A Double-Edged Sword:
Nuclear Deterrence and Nuclear Caution

Ali Sanaei
University of California, Berkeley

Abstract
Most empirical studies of the effects of nuclear weapons treat these weapons as a binary variable, but there are strong reasons to think that the effects of a state’s nuclear arsenal are conditioned by that state’s policies. I argue that the effect of nuclear weapons in non-existent disputes is determined by what I call nuclear risk: the overall likelihood that these weapons may be used without authorization by political leaders. Using a formal model, I hypothesize that higher nuclear risk leads to greater deterrent power, but it also makes the leaders of the nuclear state more cautious if their deterrence fails. I test these two hypotheses using the Correlates of War data. I first measure nuclear risk using a simple index, and then propose a novel technique for imputing risk. Both hypotheses are borne out by empirical results.
Introduction

How do nuclear weapons change the bargaining powers of states? We recently passed the 70th anniversary of the first use of nuclear weapons but we still have not found satisfactory answers to this question. If we do not know what happens the day after a state acquires nuclear weapons, we can neither have an understanding of why states might be interested in acquiring these weapons, nor can we navigate the routes that lead to deterring nuclear proliferation. This question pertains to a subgame of almost any policy game regarding nuclear weapons.

The answer is far from trivial. States might seek nuclear weapons to prevent a doomsday scenario, or to assure they can survive. The relevance of nuclear weapons for the more commonplace militarized conflicts is not clear, especially once we consider that much effort has been spent on making the threat of the use of nuclear weapons credible and that these weapons have not been fired during any conflict since 1945.\textsuperscript{1} There are opposing theoretical expectations and ambiguous empirical evidence (Gartzke and Jo 2009).

I argue that to understand the effect of nuclear weapons we need to look beyond a dichotomous measurement that only shows whether a given state possesses nuclear arsenal or not. What gives a nuclear state deterrent power is the chance that these weapons might actually be used in a militarized conflict. I will show that this deterrence comes at a great cost. A nuclear state that takes riskier postures to achieve more deterrence, or for other reasons has poorer control over its nuclear arsenal, can be expected to become more circumspect if its deterrence fails. The reason is that the political leaders of a nuclear state with higher risk are more apprehensive of their own nukes being used in the fog of war.

I use low-level militarized interstate disputes (MIDs) to empirically test the predicted effects of nuclear risk, similar to the work of Schultz (1999) and Horowitz (2009). To measure risk, an all-inclusive variable of risk should be created that, at any given time, shows how likely each nuclear-armed state is to turn a conventional conflict into a nuclear one. Measuring this variable directly is virtually impossible and I will try two different indirect avenues for measuring it. First, I create an index of nuclear risk, based on the idea that certain characteristics should lead to higher risk. Second, I try to estimate nuclear risk as a latent variable in a Bayesian estimation. This approach allows us to use our qualitative understanding of risk in order to obtain better imputed values. The results of both methods are in line with each other and corroborate the theoretical expectation that higher risk leads to higher caution when deterrence fails.

The paper is organized as follows: a brief review of the literature; presentation of the model and its empirical implications; research design and empirical results; and conclusion.
The effect of nuclear weapons

Despite a celebrated legacy of scholarly work on nuclear deterrence in the Cold War era, some of the questions that have bearing on today’s policy problems do not have satisfactory answers. In this section, I will briefly review opposite theoretical expectations and the existing empirical results.

There is a general agreement among scholars that nuclear weapons, due to their prodigious destructive power, change the calculations of warfare in ways that are qualitatively different from the way conventional weapons enter into states’ calculations (Betts 1987; Jervis 1989; Mueller 1988; Waltz 1990). The qualitatively different effect of nuclear weapons, however, has been interpreted in two opposite ways.

Some scholars—“optimists” in this context—argue that the cost of being attacked by nuclear weapons is so vast that these weapons induce opponents to restrain their behavior (Waltz 1990). Powered by an inductive reasoning from observing the happy ending of the Cold War, and nuclear non-use in general, the optimists view nuclear weapons as benign forces that add to the stability of the world (Mearsheimer 1990).

Optimists argue that a limited second-strike capability and a secure command and control is enough to provide a nuclear deterrent. Because nuclear wars are not a attractive option to anybody, the shadow of a nuclear war induces restraint in conventional conflicts because the parties—presumably both the states that possess nuclear weapons and those who do not—do not want to escalate the conflict to a point of no return. Nuclear weapons are credited with creating the longest peaceful period between major powers in modern history (Waltz 1995).

Pessimists identify several problems with this line of argument. One concern is the assumption that states can have an effective control over their nuclear arsenal. Sagan argues that this is no small feat and even the United States has had a history of mishaps that could have led to disastrous outcomes (Blair 1993; Sagan 1995). In addition to accidents, pessimists also worry about belligerence of nuclear states because of their weapons (Feaver 1992; Kapur 2005).

While the theoretical debate between the pessimists and optimists continues (Sagan 2011), an empirical literature has emerged that tries to assess the various hypotheses given the post-World War II behavior of nuclear and non-nuclear states. But the empirical studies have also found contradictory results (Beardsley and Asal 2009). Two prominent methodological problems have made it difficult to reach an empirical consensus: the limited number of observations and the endogeneity problem. The former exists due to the limited number of nuclear-armed states and the fortunate fact that the weapons have not been fired since 1945. The latter arises because the possession of nuclear weapons might be endogenous to conflict. In other words, dynamics in rivalries and power relations that influence interstate conflict can be seen both as causing proliferation, as being affected by proliferation.

Some studies show that possessing nuclear weapons do not have an appreciable effect on crisis occurrence (Beardsley and Asal 2009; Gartzke and Jo 2009; Sechser and Fuhrmann 2013). Others
have found support for the stability-instability paradox, which argues that nuclear weapons provide improve strategic stability, but encourage more risk-taking behavior in lower intensity disputes (Rauchhaus 2009).

Nuclear weapons have been found to have provide deterrence (Beardsley and Asal 2009) and extended deterrence to formal allies (Huth 1990; Fuhrmann and Secher 2014). Sobek, Foster, and Robison (2012) show that states are more likely to be attacked before weapons are acquired but once acquired, the chance of being attacked drops considerably.

In a recent re-evaluation of existing evidence, Bell and Miller (2013) find that there is little evidence that nuclear dyads are associated with lower risks of war or higher risks of MIDs, which calls into question both pessimist and optimist arguments as well as the stability-instability paradox. In asymmetric dyads, they find an nuclear states are more likely to target weaker opponents.

The empirical literature has by and large treated nuclear weapons in a dichotomous way: some states possess such weapons and some do not. Such dichotomous treatments ignores analyses that suggest that the way nuclear weapons are managed, the way difficult matters of command and control are resolved, and the posture that a nuclear-armed state takes greatly influence the effect of nuclear weapons (Feaver 1992). Perhaps part of the reason for this dichotomous treatment of nuclear weapons has been lack of data.

Two prominent exceptions are the work by Horowitz (2009) and Narang (2010, 2014). Horowitz examines how the experience of nuclear states affects reciprocation of militarized disputes and Narang makes a distinction between the nuclear weapon states based on the postures that they take.

