# DISCUSSION PAPER SERIES

DP13879
(v. 2)

## THIRD PARTY INTERVENTION AND STRATEGIC MILITARIZATION

Adam Meirowitz, Massimo Morelli, Kristopher Ramsay and Francesco Squintani

**PUBLIC ECONOMICS**

ISSN 0265-8003

# THIRD PARTY INTERVENTION AND STRATEGIC MILITARIZATION

*Adam Meirowitz, Massimo Morelli, Kristopher Ramsay and Francesco Squintani*

Discussion Paper DP13879
First Published 20 July 2019
This Revision 17 June 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Adam Meirowitz, Massimo Morelli, Kristopher Ramsay and Francesco Squintani

# THIRD PARTY INTERVENTION AND STRATEGIC MILITARIZATION


## Abstract

Codified at the 2005 United Nations World Summit, the doctrine of Responsibility to Protect articulates an ideal of international interventions motivated by compassion for victims and a desire to bring stability to hot-spots around the world. Despite this consensus, practitioners and scholars have debated the importance of unintended consequences stemming from the expectation of third party intervention. We analyze how third party intervention shapes the incentives to arm, negotiate settlements, and fight wars in a parsimonious game theoretic model. Among the unintended consequences we find: interventions that indiscriminately lower the destructiveness of war increase the probability of conflict and increasing the cost of arming makes destructive wars more likely. Other interventions, however, can have much more beneficial effects and our analysis highlights peace-enhancing forms of third party intervention. From a welfare perspective, most interventions do not change the ex ante loss from war, but do have distributional effects on the terms of peace. As a result R2P principles are hard to implement because natural forms of intervention create incentives that make them largely self-defeating.



JEL Classification: N/A

Keywords: Strategic Militarization, Intervention Policies, conflict, asymmetric information



Adam Meirowitz - adam.meirowitz@eccles.utah.edu
*University of Utah*

Massimo Morelli - massimo.morelli@unibocconi.it
*Bocconi University and CEPR*

Kristopher Ramsay - kramsay@princeton.edu
*Princeton University*

Francesco Squintani - f.squintani@warwick.ac.uk
*Warwick University*



Acknowledgements
For helpful comments and discussions we thank audiences at several seminars/conferences, and Angelica Bozzi for research assistance. Massimo Morelli acknowledges support from the European Research Council, ERC Advanced Grant 694583.

# Third Party Intervention and Strategic Militarization*

Adam Meirowitz[†] Massimo Morelli [‡]
Kristopher W. Ramsay[§] Francesco Squintani[¶]

June 15, 2020

**Abstract**

Codified at the 2005 United Nations World Summit, the doctrine of *Responsibility to Protect* articulates an ideal of international interventions motivated by compassion for victims and a desire to bring stability to hot-spots around the world. Despite this consensus, practitioners and scholars have debated the importance of unintended consequences stemming from the expectation of third party intervention. We analyze how third party intervention shapes the incentives to arm, negotiate settlements, and fight wars in a parsimonious game theoretic model. Among the unintended consequences we find: interventions that indiscriminately lower the destructiveness of war increase the probability of conflict and increasing the cost of arming makes destructive wars more likely. Other interventions, however, can have much more beneficial effects and our analysis highlights peace-enhancing forms of third party intervention. From a welfare perspective, most interventions do not change the *ex ante* loss from war, but do have distributional effects on the terms of peace. As a result R2P principles are hard to implement because natural forms of intervention create incentives that make them largely self-defeating.

**Keywords:** Strategic Militarization, Intervention Policies, Conflict, Asymmetric Information.

*For helpful comments and discussions we thank audiences at several seminars/conferences, and Angelica Bozzi for research assistance. Massimo Morelli acknowledges support from the European Research Council, ERC Advanced Grant 694583.

[†]University of Utah, David Eccles School of Business; adam.meirowitz@eccles.utah.edu

[‡]Bocconi University, IGIER and CEPR; massimo.morelli@unibocconi.it

[§]Princeton University, Department of Politics; kramsay@princeton.edu

[¶]Warwick University, Department of Economics; f.squintani@warwick.ac.uk

"L'enfer est plein de bonnes volontés ou désirs."
*Saint Bernard de Clairvaux*

# 1 Introduction

Intervention by outsiders in both inter and intra state conflicts is a prevalent and important element of international politics, with the sovereignty of some states more honored in the breach than the observance. Historically the primary purpose of domestic interventions, or extended deterrence, was geo-strategic. In the nineteenth century interventions were a part of the building of empires, the containment of eastward Russian expansion, or the execution of the Monroe Doctrine. In the Cold War the US and USSR chased each other around the globe supporting competitions between governments and rebel groups in a series of proxy wars, hot or cold, pitting communism versus capitalism. Today such strategic interventions are still relevant, but there are also new classes of humanitarian and counter-terrorism based interventions. One need look only at Syria, Afghanistan, Bosnia, Somalia, the Sudan, and Libya to see that leaders of many large states regularly make decisions about how and when to intervene in internal intrastate conflicts.

Largely as a result of these conflicts, in 2005 all members of the United Nations agreed that the world had a responsibility to protect target minorities from grievous human rights violations, even if the intervention violates state sovereignty (UN World Summit, 2005). Yet academics and policy makers voiced concerns that well intentioned interventions by state and non-state actors in these cases may have unintended and undesirable consequences, much as defensive alliance commitments had in the past (Leeds 2003, Benson 2012). On the one hand, the international community has agreed that major powers have a responsibility to protect minorities in other countries from extreme human rights violations, like genocide, but on the other hand the full scale implications of the decision to intervene are complicated and are receiving careful attention.

In the discussions leading up to, and following, the UN agreement many have argued that protection of minorities might create incentives that lead to more conflict, not less (Rowland

and Carment 1998, Cetinyan 2002, Kuperman 2002, Kuperman and Crawford 2005). Central to these arguments is the realization that rational parties, anticipating aid in the event of conflict, will be less deterred by the threat of defeat and undertake riskier or more aggressive actions. The common analogy is to the insurance market, where an insured individual may engage in riskier behavior creating a situation of moral hazard. The fear then is that the expectation of third-party remedies might create more conflicts in the first place.

This study takes seriously the strategic nature of these upstream choices that influence the likelihood of dispute emergence and the military preparedness of parties to a dispute. We build on the idea that rational expectations of how disputes will play out, and how third-parties will intervene, shape the incentives of leaders to undertake actions that influence the odds of a dispute emerging and the willingness of parties to fight and negotiate aggressively within a given conflict.

In other policy areas the implications of various forms of intervention can be evaluated based on policy experimentation or randomized trials. Such alternatives are not available, or ethical, in the case of violent conflict. In this case the total effect of conflict intervention can instead be assessed by considering a model that evaluates many stages of the crisis process. Here we model the strategic preparation in anticipation of a potential crises, negotiation and diplomatic strategy within a crisis, war initiation, third party intervention, and final crisis outcomes. At each stage actors are forward looking and strategic in their decision-making.

Rather than explicitly modeling a third party as a player who chooses the type of intervention, we opt to evaluate the pros and cons of potential types of third party intervention simply by evaluating the comparative statics with respect to the parameters of a bilateral dispute that could be altered by a third party as a matter of foreign policy. We find that many forms of interventions, possibly motivated by the responsibility to protect, generate negative unintended consequences. For example, interventions that reduce the likelihood that an ongoing dispute ends in costly fighting raise the expected payoffs to entering the dispute in a strong position. Additionally, interventions that reduce the losses from conflict also raise the expected payoffs to being strong. This is important because the odds that a given dispute ends in war, and the level of destruction if war occurs, are higher when leaders are strong.

This implies there is a critical trade-off, the benefit found from reducing the likelihood of war given a conflict has begun can be undone by the increased risk that a conflict begins.

To explore the effects of different possible interventions we start with a simple model. We build on the standard take-it-or-leave-it bargaining model with war as a outside option and endogenize the strength and information of the actors by allowing unobserved militarization of one side. Since militarization affects strength, this choice affects the payoffs to fighting for both actors. We consider, in particular, the different effects of various forms of "hard" and "soft" third party intervention–including subsidizing war costs, and the use of third party inspectors to reduce uncertainty –on the likelihood of conflict and distribution of payoffs. We ask: how do different forms of intervention affect incentives to militarize, the level of aggressiveness in negotiations and, ultimately, the probability of war?

We show that humanitarian intervention that lowers the damage from war-fighting, regardless of the arming strategy of the challenger, leads to more militarization and increases the risk of armed conflict. Other interventions, however, can reduce the risk of fighting. For example, committing to intervene on the side of a challenger in ways that are complementary to the challenger's own efforts is, in general, a peace enhancing policy.

