

# Dispute Resolution Institutions and Strategic Militarization

Adam Meirowitz\*   Massimo Morelli†   Kristopher W. Ramsay‡  
Francesco Squintani§

October, 2013

## Abstract

A central question in political science is how to best manage information asymmetries and commitment problems when disputes arise between states or nations. We argue that common framings of this problem miss an important feature: the institutions determining how disputes are resolved shape the incentives for nations to enter disputes. Because war can be sometimes understood as the down-side risk from entering a dispute, institutions that reduce the chances of war-fighting may induce perverse incentives to enter into disputes and militarize. We develop a simple crisis model that captures both the militarization decisions and bargaining behavior. We examine how features like direct communication and third-party involvement alter the incentives. Seemingly effective institutions that improve the chance of peace for a given distribution of military strength, can actually lower the chance of peace once one accounts for distortions to militarization decisions. To illustrate the value of this broader perspective we show how a form of intervention by a mediator concerned only with resolving the current crisis, turns out to create optimal militarization and bargaining incentives.

---

\*Princeton University, Department of Politics.

†Columbia University, Departments of Political Science and Economics.

‡Princeton University, Department of Politics.

§Warwick University, Department of Economics.

# 1 Introduction

A central question in politics is how institutions shape political decision-making and policy outcomes. The study of institutions that are used to resolve disputes which might devolve into military conflict is central to the field of international relations. The types of disagreements studied range from international disputes to civil and ethnic conflicts. Moreover, an important portion of this work seeks to make policy-relevant pronouncements about how particular institutional features affect the likelihood that fighting will break out.

For example, a growing body of theoretical work considers how various forms of direct diplomacy can influence the probability of bargaining breakdown (Baliga and Sjöström 2004, Ramsay 2011, Smith 1998, Sartori 2002, Traeger 2011). Others explore the role of mediation and delineate conditions where such third party actions can influence outcomes (Bester and Wärneryd 2006, Fey and Ramsay 2010, Horner, Morelli, and Squintani 2010, Kydd 2003). Still others look at how standing institutions, like the United Nations or ICJ, can influence state actions and the probability of cooperation (Fang 2010, Chapman and Wolford 2010, Gilligan, Johns, and Rosendorf 2010). Similarly, there exists a detailed empirical literature on mediation and third party intervention. Rauchhaus (2006) provides quantitative analysis showing that mediation is especially effective when it targets asymmetric information and Savun (2008) shows that hard sources of information are very influential. Bercovitch and Houston (2000), and Bercovitch et al. (1991), also point out the particular relevance of mediation for peace outcomes when the uncertainty about the disputants' strength is high.<sup>1</sup> Related empirical research on various forms of third party intervention, such as Dixon (1996) and Frazier and Dixon (2006), finds that mediation can be effective at increasing the probability of settlement and de-escalation of disputes.

Common to all these studies is the idea that features of the institution within which the leaders of opposing sides bargain can have important consequences for the likelihood of peace, the terms of the settlement, and whether the settlement is enduring. Furthermore, each of these studies takes the crisis as the relevant starting point and is typically focused on

---

<sup>1</sup>More generally, for an overview of the literature on effective forms of negotiations, with and without mediators, see Bercovitch and Jackson (2001), Wall and Lynn, (1993).

what might be called the crisis specific incentives. Scholars framing the problem in this way ask questions like: given a dispute, how will the features of the institution—say commitment, information transmission, the presence of a mediator, external motivations, an audience, trade spillovers—influence the likely resolution? Following this analysis scholars prescribe what they take to be the best combination of institutional features to resolve the disputes that emerge. In this paper we argue that, because institutions can influence the ways that a given dispute is resolved, they can also have important effects on the types of disputes that emerge. Intuitively, a bloody, costly or wasteful conflict can be understood as the possible ‘punishment’ for militarizing and entering a dispute in the first place. Hence, institutions that minimize the chances that disputes lead to conflict may lead to more militarization and disputes. Although this summary is accurate it is still too simple. It is not sufficient to think about whether institutional features increase the risk of equilibrium bargaining failure and war. A proper study must account for how features change the likelihood that different configurations of capacity result in war. The payoff to our theoretical approach is that it offers a framework to handle these nuanced considerations.