**Theory**

Schelling’s famous “the threat that leaves something to chance” is a way to make an unthinkable disaster possible without committing to carrying it out after any particular red line has been crossed (Schelling 1960, 187). In the same vain, Powell (1990) shows that one way to make a nuclear threat credible is to tie the use of the weapons to something truly exogenous. An exogenous trigger is required for an actor to utilize the threat of an option she would never intentionally undertake. This was the big problem with which early nuclear policy experts had to wrestle when the superpowers acquired second strike capabilities. The other solution is the step-by-step, or limited response strategy. As Powell shows, these are similar in that they are attempts at bridging the gap between no response and total annihilation in a way that makes the threat of the use of nuclear weapons credible: one by using a mixed strategy and the other by using small increments that turn a dichotomy into a continuum.

The problem of leaving an important decision at the mercy of an exogenous trigger or creating policies that prescribe gradually increasing use of nuclear weapons is not limited to the Cold War dynamics between the superpowers. It arises because political leaders have inhibitions against using
nuclear weapons; and the non-use of nuclear weapons has only increased this inhibition (Tannenwald 2005). This inhibition means that leaders can make better use of their nuclear arsenal by reducing their own control over their weapons. In a conflict with at least one nuclear-armed state, as long as the opponents stay firm, there exists a non-zero likelihood of escalation to a nuclear conflict. This likelihood is higher when leaders rely more on chance and increases as the conflict is escalated in conventional levels, even when the weapons are tightly controlled. Managing nuclear weapons, which includes managing the exogenous chance that they will be used, is an important part of nuclear policy of any state that acquires nuclear weapons.

The “arrangement of facilities, personnel, procedures, and means of information acquisition, processing, and dissemination used by a commander in planning, directing, and controlling military operations” is referred to as command and control (Bracken 1983). From the inception of nuclear weapons, states have treated them as an exceptional category of weapons that requires its own command and control. For example, until the Korean War, the United States military did not have access to the nuclear material of atomic bombs which was handled by a civilian agency and was to be provided to the armed forces—which possessed the non-nuclear material of atomic bombs—by direct orders from the president.

Nuclear command and control (NCC) has taken various shapes at different times and by different nuclear-armed states. NCC depends on a host of issues. On a technical dimension, it depends on the number, type, and methods of delivery of nuclear warheads, the electronic early warning systems, and the computational and telecommunication capabilities of the nuclear-armed state. On a political dimension, it depends on organizational culture, civil-military relationships, and domestic institutional stability. On a strategic dimension, it depends on the types of threats that a state considers high on the list of potential threats.

With nuclear weapons, there is always a risk even in low-level types of military conflicts. A score of incidents can inform any overly optimistic observer that even the United States, which probably has the most reliable technology at the service of its NCC, has had narrowly avoided nuclear accidents multiple times (Betts 1987; Sagan 1995). A common theme in all the accidents that nearly happened is that the higher the level of alert, the less likely it is to dismiss what is ordinarily dismissed and the closer the possibility of a catastrophe.

The events of November 5, 1956 provide an illustrative example. At the time, the Soviet forces were involved in Hungary and there was also heightened alert because of the Suez Crisis where the British and French forces were deployed. Within a short span it was signaled that one hundred Soviet MiGs had been flying over Syria, a high-altitude British Jet was shot down over Syria, and that unidentified jets were flying over Turkey. Fortunately no action was taken before it was realized that the jets over Turkey were a flock of geese, the British jet had come down because of its own technical malfunctioning, and the 100 jets over Syria were a few jets that were escorting the Syrian president on his flight back from Moscow! That these independent signals were assumed to be
correlated and that interpretations erred on the side of exaggeration of the threat was attributed to
the heightened alert level at the time (Bracken 1983).

This and more than a handful of other examples show how the combination of cognitive bias
at the time of war and normal technological errors can set the stage for disastrous outcomes. The
point is not such disastrous outcomes have large probabilities, but it is that they may happen with
a non-zero probability and this non-zero probability is affected by NCC.

The optimal chance of the disaster

The main goal of NCC is to assure that nuclear weapons can be used if needed (according to the
state’s strategy), in a timely fashion, and also to prevent the weapons from being used otherwise.
Any NCC is an answer to what Feaver has called the always/never dilemma: always ready to be
used and never used without authorization (Feaver 1992).

According to Feaver, if a state is more concerned with timely responses and uncertain about
preserving its chain of command during a war, it is more likely to sacrifice the ‘never’ side for the
‘always’ side and choose a delegative NCC. On the other hand, if a state suffers from intermittent
spasms in its civil-military relationships, it would find it more prudent to assuage the ‘never’ concerns
by giving up some of the ‘always’ concerns, and choose a more assertive NCC: one in which the use
and deployment of nuclear weapons requires express civilian authorization.

Important in understanding the always-never tradeoff is that the Pareto front of this tradeoff
depends on a number of factors such as military strategy, and technological sophistication of the
state. For example, a state with a defensive posture that has advanced weapons-delivery technology
and access to reliable radar systems and satellite communications can afford to use a more assertive
control of its weapons even in the face of time-urgent threats. So, even similar NCCs might result in
different chances of escalation of a conventional conflict to a nuclear one. Narang (2014) considers
both France and post-1998 Pakistan as ‘asymmetric escalators’ in their nuclear posture, but the
risk of escalation to a nuclear war is different. As much as the French government might want to
emphasize their asymmetric escalation policy, it strains credulity to believe that they would use
nuclear weapons in small conflicts, or even in a a French version of the Falkland War. Similarly,
Nixon failed in his wielding of the ‘atom card’ despite his expectation that it would bring a fast end
to the war in Vietnam, ostensibly because it was not deemed credible (Ellsberg 2009).

The likelihood that nuclear weapons may be used depends on more than just the nuclear com-
mand and control. A myriad other factors affect how risky a state’s nuclear weapons really are. I
define nuclear risk as the latent variable that measures the chance that nuclear weapons may be
used in a time of conflict without explicit authorization by the political leadership.
The effect of nuclear risk

We are interested in how the effect of varying degrees of nuclear risk on the behavior of states in a MID. Figure 1 illustrates a simple game that shows the strategic interaction of two states during a MID.

There are two players: $S_1$ (initiator) and $S_2$ (target). At the start of the game, $S_1$, which is a non-nuclear state, can either accept the status quo or initiate a threat. The $S_2$, which is a nuclear-armed state can either concede the prize, valued at $v$, or reciprocate the threat. If the threat is reciprocated, $S_1$ can either back down and pay an audience cost (domestic and International) of $-a_1$, or stand firm and materialize the threat. If $S_1$ backs down, $S_2$ gains a payoff of $a_2$. If the threat is materialized, nature determines whether the war will be conventional (with probability $1 - \rho$) or nuclear (with probability $\rho$). In other words, $\rho$ is $S_2$’s level of nuclear risk.

If a conventional war happens, $S_i$ ($i \in \{1, 2\}$) has an expected payoff of $w_i$. If a nuclear war happens, $S_i$ has a payoff of $-C_i$ where $C_i$ is assumed to be large compared to other parameters ($w_i$ and $v$). This setup means that the stakes are small enough that even the nuclear-armed $S_2$ is not willing to use nuclear weapons.