In the next section we discuss the literature on third party intervention and outline how our analysis speaks to existing work. Next we present our model, characterize the equilibrium, and review how changes to key parameters change incentives and equilibrium behavior. Although the basic model may seem to lack particular forms of third party interventions, it is specified in such a way that a substantial set of distinct types of interventions can be interpreted as changes to particular parameters in the game. Importantly, we follow the presentation of the main results with a discussion of different kinds of intervention a third party might pursue and trace out the consequences of these interventions for war onset, welfare and distribution of benefits in peace. Key results are descriptions of equilibria or equilibrium comparative statics. The key conclusions from our analysis are highlighted in the formal propositions.

## 2 Literature on third-party Interventions

There is a large literature on third party interventions and the possibility of moral hazard or challenger "emboldenment" in international relations. The topics analyzed run from research on extended deterrence and alliance formation (Snyder and Diesing 1977, Snyder 1984, Snyder 1997, Zagare and Kilgoure 2003, 2005, Yuen 2009, Benson 2012), to studies of common post Cold War strategic and humanitarian interventions in civil conflicts (Crawford 2003, Kuperman 2002, Lischer 2003, Grigoryan 2010).

The first large literature on strategic intervention focused on alliances and extended deterrence. Snyder and Diesinger (1977) and Snyder (1984) locate the challenge of deterring an adversary while still restraining the adventurism of the protégé at the center of the problem of potential intervention. More recently, numerous scholars have explored a number of aspects of this form of intervention in international rivalries. For example, Zagare and Kilgoure (2006) specifically consider how third party support can shape the eventual backing down of a challenger. Benson (2012) explores the structure of international alliances, noting that many seem to be intentionally ambiguous, thus decreasing the incentive of junior partners to provoke conflicts that might activate a defensive alliance. Yuen (2009) explicitly considers how the possibility of later intervention might influence bargaining.

When it comes to humanitarian intervention, many have made moral arguments about the responsibility of capable leaders to protect the human rights of minorities, especially against the most extreme forms of violence like genocide (Hoffman 1996, Holzgrefe and Keohane 2003, Finnamore 2003).

Starting with Rowland and Carment (1998) decision makers began to worry about the strategic implications of expected intervention, even if it was morally justified at the time of action. The first research pointed to the potential for an expectation of intervention to generate a moral hazard for challengers who were, absent outside efforts, unlikely to succeed in reaching their goals. A compelling argument for these effects on risk taking in civil conflict is made by Kuperman (2002), who argues that the expectation of Western intervention played a very important role in the escalation of tensions between the Serbs, Bosnians, and Kosovars

in the former Yugoslavia.

Others have since advanced and expanded Kuperman's argument, (Crawford 2005, Crawford and Kuperman 2005, Grigoryan 2010, Jenne 2004, Thyne 2006) while still others have made counter-arguments (Rauchhaus 2005, Wagner 2005, Western 2005). One of the most prominent critiques comes from Cetinyan (2002), who argues from the perspective of the bargaining model of war that third party interventions should mainly effect the terms of the settlement, but not the probability that a minority and the central government escalate a conflict to war.

A recent article on the effect of interventions on challenger incentives by Kydd and Straus (2013) takes Cetinyan's criticism of the moral hazard framework seriously and integrates a simple bargaining model into an intervention game for a more integrated approach to studying the strategic consequences of the expectation of intervention. This framework generates similar insights as Yuen (2009), and shows how bargaining dynamics complicate what exactly "emboldening" means.

Our analysis compliments the approach to Kydd and Straus (2013), but differs in important ways. First, the interventions studied by Kydd and Straus are focused on a third party that can impose a sanction or join the war themselves, and assumes the third party could win as an independent force. In their framework the outside country does not affect the relative probabilities of victory for the challenger, and can only affect the costs of atrocities carried out by the leader toward the challenger (see also Esteban et al, 2017). They cannot make the challenger more likely to win, cannot decrease their cost of war, or influence their incentives to arm–which is a crucial step in many crises.[1] Our approach considers a model where bargaining outcomes–relating to both peace and war–and arming decisions by the challenger are made with an eye toward third party intervention. Importantly, our model explores a wider variety of potential interventions, allowing us to consider the effects of subsidizing arming, direct military subsidies to enhance the fighting abilities of one side, sharing the cost of conflict, and conditioning post-conflict aid on crisis behavior.

[1]Kuperman's (2008) discussion of the Slovenian, Croatian, and Bosnian Muslim's pre-conflict arming strategies are examples.

Importantly, existing research focuses on a choice to support a protégé or a rebel group where this choice uniformly affects the probability of victory or cost of conflict, but there are a wide variety of alternative forms of intervention that materialize in practice, which can have differential and contingent impacts on conflict. Our model traces out potentially important consequences of considering richer forms of intervention.

Our analysis is also related to papers that treat militarization as a strategic choice taken in anticipation of negotiations and possibly war-fighting (Jackson and Morelli 2009, Meirowitz and Sartori 2008, Baliga and Sjostrom 2015). Like these models, we consider the situation where militarization choices are strategic and imperfectly observed by the challenger's opponent and uncertainty about capacity can arise endogenously. Like this previous work, in non-degenerate cases equilibrium requires uncertainty about capacity and a risk that bargaining will break down requiring countries to use their capacity.

Hörner, Morelli and Squintani (2015) and Meirowitz, Morelli, Ramsay and Squintani (2019) focus on how un-mediated peace talks and the use of a weak mediator with a mandate of brokering peace influences the probability of war. In this paper we examine instead how a broad menu of commonly debated and often employed interventions, like selling arms to the challenger, directly change the payoffs and strategies of participants. Examples include ex-post resource redistribution between winners and losers, subsidies to help ameliorate the damage from conflict, and interventions to change the relative strength of different actors. We also explore the effect of soft intervention like intelligence sharing. A key payoff from the analysis is that we can begin considering how interventions of both hard and soft forms alter incentives and equilibrium outcomes.

# 3 The Model

We begin with a simple model of strategic militarization and negotiation. The baseline model involves two stages and two players, a leader (A) and a challenger (B). In the interstate context, the leader, A, is a local power challenged by state B. In the intrastate context we may think of the leader as the national government and the challenger as an insurgent group

seeking concessions of some form.

In the initial militarization stage the challenger B decides whether to remain militarily weak or to arm and become strong, at a cost $k > 0$. The militarization decision is treated as a hidden action, so the leader does not observe the choice of B.[2]. The most natural interpretation of the arming choice is the purchase of military materials, though really any costly investment that improves a challenger's war fighting ability applies. This may include building command and control institutions, investing in soldier training, or recruiting activists willing to fight.

After the hidden militarization decision the players enter into a dispute and negotiate over a prize normalized to unit size. The dispute can end peacefully with the players each consuming some share of the pie. If no peaceful settlement is reached, the players fight.[3] Formerly, after the hidden investment decision, the leader selects an offer of the form $(x, 1-x)$ and then the challenger decides to accept or reject. Following acceptance the payoffs are $x$ to A and $1-x$ to B. Following rejection, payoffs derive from conflict. Conflict is treated as a costly lottery that shrinks the expected value of the pie.[4]

The odds of winning a war depend on player B's unobserved arming decision. We assume that if the challenger is strong then the war payoffs are $(1-s)\ell$ to the leader and $s\ell$ to the challenger, and if the challenger is weak then the war payoffs are $(1-w)h$ to the leader and $wh$ to the challenger. To capture the effect of arming on the probability of success we assume $s > w$. We also assume that more powerful rivals have more destructive wars so that $\ell < h$.[5] We focus on the case $s\ell > wh$, a well armed challenger has higher war payoffs than a poorly armed challenger—if this were not the case the increased destruction from a well armed war would prevent arming itself. Finally, it is natural to consider cases where even a well armed challenger is disadvantaged and $1/2 > s$.

[2]For expositional purposes, we assume that military strength of B is entirely private information. Our results hold also if the players have some information on each other's military, as long as there is also some residual uncertainty.

[3]The analysis of the ultimatum, or take it or leave it, game with one-sided private information is standard in the literature (Fearon 1995) and so we follow suit for comparability. Similar results follow from other bargaining procedures, like the Nash Demand game.

[4]Conflict destruction need not only consist of physical war damages, but may also include foregone gains from trade and increased military flow costs.

[5]In an extension where war fighting is represented by a strategic contest, this is a natural result.

A strategy of this multistage game of hidden actions for player $B$ has two components. It includes a mixed arming strategy $q \in [0, 1]$, the probability of arming and becoming strong, and a contingency plan for accepting or rejecting offers that depends on its arming choice ($H$ or $L$) and its share of the pie from the offer on the table, $1 - x$. A mixed strategy for the leader is simply a lottery over offers. The consideration of mixed strategies should not be taken as literally claiming that players are indifferent and randomize when making their choices. As shown by Harsanyi (1973), mixed strategy equilibria can be explained as the limit of pure strategy equilibria of a game in which players are not precisely sure about the payoffs of opponents. Further, an alternative interpretation of our model is that $q$ is the level of investment in uncertain military technologies that lead to high military strength with probability $q$, and low strength otherwise, and that such investments bear the linear cost $qk$.

