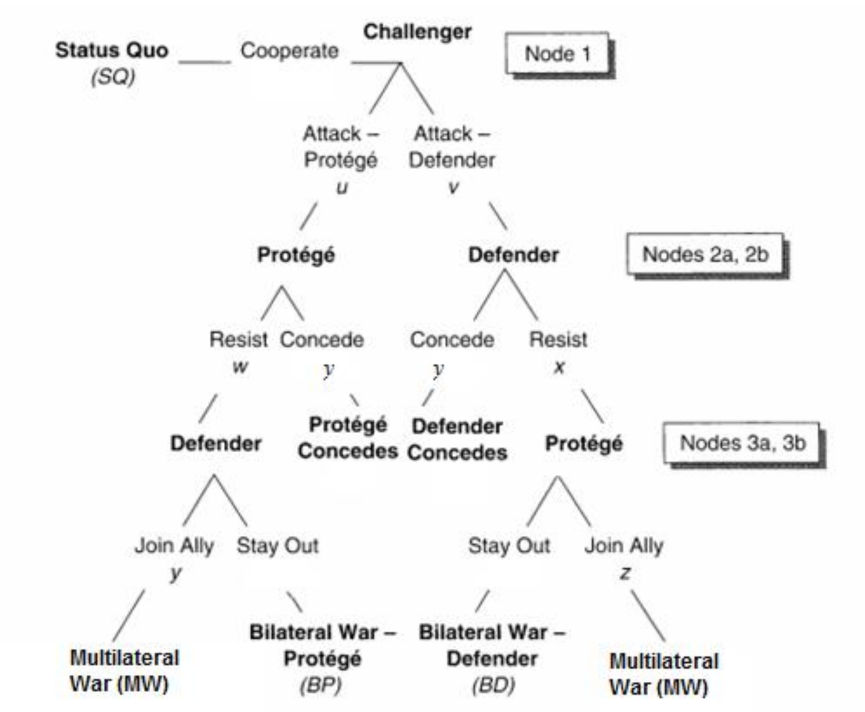

FIGURE 1: QUACKENBUSH'S MODEL ADAPTED VERSION

TWO STATE NUCLEAR DETERRENCE GAME

The two person nuclear deterrence game is a sequential game in which the second player reacts to the actions of the first player. Powell publishes the deterrence game in his book *Nuclear Deterrence Theory: The Search for Credibility* (1990). The advantage of the two-person nuclear deterrence game is that it shapes this into a formal model. Moreover, it is a pure form of direct deterrence (Morgan, 2003). This means that the defender is focussed only on protecting himself via deterrence. The two-person nuclear deterrence game can be seen in figure 2.

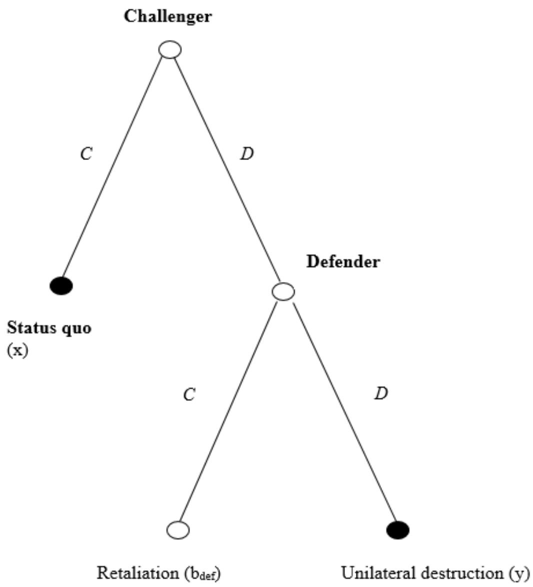

FIGURE 2: TWO STATE NUCLEAR DETERRENCE MODEL

THREE-STATE NUCLEAR DETERRENCE GAME

The three-state nuclear deterrence game has been inspired by the Powel's nuclear deterrence game. It allows for a third bloc or actor to function within the nuclear deterrence game, allowing it to emulate the current world system better. The fundamental change is that this model accounts for that no state wants to be affected by the nuclear battles of two other states. If state C would attack state P with a nuclear strike, the odds for an effect on state D are significant. This third state nuclear collateral damage could affect the state in terms of, for example, a nuclear pollution due to the spread of particles, an economic downfall due to the loss of a trading partner, or refugee crises. It emulates the willingness of non-threatened states to safeguard the protégé (P) from being struck. The challenger has the option to maintain the status quo, attack the defender or attack the protégé. This is illustrated in figure 2. The struck actor can then decide whether or not to retaliate upon the challenger. The defender can strike back and cause mutual destruction (b_{def}) or stay passive and allow for its unilateral destruction (y). If the protégé were to be struck, the protégé can choose either to allow for unilateral destruction (y) or to fight back (b_{pro}).

In this model the protégé does not have nuclear capabilities. The defender does have nuclear capabilities. This causes a retaliation by the protégé to be far less effective than that of the defender. Therefore, it is likely that the challenger is not concerned with the retaliation with conventional forces by the protégé if it were to strike the defender. The defender's own retaliation is far more concerning.