It is clear that it is in the interest of the nuclear-armed state to maintain a very small $\rho$ but pretend that $\rho$ is quite high. To capture this strategic aspect, we assume that only $S_2$ knows the true value of $\rho$. What $S_1$ knows at the beginning of the game is the probability distribution of $\rho$.

The Bayesian perfect equilibria of the game are studied in the Appendix A. It is shown in the appendix that if the cost of a nuclear war is large, $S_1$’s expected payoff from initiation is decreasing in $E[\rho]$, and if $S_1$ initiates, $S_2$’s expected payoff from reciprocation is also decreasing in $\rho$. The reason that one of the comparative statics is with respect to $E[\rho]$ and the other is with respect to $\rho$ is that $S_1$ does not know $\rho$ but $S_2$ does.

This shows that better deterrence is one of the results of an increased nuclear risk. We expect higher levels of nuclear risks to bring down the overall expected utility of initiating a conflict, hence resulting in lower probability of MID initiation. But the same risk that brings about deterrence has another effect after an MID has been initiated. When a nuclear-armed state is the target of an MID, a similar logic holds that conditioned on being targeted, the leaders who have poorer control over their weapons are less likely to reciprocate the dispute. This time, it is the leaders of the target state that are apprehensive of an unwanted nuclear disaster. The fear of a nuclear fait accompli makes leaders less likely to reciprocate a militarized dispute when nuclear risk is high.

**Hypothesis 1:** The higher the risk of escalation of conflicts to a nuclear war by a nuclear state, the lower the probability of that state being the target of a militarized dispute.

**Hypothesis 2:** Compared to a nuclear-armed state with more tightly managed weapons, a nuclear-armed state that has higher risks of its nuclear weapons being used is less likely to reciprocate an MID.
Research Design

The goal is to test hypotheses 1 and 2. In an observational setting, these hypotheses lend themselves directly to a particular setup, similar to other works studying various covariates of conflict initiation and reciprocation; most notably the dyadic studies of the democratic peace. For example, Maoz and Russett (1993) is concerned with all dyad-years of countries and existence of disputes is their dependent variable, whereas Schultz (1999) looks at dyad-years where a dispute is initiated, and takes reciprocation of a dispute as dependent variable.

The similarity of the dependent variable to that of a considerable number of scholarly works make this effort straightforward. There are two problems that have to be addressed. One is the selection problem that has been largely ignored in previous studies, and the other is measuring nuclear risk. In this section I discuss the research design assuming we do not have any measurement problems and then dedicate the following section to the problem of measuring the nuclear risk variable.

Setting the stage

The dependent variables are dispute initiation (initiation) for testing Hypothesis 1 and dispute reciprocation (reciprocation) for testing Hypothesis 2. For initiation (hereinafter first stage), I look at all possible dyads of states after the Second World War (until data is available). The Data is available annually, and the unit of observation is directed dyad-year of states.

The dyads are directed because we are interested in who initiates any dispute. So for a given pair of states in year $t$, there are two observations. Side $A$ is the potential initiator. Side $B$ is the potential target.

For reciprocation (hereinafter second stage), I look at disputes that have already been initiated. So, the unit of analysis is militarized interstate dispute.

In the first stage, I look at all dyads, not the so-called “relevant” dyads because of the possible biases that might result from systematically not looking at a part of the data (Lemke and Reed 2001).³

Data

The MID data set is obtained from EUgen software (Bennett and Stam 2000), which relies on the Correlates of War (COW) project for most of the variables (Singer and Small 1994). The COW project provides data on the initiation and reciprocation of MIDs. It distinguishes between four types of revisions that are at the root of the conflict: territory, policy, regime/government change, and “other”. It provides primary and secondary revision issues for the initiator and the target. Here, I only look at the primary revision of the state initiating the conflict. Moreover, COW provides data on the level (severity) of the dispute which I ignore in the present analysis.
In the second stage, I include dummies that control for the revision issue, similar to Schultz 1999 and Horowitz 2009. In the first stage, however, controlling for the issues would be tantamount to using outcome as an explanatory variable.

**Explanatory variables**

A number of variables are suggested in previous works that I include here, in both stages, to avoid omitted variable bias. Notice that the main explanatory variable is correlated with all of the following because it has elements of strategic setting, balance of power, and institutional strength. To make sure that the results are robust, I conduct the analysis with a number of different specifications.

The explanatory variable of interest is the risk of nuclear weapons being used, which I refer to as nuclear risk (shown by $\rho_A$ and $\rho_B$ for sides A and B). Although our hypotheses are only concerned with $\rho_B$. Nuclear risk is zero for nonnuclear states and varies for different nuclear states (and over time). This variation is of interest, not the fixed effect of having nuclear weapons. For this reason, I include dummy variable indicating possession of nuclear weapons by states on both sides. I also include a jointnuke variable which is the interaction of these dummies showing both states in a dyad-year have nuclear weapons. Controlling for these variables separates the effect of the baseline effect of nuclear weapons from the effect that is due different values of nuclear risk.

The models also include splines of peace-years between dyads (Beck, Katz, and Tucker 1998), dyadic satisfaction (Signorino and Ritter 1999; Bennett and Stam 2004), and relative balance of power ratio of the two states as coded by COW. In the second stage, as mentioned above, I also include indicators of issue types as suggested by Schultz (1999), because different issues are valued differently and have possibly different rates of reciprocation, and possibly different risks of ‘something going wrong’ in the case of reciprocation. To control for the effect of the democratic peace, I include indicators of democracy for both sides. I use a dichotomous version of the polity score for democracy: states with a polity score greater than 6 are coded as democratic (Marshall and Jaggers 2011).

Finally, I also include a dichotomous variable that shows whether the dispute happens during the Cold War or not. A bipolar world is expected to be different from a unipolar or multi-polar world and we can expect the rate of reciprocation to be different during the Cold War and after it. Since countries enter the nuclear-armed club at different times, absence of this might potentially bias our results. This variable takes the value of 1 if the dispute happens after 1988.

**Estimation**

The dependent variables in both stages are dichotomous and I use probit regressions for all the analysis. Assume that the difference in the utility of initiating and not initiating an MID is as follows

$$U_{i,j,t}(\text{initiate}) - U_{i,j,t}(\text{status quo}) = \alpha_1 \rho(i,t) + \alpha_2 \rho(j,t) + X'_{i,j,t} \beta + \epsilon_{i,j,t} ,$$

where $\rho(i,t)$ and $\rho(j,t)$ refer to the nuclear risk for sides A and B, respectively. $X'_{i,j,t}$ is a vector of other explanatory variables, and $\beta$ is the vector of coefficients.
where $i$ and $j$ are indices of the potential initiator and potential target in the dyad (side A and side B, respectively), $t$ is the year. $X_{i,j,t}$ is the vector of covariates discussed above, $\rho(i, t)$ and $\rho(j, t)$ are the nuclear risk of sides $A$ and $B$ at year $t$, respectively. The main coefficient in which we are interested is $\alpha_2$ which Hypothesis 1 predicts to be negative.