Our welfare analysis is based on a variety of different measures that may be relevant when choosing from the policy options. We consider the implications for intervention for the equilibrium militarization probability of the challenger ($q$), the overall probability of peace, defined as ($V$), the players' expected payoffs ex-ante ($U_A$ and $U_B$), and the ex-ante utilitarian welfare ($W$), defined as the sum of the expected payoffs.

# 4 Equilibrium Results

In equilibrium, the leader offers either $(1 - wh, wh)$, which is accepted by only the weak challenger, or $(1 - s\ell, s\ell)$, which is accepted by both the strong and weak challenger, or assigns probability $r$ to the first (screening) offer and $(1-r)$ to the offer all types of challenger accept.

The weak challenger's *ex ante* payoff is thus $rwh+(1-r)s\ell$, whereas the strong challenger's payoff is $s\ell - k$. There are two things worth noting here. First, the challenger arming with probability 1 cannot be part of an equilibrium to the bargaining and arming game because arming is a hidden action. To see why, notice that if the leader believes the challenger is well armed with probability 1, she proposes a settlement of $(1 - s\ell, s\ell)$. Both a well armed and a weak challenger will accept this offer. If the challenger knows the leader expects him to be

well armed and is sure to offer $(1 - s\ell, s\ell)$, then he could improve his payoff by $k$ if he chose not to arm, and since this is a hidden action the leader would be none-the-wiser and peace would prevail. This profitable deviation means it is never an equilibrium for the challenger to arm with probability 1 when $k > 0$.

Second, never arming at all produces peace for sure and the leader offers $(1-wh, wh)$, and the challenger can do better by deviating, arming and rejecting the offer if $s\ell - k > wh$. This deviation is profitable if $k < s\ell - wh$. In other words, if the cost of arming is large enough relative to the impact of militarization on war-payoffs then there is a peaceful equilibrium in which the challenger remains weak. Throughout we maintain the assumption

**Assumption 1:** $k < s\ell - wh$.

Assumption 1 implies that the challenger randomizes, $q \in (0, 1)$, generating strategic uncertainty. The probability of war is strictly positive. Given this, the unique equilibrium is characterized by analyzing two indifference conditions.

For the challenger to play a mixed strategy at the arming stage we have the equilibrium condition

$$rwh + (1 - r)s\ell = s\ell - k. \tag{1}$$

This holds for a mixture between the high and low offers where

$$r = \frac{k}{s\ell - wh},$$

which is admissible (i.e., which satisfies $0 \leq r \leq 1$) under assumption 1.

To complete the equilibrium description, note that it must be that the militarization strategy $q$ makes the leader indifferent between the offer that both types of challenger accept and the offer that screens and results in war against a strong challenger. This equilibrium condition can be written as

$$1 - s\ell = q(1 - s)\ell + (1 - q)(1 - wh). \tag{2}$$

Solving yields the militarization mixed strategy

$$q = \frac{s\ell - wh}{1 - \ell + s\ell - wh},$$

which is always admissible. Thus, under assumption 1 the unique equilibrium is given by these solutions.

**Proposition 1.** *Assume that the cost of arming $k$ is smaller than $s\ell - wh$. The unique equilibrium involves the challenger arming with probability*

$$q = \frac{s\ell - wh}{1 - \ell + s\ell - wh}$$

*and the leader offering $(1 - wh, wh)$ with probability*

$$r = \frac{k}{s\ell - wh},$$

*and offering $(1 - s\ell, s\ell)$ with probability $1 - r$.*

First consider the incentive to arm. We can see how rational expectations about future intervention can shape how a challenger chooses to invest in military capabilities. As we shall see certain interventions that make war more attractive to B will increase the odds that B is strong. Notably absent is a dependence between the cost of arming and the strategy to militarize. The effect of increased costs is neutralized in equilibrium, and results in no effect on the probability of arming, but rather increased arming costs make the leader more aggressive in their proposal strategy (unless the increased arming cost leads to a violation of Assumption 1). This increased aggressiveness by A decreases the expected value of peace to B and increases the incentives for B to be well armed.

The arming incentives can be understood by thinking about two quantities of interest, the total loss from the destruction of war in heavily militarized conflicts $(1 - \ell)$ and the marginal increase in the payoff to war between a well armed and a less well armed challenger, $(s\ell - wh)$. Some interventions, like increasing the benefit of being strong or the probability of victory for the weak, have straightforward effects through the marginal return to arming. When interventions increase the advantage of being strong, the rate of militarization increases, and when weak types are less likely to lose, militarization decreases.

In other circumstances, like when an intervention decreases the losses from war, the result is not obvious. If there is an expectation that outside actors will aid in post-war reconstruction

when war is particularly destructive because of high levels of challenger arming, then the anticipation of that aid encourages arming and decreases the chance of war, but the wars that do happen will be particularly bloody and destructive. One implication of this result is that in the early post-Cold War period, when the U.S. and U.S.S.R. were no longer vetoing each other's international interventions, the growing expectation of large sums of international humanitarian aid for civil war reconstruction, like those sent to Rwanda, created an incentive for increased arming by rebel groups in the following 15 years of increasingly destructive civil wars world wide.

On the other hand, decreasing the destructiveness of war with unarmed challengers also decreases the benefit of arming, reducing the probability of militarization. But because the incremental benefit to arming is greater than the cost whenever arming occurs, this trade-off is resolved in the direction of increasing militarization when an intervention decreases the costs of war, regardless of the challenger's militarization choice. Increasing a challenger's probability of victory, regardless of their arming choice, makes arming less attractive because the unarmed challenger's payoff from peace increases faster than the armed challenger's payoff from peace.

As $k$ increases within the range of assumption 1, bargaining must be more aggressive, meaning the leader more frequently makes the low offer, formally the result of a larger $r$. If the international community could make the arming cost (economic or political) so high to violate assumption 1, then of course the welfare consequences would be unambiguously positive in terms of probability of peace. This would occur because when everyone knows arming is prohibitively costly, there is no question what the Pareto peaceful settlement must be for both sides to agree. This would be true regardless of how one models the bargaining process.

Interventions that increase the difference in the payoff to war between well armed and less well armed challengers make A less aggressive. Thus decreasing the cost of wars with weak challengers and bolstering the weak challenger's chance of victory make negotiations more contentious while increasing the chance of victory for the well armed challenger and decreasing the destructiveness of their wars make negotiations more likely to end in peace.

Uniformly decreasing the destructiveness of conflict without respect to the level of arms chosen by the challenger, generates competing effects for armed and unarmed challengers, but overall it makes bargaining behavior more accommodating.

Determining how welfare or the probability of war change requires determining how these sometimes complementary effects balance out. The equilibrium probability of peace, denoted by $V$, is

$$\begin{aligned} V &= 1 - qr \\ &= 1 - \frac{k}{1 - \ell + s\ell - wh}. \end{aligned}$$

This expression allows us to evaluate the comparative statics on the probability of peace following different kinds of interventions. Interestingly interventions that increase the probability of victory for strong types make war less likely in equilibrium, while interventions that help the weak have the effect of making war more likely. So interventions like sending special military advisors and advance military training, that increase the fighting abilities of only well armed challengers, have the effect of promoting peace in equilibrium. While such policies do not ensure peace, they have been used frequently, examples being the US Military Assistance Advising Group in the 1950s and 1960s, or more recent support for Kurds in Iraq.

Returning to the question of moral hazard and emboldenment we see that indiscriminate or unconditional intervention to help the challenger in war results in an increased risk of war. Whether the third party decreases the final destructiveness of war or provides military aid that makes both strong and weak challengers more likely to win, in both instances the effect is clear war becomes more likely. In other words, there is an equilibrium moral hazard or emboldenment effect from interventions that are not correlated with specific behavior by the challenger. Counter-intuitively it is selective support of armed challengers and not unarmed challengers that makes peace most likely precisely because such an intervention discourages aggressive bargaining behavior more than it increases the incentive for the challenger to arm. It is worth mentioning that these expected effects that discriminate between the well armed and the less armed challenger do not have to be an intentionally targeted response. This finding underscores the fact that equilibrium analysis requires us to compare how the incentives of

both the challenger and the leader are impacted by interventions.