This results in the defender acting as a safeguarding actor. It is the actor that can deter the challenger from attacking the protégé. In reality, the safeguarding actor would be a state that has nuclear capabilities. The safeguarding state can act as an increase in credibility for the strike back. This will, in the end, make successful deterrence more likely.

The disincentive is caused by the option of the safeguarding state or defender to retaliate once the protégé has chosen not to retaliate for the first strike by the challenger. The defender has the option to cooperate, and allow the unilateral destruction of the protégé. This results in the outcome of 'y'. The other option is to defect. This means that the defender will retaliate for the strike on the protégé. This will lead to outcome 'w'. These options are illustrated in the lower part of the tree in figure 3. The attacking state might prefer outcome 'y' over x, but not over 'w' or 'z'. In order to be able to fully explain any case in this deterrence game, a distinction has to be made between the outcomes 'w' and 'z'.

The trustworthiness or dependability on the defender is demonstrated by the concept of safeguarding credibility. This concept is similar to the concept of the credible strike back. Both concepts qualify states as being able and willing to retaliate with nuclear arms if the defender is struck by a first strike.

However, safeguarding credibility concerns the likelihood that the defender will retaliate if the protégé has had the capability to strike back. The safeguarding credibility is heavily characterized by the stakes of the defender in the game. The more interdependent or near the defender and the protégé are, the higher the odds are that the former retaliates. A high safeguarding credibility influences both the challenger as well as the protégé. The higher the odds are that the safeguard retaliates, the more likely the challenger will cause the annihilation of both itself and the defender. The effect on the protégé works differently. The protégé could choose to prefer a retaliation by the safeguard over one by itself. This preference relies heavily on the credibility of the safeguard. For example, if the safeguard is not deemed credible, the odds are high that the defender will choose to defect and cause outcome 'z'. The defender could be taking an unnecessary risk if it choose to cooperate.

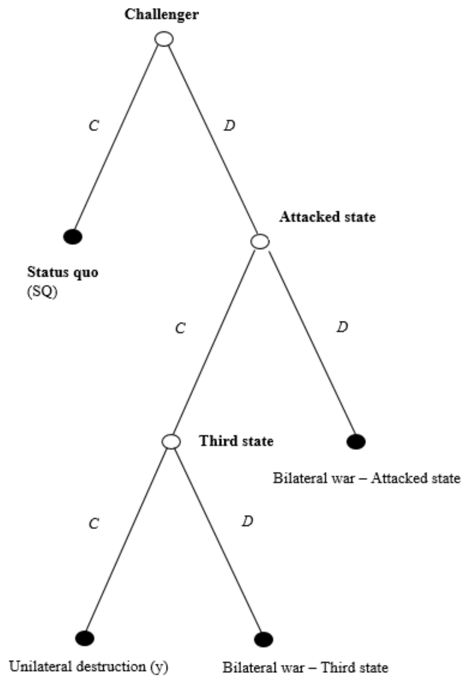

FIGURE 3: ORIGINAL NUCLEAR DETERRENCE MODEL

On the other hand, both the safeguard and the defender can deliberately choose outcome 'y'. The result 'y' can be chosen by both actors by cooperating in the game. They are deemed equal as the result is the same. The outcomes are identical as they result in the unilateral destruction of the defender. Moreover, there is no difference in their meaning in the preference orderings for both actors. For both the safeguard and conjointly the defender, the 'y'-outcome means that the defender will be unilaterally destroyed. There is no action required by either state.

METHODOLOGY

In order to determine the differences between the protégé mode and the safeguard model, a comparison is made between the two games which correspond with a non-nuclear protégé and a nuclear protégé. This means that is solely based game theory on a game theoretical analysis. An analysis as such allows the researcher to construct a theory in its purest form, as a formal model. No white noise, such as irrelevant historical data will interfere with the conclusion. Moreover, a game theoretical comparison allows for a clear and fast analysis of the differences between the games. Unlike historical analyses, game theory follows rational choice theory, thus allowing its games to predict the outcomes of future cases of equal games.

There are specific minimal requirements for these games in order to correctly define a game within game theory (Rusmussen, 2006). These requirements are the actors of the game, and the available actions and information to each player at each player at each decision point. These requirements are specifically intended to define and select usable right cases. For the protégé model the three actors are the challenger, the protégé and the defender. For the three-state deterrence game the three actors are the challenger, the defender and the safeguarding state. The available actions in the safeguard model are of the challenger, defender, protégé and safeguard are illustrated in figure 3.

In order to determine the results and actions of each game the method of backwards induction is utilised. Backwards induction is a bottom-up method that first determines the actions of the last player. It then works its way up to the top. Each node in the sequential game allows for the options of cooperating or defecting. If an actor rationally chooses the option of cooperating, the actor above will have a different set of options to select from than when the first actor would have chosen to defect. Backwards induction in this way allows the outcomes of the cases to be predetermined. The selected cases are based on their likelihood in the contemporary world system. The defined cases are the nominal deterrence case and the reluctant target case. These cases are elaborated on below.