Similarly, we have the following for the second stage

$$U_{i,j,t}(\text{reciprocate|initiate}) - U_{i,j,t}(\text{concede|initiate}) = \eta_1 \rho(i, t) + \eta_2 \rho(j, t) + X'_{i,j,t} \theta + \epsilon_{i,j,t}^{(2)}$$

(2)

This is different from Eq. 1 in that $(i, j, t)$ does not cover all dyad-years but only those in which an MID is initiated by side A. Again, we are mainly interested in the coefficient of the nuclear risk variable for the target side, which is predicted by Hypothesis 2 to be negative.

Given that the second equation is only applied to a selection that is governed by the first equation, we potentially have selection biases. Let us turn to this problem.

**Selection problem**

In Eq. 1 and Eq. 2, we have no reason to believe that the disturbances are not correlated, i.e.,

$$\text{cov} \left( \epsilon_{i,j,t}^{(1)}, \epsilon_{i,j,t}^{(2)} \right) \neq 0$$

because of our limited knowledge about conflict initiation and a host of covariates that are omitted in our model.

The selection problem here is an example of a notorious class: selection problems that occur as a result of strategic interaction of agents. Because the explanatory variables for agents’ utilities are similar if not the same, one is often prohibited from using the Heckman correction (Heckman 1979), because the correction proposed by that model would suffer from collinearity (the correction term is a function of the explanatory variables) and only identified thanks to nonlinearity of the Inverse Mills Ratio, unless one can find a covariate that shows up in the selection stage but not in the second stage.

In order to make the Heckman selection model work, we need to find a variable that is important only in the selection stage. One variable that comes to mind is a measure of political relevance of the dyad. Typically, however, political relevance is constructed in such a way that it contains elements of power status of states, for example, major powers are assumed to form politically relevant dyads with all other states. Hence, even though political relevance is expected to matter for the selection stage not the reciprocation stage, the way the measures are constructed is not helpful. An element of political relevance, however, can be used: territorial contiguity.

In the present work, I use a dummy variable for land contiguity as a covariate in the first stage but not in the second stage. This is a defensible choice because on the one hand, contiguous states are more likely to have MIDs. On the other hand, once an MID has been initiated, the relevance
of the dyad is not as important. In other words, contiguity presents an opportunity for resolving the selection problem because it is a factor only in the first stage. I will also use jointdemocracy as another dummy because it also is expected to be an important factor only in the selection stage (Schultz 1999).

In the models where I use the Heckman correction, the only change is another additive term $g(\hat{y}_{i,j,t}^{(1)})$ in Eq. 2, where $g$ is the inverse Mills ratio and $\hat{y}_{i,j,t}^{(1)}$ is our estimation of the right hand side in Eq. 1. This term reaches statistical significance in the second stage, which corroborates that there are selection issues at work.

**Empirical results**

We do not have direct access to the details of nuclear policies for any state, although some aspects of nuclear policies of different states have now been declassified. But even if we knew such details, we would still need to take into account how much of those policies have been known to each state’s opponents. Here I take two routes for measuring risk. First, I build an index which relies on more readily available data, but is somewhat arbitrary. Then, I propose another method designed to indirectly measure nuclear risk and discuss a way for imputing the latent variable using Bayesian statistics. This proves much more flexible, albeit at a heavy computational price.

**Nuclear risk index**

The concept of the risk of nuclear escalation cannot be directly observed, but, as previously stated, a rough estimate can be inferred from the constraints that a nuclear state faces (Feaver 1992). I use five questions and sum up the answer to these questions as an index of nuclear risk (NRI). For states that do not possess nuclear weapons, NRI is zero. Table 1 lists these questions. Questions 1 and 2 are about institutional stability. Questions 3 and 4 ask about technological aspects of the NCC. The last question asks about whether or not a state is facing time-urgent threats and needs to be able to deliver its weapons in the shortest time possible. This way of making the risk index is inherently somewhat arbitrary.

The index is made in such a way that it is always 0 for non-nuclear states, takes the value of 0 for nuclear states with weapons that are unlikely to be used in an accidental or unauthorized manner, and 5 for states with the riskiest NCC. Table 2 shows an example of how the NRI is constructed.

Given the stability of some of the components of NRI (both structural and strategic components) and the relative nature of the rest (the relative meaning of technological sophistication, for example), NRI is expected to not change much over a few years. If we fail to capture some of the nuance of nuclear risk with this index, the loss will be in the direction of ignoring part of the variation of the explanatory variable of interest which is not necessarily going to work in one direction or the other,
Figure 1: A dispute bargaining model where State 2 possesses nuclear weapons. $\rho$ is the risk associated with State 2’s nuclear weapons.

Table 1: Components of the nuclear risk index.

<table>
<thead>
<tr>
<th>label</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>q1</td>
<td>does the nuclear-armed state have a history of volatile civil-military relations?</td>
</tr>
<tr>
<td>q2</td>
<td>does the nuclear-armed state lack regular government turnover?</td>
</tr>
<tr>
<td>q3</td>
<td>is the nuclear-armed state technologically disadvantaged in its warning systems such as radars, satellites?</td>
</tr>
<tr>
<td>q4</td>
<td>does the nuclear-armed state lack access to parallel channels of communication for preserving the chain of command after being attacked?</td>
</tr>
<tr>
<td>q5</td>
<td>does the nuclear-armed state face time-urgent challenges?</td>
</tr>
</tbody>
</table>
except that it is probably going to dampen the significance levels of our variable.

Results

There are 1,187,735 lines of data, from which 1,063,019 observations are usable. As mentioned before, all dyad-years with an MID are retained and 50,000 dyad-years without a MID are randomly selected and retained. Table 3 shows the results of testing four models that all use the NRI as the main explanatory variable. We are mostly interested in how the risk variables for side B appears in the results. This is shown in the table by $\rho_B$.

Models 1 and 2 are regular probit models that ignore the selection problem. Model 3 is a Heckman selection model with binary outcome variables. A dichotomous measure of territorial contiguity is included in the first stage but not in the second stage to make the model identifiable. Model 4 is similar to Model 3 except that we rely on a measure of joint democracy for making the model identifiable. We see that both hypotheses are borne out by empirical evidence across all models. The effect of nuclear risk for the target state is negative, and statistically significant across all estimations.

Because of the non-linear nature of the estimation procedure, the practical significance of the estimated coefficients are not immediately clear. To show the practical effect of $\rho_B$, two sets of post-estimation analyses are reported here. Figure 2 shows predicted probabilities as a function of $\rho_B$ both for initiation and reciprocation based on Model 4. For the purpose of this estimation, it is assumed that the issue is a territorial issue in the post-Cold-War era, and that Side A is non-nuclear and Side B is nuclear-armed; for all other variables their averages are averaged. The graph also shows the predicted probabilities if the target was not a nuclear state. This is different from the case of $\rho_B = 0$ because the estimation in Table 3 includes a dummy variable for the existence of nuclear weapons in addition to nuclear risk. It is clear that depending on the level of nuclear risk, a nuclear-armed state may be more or less likely than a non-nuclear state to be the target of an MID and also to reciprocate it.