We can use the indifference conditions (1) and (2),to obtain simple expressions for the players' expected payoffs,

$$U_A = q(1-r)(1-s\ell) + qr(1-s)\ell + (1-q)(1-wh) = 1 - s\ell$$
$$U_B = (1-q)[r(s\ell - k) + (1-r)(rwh + (1-r)s\ell)] = s\ell - k.$$

The challenger can always expect to get $s\ell - k$ by arming regardless of the leader's strategy because, in equilibrium, she must be indifferent between all the strategies she plays with positive probability. Thus the challenger's equilibrium expected utility must coincide with the expected war payoff $s\ell$ minus the militarization cost $k$, and the leader captures the residual $1 - s\ell$. Thus, making player $B$ stronger by raising $s$ increases his ex-ante share of the pie, $s\ell$, net of the militarization cost, $k$. Less obvious is that humanitarian intervention that takes the form of increasing the size of the pie, $\ell$, that survives a destructive war also increases the ex-ante share of the pie for player $B$.

The equilibrium social welfare is obtained by considering the total utility of the game

$$\begin{aligned} W &= U_A + U_B \\ &= 1 - k, \end{aligned}$$

which does not depend on the probability of war. This is not because arming is the only welfare loss–in fact since arming only occurs with probability $q < 1$ the welfare loss from arming is only $qk$. As expected there is also a loss from war when it happens. But in equilibrium this expected loss is exactly equal to the remaining $(1-q)k$.

The simple message of the equilibrium welfare $W$ derivation is that the sum of the ex-ante payoffs of players $A$ and $B$ is influenced (negatively) only by the cost of arming $k$. Any intervention that alters the other parameters of the game are washed out because of the strategic behavior of players $A$ and $B$ in response to these parameter changes. Put differently, the effect of interventions changing the parameters of the game other than $k$ are distributive.

The above comparative statics analysis is summarized in the following result.

**Proposition 2.** *In the baseline model, the peace probability $V$ is decreasing in the arming costs $k$, in the war payoffs $\ell$ and $h$ and in the victory probability of a weak challenger $w$, it is increasing in the victory probability $s$ of a militarized challenger.*[6] *The utilitarian welfare $W$ is decreasing in the arming costs $k$, and constant in the war payoffs $\ell$ and $h$, and in the challenger's victory probabilities $s$ and $w$. Ex-ante welfare of the leader $U_A$ is decreasing in $s$ and $\ell$ and constant in $k$, $w$ and $h$. Challenger's welfare $U_B$ is increasing in $s$ and $\ell$, constant in $w$ and $h$ and decreasing in $k$.*

When wars between a well armed challenger and the leader become more costly, the value of arming decreases, but not because the price of guns increases, but from changes in the nature of conflict. This makes the probability of war go up because the resulting change in the aggressiveness of the leader's offers dominate the reduction of the arming rate. Thus the equilibrium payoff to the challenger goes down. In equilibrium she is indifferent between arming and not arming and a reduction in her payoff from arming must then correspond to a reduction in her equilibrium payoffs. The leader, however, obtains a higher equilibrium payoff as she expects to keep the difference between the value of the prize and the strong challenger's war payoff.

The comparative statics from changing the destructiveness of war with a weak challenger is more complicated. First it is clear that the probability of peace is decreasing because, while the rate of arming goes down, in equilibrium the increased aggressiveness of the leader's bargaining position more than off-sets that change. The equilibrium payoffs of the leader and challenger do not, however, depend on $w$ and $\ell$ because the indifference conditions pin down the payoffs at the value of investing and fighting (for the challenger) and at the value of making a concession that the strong challenger accepts (for the leader).

**Remark** In some contexts while third parties may wish to increase the peace probability, $V$, they also may have some affinity for the challenger. Inspection of the result illustrates that

[6] However, if arming costs $k$ rise above $s\ell - wh$, then the equilibrium of Proposition 1 discontinuously changes into an equilibrium in which player $B$ does not arm, $A$ offers $wh$ and $B$ accepts. War probability is 0, and there is no utilitarian welfare loss.

Table 1: Comparative Statics and Intervention Effects on Peace and Agreements

| Variable | Prob. of Peace | Redistributive Effect on Settlement |
|---|---|---|
| increase $s$ | increase | favors the challenger |
| increase $w$ | decreases | favors the leader |
| increase $s$ and $w$ uniformly | decreases | favors the challenger |
| increase $l$ | decrease | favors the challenger |
| increase $h$ | decreases | favors the leader |
| increase $l$ and $h$ uniformly | decreases | favors the challenger |
| increase $k$ | decrease to increasing | favors the leader |

*Note*: The effect of a variable on the probability of peace is either the partial derivative of the peace probability $V$ with respect to a variables or the directional derivative of $V$ in the direction of equal change in the subset of variables. The redistributive effect is measured with respect to the probability that the leader offers the challenger the larger offer, and is calculated in the same way.

both $V$ and $U_B$ are increasing in $s$ and decreasing in $k$. Accordingly, it is in moving these parameters that there is no tension between reducing war risk and benefiting the challenger. Moreover, the fact that $U_B$ is constant in $w$ and $h$ allows for $V$ to be controlled by moving these parameters without negatively effecting $U_B$. All other changes will involve a tradeoff between the goals of reducing conflict and aiding the challenger. In other contexts the third party may be sympathetic to the leader. The goals of increasing $V$ and $U_A$ are promoted only by reducing $\ell$. It is possible to increase $V$ without reducing $U_A$ by reducing $k, h$ or $w$.

Table 1 collects the comparative static effects of different interventions on *ex ante* distribution of benefits and on the probability of peace.

**Robustness.** Given that we are making policy related claims from a stylized model's comparative statics, it is natural to want to consider how robust the theoretical predictions are to relaxing some assumptions of our analysis. In the appendix we analyze a more general model with continuous arming levels $a$. The cost of arming $c(a)$ is determined by a smooth increasing and weakly convex function $c$, with $c(0) = 0$, the probability of winning is denoted by $p(a)$ and the size of the pie in case of war by $y(a)$. We assume that the challenger's war payoff $p(a)\,y(a)$ is increasing and weakly concave in $a$.

To make the continuous arming model comparable with our simpler binary level of arming

model, let $k$ be the maximum arming spending level of the challenger. Hence, the challenger chooses $a$ out of the interval $[0, \hat{a}]$, where $\hat{a} \equiv c^{-1}(k)$, and the binary model notation is recovered through the definitions $w \equiv p(0)$, $s \equiv p(\hat{a})$, $h \equiv y(0)$, $\ell \equiv y(\hat{a})$. We show in the appendix that also in the general continuous model, there is no pure strategy equilibrium, and that our results on the players' expected payoffs, $U_A = 1 - s\ell$ and $U_B = s\ell - k$, and on the *ex ante* welfare, $W = 1 - k$, carry through.

We explicitly calculate the unique mixed strategy equilibrium and probability of peace $V$. A natural continuous benchmark is the linear case, where both the war payoff and the cost of arming, as functions of the arming choice, are smooth linear functions. Here the equilibrium solution is for the challenger to mix over $a \in \{0, \hat{a}\}$ with exactly the same probabilities as in our model above. The leader's response is also identical. Hence, all the results of the binary level of arming model generalize.

What about the robustness of our results on the probability of peace for the general nonlinear model? While a general analytical result is not available, we numerically show that for a natural parametric family of functions all the results from the two action and linear cases generalize. Specifically, we describe spoils of war as a function $y(a) = \beta(1 - a^2)$, the war technology as a function of $a$ is $p(a) = \alpha + (1/2 - \alpha \times 1.01)a$, and the cost of arming as $c(a) = \kappa a^2$

# 5 Hard Interventions

We draw on the comparative statics above to evaluate the equilibrium effects of third party interventions. We classify interventions into two different classes. First, we use the term "hard interventions" for actions that directly affect the parameters of the conflict: $s, w, h, \ell$ or the cost of arming that the challenger internalizes, $k$. Many widely debated and often employed forms of intervention by powerful states influence the costs of arms, the costs born by different actors from armed conflict or the odds that one party wins certain types of conflict and can be interpreted as movements of these parameters. In considering these interventions it is important to recognize that the model is focused on the equilibrium effects of anticipated

interventions. A change in the war payoffs that is not anticipated at the time of arming or negotiations will not have the equilibrium effects captured here. Changes to $s, w, h$ and $\ell$ can occur because of transfers or policies that are implemented before a potential war begins or they can stem from implicit or explicit promises or even anticipation of international norms.