CASE 1 – NOMINAL NUCLEAR DETERRENCE

Quackenbush's model

In Quackenbush's game it is presumed that the Challenger has the preference of unilaterally destroying the protégé. Note that the protégé in this game has no nuclear capabilities. The actor also prefers a bilateral war with the protégé over the status quo. This is mainly due to the fact that the protégé only has conventional armed forces. Moreover, most of the protégé's forces have probably been wiped out by use of the first nuclear strike. This leads to the preference of the bilateral war over the maintaining of the status quo. As the defender has nuclear capabilities in this game, the challenger prefers not to be struck by its nuclear retaliation. This causes the challenger to prefer the maintaining of the status quo over a nuclear retaliation by the defender (w).

The protégé above all prefers to maintain the status quo (x). If it were to be attacked by the challenger with a first strike by the challenger, it would prefer the defender to retaliate and disable the challenger (w). If the defender would not execute this retaliation, the protégé would prefer to start a bilateral war (b). It will send what is left of its conventional armed forces in order to deter the challenger from doing more damage and from annexing its resources. The least preferred outcome is its unilateral destruction (y).

Lastly, the defender in Quackenbush's model prefers to maintain the status quo. As a second preference it wants to see the defender make an effort before it will, as the usage of nuclear weapons comes with a large amount of political turmoil. The least preferred outcome, however, would be for the protégé to be annihilated unilaterally. These aforementioned motives determine the preference of this case orderings as follows:

$$[C]\ y > b_{pro} > x > b_{def} = mw$$

$$[P]\ x > b_{def} > MW > b_{pro} > y$$

$$[D]\ x > b_{def} = MW > b_{pro} > y$$

Original model

In the nominal nuclear deterrence preferences it is assumed that the challenger prefers unilateral destruction over maintaining the status quo, yet sees mutual destruction as the least preferred outcome. It also prefers a bilateral war with the weak protégé over the status quo.

The preference ordering of the protégé is in contrast with that of the challenger. The protégé prefers the status quo above all other options. The second preference is that of mutual destruction, while its own unilateral destruction is the least preferred. With the addition of the third safeguarding the states have to add the additional outcome of w to their preference ordering. Moreover, the preferences of the protégé have to be described corresponding to its motive. The defender prefers mutual destruction over unilateral destruction. However, the distinction between the outcomes w and z is relevant for the defender. It is assumed that the protégé prefers the retaliation by the defender, as it will alleviates the defender of having to maintain its strike back credibility.

The safeguarding state prefers the maintaining of the status quo above all. The defender is a state that has only disadvantages from the destruction of the protégé. The motives behind this preferences are exemplified by refugee crises, economic backlashes and political instability. It is evident that the safeguard prefers the mutual destruction above the unilateral destruction, as the safeguard wants to deter the challenger from attacking the defender. Similarly to the defender, the distinction between w and z is relevant to the safeguard. The disposition of the safeguard is that it prefers the retaliation by the defender. By not executing the nuclear arms operation, the safeguard's government stays clear from any political instability caused by its interference in another state's conflict. However, the defender does realize that a retaliation by the protégé is ineffective and will not be deemed credible by the challenger. In conclusion, the preference orderings of the three state actors are as follows:

[C] $y > b_{pro} > x > b_{def}$

[P] $x > b_{def} > b_{pro} > y$

[D] $x > b_{def} > b_{pro} > y$

CASE 2 – NONVIOLENT PROTÉGÉ

Quackenbush's model

The pacifistic defender case is a slightly different version of the nominal deterrence case. In this case the protégé prefers not to start a bilateral war at al. It prefers unilateral destruction over it. Several motives that can be used to explain this behaviour are that the challenger does not intend to annex the protégé's resources or that the protégé does not want to make its forces suffer more. The preference orderings of the other actors stay the same. This results in the following preference orderings:

[C] $y > b_{pro} > x > b_{def} = MW$

[P] $x > b_{def} > y > MW = b_{pro}$

[D] $x > b_{def} = MW > b_{pro} > y$

Original model

The second case is again an alternative version of the first case and analogous to the protégé case mentioned above. In this scenario the defender does not prefer executing a retaliation. It prefers not to strike back. All other preferences stay equal. This scenario is possible in reality as not all states will see the necessity to strike back, as their deterrence has failed. It is not unlikely that a state will not see the necessity to strike back on the innocent civilians of the challenger. Moreover, its nuclear arms systems can fail to operate, leading to no retaliation. Next to this, the safeguard has the same preference orderings as in the last case. The state prefers to maintain the status quo, as it yields a utility of 4. The retaliation by the defender is still preferred above the execution of a retaliation by itself. The unilateral destruction of the defender is the least preferred outcome, as it yields a utility of 1. The defender will still prefer a retaliation by the safeguard over a retaliation by itself. The rationale behind this is that if the defender chooses to not retaliate, it knows that the safeguard will instead. This leads to the following preference orderings:

[C] $y > b_{pro} > x > b_{def}$

[P] $x > b_{def} > y > b_{pro}$

[D] $x > b_{def} > b_{pro} > y$

ANALYSIS OF THE CASES

In this section the analysis of the two models is discussed. The two selected cases are each used as input for the extended nuclear deterrence game and the original nuclear deterrence game. Firstly, the standard analysis questions are introduced. Secondly, the case of the nominal nuclear deterrence case is discussed. Next to this, the case of the nonviolent defender case is elaborated upon.