Averaging over all variables may not be very informative because the average dyad may be far from any real pair of states. Another way to gauge the effect nuclear risk is to select a case and see how changing $\rho_B$ leads to different predicted probabilities. The predicted probabilities for the MID that was initiated by Sri Lanka against India in 1992 are shown in Figure 3. In reality, by COW’s coding, this MID was not reciprocated.
Figure 2: Predicted probabilities as nuclear risk changes from Model 4 in Table 3: The graph are (a) the probability of initiation for mean values of parameters in a territorial issue (b) the probability of reciprocation for mean values of parameters in a territorial issue. For comparison, the horizontal dotted lines show the predicted probability if the target were not a nuclear state.
Figure 3: Predicted probabilities as nuclear risk changes from Model 4 in Table 3: The graph are (a) the probability of initiation for the case of Sri Lanka vs. India in 1992 (b) the probability of reciprocation for the case of Sri Lanka vs. India in 1992.
Table 2: An example of how the nuclear risk index is constructed. This example shows the year 2000. NRI is 0 for all other states. Questions are referred to by the labels assigned in Table 1.

<table>
<thead>
<tr>
<th>question</th>
<th>US</th>
<th>Russia</th>
<th>UK</th>
<th>France</th>
<th>China</th>
<th>Israel</th>
<th>India</th>
<th>Pakistan</th>
</tr>
</thead>
<tbody>
<tr>
<td>q1</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>q2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>q3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>q4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>q5</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>nucl. risk index</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>
Table 3: Effect of nuclear risk on MID initiation and reciprocation.

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_A$</td>
<td>-0.14</td>
<td>-0.20</td>
<td>-0.14</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.086)</td>
<td>(0.041)</td>
<td>(0.084)</td>
</tr>
<tr>
<td>$\rho_B$</td>
<td>-0.19</td>
<td>-0.30</td>
<td>-0.19</td>
<td>-0.26</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.10)</td>
<td>(0.050)</td>
<td>(0.099)</td>
</tr>
<tr>
<td>joint nuclear</td>
<td>0.13</td>
<td>-0.098</td>
<td>0.12</td>
<td>-0.075</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.23)</td>
<td>(0.15)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>nuclear A</td>
<td>0.73</td>
<td>0.49</td>
<td>0.72</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.21)</td>
<td>(0.10)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>nuclear B</td>
<td>0.55</td>
<td>0.81</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.22)</td>
<td>(0.11)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>capability ratio</td>
<td>0.35</td>
<td>0.19</td>
<td>0.34</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.14)</td>
<td>(0.052)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>capability A</td>
<td>3.22</td>
<td>3.29</td>
<td>3.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.53)</td>
<td>(0.52)</td>
<td>(0.52)</td>
<td></td>
</tr>
<tr>
<td>capability B</td>
<td>6.08</td>
<td>5.94</td>
<td>5.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.56)</td>
<td>(0.56)</td>
<td>(0.57)</td>
<td></td>
</tr>
<tr>
<td>cold war</td>
<td>-0.038</td>
<td>0.38</td>
<td>-0.038</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.082)</td>
<td>(0.033)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>Sdy</td>
<td>-0.29</td>
<td>-0.0099</td>
<td>-0.29</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.12)</td>
<td>(0.053)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>democracy A</td>
<td>0.039</td>
<td>-0.14</td>
<td>0.039</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.11)</td>
<td>(0.044)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>democracy B</td>
<td>0.34</td>
<td>-1.00</td>
<td>0.34</td>
<td>-0.18</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.097)</td>
<td>(0.040)</td>
<td>(0.094)</td>
</tr>
<tr>
<td>joint democracy</td>
<td>-0.20</td>
<td>-0.13</td>
<td>-0.21</td>
<td>-0.056</td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.19)</td>
<td>(0.070)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>contiguity</td>
<td>2.06</td>
<td>0.54</td>
<td>2.06</td>
<td>2.06</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.089)</td>
<td>(0.038)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>intercept</td>
<td>-2.34</td>
<td>-0.45</td>
<td>-2.33</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.18)</td>
<td>(0.063)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>Heckman’s $\lambda$</td>
<td>-0.35</td>
<td>-0.35</td>
<td>-0.35</td>
<td>-0.35</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>peace splines</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>revision type</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>observations</td>
<td>51567</td>
<td>1567</td>
<td>51567</td>
<td>1567</td>
</tr>
<tr>
<td>log likelihood</td>
<td>-3978.2</td>
<td>-889.7</td>
<td>-4866.5</td>
<td>-4866.5</td>
</tr>
</tbody>
</table>

Standard errors are reported in parentheses.
Imputing values of nuclear risk

The NRI suggested above was a simple latent measure of nuclear risk. The problem with this measure is twofold: the arbitrary nature of choosing the indicators and their weights and that the indicators are themselves measured in a subjective way. Measurement difficulty is a fact of life for many scientific endeavors and that the ideals cannot be obtained does not mean we have to accept results that we expect to be problematic. Part of the progress of any research program is in finding intelligent ways to circumvent the measurement hurdles. Here, I propose a novel technique for combining hypotheses 1 and 2 and performing a joint-test on them.

The core of the idea is to assume one of the hypotheses true and impute the variable of interest from the model. Then the estimated variable can be plugged in to test the second hypothesis. The result is weaker than if we had the opportunity of testing the hypotheses separately, but it has two important benefits: it allows us to incorporate our qualitative knowledge about the underlying measure without making arbitrary assumptions.

We assume there are four categories of nuclear risk, going from low-risk to high-risk. For state-years where we know which category is correct, we can fix the category (like post-1998 Pakistan which takes the highest category). For other states, we can allow the nuclear risk category to be imputed. Then, for each category of risk, we impute a specific numerical value (this is similar to cut-offs in an ordinal logit model). The details are presented in Appendix B.

Results using Bayesian estimation

The choice models for the two stages, as described in in the appendix are estimated simultaneously using a Markov chain Monte Carlo method (MCMC). Like before, initiation of MID's is the dependent variable for the first stage and reciprocation of an initiated MID is the dependent variable for the second stage.

The Model is estimated with three MCMC chains. For each chain, the first 10,000 samples are discarded for burn-in round, and then 20,000 more samples are obtained which are thinned by a factor of 1 to 20. As we saw in the results of the previous section, there is significant selection effect in the model.

Table 4 presents the results. I use a probit as required by Heckman correction, because of the assumption of Gaussian distribution for the disturbances, and include ‘territorial contiguity’ as another explanatory variable in the first stage as discussed before. We can see that $\rho_B$ has a negative and statistically significant effect, which corroborates the results from the previous section.

In order to make the estimation of the Bayesian model faster, I dropped the nuclear dummy variables for sides A and B, and instead allowed the first level of risk to also take negative values. The posterior values of risk for each category ($R$ function in Appendix B) are shown in Figure 4. The first category has indeed taken mostly negative values. This is the baseline effect of nuclear weapons, which is apparently making states more aggressive; this result is substantively similar to
Table 4: The effect of nuclear risk on initiation and reciprocation. Estimates are obtained using the Bayesian model with Heckman correction.