## 5.1 Subsidized Arms Purchases

One avenue for intervention involves manipulating the cost of acquiring fighting capacity. Policies that reduce the real cost of acquiring arms can come in many forms: A third party might opt to sell small arms or other military equipment to an insurgent group at a discounted price or it might facilitate transport. States might make excess equipment available if the leader can manage to arrange for transport. Governments may ease trade restrictions or selectively not enforce trade limits or tariffs. Alternatively, to increase costs they may impose unilateral or multilateral embargoes. The key is that a movement in $k$ amounts to changing but not zeroing out the marginal cost of fighting capacity. Reducing the cost $k$ has unambiguous effects, it increases the welfare of the players in the game. Of course it may reduce the welfare of the third-party or others outside the model. Moreover, reducing $k$ increases the probability of peace. the mechanism by which reducing $k$ increases welfare is subtle and discussed above. Reducing $k$ has no direct equilibrium effect on arming, but does result in a more accommodating bargaining strategy from the leader. This increases the probability of peace. Interestingly then, reducing $k$ has a direct and indirect effect, and both benefit the challenger. The direct effect is a reduction in the cost of arming if the challenger arms, The indirect effect is an increase in the likelihood that the leader makes the larger offer that both types of challenger accept. This effectively transfers rents during peace from the leader to the challenger. In light of this result, Taiwan's discounted military purchases from the U.S. (see Horton, 2019) can be seen as a force for peace and a method of tilting the political process toward the island and away from the mainland.

## 5.2 Military Aid

States may provide equipment or personnel to the challenger or to the leader/host country. In the model the payoffs from war include two types of terms, the total post-war resources available when the challenger is strong $\ell$ and the resources available when the challenger is weak, $h$, as well as the relative strength of the challenger, $s$ if strong and $w$ if weak. In linking military aid to the four parameters of the war payoffs there are two facets to consider. First, who is being helped? Military aid can be targeted at either the challenger or the leader. The former includes support for insurgents or revolutionaries and the ladder include support for the national government. Second, does the assistance depend on the challenger's investment decision. We treat investment by the challenger as an action that cannot be observed by the leader when it makes its offer. This does not mean that the effect of military aid cannot depend on the challenger's investment. Two natural channels for conditionality come to mind. First, a third party can distribute forms of military support that exhibits complementarities with the challenger's capacity. Non-combatant military advisors may increase the effectiveness of even a small and poorly equipped militia, but arguably the return on these advisors is larger if there are more combatants and they are better armed. Alternatively, sending in equipment that requires skilled and organized personnel to coordinate and conduct maintenance will have a larger effect if the requisite number of combatants and relevant communications systems are in place. We might then think that this form of aid increases $s$ more than it increases $w$.

Alternatively, technology transfer or provision of surgical weapons may transform the mode of battle from one that involves a high degree of collateral damage to more targeted surgical strikes. This may reduce damage, thus increasing $\ell$ and $h$. But as in our previous discussion it is natural to think that the benefits of this form of assistance can depend on the latent strength of the challenger, and thus they may impact $\ell$ and $h$ differently. This is true for aid that is provided to the challenger or the leader. Finally this form of aid may also impact the terms $s$ and $w$.

We recall that the welfare of the leader and challenger are invariant to changes in the parameters that we associate with military aid. The probability of peace, however, is not.

The probability of peace is decreasing in $\ell, h, w$. It is increasing in $s$. For this discussion it is worth noting that the probability of peace is increasing in $s-w$. Accordingly aid that exhibits strong complimentarities with challenger investment helps reduce the risk of war but support that has the same impact on the distributional term regardless of challenger strength has no impact. On the other hand aid that increases the destructiveness of fighting has the benefit of reducing the probability that fighting occurs.

## 5.3 Humanitarian or Economic Aid

In many ways a natural manifestation of responsibility to protect is a sense of duty to provide humanitarian aid after a conflict. We think of this intervention in terms of its impact on the overall level of suffering following a conflict, $\ell$ and $h$. It might also be natural to think that aid is sometimes targeted. NGO's don't necessarily provide support equitably. Aid can be targeted at locations that tend to be highly concentrated with particular enthnic or religious groups in ways that provide the aid primarily to the challenger. Alternatively, aid can be provided to the government in which case it may be natural to expect that surviving challengers receive less. Thus even humanitarian aid that only arrives after fighting may still have a distributional impact. What then do we learn about humanitarian aid from the model. Sadly, Proposition 2 tells us that when the participants (challenger and leader) anticipate that third parties will feel a sense of duty or responsibility to provide aid the probability of conflict goes up. The only way that aid of this form can lower the probability of conflict is if there is an expectation that more aid will come if the challenger was strong. This expectation runs counter to an ethos of providing help when the underdog or challenger really took a beating. That said, our model of conflict ties a rebel's strength to the level of destruction in the war. Recall that a conflict where the challenger is well armed does more damage. Therefore, an aid strategy that targets the most destructive wars will decrease the probability of conflict, even though the conflicts that do occur will be more violent. Our welfare result says this happens without making the total expected loss larger, though it does redistribute the benefits of peace away from the leader and toward the challenger.

### 5.4 Investment and market reactions

Another pathway for third-party involvement is through market forces. Investors in politically unstable regions do not ignore changes in political risk. The interest rates respond to information that changes expectations about regime stability. In our context then, upon seeing war-fighting or any signals that suggest the challenger is stronger than expected, one would anticipate interest payments on investment to go up. This hurts the leader. If the challenger does not internalize the benefits of the investment to the same degree as the leader than the shock following a market update that the challenger is strong is analogous to a distributional shock against the leader with less or no effect on challenger payoffs. One might argue that a talented leader would pass the costs on to citizens and challengers but it is unlikely in the presence of rent-seeking or corruption that the pass-through is 1. Overall, the market adjustment from learning that the challenger is stronger than expected would have the effect of reducing the leader's payoff more than it reduces the challengers payoff. It might then be viewed as an increase in $s - w$ and a decrease in both $\ell$ and $h$. Thus this market reaction may be viewed as having a positive unintended consequence. If the adjustment is anticipated then its expectation raises the probability of peace.

## 6 Soft Interventions: Inspections

We conceive of soft intervention by a third-party not as changes to the parameters of the crisis bargaining game, but rather as changes that reduce the informational frictions driving bargaining failure and potential conflict. Here, we consider inspections.

Consider a modification of the model that allows for a third party to provide information from their own intelligence apparatus. Players engage in the baseline game with the modification that if third party espionage is successful the challenger's militarization choice will be revealed prior to bargaining. This means that either bargaining will look as in the baseline model or it will involve perfect information (and the proposer will extract all rents with 0 risk of war). This lottery alters the militarization incentives and changes the indifference condition, which does not alter the militarization probability $q$ but does alter the proposers

strategy $r$. Note that it does not matter if the challenger learns that the government has learned her strength (i.e successful inspections are a public event) or if the challenger faces uncertainty about whether the government learns. This is because the challenger's decision to accept or reject an offer is independent of what she believes the government knows.

Formally, we assume inspections are a public event. With probability $z < 1$, the type of the challenger becomes common knowledge and with probability $1 - z$, the inspection fails, and this is common knowledge. The indifference condition for the arming decision becomes

$$whr(1 - z) + s\ell(1 - r)(1 - z) + zwh = s\ell - k.$$

The indifference condition for the leader's offer at the history in which the inspections did not reveal any information is the same as the leader's indifference condition in the baseline model

$$1 - s\ell = q(1 - s)\ell + (1 - q)(1 - wh).$$

Accordingly, the monitoring will not change $q$ but it will change $r$ (in a desirable direction). The new solution is

$$r(z) = \frac{k - z(s\ell - wh)}{(1 - z)(s\ell - wh)}$$

Note that as $z$ vanishes this expression converges to $r$ from the baseline model. When $z$ gets close enough to 1, because $s\ell - wh > k$, the solution becomes negative. And so when $z$ is too large, $r(z) = 0$, the leader makes the pooling offer with probability one when the inspections do not reveal information. When this happens there is zero risk of war. Else, the peace probability is $V = 1 - (1 - z)qr(z)$, increasing in $z$.

The challenger's welfare $U_B = s\ell - k$ is unaffected by inspections. In other words, although observability reduces the probability of war it does not provide any rents to B. Based on the welfare effects of this form of monitoring technology we see that ex ante B would not support or oppose this technology. However, ex interim incentives to manipulate the service could exist for B. A, or a third party, however, stands to gain as long as the cost of inspections is not too high and so we might expect them to exert energy establishing this monitoring technology.

The leader's welfare is derived using the indifference condition (2):

$$\begin{aligned} U_A &= (1-z)(1-s\ell) + z(q(1-s\ell) + (1-q)(1-wh)) \\ &= (1-z(1-q))(1-s\ell) + z(1-q)(1-wh) \end{aligned}$$

and increases in $z$. This leads to the welfare:

$$\begin{aligned} W &= s\ell - k + z(1-q)(1-wh) + (1-z(1-q))(1-s\ell) \\ &= 1 - k + z(1-q)(s\ell - wh). \end{aligned}$$

Inspections increase welfare by adding the "bonus" $(s\ell - wh)$ with probability $z(1-q)$.

As in the baseline model there cannot be an equilibrium with $q = 0$ or $q = 1$ and thus these solutions characterize the equilibrium. We wrap up the analysis of monitoring with the following summary result.