We show that failures to account for the possible upstream connections between dispute resolution and militarization incentives may lead scholars to get the relationships between institutions and conflict backwards and can result in policy prescriptions that increase the likelihood of conflict, not decrease it. Specifically, some celebrated institutions such as unmediated peace talks that reduce conflict in the short-run actually raise the probability of war by changing incentives to militarize, and thus the type of crises that arise. This paper pushes for a reformulation of the study of conflict and institutional design and draws a finer distinction between different ways of fostering peace for a given crisis; some are also beneficial at reducing militarization while others generate offsetting negative incentives.

To flesh out the relevance of militarization incentives, and the potential tradeoffs involved across dispute settlement mechanisms, we consider a simple setting in which two nations decide whether to become strong by investing in military capabilities. These decisions are not publicly observed and the choice to become strong involves paying a cost. After the investment decisions are made the players bargain over the division of a prize

where bargaining failure leads to war. In this game each side makes a demand and if the demands are compatible—i.e., the sum of their demands is less than the size of the pie—the settlement that splits the difference between the two demands is implemented. If, on the other hand, the demands are not compatible, the nations fight.

This bargaining game is canonical and serves as a useful benchmark.<sup>2</sup> To study how institutional changes can influence militarization incentives, we first investigate the effects of adding a pre-play cheap-talk stage to the bargaining game. In this version, after the nations arm, but before they make their demands, they have the opportunity to meet and conduct direct diplomacy. The peace talks serve two purposes: first, the delegates of the nations have the opportunity to share information regarding their arming decisions and whether they are "strong" or "weak." The information relayed by the delegates is unverifiable and costless. Second, the meetings may also allow the nations to coordinate future play: with some probability the meeting is successful and yields a peaceful resolution of the dispute, otherwise the conflict escalates, ultimately leading to war.

As has been shown e.g. in Baliga and Sjöström (2004), the addition of this type of cheap-talk stage can be consequential in crisis bargaining. For a fixed probability that each nation has armed, the model with cheap talk can result in a strictly lower probability of war than the model without cheap-talk. This fact shows that institutions can shape the short-run incentives to fight. Most importantly, we find that, in some circumstances, the anticipation that bilateral peace talks will occur in a crisis raises the incentives to arm, because this pre-crisis diplomacy is effective in improving the chances of a peaceful resolution of conflicts involving strong nations. As a result, the bilateral peace talks reduce the down-side risk from arming, and thus increase the incentives to arm. The equilibrium of the cheap talk extension of our bargaining game with strategic militarization also has a higher *ex ante* probability of war than the equilibrium that is played when there is no communication during bargaining. In other words, the effects on strategic militarization trump the gain in odds that players of particular strengths will negotiate a peaceful settlement.

---

<sup>2</sup>The game is purposely simple, but captures key tradeoffs and is sufficiently tractable to allow for extensions to consider how cheap-talk communication and third party intervention influence the short-run and long-run incentives.

To provide a fuller account of how institutions and militarization incentives interact we contrast the way that cheap-talk communication helps resolve conflict with an alternative institution, the introduction of a third-party with a mandate to facilitate a peaceful settlement. Such an intervention can have very different effects on the long-run incentives. Although cheap-talk and dispute resolution by a mediator with no private information or means to provide incentives are both in principle aimed at allowing for information sharing, it turns out that the ways these processes foster information transmission can create decidedly different militarization incentives. We find that a third-party facilitating negotiations can broker peace while avoiding the perverse militarization incentives that surface when deals are reached through direct communication. Interestingly, this is despite the fact that we do not allow the third party's mandate to include imposing costs on nations or control of the militarization incentives. In our framing we conceive of a mediator motivated only to do the best she can at brokering a settlement of the dispute at hand—one for which the militarization decisions have already been made, when the third party is called to mediate the dispute. Even in this “worst-case scenario” of a third-party with no information, or resources and a narrow mandate, conflicts can be better resolved without increasing the upstream incentive to arm.