<table>
<thead>
<tr>
<th>Variable</th>
<th>First Stage</th>
<th></th>
<th>Second Stage</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>conf. int.</td>
<td>mean</td>
<td>conf. int.</td>
</tr>
<tr>
<td>( \rho_A )</td>
<td>-1.575</td>
<td>[-2.26, -1.091]</td>
<td>0.022</td>
<td>[-0.36, 0.422]</td>
</tr>
<tr>
<td>( \rho_B )</td>
<td>-1</td>
<td>fixed</td>
<td>-0.391</td>
<td>[-0.93, 0.292]</td>
</tr>
<tr>
<td>joint nuclear</td>
<td>-0.200</td>
<td>[-0.64, 0.261]</td>
<td>-0.106</td>
<td>[-0.51, 0.292]</td>
</tr>
<tr>
<td>democracy A</td>
<td>-0.023</td>
<td>[-0.15, 0.100]</td>
<td>-0.146</td>
<td>[-0.31, 0.022]</td>
</tr>
<tr>
<td>democracy B</td>
<td>0.341</td>
<td>[0.24, 0.449]</td>
<td>-0.206</td>
<td>[-0.36, -0.051]</td>
</tr>
<tr>
<td>joint democracy</td>
<td>-0.138</td>
<td>[-0.32, 0.060]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>capability A</td>
<td>11.912</td>
<td>[9.12, 14.722]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>capability B</td>
<td>11.460</td>
<td>[9.48, 13.629]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>capability ratio</td>
<td>0.406</td>
<td>[0.27, 0.553]</td>
<td>-0.026</td>
<td>[-0.26, 0.200]</td>
</tr>
<tr>
<td>cold war</td>
<td>-0.025</td>
<td>[-0.11, 0.064]</td>
<td>0.402</td>
<td>[0.24, 0.561]</td>
</tr>
<tr>
<td>Sdy</td>
<td>-0.251</td>
<td>[-0.40, -0.105]</td>
<td>0.174</td>
<td>[-0.03, 0.372]</td>
</tr>
<tr>
<td>contiguity</td>
<td>2.406</td>
<td>[2.30, 2.519]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>territory</td>
<td></td>
<td>[0.03, 0.411]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>policy</td>
<td>-0.764</td>
<td>[-0.93, -0.600]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>regime/government</td>
<td>-0.020</td>
<td>[-0.35, 0.314]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intercept</td>
<td>-1.840</td>
<td>[-2.02, -1.670]</td>
<td>-0.275</td>
<td>[-0.59, 0.027]</td>
</tr>
<tr>
<td>Heckman’s ( \lambda )</td>
<td></td>
<td>✓</td>
<td>-0.613</td>
<td>[-0.85, -0.372]</td>
</tr>
<tr>
<td>peace splines</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>observations</td>
<td>11567</td>
<td></td>
<td>1567</td>
<td></td>
</tr>
</tbody>
</table>

The posterior 95% confidence intervals of the coefficients are reported in brackets.
the positive coefficients for the dummy variable \textit{nuclear B} in Table 3.
Figure 4: Box plot of the latent level of risk for each risk category. The boxes cover the middle quartiles of each random variable’s posterior distribution.
**Conclusion**

It was hypothesized that nuclear-armed states that have policies, characteristics, and command structures that increase the likelihood of the unauthorized or accidental use of nuclear weapons are less likely than other nuclear states to be targets of militarized interstate disputes. If targeted, however, these states are themselves less likely to reciprocate those disputes because their political leaders are apprehensive of the prospect of a nuclear fait accompli.

The biggest challenge in testing these hypotheses is obtaining valid and reliable measures of nuclear risk. The problem is all the more intransigent because of the small number of nuclear states. The solution that was developed in this work takes one of the hypotheses—naturally, the one which can be tested with more observations—as given and uses that to create a latent variable of risk. To achieve this, it was shown that using a Bayesian model and an MCMC estimation, we are afforded much flexibility regarding incorporating what we know into our model.

The empirical results presented here show that keeping increasing the exogenous risk that nuclear weapons may be used has substantially large and statistically significant effects both in increasing deterrence and in making states more cautious when they are targets of militarized disputes. The results also show an important way in which previous studies have reached contradictory results about the effects of nuclear weapons: in both sets of analyses, it was found that the base-line, before we consider the effect of nuclear risk, is positive. This means that without taking into account the effect of nuclear risk, we are not able to correctly understand the effect of nuclear weapons in militarized disputes.

**Appendix A   Equilibria of the game**

Let \( f(\rho) \) denote the probability density function of \( \rho \), which is \( S_1 \)'s belief about \( \rho \), at the start of the game. If \( S_1 \) initiates and \( S_2 \) reciprocates, \( S_1 \)'s belief about \( \rho \) may change. Let \( g(\rho) \) denote the posterior distribution of \( \rho \) if \( S_2 \) reciprocates, which is obtained using Bayes' rule. Let \( F(\rho) \) and \( G(\rho) \) be the cumulative density functions corresponding for \( f \) and \( g \). In parts of the analysis below, for simplicity, I assume that \( \rho \) is uniformly distributed between 0 and \( \rho_{\text{max}} \).

We use backwards induction to find the Bayesian perfect equilibria the game. There are three possible equilibria that are derived below; for each equilibrium, it is shown why Hypotheses 1 and 2 are deducted from this model.

Consider the last decision node in the game where \( S_1 \) is choosing between standing firm (SF) or backing down (BD). The payoff from BD is \( -a_1 \). The expected payoff from SF is \( E_g[(1-\rho)w_1 - \rho C_1] \), where the subscript indicates that the expected value should be calculated using \( S_1 \)'s posterior belief about \( \rho \). This means that \( S_1 \) stands firm if

\[
w_1 - E_g[\rho](w_1 + C_1) > -a_1
\]
In the previous step of the game, $S_2$’s payoff from concession is $-v$ and its payoff from reciprocation is $a_2$ if $S_1$ is backing down.

**Equilibrium I:** If (3) does not hold, then $S_1$’s strategy must be to back down. In this case $S_2$’s strategy will be to always reciprocate, which means $f = g$, and therefore $E_g[\rho] = E_f[\rho]$. Because initiating and backing down with certainty is costly, $S_1$ will not initiate in the first place if $w_1 - E_f[\rho](w_1 + C_1) < -a_1$. This implies that if
\[
E_f[\rho] > \frac{w_1 + a_1}{w_1 + C_1},
\]
then $S_1$ will always choose the status quo and not initiate. It is clear that in this case reducing $E_f[\rho]$ will at some point break this condition.

**Equilibrium II:** If $S_1$’s strategy is to stand firm, then $S_2$’s payoff from reciprocation is $(1 - \rho)w_2 - \rho C_2$. Because this is decreasing in $\rho$, let $\rho^*$ denote the threshold of $\rho$ above which $S_2$ prefers to concede. This means $(1 - \rho^*)w_2 - \rho^* C_2 = -v$ which gives
\[
\rho^* = \frac{w_2 + v}{w_2 + C_2}.
\]

It is clear from this decision rule that higher $\rho$ makes $S_2$ more circumspect about reciprocation of an MID.