**Proposition 3.** When the militarization becomes public with probability $z$, militarization occurs with probability

$$q = \frac{s\ell - wh}{1 - \ell + s\ell - wh}.$$

When it becomes public the leader makes the screening offer, of $s\ell$ if the challenger has armed and $wh$ if it has not. If the arming choice is not revealed then the leader offers the former with probability $r$ and the latter with probability $1 - r$ where

$$r = \max\{0, \frac{k - z(s\ell - wh)}{(1-z)(s\ell - wh)}\}.$$

The peace probability takes the value

$$V = 1 - \max\{0, \frac{k - z(s\ell - wh)}{1 - \ell + s\ell - wh}\}.$$

The leader's welfare $U_A = s\ell - k$ is unaffected by inspections, the challenger's welfare

$$U_B = 1 - s\ell + z(1-q)(s\ell - wh)$$

increases in $z$, together with the utilitarian welfare

$$W = 1 - k + z(1-q)(s\ell - wh).$$

The intuition behind the change in $r$ is that when type is revealed the weak leader loses rents and has a greater incentive to arm. But if the challenger is arming more, the leader wants to increase the frequency of the generous offer when inspections fail. These two forces form an equilibrium when the challenger is again indifferent between arming and not and generates sufficient strategic uncertainty to induce such an $s(z)$. Thus conditional on inspections failure and arming decisions the probability of war goes down.

The special case of $z = 1$ corresponds to fully observable investments. In this case A will make the best offer that gets accepted and the arming indifference condition cannot be satisfied for $k < s\ell - wh$. In this case $W = -k$ and $V = 1$. Observe, however, that the payoff to B from the equilibrium to this game is $s\ell - k$ which is the payoff she obtains in the equilibrium of the baseline model.

# 7 Conclusion

Intervention by third parties in internal, and sometimes international, conflicts is a common component of the foreign policy considerations of many powerful states. Sometimes these interventions are geo-strategic, other times they are humanitarian, and sometimes they are simply the result of internal political pressures. But whatever their motivation, it is important to understand how intervention can shape outcomes and influence the likelihood of costly war-fighting. In the absence of significant amounts of data and plausible ways to obtain the identification of causal effects, theoretical models represent a natural way to address this question. Here we build a parsimonious game theoretic model of intervention with strategic militarization and bargaining. We consider a wide variety of possible interventions that range from a blanket commitment to military assistance in case of war, to conditional military support, to subsidizing a challenger's militarization.

A key virtue of our approach is that by including a richer set of possible ways to influence the payoffs of key actors and tracing out the equilibrium effects of these changes from strategic arming decision through the bargaining process we can speak to both the broader incentives surrounding a crisis and gain insights about alternative forms of third party intervention.

We find some unexpected results. First subsidizing the militarization of the challenger can increase the chance of peace because, even though the subsidy increases the challengers incentive to arm, this increased incentive to arm generates a strategic response by the leader of decreasing the aggressiveness of their negotiation strategy. In our model, the net effect of these competing forces is peace enhancing. Relatedly, we also find that conditional support that creates positive synergies with a challenger's decision to militarize makes negotiation strategies more peaceful and decreases the probability of war. For example, providing certain forms of military support, like advising or advance communication technologies, that are complementary to challenger investment is a policy that increases the chance of peace whereas blanket military support or efforts that are substitutes for challenger investment decrease the chance of peace.

Broadly, our findings are consistent with existing theory and data that find that interventions aimed at increasing the probability of victory for the challenger or decreasing the destruction of war have systematic negative effects on the likelihood of peace. Here the moral hazard effect is robust. The more nuanced understanding that comes out of the analysis indicates that this intuition and empirical findings likely stem from interventions that are uniform or differentially benefit weak challengers. The presence of counter-examples or mixed empirical findings is likely when also considering interventions that are targeted in the right direction. Thus the theory here lays out potential ways to refine empirical work when data is available.

Finally, we also explore a form of what we call soft intervention in this framework. Namely, we think about the possible role of weapons inspections. Here a third party can reveal the militarization decision of challengers with some probability. We see that such an inspection regime decreases the probability of war and increases the challengers welfare. Interestingly, the challenger may not exert resources to block such inspection (ex-ante). Although ex interim incentives might be more problematic.

We conclude that third party intervention in conflict can have many different and conflicting effects. Importantly, supporting an ally, a political important ethnic group or people who are the target of human rights abuse, need not always generate the problem of moral hazard.

Well designed interventions can shape incentives that satisfy the third party's goals and even support self-enforcing settlements, without risking greater conflict. But these interventions need to be chosen carefully.

# References


[1] Baliga, S. and T. Sjöström (2015): "The Strategy of Conflict and the Technology of War", mimeo, Northwestern University.

[2] Benson, B.V. (2012). Constructing International Security: Alliances, Deterrence, and Moral Hazard. (New York, NY: Cambridge University Press).

[3] Esteban, J, M. Morelli and D. Rohner (2016): "Incentives and Constraints for Government Mass Killings: a Game-Theoretic Approach,." in Anderton and Brauer (eds) *Economic Aspects of Genocide, Mass Atrocities, and Their Prevention*, Oxford University Press.

[4] Fearon, J. D. (1995). "Rationalist explanations for war." International organization, 49(3), 379-414.

[5] Fey, M., A. Meirowitz and K. Ramsay. (2013) "Credibility of Peaceful Agreements in Crisis Bargaining " , Political Science Research and Methods: the Journal of the European Political Science Association 1(1).

[6] Finnemore, Martha. (2003) The Purpose of Intervention: Changing Beliefs About the Use of Force. (Ithaca, NY: Cornell University Press).

[7] Grigoryan, Arman. 2010. "Third Party Intervention and the Escalation of State-Minority Conflicts." International Studies Quarterly 54(4): 1143-74.

[8] Harsanyi, J. C. (1973). "Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points." International journal of game theory, 2(1), 1-23.

[9] Hoffman, Stanley. (1996) The Ethics and Politics of Humanitarian Intervention. Notre Dame, IN: Notre Dame University Press.

[10] Holzgrefe, J. L., and Robert Keohane, eds. (2003) Humanitarian Intervention: Ethical, Legal, and Political Dilemmas. Cambridge, UK: Cambridge University Press.

[11] Hörner, J., M. Morelli and F. Squintani (2015): "Mediation and Peace", Review of Economic Studies, 82(4): 1483–1501.

[12] Jackson, M.O., and M. Morelli (2009): "Strategic Militarization, Deterrence and Wars," Quarterly Journal of Political Science, 4(4): 279-313.

[13] Jackson, M.O. and M. Morelli (2011): "Reasons for War: an Updated Survey," in Handbook on the Political Economy of War, Chris Coyne (Ed.), Elgar Publishing.

[14] Jenne, E. (2004). "A bargaining theory of minority demands: Explaining the dog that did not bite in 1990s Yugoslavia. International Studies Quarterly, 48(4), 729-754.

[15] Kuperman, A. J. (2002) Tragic Challenges and the Moral Hazard of Humanitarian Intervention: How and Why Ethnic Groups Provoke Genocidal Retaliation. Unpublished Doctoral Dissertation. MIT

[16] Kuperman, A.J. (2008): "The Moral Hazard of Humanitarian Intervention: Lessons from the Balkans." International Studies Quarterly, 52(1): 49–80.

[17] Kydd, A.H. and S. Straus (2013): "The Road to Hell? Third-Party Intervention to Prevent Atrocities", American Journal of Political Science, 57(3): 673–684.

[18] , Leeds, B. A. (2003) "Do alliances deter aggression? The influence of military alliances on the initiation of militarized interstate disputes." American Journal of Political Science, 47(3): 427–439.

[19] Lischer, S. K. (2003). "Collateral damage: Humanitarian assistance as a cause of conflict." International Security, 28(1), 79-109.

[20] Meirowitz, A.H., M. Morelli, K.W. Ramsay, and F. Squintani (2019): "Dispute Resolution Institutions and Strategic Militarization," *Journal of Political Economy* vol. 127, issue 1, 378-418;

[21] Meirowitz, A.H. and A.E. Sartori (2008): "Strategic Uncertainty as a Cause of War," *Quarterly Journal of Political Science*, 3: 327–352.

[22] Horton, Chris (2019): "Taiwan Set to Receive $2 Billion in U.S. Arms, Drawing Ire From China." *New York Times*, July 9 2019. https://www.nytimes.com/2019/07/09/world/asia/taiwan-arms-sales.html

[23] Rauchhaus, R. W. (2005). "Conflict management and the misapplication of moral hazard theory." Ethnopolitics, 4(2), 215-224.

[24] Rauchhaus, R. W. (2009). "Principal-agent problems in humanitarian intervention: moral hazards, adverse selection, and the commitment dilemma." International Studies Quarterly, 53(4), 871-884.