Our results are also related to the study of contests in economics. In this literature it has become common to model militarization as a contest, (see Garfinkel and Skaperdas, 1996, for a survey). Here strategic militarization is treated as an arms race to prepare for a sure conflict and the military capacities of each country influence the war-payoffs through a contest function. For us, militarization increases the expected payoff to fighting in a crisis, but the goal is expected utility maximization in a game in which war is not a foregone conclusion. Players may reach a settlement or fight and the militarization choice can have strategic affects on settlements as well as the odds of fighting. There is a recent body of formal theory on endogenous militarization in the shadow of bargaining and war fighting. Meirowitz and Sartori (2008) connect militarization with bargaining behavior but provide only results on the impossibility of avoiding conflict. They do not study optimal mechanisms or make comparisons across different institutions. Jackson and Morelli (2009) consider militarization and war but in their analysis states observe each other's investment

decisions prior to bargaining and thus there is no room for communication of any kind to solve information problems. Akcay, Meirowitz and Ramsay (2012) study investment that influences the value of agreement for one player and disagreement for the other prior to the play of a mechanism and provide a characterization of the equilibrium relationship between the probability of bargaining failure and investment levels. In that paper, however, the investments are over a private value component whereas here investment by either player influences the payoffs of both players.

In a less closely connected literature observable armament are viewed as a deterrent or a signal. Garfinkel (1990) focuses on a simple repeated armament setting that illustrates this role, and Powell (1993) provides an example of the conditions under which endogenous armament can lead to effective deterrence, in a world of perfect information. Fearon (1997) argues that observed military expenditures can be a form of costly signaling. Collier and Hoeffler (2006) consider the signaling and deterrent effects of armaments as competing mechanisms explaining levels of post-conflict military expenditures in countries recently experiencing a civil war. Chassang and Padro (2008) show that while weapons have deterrence effects under complete information in a repeated game, when adding strategic uncertainty it is no longer true that there is a monotonic relationship between the size of an arsenal and the degree of deterrence in equilibrium. For very different reasons, even in our model the effects of endogenous militarization incentives on peace are non-monotonic.

## **2 Unmediated Peace Talks and Militarization**

In this section we develop the baseline model and augment it to allow for unmediated communication. The analysis focuses on the comparison of these structures and shows that unmediated peace talks may breed perverse incentives for militarization. Unmediated talks help resolve disputes by allowing for settlements that would not otherwise be possible when the nations are strong. But this feature of what cheap talk does to expectations of how bargaining will unfold, increases the incentive to become strong. We then show that for some model parameters, unmediated peace talks can increase the overall war probability when militarization decisions are taken into account.

**The Crisis Bargaining Game** Two players,  $A$  and  $B$ , dispute a prize or pie normalized to unit size. If no settlement is reached the disputants fight. We adopt the simple bargaining model with private information to describe how players negotiate over the prize. The game is also called the Nash demand game (see, e.g. Nash, 1953, and Matthews and Postlewaite, 1989).<sup>3</sup> When players bargain they simultaneously make demands  $x_A$  and  $x_B$  both in  $[0, 1]$ . So like giving a diplomat a set of instructions to “...bring back at least a share —,” players chose a bargaining strategy that will influence both the settlement (if reached) and the probability of war. We assume that if one player’s demands are sufficiently generous to accommodate the minimal demands of the other, i.e., if the sum of the demands is less than 1, then the split that emerges from bargaining gives each disputant her demand and half the surplus. If the demands  $x_A$  and  $x_B$  are incompatible, meaning they sum to more than the whole prize, then no split is feasible and the outcome is war.<sup>4</sup>

War is treated as a lottery which shrinks the value of the pie to  $\theta < 1$  but the odds of winning a war depend on the configuration of arming decisions. Each player can possess one of two possible arming levels,  $H$  or  $L$ . We will often refer to type  $H$  as a “hawk” and to a  $L$  type as a “dove”. We will describe the process by which nations select their arming levels subsequently. When the two players are of the same type, each wins the war with probability  $1/2$ . When a type  $H$  player fights against a  $L$  type, her probability of winning is  $p > 1/2$ , and hence her expected payoff is  $p\theta$ , which we assume to be larger than  $1/2$ , otherwise the dispute can be trivially resolved by agreeing to split the pie in half. In the case of different types fighting the expected payoff to the  $L$  type is  $(1 - p)\theta$ . Note that whether or not country  $A$  is type  $H$  or  $L$  influences country  $B$ ’s payoff from fighting, hence we are in a context of “interdependent values”.