At the start of the game, $S_1$’s expected payoff from initiation is
\[
E[U_{S_1}(\text{initiate})] = (1 - F(\rho^*))v + F(\rho^*)(w_1 - E_g[\rho](w_1 + C_1))
\]
(5)

With a uniform distribution of $\rho$, we have $g(\rho) = \frac{1}{\rho^2}$, which means $E_g[\rho] = \frac{w_2}{w_2 + C_2}$. Simplifying (5) we obtain
\[
E[U_{S_1}(\text{initiate})] = \frac{1}{\rho^*_{\text{max}}} \left( \frac{C_2 - v}{w_2 + C_2} v + \frac{w_2 + v}{w_2 + C_2} \left( w_1 - \frac{w_2 + v}{2(w_2 + C_2)}(w_1 + C_1) \right) \right),
\]
which gives,
\[
\frac{\partial E[U_{S_1}(\text{initiate})]}{\partial \rho_{\text{max}}} = -\frac{1}{\rho^*_{\text{max}}} \left( \frac{C_2 - v}{w_2 + C_2} v + \frac{w_2 + v}{w_2 + C_2} \left( w_1 - \frac{w_2 + v}{2(w_2 + C_2)}(w_1 + C_1) \right) \right).
\]

This can be positive or negative, but if $C_i$s are large enough, it is always negative regardless of the value of other parameters. To see this, it suffices to assume $C_2 = \frac{1}{k} C_1$ and let $C_2 \rightarrow \infty$, which gives
\[
\lim_{C_1 \rightarrow \infty} \frac{\partial E[U_{S_1}(\text{initiate})]}{\partial \rho_{\text{max}}} = \frac{-v}{\rho^2_{\text{max}}} < 0.
\]

So, higher expected values of nuclear risk means better deterrence for $S_2$.

**Equilibrium III:** Assume that $S_1$ is indifferent between backing down and standing firm and uses
a mixed strategy where it stands firm with probability $p$. This requires

$$E_g[\rho] = \frac{w_1 + a_1}{w_1 + C_1}. \quad (6)$$

Knowing that $S_1$ stands firm with probability $p$, $S_1$ chooses to reciprocate if $\rho < \rho^*$, where for $\rho^*$ we must have $p((1 - \rho^*)w_2 - \rho^*C_2) + (1 - p)a_2 = -v$ which yields

$$p = \frac{a_2 + v}{a_2 - (1 - \rho^*)w_2 + \rho^*C_2}. \quad (7)$$

Assuming a uniform distribution for $\rho$, we have $E_g[\rho] = \frac{1}{2}\rho^*$. Using this, (6) means $\rho^* = \frac{2(w_1+a_1)}{w_1+C_1}$. Plugging this in (7) results in

$$p = \frac{(a_2 + v)(w_1 + C_1)}{a_2(w_1 + C_1) - w_2(C_1 - w_1 - 2a_1) + 2C_2(w_1 + a_1)}. \quad (8)$$

Finally, in the first node of the game, $S_1$ initiates if its expected payoff is greater than zero. Knowing that the $S_1$’s expected payoff from war is $-a_1$, (because of the assumption of mixing), we have

$$E[U_{S_1}(\text{initiate})] = (1 - F(\rho^*))v + F(\rho^*)(-a_1) = \frac{1}{\rho_{\text{max}}} \times \frac{c_1 - w_1 - 2a_1(a_1 + w_1 + 1)}{w_1 + c_1},$$

which yields,

$$\frac{\partial E[U_{S_1}(\text{initiate})]}{\partial \rho_{\text{max}}} = -\frac{1}{\rho_{\text{max}}^2} \times \frac{c_1 - w_1 - 2a_1(a_1 + w_1 + 1)}{w_1 + c_1}. \quad (9)$$

As before, assuming large values for $C_1$, this is guaranteed to be negative, independent of the value of other parameters:

$$\lim_{C_1 \to \infty} \frac{\partial E[U_{S_1}(\text{initiate})]}{\partial \rho_{\text{max}}} = -\frac{1}{\rho_{\text{max}}^2} < 0.$$

Again, we see that increasing $E[\rho]$ decreases $S_1$’s payoff from initiation, and $S_2$ uses a threshold rule on the value of $\rho$ to decide whether it should reciprocate an MID or not.

**Appendix B  Estimation technique**

Let us call the statement posited in Hypothesis 1 Statement $A$ and the one in Hypothesis 2 Statement $B$. Let $\neg A$ and $\neg B$ mean the logical complements of theses statements, i.e., not $A$ and not $B$, respective. Since we want to assume $A$ and test $B$, one might think that we are testing $\neg A \lor B$, but this is not the case because it ignores the relationship between $A$ and $B$. I create two statements corresponding to $A$ and $B$ as follows:

$\tilde{A} \equiv \text{States with riskier nukes are targets of MID}s more often than states with less risky nukes.
States with riskier nukes are more likely than states with less risky nukes to reciprocate. These alternative statements are more specific than the negation of $A$ and $B$, respectively. They purport that not only are $A$ and $B$ false, but that the nuclear risk has the opposite effect of what was hypothesized for both initiation and reciprocation.

Now, consider three states of the world. First, when we have $A \land B$. Second, when we have $\tilde{A} \land \tilde{B}$. And third, when neither of theses cases is correct: $\neg((A \land B) \lor (\tilde{A} \land \tilde{B}))$.

If the values we impute to risk in the first stage are highly correlated with the true values of risk, then positive results of the joint-test show that the third state of the world imagined above is not a correct description of the real world, for which we have

$$C \equiv (A \land B) \lor (\tilde{A} \land \tilde{B})$$

Let us call Statement $C$ Hypothesis 3. This is a weaker hypothesis than the combination of Hypotheses 1 and 2. If our test rejects $C$, we can say that the theory is falsified. If, however, the test is corroborated, then we are logically limited to see this as either showing $A \land B$ or $\tilde{A} \land \tilde{B}$. Given that $\tilde{A}$ is contrary to much prior empirical evidence and there is absolutely no reason to accept it, we are on a much stronger footing if we attribute the success of our test to $A \land B$. Again, notice that $\tilde{A} \land \tilde{B}$ is much less unlikely to be true than simply $\neg A \land \neg B$.

Before moving to the details of how this technique can be implemented, a more careful look at what is to be done here deserves attention.

To reiterate, I use a research design suitable for Hypothesis 1, but instead of estimating the effect of risk, I hold the effect of risk constant and estimate the risk itself. Then I use this estimated risk in a setup suitable for hypothesis 2. Here, there exist three conspicuous causes of concern. First, many a time states have expressly different procedures regarding how to respond to a conflict depending on who has initiated the conflict. In the analysis presented here, this dyadic feature of deterrence is lost because I can only estimate an overall risk for a given state at a given time. This is an innate problem that cannot be alleviated with more sophisticated techniques (unless we gather data for the policies of each nuclear state regarding every other state). This possibly leads to biased estimates but we cannot know whether that bias is immaterial or not.

Second, there is a subtle difference between what is posited as Hypothesis 3 and what the tests in this research design tell us: there is nothing in the estimation process that forces the estimation to be exclusively about the ‘nuclear’ risk; any risk that is unaccounted for in the control variables might be at work here. To the defense of the research design, however, we expect the most important factor driving the estimation of risk to be nuclear weapons and the variable is also only estimated for nuclear states. Moreover, in the Bayesian estimation below, we have the opportunity to carefully delineate what we expect from our latent variable and reduce the possibility of a catch-all latent variable. The strongest defense of this way of testing Hypothesis 3 is that unlike nuclear weapons, stronger conventional power—which brings better deterrence—makes the strong states more likely
to reciprocate conflict. In other words, conventional factors are expected to work against a strong result.