[25] Rowlands, D., and D. Carment. (1998) "Moral Hazard and Conflict Intervention. " in The Political Economy of War and Peace, edited by Murray Wolfson. Boston: Kluwer Academic Publishers.

[26] Snyder G.H. and P. Diesing. (1997). Conflict among Nations: Bargaining , Decision Making and System Structure in International Crises. (Princeton, NJ: Princeton University Press).

[27] Snyder, G. H. (1984). "The security dilemma in alliance politics." World politics, 36(4), 461-495.

[28] Thyne, C. L. (2006). Cheap signals with costly consequences: The effect of interstate relations on civil war. Journal of Conflict Resolution, 50(6), 937-961.

[29] Wagner, H. R. (2005). "The hazards of thinking about moral hazard." Ethnopolitics, 4(2), 237-246.

[30] Western, J. (2005). "Illusions of moral hazard: a conceptual and empirical critique." Ethnopolitics, 4(2), 225-236.

[31] Yuen, A. (2009). "Target concessions in the shadow of intervention." Journal of Conflict Resolution, 53(5), 745-773.

[32] Zagare, F. C., and Marc Kilgour, D. (2003). "Alignment patterns, crisis bargaining, and extended deterrence: A game-theoretic analysis." International Studies Quarterly, 47(4), 587-615.

[33] Zagare, F. C., and Kilgour, D. M. (2006). "The Deterrence-Versus-Restraint Dilemma in Extended Deterrence: Explaining British Policy in 1914." International Studies Review, 8(4), 623-641.

# Appendix: Continuous militarization

**Set up.** Two contestants bargain over a pie of size one. If the challenger picks a level of arming $a \geq 0$, it pays cost $c(a)$, where $c$ is smooth, increasing and weakly convex in $a$, with $c(0) = 0$, $c'(0) = 0$, and $\lim_{a\to\infty} c(a) = +\infty$. Let the challenger maximum arming spending level be $k$, and define $\hat{a} \equiv c^{-1}(k)$. Hence the challenger may choose a level of arming $a \in [0, \hat{a}]$. Let the challenger winning probability $p$ be a smooth and increasing function of $a$. Let $y$ be the size of the pie in case of war, $y$ is a smooth and decreasing function of $a$. We assume that $p(a)y(a)$ is increasing, with $\partial(p(a)y(a))/\partial a > 0$ at $a = 0$, and weakly concave.

If the challenger picks a level of arming $a$, the ensuing leader and challenger war payoffs are:

$$v_A(a) = (1 - p(a))y(a)$$

$$v_B(a) = p(a)y(a) - c(a).$$

Because $0 \leq p \leq 1$ and $0 \leq y \leq 1$, whereas $c(0) = 0$ and $c$ is unbounded above, there exists an interior maximum $a^*$ of the challenger war payoff $v_B(a)$. To generalize Assumption 1, suppose that the challenger is constrained in its spending capability in militarization below the optimal level $a^*$, that is, $\hat{a} \equiv c^{-1}(k) < a^*$. In the binary arming levels notation: $p(0) = w$, $p(\hat{a}) = s$, $y(0) = h$, $y(\hat{a}) = \ell$. Assumption 1 ($k < s\ell - wh$) is implied by the assumption of constrained militarization spending capability.

**Pure strategy equilibrium.** Again, there is no pure strategy equilibrium. Suppose by contradiction that the challenger played a pure strategy $a > 0$. Then, the leader would offer $1 - x = p(a)y(a)$, and the challenger would accept. In anticipation of this offer, the challenger has an incentive to deviate and play $a = 0$ to avoid paying the cost $c(a) > c(0) = 0$. If the challenger plays $a = 0$, the leader offers $1 - x = v_B(0) = wh$. This cannot be an equilibrium, because $v_B(a)$ is increasing at $a = 0$ and by deviating to $a > 0$ small, the challenger gets war payoff $v_B(a) > v_B(0)$.

**The ex-ante results.** Some of our main results concern the players' ex-ante payoffs, $U_A = 1 - s\ell$, $U_B = s\ell - k$, and ex-ante welfare defined as the sum of the players' ex-ante payoffs

$W = U_A + U_B = 1 - k$.

Without calculating the challenger's mixed strategy c.d.f. $Q$, let $\bar{a}$ be the supremum of $Q$. Let $x(a) = 1 - p(a)y(a)$ be the offer that makes a challenger of type $a$ indifferent between accepting or rejecting. Of course, the leader would never make any offer that leaves $x < x(\bar{a})$ for himself. Because the challenger war payoff $\tilde{v}_B(a) = p(a)y(a)$ net of the sunk militarization cost $c(a)$ increases in $a$, it follows that any type $a$ of challenger such that $a$ belongs to the support of $Q$ would accept the offer $1 - x(\bar{a})$. As a consequence, the challenger's equilibrium payoffs $u_B(a)$ for playing any strategy $a$ on the support of $Q$ is such that $u_B(a) = u_B(\bar{a}) = v_B(\bar{a}) = p(\bar{a})y(\bar{a}) - c(\bar{a})$ (using continuity of $v_B(a)$ in $a$). And for any other strategy $a$, it is the case that $u_B(a) \leq u_B(\bar{a}) = v_B(\bar{a})$.

To generalize the ex-ante results, we prove that $\bar{a} = \hat{a}$. Note that by playing $\hat{a}$, the challenger can guarantee himself at least $v_B(\hat{a})$ by fighting, i.e., that $u_B(\hat{a}) \geq v_B(\hat{a})$. By construction, $v_B(\bar{a}) < v_B(\hat{a})$ for all $\bar{a} < \hat{a}$. Hence, $\bar{a} < \hat{a}$ would contradict $u_B(a) \leq u_B(\bar{a}) = v_B(\bar{a})$ for $a = \hat{a}$.

Now, the indifference condition $u_B(a) = u_B(\bar{a}) = u_B(\hat{a})$ for all $a$ in the support of $Q$ implies that the challenger's ex-ante equilibrium payoff is:

$$U_B = u_B(\hat{a}) = v_B(\hat{a}) = p(\hat{a})y(\hat{a}) - c(\hat{a}) = p(\hat{a})y(\hat{a}) - c(\hat{a}).$$

The leader can guarantee herself

$$v_A(x(\bar{a})) = v_A(x(\hat{a})) = 1 - s\ell,$$

by offering $1 - x(\bar{a}) = s\ell$ to the challenger and by retaining $x(\bar{a}) = 1 - s\ell$ to herself. As seen above, this is an offer the challenger always accepts. Hence, the leader's ex-ante payoff is $U_A = u_A(x(\bar{a})) = v_A(x(\bar{a}))$.

Recalling our assumptions above, the players' ex-ante payoffs are then:

$$U_A = 1 - s\ell, \; U_B = s\ell - k,$$

and the ex-ante welfare is:

$$W = U_A + U_B = 1 - k.$$

**The linear case.** To explicitly calculate (mixed strategy) equilibrium, we begin by supposing for simplicity that the functions $c$, $p$, and $y$ are linear. The challenger can pick any level of arming $a \in [0, \hat{a}]$ at cost $ak$.

Consider the mixed strategy equilibrium $(r, q)$ calculated in the paper. The challenger arms with probability $q = \frac{s\ell - wh}{1-\ell+s\ell-wh}$ and the leader offers $(1 - wh, wh)$ with probability $r = \frac{k}{s\ell - wh}$, and $(1 - s\ell, s\ell)$ with probability $1 - r$. Let's show that this is still an equilibrium. The challenger accepts offer $s\ell$ for all $a \in [0, \hat{a}]$ and accepts offer $wh$ if and only if $a = 0$. Hence, her payoff is

$$u_B(0) = (1 - r)\, s\ell + rwh$$

$$\begin{aligned} u_B(a) &= (1 - r)\, s\ell + r[as\ell + (1 - a)wh] - ak \\ &= \left(1 - \frac{k}{s\ell - wh}\right) s\ell + \frac{k}{s\ell - wh}\,(as\ell + (1 - a)wh) - ak = s\ell - k \end{aligned}$$

$$u_B(\hat{a}) = s\ell - k.$$

For the leader, it is immediate that there is no profitable deviation.

Below, we show that the mixed strategy equilibrium $(r, q)$ is unique.

**Mixed strategy equilibrium in the non-linear case.** The calculation of the mixed strategy equilibrium for the case in which $c$ is strictly convex, and $py$ strictly concave is significantly more involved.

Let us consider a mixed strategy equilibrium $(R, Q)$, where $Q$ is the c.d.f. of the challenger militarization level $a$, and $R$ is the c.d.f. of the militarization level $a$ associated with the leader's offer $x(a)$ to herself. Earlier arguments imply that the upper bound of the supports of $R$ and $Q$ is $\hat{a}$.