ANALYSIS QUESTIONS

In order to be able to compare the cases to each other, several two are posed. The answers to these questions are answered both dichotomously, by the use of yes and no, and in a more elaborate manner. The latter resulting in the motivation behind the answer. These questions are as follows:

1. Has the challenger been deterred from executing a first strike on the defender?
2. By which actor has the challenger been deterred?

ANALYSIS – NOMINAL DETERRENCE CASE

QUACKENBUSH'S MODEL

In the nominal deterrence case the preference orderings of the challenger, defender and protégé are as follows:

[C] $y > b_{pro} > x > b_{def}$

[P] $x > b_{def} > y > b_{pro}$

[D] $x > b_{def} > b_{pro} > y$

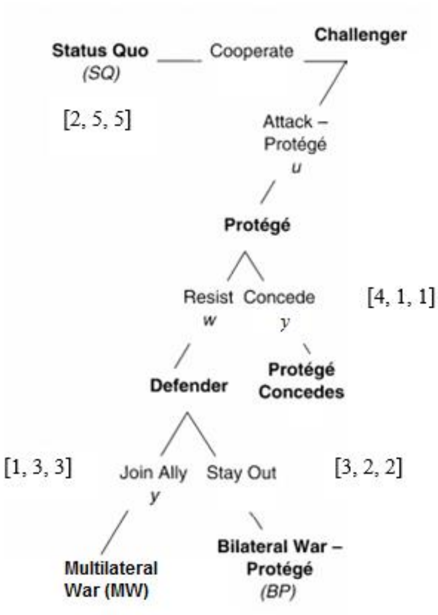

FIGURE 4: QUACKENBUSH'S GAME – PROTÉGÉ ATTACKED CASE 1

If the challenger would strike first on the protégé, the backwards induction for that tree will yield the following results. The game tree is illustrated in figure 4. Firstly, the defender has to choose whether to do a nuclear retaliation (b_{def}) or stay passive and let the protégé handle the fight himself with conventional warfare (b). The former yields a utility of 4, while the second yields a utility of 2. As the retaliation yields more utility, it will be the option chosen by the defender. Consecutively, the protégé has to select whether to let the defender to retaliate (bdef) or to concede (y). As the outcome of bdef yields more utility, the protégé will choose to defect and fight back. However, the protégé does not have the capability to fight back after a nuclear strike. This causes the protégé to be forced to select the option of conceding. This results in the outcome y in this decision node. Now, the challenger has to choose whether to cooperate, and maintain the status quo (x), or to defect and unilaterally destroy the protégé (y). The latter option yields the most utility and, thus, has the preference of the challenger.

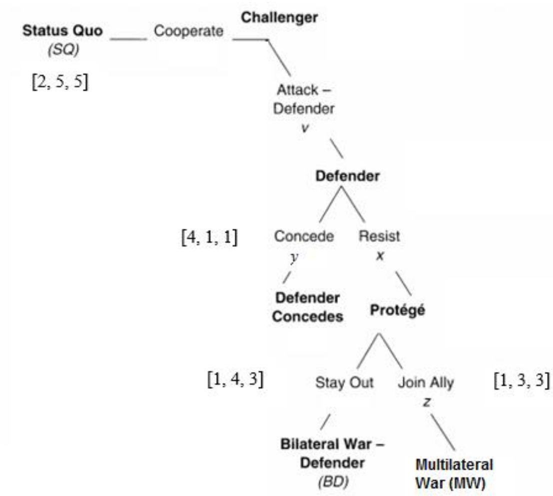

FIGURE 5: QUACKENBUSH'S GAME – DEFENDER ATTACKED CASE 1

On the other hand, if the challenger would strike the defender first, the backwards induction starts at the protégé. The game tree is illustrated in figure 4. The protégé has to choose whether to let the defender fight back alone (b_{def}), or to start a multilateral war (MW). The As the multilateral war yields more utility, the protégé will aid the defender in its battle with the challenger. Consecutively, the nuclear armed-defender chooses either the option of the multilateral war (MW) with the challenger or to fight back alone in a bilateral war (b_{def}). The defender chooses the multilateral war. This leads to the challenger having to determine whether to defect and start a multilateral war (MW) and being retaliated upon by the defender or maintaining the status quo (x). Evidently, the challenger embraces the option of maintaining the status quo as it yields more utility over being annihilated by the defender.