**Bayesian estimation of risk**

We have to create a structure for the nuclear risk variable to be estimated from the first stage. Given that the risk variable is not expected to vary much from year to year and based on our qualitative understanding of countries’ nuclear policies and their strategic challenges, we can partition the time that each state has nuclear weapons into periods that we think the state has had different policies. Let us assign natural numbers to these partitions and show them with $\lambda(j, t)$ which is a function of the state on side B of a dyad and time (State $j$ at time $t$).

So, for example, all states without nuclear weapons are assigned $\lambda(j, t) = 1$; US has a $\lambda$ of 2 from 1945 to 1949 (when the Soviets acquire the bomb), 3 from 1950 to 1960 (The Kennedy Administration’s revision of the policy), etc; the USSR has a $\lambda$ of 6 from 1949 to 1960, and so on.

In the empirical results presented here, I have assumed 16 different partitions for nuclear states, which means that $\lambda$ ranges from 1 to 17. The results seem to be robust against different choices of $\lambda$ assignments.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>...</td>
</tr>
</tbody>
</table>

Table 5: $\lambda$ assignments example

Let us further assume that there are four possible distinct policies. Compared to a continuous case, this discrete treatment has two advantages. First, it reflects the nature of the problem: if there are distinct policy options, why should we treat it in a continuum? Second, we are also constraining the latent variable more which means it is not going to simply capture what is left in the error term. Let $\phi$ represent the mapping from $\lambda$ to the policy category:

$$\phi(\lambda) : \{1, \ldots, 17\} \rightarrow \{0, 1, \ldots, 4\},$$

where $1 \ldots 17$ is the range of $\lambda$ and $0 \ldots 4$ is the range of possible policies with 0 showing no nuclear weapons.

We represent the risk associated with these policies by $R(k), k \in \{1, 2, 3, 4\}$ where $R(4)$ is the highest risk. For convenience with notation we extend the definition of $R$ to include nonnuclear states and assume $R(0) = 0$ for all non-nuclear states (their $\lambda$ is 0 and $\phi(0) = 0$ by assumption).

We want to assume Hypothesis 1 and estimate risk. We do this by using Markov Chain Monte Carlo to simultaneously estimate the effect of a number of covariates as well as imputing the risk variable.
The utility difference can be specified as follows

$$U_{i,j,t}^{(\text{initiate})} - U_{i,j,t}^{(\text{status quo})} = X_{i,j,t}' \beta + I(j,t) \left( \alpha_0 + \alpha_1 R(\lambda(j,t)) \right) + \epsilon_{i,j,t},$$

(10)

We assign $\lambda()$ from our qualitative understanding and estimate $\phi()$ and $R()$ along the coefficients for the explanatory variables. We can fix specific points of $\phi$ when we have some knowledge about the ordering of nuclear risk. For example, based on our qualitative evidence, we know that Israel is a nuclear state that has the lowest risk of ever using its weapons. I set $\phi(\lambda(\text{Israel, 1967} - 2003)) = 1$: the partition for Israel for the entire period that it has had nuclear weapons is assigned the lowest possible risk for a nuclear state. As another example, we have similar information that post-1998 Pakistan is the highest risk nuclear state which leads us to set $\phi(\lambda(\text{Pakistan, 1998} - 2003)) = 4$. These ordinal assignments are done while the values of $R(1)$ and $R(4)$ themselves are left to be estimated. It is then clear how this method allows us to utilize the information we possess without forcing any arbitrary constraints on the estimands.

The coefficients are undetermined the way it was presented above. To make them determined, I fix $R(1) = 0$, and $\alpha_1 = -1$ in (10). To see that we are not losing any information here, assume that the utilities are defined as

$$U = \cdots + I(j,t)(\alpha_0 - \alpha_1 \rho(j,t)), \quad \alpha_1 \geq 0.$$

We are modeling the above with

$$U = \cdots + I(j,t)(\hat{\alpha}_0 + (-1.0) \times \hat{\rho}(j,t)),$$

which, assuming an unbiased estimation, gives $E(\hat{\alpha}_0) = \alpha_0$ and $E(\hat{\rho}(j,t)) = \alpha_1 \rho(j,t) + c$—for nuclear states. We are tolerant of multiplication by a positive factor and addition of an intercept as they preserve the correlation structure we want to study.

We construct the Bayesian model as follows. Parameters of the probit models are assumed to have uninformative Gaussian priors with mean 0. $\phi$ is a categorical variable, whose prior probability is determined by a Dirichlet distribution, which chooses categories with equal probabilities. For the risk in each category, we have $R(1) = 0$, and for the rest $R(k), i \in \{2, 3, 4\} = R(k - 1) + \psi_{k-1}$, where $\psi_{k-1}$ is a random variable with exponential distribution, with a prior mean of $\frac{1}{2}$. Finally, we can incorporate the second stage into this model. This provides the opportunity to include the Heckman correction term and also improves the efficiency of our estimation.
Notes

1. Firing a weapon is the last stage in using it. When rational actors interact strategically, the most effective weapons need not be used. To demonstrate this important difference Ellsberg compiles a list of the most well known uses of nuclear weapons by American presidents and finds that every president from Truman to George W. Bush, except Reagan, has made use of the nuclear weapons in their arsenal (Ellsberg 2009). Some with success, like George H.W. Bush’s preventing Iraq from resorting to chemical weapons, and some without results, like Nixon’s threats in the final stages of the Vietnam War.

2. Following the same logic in these hypotheses, it seems plausible to argue that when a nuclear state initiates a conflict against another state, higher risks associated with initiator’s nuclear weapons should reduce the probability of reciprocation by the target. But the goal here is the study of deterrence and the case where \( S_1 \) is a nuclear-armed state is left out of the current analysis for the sake of simplicity, although the empirical results presented here have the potential to corroborate or reject this.

3. Selecting on the dependent variable provides a more secure way of deflating the zero observations in the data (King and Zeng 2001). For some of the results presented here, this is actually what I am doing, where for the first stage, I choose all dyad-years where the dependent variable is 1 and then select a random sample of size \( 5 \times 10^4 \) or \( 1 \times 10^4 \) of dyad-years without an initiated MID.

4. Which countries have nuclear weapons at any given time? There are different estimates in various previous works as it is not an easy task to determine when exactly each state has acquired nuclear weapons. Fortunately, the more delegative types of NCC, in order to be effective, rely on public shows of force. In this sense, an estimate based on public ‘understandings’ is not far off the mark for the present analysis. I follow Horowitz (2009) and Gartzke and Kroenig (2009).

5. Shifting the cutoff year for the end of the Cold War by as much as two years does not change the results.

6. South Africa is the only state that acquired nuclear weapons and then disposed of them. In addition to missing covariates, South Africa is dropped from this analysis.

7. A set of diagnostic tests that indicate convergence are performed. Both Heidelberger test and the Gelman and Rubin method indicate the stationarity of the chain of estimands (Jackman 2009).

8. This is chosen as Narang (2010) has three postures for smaller nuclear powers and we assume one more to make sure we are not unduly constraining the universe of possible policies.
Bibliography