The challenger's equilibrium payoff is

$$\begin{aligned} u_B(a) &= p(a)y(a)\,R(a) + \int_a^{\hat{a}} [1 - x(t)]dR(t) - c(a) \\ &= p(a)y(a)\,R(a) + \int_a^{\hat{a}} p(t)y(t)\,dR(t) - c(a) \end{aligned}$$

The leader's equilibrium payoff is

$$u_A(x(a)) = x(a)Q(a) + \int_a^{\hat{a}} [1 - p(t)]\, y(t)dQ(t)$$

$$= \quad [1 - p(a)y(a)]\, Q\,(a) + \int_a^{\hat{a}} [1 - p(t)]\, y(t) dQ(t).$$

The first round of arguments assesses whether the c.d.f.s $R$ and $Q$ have atoms or gaps.

Suppose by contradiction that $Q$ has an atom at some militarization level $a > 0$. Then, the leader cannot be indifferent between offering $x\,(a)$ and $x\,(a - \varepsilon)$ to herself for $\varepsilon > 0$ small enough. For an $\varepsilon$-order payoff increment in case of no war, the leader increases the risk of war by a fixed amount, $dF\,(a)$.

Suppose by contradiction that $Q$ has a gap $(a, a')$. Then, the leader has no reason to make any offer $x\,(a'')$ to herself such that $a'' \in (a, a')$. Such offers are accepted if and only if the challenger invests $a$, but then the leader prefers to make the offer $x\,(a)$ to herself.

Now suppose by contradiction that the leader's c.d.f. $R$ has a gap $(a, a')$. Then it must be the case that $u_B\,(a) = u_B\,(a')$. For any $a'' \in (a, a')$, the challenger faces the same risk of war as with playing $a$. Hence, because the war payoff $v_B$ increases in $a$, $u_B\,(a'') > u_B\,(a)$, a contradiction.

We have concluded that $Q$ has no atom at any $a > 0$ and neither $R$ nor $Q$ have any gaps on $[0, 1]$. We cannot rule out that $R$ has atoms, instead, because the challenger's payoff $u_B\,(a)$ is continuous in $R$ even if $R$ has atoms. We assume as is usual in these exercises that $Q$ is smooth when continuous.

Turning to calculate the leader's mixed strategy c.d.f. $R$, consider the challenger's indifference condition. The payoff

$$u_B(a) = p(a)y\,(a)\, R(a) + \int_a^{\hat{a}} p(t)y\,(t)\, dR(t) - c(a)$$

must be constant on $a \in [0, \hat{a}]$.

Differentiating $u_B$ (or taking differences $u_B(a + \Delta) - u_B\,(a)$ with $\Delta \to 0$), we obtain:

$$p'(a)y\,(a)\, R(a) + p(a)y'\,(a)\, R(a) + p(a)y\,(a)\, r(a) - p(a)y\,(a)\, r(a) - c'(a) = 0.$$

that is:

$$p'(a)y\,(a)\, R(a) + p(a)y'\,(a)\, R(a) - c'(a) = 0.$$

Even if $R$ is not differentiable and not even continuous, the same cancellation takes place by taking differences, because $p(a)\,y(a)$ is continuous.

We obtain:

$$R(a) = \frac{c'(a)}{p'(a)y(a) + p(a)y'(a)},$$

the quantity is positive because $c'(a) > 0$ and because $p(a)y(a)$ increases in $a$.

It is also increasing:

$$\frac{\partial}{\partial a}R(a) = \frac{c''(a)\left[p'(a)y(a) + p(a)y'(a)\right] - c'(a)\,\partial\left(p'(a)y(a) + p(a)y'(a)\right)/\partial a}{\left[p'(a)y(a) + p(a)y'(a)\right]^2} > 0,$$

because $c$ is convex, $c'' > 0$, and $p(a)\,y(a)$ is concave, $\partial\left(p'(a)y(a) + p(a)y'(a)\right)/\partial a < 0$. Finally, because $c'(0) = 0$ and $p'(0)y(0) + p(0)y'(0) > 0$, it is the case that $R(0) = 0$ and because $p'(\hat{a})y(\hat{a}) + p(\hat{a})y'(\hat{a}) > c'(\hat{a})$, it is the case that $R(\hat{a}^-) < 1$. The function $R$ is verified to be a c.d.f., with a unique atom at $\hat{a}$ and smooth elsewhere.

Also note that in the linear case seen above, both the marginal benefit of militarization $p'(a)y(a) + p(a)y'(a) = s\ell - wh$ and the marginal cost $c'(a) = k$ are constant. Hence the c.d.f. $R$ is constant on $a \in (0,1)$. It has two atoms, at $a = 0$ and $a = \hat{a}$, thereby reproducing the mixed strategy equilibrium $(r, q)$ found in the paper. That equilibrium is proved to be unique.

For the challenger's mixed strategy c.d.f. $Q$, consider the leader's indifference condition. The equilibrium payoff

$$\begin{aligned} u_A(x(a)) &= x(a)Q(a) + \int_a^{\hat{a}} \left[1 - p(t)\right] y(t)dQ(t) \\ &= \left[1 - p(a)y(a)\right] Q(a) + \int_a^{\hat{a}} \left[1 - p(t)\right] y(t)dQ(t). \end{aligned}$$

must be constant on $a \in [0, \hat{a}]$.

Differentiating $u_A$, we have :

$$0 = (1 - p(a)y(a))q(a) - p'(a)y(a)Q(a) - p(a)y'(a)Q(a) - (1 - p(a))y(a)q(a)$$

hence

$$Q(a)\left[p'(a)y(a) + p(a)y'(a)\right] = q(a)\left(1 - y(a)\right).$$

This is a differential equation in $Q$. It can be solved with standard techniques: The particular solution is:

$$Q(a) = \exp\left(-\int_a^{\hat{a}} \frac{p'(t)y(t) + p(t)y'(t)}{1 - y(t)} dt\right).$$

This quantity belongs to $(0,1)$, because $y(t) \in (0,1)$ and $p(t)y(t)$ is increasing, so that $p'(t)y(t) + p(t)y'(t) > 0$ and hence $\frac{p'(t)y(t)+p(t)y'(t)}{1-y(t)} > 0$. The exponential of a negative term is in $(0,1)$. The quantity $Q(a)$ is increasing, with density

$$q(a) = \exp\left(-\int_a^{\hat{a}} \frac{p'(t)y(t) + p(t)y'(t)}{1 - y(t)} dt\right) \frac{p'(a)y(a) + p(a)y'(a)}{1 - y(a)} > 0.$$

Hence $Q(a)$ is a c.d.f., with a unique atom

$$Q(0) = \left(-\int_0^{\hat{a}} \frac{p'(t)y(t) + p(t)y'(t)}{1 - y(t)} dt\right) > 0.$$

The war event occurs when the challenger's arming choice is $a$ and the leader offers to herself $x(a')$ with $a' < a$. The probability that the leader's offers to herself $x(a')$ with $a' < a$ is $R(a^-) = R(a)$ by continuity. The density of the challenger's $a$ is $q(a)$. The event that $a = 0$ occurs with probability $Q(0) > 0$, but it is irrelevant because $a'$ cannot be strictly smaller than 0. Hence the peace probability is:

$$\begin{aligned} V &= 1 - \int_0^{\hat{a}} R(a) q(a)\, da \\ &= 1 - \int_0^{\hat{a}} \frac{c'(a)}{p'(a)y(a) + p(a)y'(a)} \exp\left(-\int_a^{\hat{a}} \frac{p'(t)y(t) + p(t)y'(t)}{1 - y(t)} dt\right) \frac{p'(a)y(a) + p(a)y'(a)}{1 - y(a)} da \\ &= 1 - \int_0^{\hat{a}} \frac{c'(a)}{1 - y(a)} \exp\left(-\int_a^{\hat{a}} \frac{p'(t)y(t) + p(t)y'(t)}{1 - y(t)} dt\right) da. \end{aligned}$$

The expression $V$ characterizes the equilibrium probability of peace in the general setting. For the parametric family of non-linear environments where the spoils of war are $y(a) = \beta(1 - a^2)$, the war technology as a function of $a$ is $p(a) = \alpha + (1/2 - \alpha \times 1.01)a$, and the cost of arming is $c(a) = \kappa a^2$ we numerically calculate the derivative in the cube $[.1, .5] \times [.75, .95] \times [.01, .1]$ for $\alpha$, $\beta$ and $k$. The derivatives $\frac{\partial V}{\partial \alpha}$ and $\frac{\partial V}{\partial \beta}$ are both negative, consistent with our results in the binary arming level model. For $\kappa$ we find the derivative is non-monotonic. Like in the binary model, the probability of peace is decreasing and then increasing in $\kappa$.

The MATLAB code for the numerical analysis can be found in an appendix on the author's web site.