ORIGINAL MODEL
The preference orderings in the original model are as follows:

[C] $y > b_{pro} > x > b_{def}$

[P] $x > b_{def} > b_{pro} > y$

[D] $x > b_{def} > b_{pro} > y$

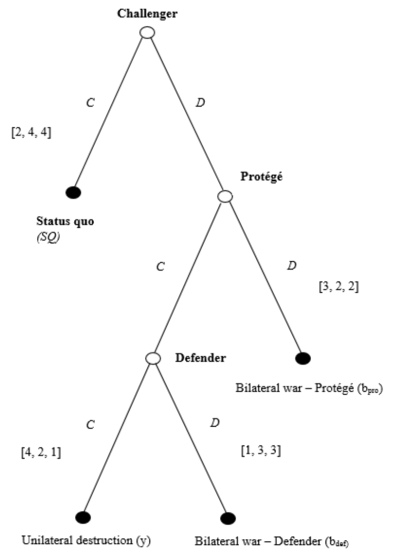

FIGURE 6: ORIGINAL GAME – PROTÉGÉ ATTACKED CASE 1

If the protégé is targeted by the challenger in the three-state variant, the defender has the options of either cooperating or letting the challenger unilaterally destroy the protégé, or to defect by retaliating onto the challenger. The corresponding utility values and the game are illustrated in figure 5. Cooperating yields a utility of 1. Defecting yields a utility of 2. The defender thus chooses to defect at that decision node and will retaliate. The protégé now has to choose between a retaliation by the defender and the retaliation via its own arsenal. Cooperating and defecting respectively lead to utilities of 3 and 2. The protégé thus chooses to let the defender retaliate instead. The challenger in turn has to determine whether to cooperate and maintain the status quo or to defect. Defecting will cause employ a first strike on the protégé, yet causes a retaliation by the defender. The challenger has thus been deterred by the defender.

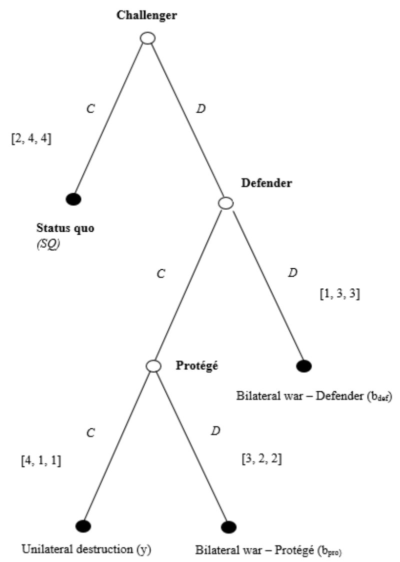

FIGURE 7 ORIGINAL GAME – DEFENDER ATTACKED CASE 1

On the contrary, the challenger can also strike the defender first. This is shown in figure 6. This would result in the protégé being the first actor to make a decision in the backwards induction analysis. The protégé has to choose whether to retaliate itself (z) or to neglect its ally and allow for the unilateral destruction (y) of the defender. The first outcome yields the protégé a utility of 2, while the second grants a utility of 1. The protégé, thus, prefers the outcome of z. The defender now has the option whether to allow for a conventional retaliation (z) by the protégé or its own retaliation that causes a mutual destruction (b_{def}). This results in the defender preferring the outcome bdef. Subsequently, the challenger has to adhere either the maintaining of the status quo (x) or the mutual destruction (b_{def}). The current status quo yields a utility of 3, while the mutual destruction yields a utility of 1. Accordingly, the challenger rationally maintains the status quo. This means that the challenger has been deterred by the threat of retaliation by the defender.

ANALYSIS – NONVIOLENT PROTÉGÉ CASE

QUACKENBUSH'S MODEL

In the Quackenbush model for the nonviolent defender the preference orderings of the challenger, protégé and defender are defined as follows:

$$[C]\ y > b_{pro} > x > b_{def} = MW$$

$$[P]\ x > b_{def} > y > MW = b_{pro}$$

$$[D]\ x > b_{def} = MW > b_{pro} > y$$

If the situation arises in which the challenger determines to strike the defender first, the backwards induction process again starts at the protégé. This is illustrated in figure 7. The protégé has to decide whether to retaliate via conventional warfare and start a multilateral war (w) or to allow the defender to retaliate on its own and start a bilateral war (b_{def}). The first option yields a utility of 2, while the latter yields a utility of 3. Therefore, the protégé chooses to take on a passive role. The defender has to promptly establish whether it prefers to retaliate and cause a bilateral war (b_{def}) or to stay nonviolent and allow for its unilateral destruction (y). The former yields a utility of 3. The unilateral destruction outcome yields a utility of 1. Therefore, the defender chooses to retaliate and establish a bilateral war. Ultimately, the challenger has to decide whether to retain the status quo (x) or to start a bilateral war with the defender (b_{def}). For the challenger the status quo yields a utility of 3, while the bilateral war-defender yields a utility of 1. Therefore, the challenger will maintain the status quo. This result causes the challenger to be successfully deterred by the defender.

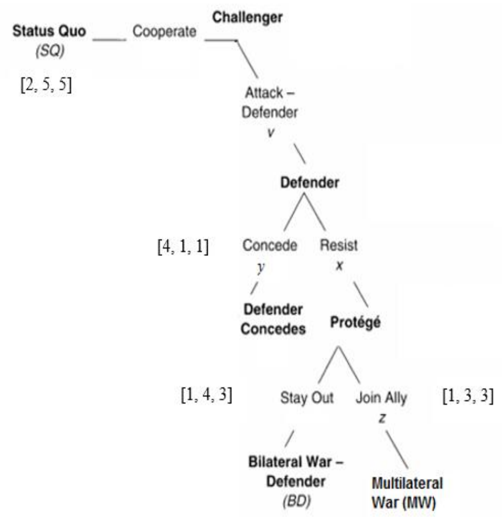

FIGURE 9: QUACKENBUSH'S GAME – DEFENDER ATTACKED CASE 2

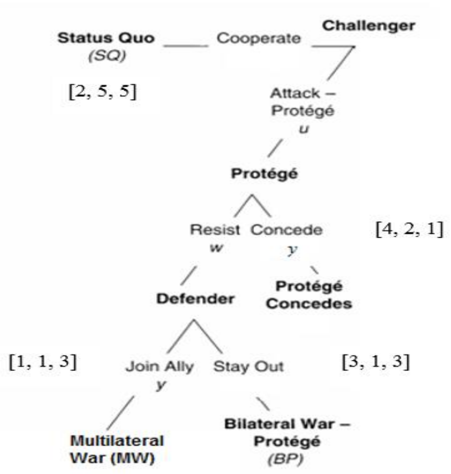

FIGURE 8: QUACKENBUSH'S GAME – PROTÉGÉ ATTACKED CASE 2

In the situation where the attack would execute a first strike on the protégé, the backwards induction starts at the third decision node of the right side of the game tree. This is illustrated in figure 8. Here the defender has to decide whether to retaliate onto the challenger (w) and start a multilateral war or to stay passive and allow the protégé to fight a bilateral war with the challenger (b_{pro}). The defender chooses to retaliate and start a multilateral war as this yields him utility of 4 instead of 2. This leads to the protégé-actor that has to distinguish between fighting a multilateral war (w) and allowing for its own unilateral destruction (y), and therefore cooperating. The latter option yields the most utility towards the protégé. This causes the protégé to surrender to the challenger and allow for the outcome of unilateral destruction. The challenger now has to determine whether to gain a utility of 4 by unilaterally destructing the protégé or gaining a utility of 3 by maintaining the status quo. The challenger will rationally choose to unilaterally destruct the protégé. This means that protégé will unilaterally be destructed as deterrence has failed.

ORIGINAL MODEL

In this case for the original model the preference orderings are determined as follows:

[C] $y > b_{pro} > x > b_{def}$

[P] $x > b_{def} > y > b_{pro}$

[D] $x > b_{def} > b_{pro} > y$

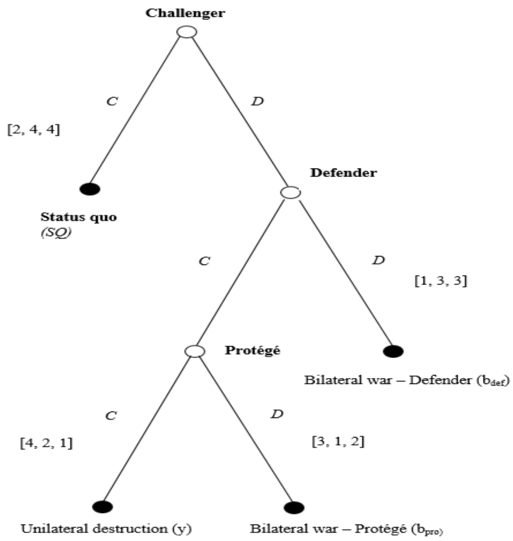

FIGURE 10: ORIGINAL MODEL – PROTÉGÉ ATTACKED CASE 2

In this case the protégé is struck first. This is depicted in figure 9. Note that the protégé does not have nuclear arms. The defender at the third node has to decide whether to retaliate onto the challenger (w) or to stay passive and allow for the unilateral destruction (y) of the protégé. The retaliation allows for more utility. Thus, the defender retaliates onto the challenger. The protégé now decides whether to allow the defender to retaliate (w) or to retaliate himself (z). The retaliation by the defender yields the protégé more utility. Therefore, it chooses to stay passive itself and let the defender retaliate in its place. Finally, the challenger has to decide whether to maintain the status quo (x) or to cause mutual destruction (w) via a nuclear retaliation by the defender. The maintaining of the status quo yields a utility of 3, while the yielded utility of mutual destruction by the defender is 1. Therefore, the challenger decides to retain the status quo. This means that the challenger has been deterred from striking upon the protégé in the case of the nonviolent defender.

In the event of a first strike upon the defender, the protégé is the first actor to determine its actions. This is illustrated in figure 10 for this case. The protégé chooses between retaliating (z) and allowing for the unilateral destruction (y) of the struck actor, the defender. At this decision node, the protégé-actor chooses to allow for the unilateral destruction of the defender, as it yields a utility of 3. The retaliation yields only 1. This causes the defender to have to choose between retaliating himself (w) and permitting its own unilateral destruction (y). The defender chooses to defect, and retaliate (w). The challenger now has to determine whether to maintain the status quo (x) or to tolerate mutual destruction. The challenger chooses to maintain the status quo, as it grants a utility of 3. This is more than the mutual destruction-outcome. This causes the challenger to recognize the low possible utility and permits its deterrence via the threat of a mutual destruction.

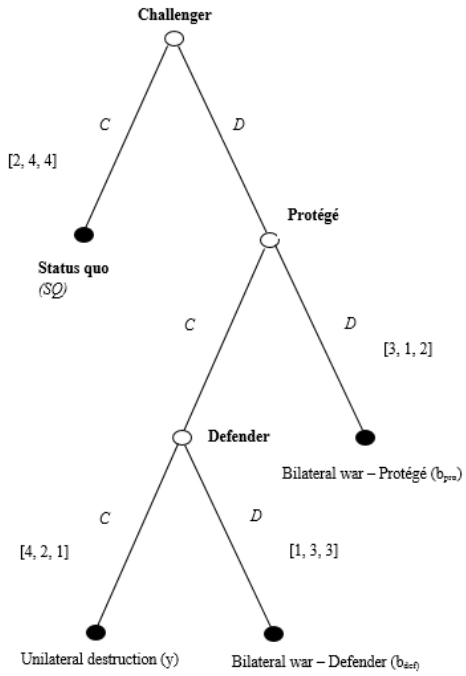

FIGURE 11: ORIGINAL MODEL – DEFENDER ATTACKED CASE 2

RESULTS OF THE ANALYSIS

Table 1 depicts the results of the analysis in a schematic manner. The columns show the case that is described. The rows illustrate the subject of the content of the cells.

	Quackenbush model case #1.1	Quackenbush model case #1.2	Quackenbush model case #2.1	Quackenbush model case #2.2	Original model case #1.1	Original model case #1.2	Original model case 2.1	Original model case #2.2
TARGETED STATE	Protégé	Defender	Protégé	Defender	Protégé	Defender	Protégé	Defender
Deterrence successful?	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Deterring actor	N/A	Defender	N/A	Defender	Defender	Defender	Defender	Defender
Outcome if SQ is not considered	w	w	y	b_{def}	w	W	w	w

TABLE 1: RESULTS OF THE ANALYSIS

CONCLUSION

The aim of this thesis is to determine the difference in the outcomes of nuclear deterrence between Quackenbush's extended deterrence game and the newly constructed 'original nuclear deterrence game'. This questions whether the success rate of deterrence is increased by allowing for this action after the protégé has conceded. Firstly, this provides an original formal model that includes this action within its game tree. Secondly, it shows that the nonviolent protégé is protected by the defender in this game tree. This permits us to examine the results concerning the most viable target of the challenger and the reliability of the alliance.

The analysis shows that the challenger's target in the case of a nonviolent protégé differs per model. This result that is found in the fifth column of table 1. When using Quackenbush's model for the determination of the outcome of the nonviolent protégé case, the protégé causes the deterrence to fail. Quackenbush's model allows an idiosyncratic feature to be reflected in the outcome. It shows that, despite the preference of the defender to retaliate onto the challenger, the preference of the protégé to concede cancels out the option of a retaliation. Hence, the challenger is not deterred in this game. The analysis shows not only whether to strike, but also whom. In this case, it is the nonviolent protégé.

In the original model, the same case has a different result. Though the protégé prefers its own unilateral destruction, the defender has the option to retaliate. This is reflected in the outcome of the game. The result is the deterrence of the challenger by the threat of a mutual destruction from the defender. This is turn is in great contrast with the result of Quackenbush's model. In this case, the challenger is deterred by the defender from attacking both the defender and the protégé. This results in that the challenger is deterred in every case in the original model.

All other comparable analysed cases have the maintaining of the status quo as their outcome. It can be stated that the odds of the challenger utilising a nuclear first strike on either state have therefore decreased in the original game compared to Quackenbush's game for these two cases. Therefore, it can be stated that the including of the option of retaliation for the defender in the case of a nonviolent protégé will result in lower odds for an act of aggression by the challenger.

DISCUSSION

The conclusion shows that there is a significant difference in the way nuclear deterrence is performed when the third state can retaliate at all times. It is evident that the addition of such a feature changes the dynamic of the nuclear deterrence game. The defender can take over the role of the deterring actor from the protégé. This is proven in the nominal nuclear deterrence case. An example of this is the reluctant protégé case in the analysis.

The defender's strike back credibility and nuclear collateral damage appear to be of use in the original three-state nuclear deterrence game. Simultaneously, the challenger and the protégé have to determine whether the third state will function as the protégé's defender. The challenger requires this information as it can be vital to its first-strike strategy. The nonviolent protégé case displays that the success of the deterrence is entirely dependent on the defender. The protégé on the other hand changes its preference ordering in the case of a credible defender. Thus, the third-state strike back credibility is evidently an important variable in the three-state nuclear deterrence game.

The impact of the knowledge of these concepts might have an impact on the contemporary international relations theories. The concepts fit seamlessly into the school of neorealism. Neorealism assumes states to be rational, act in a state of international anarchy and self-helping. Moreover, the striving for relative gains is displayed by the concept of nuclear collateral damage. If the nuclear collateral damage is deemed too great by the defender, it will retaliate onto the challenger. In other words, the defender does not allow the challenger to shift the balance of power too much towards to itself by annihilating a great power and making another one dysfunctional. Therefore, the defender will even out this advantage and retaliate upon the challenger.

Clearly, as the abovementioned concepts are original in deterrence theory, the concepts are to be researched more extensively in the future. The concepts of nuclear collateral damage, third-state strike back credibility and the new role of the defender in the nuclear deterrence game have to be elaborated upon. Nuclear collateral damage is to be researched as it can mean the difference between a nuclear conflict between two nuclear states and three nuclear states. The amount of collateral damage can essentially determine the pace at which an annihilated region modernize itself again. The third-state strike back credibility is influenced by the aforementioned concept. Nonetheless, it is probable that there are more factors that influence the tendency to defender. In international relations, knowledge of this subject can again be the disparity between a conflict with two or with three states. Overall, the significance of the multipolar world's extra state can impact the theory of nuclear deterrence.

REFERENCES

Arrow, Kenneth J. (1963): Social Choice and Individual Value, 2. Auflage, New York: John Wiley and Sons. https://www.unihamburg.de/onTEAM/grafik/1164287680/Arrow's%20Axioms.pdf

Bueno de Mesquita, Bruce. 1981. The war trap. New Haven, CT: Yale University Pre

Frank C. Zagare. (2014). A Game-Theoretic History of the Cuban Missile Crisis. Economies, 2(1), 20-44.

Huth, P. (1988). Extended deterrence and the prevention of war. New Haven ; London: Yale University Press.

Keohane, R. (1986). Reciprocity in international relations. International Organization, 40(1), 1-27.

Kraig, M. (1999). Nuclear Deterrence in the Developing World: A Game-Theoretic Treatment. Journal of Peace Research, 36(2), 141.

Kilgour, D. M., and F. C. Zagare. "Uncertainty and the Role of the Pawn in Extended Deterrence." Synthese 100.3 (1994): 379-412.

Lebow, R., & Stein, J. (1995). Deterrence and the Cold War. Political Science Quarterly, 110(2), 157-181.

Mearsheimer, J. (1984). Nuclear weapons and deterrence in Europe. International Security, 9(3), 19-46.

Mearsheimer, J. (2001). The tragedy of great power politics (The Norton series in world politics). New York, NY: Norton.

Michener, H. Andrew, Choi, Young C., & Dettman, David C. (1986). Stability by deterrence in cooperative non-side payment n-person games. Journal of Economic Behavior & Organization,7, 375. (05-04)

Morgan, P. (2003). Deterrence Now. Cambridge: Cambridge University Press.

Pellikaan, H. (2017). A new introduction to game theory with applications in international relations. Unpublished manuscript, Department of Political Science, Leiden University, Leiden, The Netherlands.

Powell, R. (1989). Nuclear deterrence and the strategy of limited retaliation. American Political Science Review, 83(2), 503-520.

Powell, Robert. "Bargaining in the Shadow of Power." Games and Economic Behavior 15.2 (1996): 255-89.

Quackenbush, S. (2006). Not Only Whether but Whom: Three-Party Extended Deterrence. The Journal of Conflict Resolution, 50(4), 562-583.

Rasmusen, E. (2007). Games and information: An introduction to game theory (4th ed.). Malden, MA : Blackwell.

Van der Meer, S. (2016). States' Motivations to Acquire or Forgo Nuclear Weapons: Four Factors of Influence. Journal of Military and Strategic Studies, 17(1). Retrieved from http://jmss.org/jmss/index.php/jmss/article/view/585

Waltz, K. (1964). The Stability of a Bipolar World. Daedalus, 93(3), 881-909.

Waltz, K. (1979). Theory of international politics. Long Grove, I.: Waveland Press.

Waltz, K. (1981). The spread of nuclear weapons: More may be better (Adelphi papers ; 171. 822386682). London: International Institute for Strategic Studies.

Wagner, R. (1991). Nuclear deterrence, counterforce strategies, and the incentive to strike 1st. American Political Science Review, 85(3), 727-749.

Werner, Suzanne. "Deterring Intervention: The Stakes of War and Third-Party Involvement." American Journal of Political Science 44.4 (2000): 720-32. Web.

Saperstein, A. (1991). The "Long Peace" RESULT OF A BIPOLAR COMPETITIVE WORLD? The Journal of Conflict Resolution (1986-1998), 35(1), 68.

Stein, J. (1987). Extended deterrence in the Middle East. World Politics, 39(3), 326-352.

Snyder, G.H. (1961). Deterrence and Defense. Princeton UP.

Szilagyi, M. (2012). Analytical solutions of N-person games. Complexity,17(4), 54-62. (05-04)

UN General Assembly. (2016). Taking forward multilateral nuclear disarmament negotiations. Retrieved October 08, 2017, from http://reachingcriticalwill.org/images/documents/Disarmament-fora/1com/1com16/resolutions/L41.pdf

Wu, Samuel S.G. (1990). To attack or not to attack - a theory and empirical assessment of Extended Immediate Deterrence. Journal of Conflict Resolution, 34(3), 531-552.

Zagare, F., & Kilgour, D. (2006). The Deterrence-Versus-Restraint Dilemma in Extended Deterrence: Explaining British Policy in 1914 1. International Studies Review, 8(4), 623-642.

APPENDIX – QUACKENBUSH'S EXTENDED-DETERRENCE MODEL