In the initial militarization stage,  $A$  and  $B$  each decide whether to remain doves or to arm and become hawks at a cost  $k > 0$ . We characterize a mixed arming strategy by  $q \in [0, 1]$ , the probability of arming and becoming a hawk. The militarization decisions are treated as hidden actions, so neither player observes the choice of the other nation,

---

<sup>3</sup>Wittman (2009) and Ramsay (2011) use the same bargaining game in their studies of crisis bargaining.

<sup>4</sup>In the economics literature this protocol is also sometimes called the  $\frac{1}{2}$ - double auction because the surplus is divided equally.

but in an equilibrium the nations hold correct conjectures of the equilibrium strategy-and thus they know the probability that their opponent has armed. For simplicity, and given the symmetry of the game, we restrict attention to equilibria that are symmetric in the militarization strategies  $q$ .

When comparing outcomes under different institutions our welfare analysis will focus on three measures: the equilibrium militarization probability of a country  $q$ , the probability of peaceful conflict resolution  $V$ , and the utilitarian ex-ante welfare  $W = \theta(1 - V) + V - 2kq$ . Before presenting our results, we define the parameter  $\gamma \equiv [p\theta - 1/2]/[1/2 - \theta/2]$ , representing the ratio of benefits over cost of war for a hawk: the numerator is the gain for waging war against a dove instead of accepting the equal split, and the denominator is the loss for waging war against a hawk rather than accepting equal split. It subsumes the two parameters  $\theta$  and  $p$  in a single parameter, and allows a more parsimonious representation of the results. To simplify the exposition, we will henceforth assume that  $\gamma \geq 1$ , i.e., the benefit of war for a hawk is sufficiently high relative to the cost. In terms of the deep parameters,  $\theta$  and  $p$ , this is equivalent to assuming that war is not too destructive and that the hawk's advantage over the dove is significant. But one of our main results (Theorem 1) also holds for  $\gamma < 1$ , as explained in the Appendix. In the formal proofs, as we show in the Appendix, it is often convenient to work with the odd ratio  $\lambda = q/(1 - q)$  associated with the arming probability  $q$ , instead of working with  $q$ .

**Benchmark equilibrium in the crisis bargaining game** Let us start by solving the crisis bargaining game when  $q$  is fixed. This game is characterized by broad equilibrium multiplicity. For example, any demand strictly larger than  $1 - (1 - p)\theta$  leads to war with probability one. This claim is so excessive that even a dove knowing that the opponent is stronger would prefer to fight rather than making such a drastic concession. As a result, there always exists an equilibrium in which war occurs with probability one in the benchmark game. In these equilibria the demands of both types of both players are larger than  $1 - (1 - p)\theta$ ; i.e., in this equilibrium the players always coordinate on war.

Given this multiplicity, we restrict attention to equilibria in which, once the dispute has

arisen, the disputants coordinate on the equilibrium of the crisis bargaining game which maximizes the chances of peaceful resolution.<sup>5</sup>

Proposition 1 finds the equilibrium of the crisis bargaining game that maximizes the peace probability  $V$ , given any arming strategy  $q$ . Then, we find the equilibrium arming strategy  $q$  in the overall game without communication, given the solution of the crisis bargaining game from Proposition 1.

**Proposition 1** *The optimal equilibrium of the crisis bargaining game, as a function of the arming probability  $q$  is as follows. For  $q \geq \gamma/(\gamma + 1)$ , the players always achieve peace, by playing  $x_A = x_B = 1/2$ . For  $\gamma/(\gamma + 1) > q \geq \gamma/(\gamma + 2)$ , peace is achieved unless both disputants are hawks; hawk players demand  $x_H \in [p\theta, 1 - (1 - p)\theta]$  and dove players demand  $x_L = 1 - x_H$ . For  $q < \gamma/(\gamma + 2)$ , peace is achieved only if both disputants are doves, the demand of doves is  $x_L = 1/2$ , whereas hawks trigger war by demanding  $x_H > 1/2$ .*

This result has fairly natural intuition. When  $q$  is large, each player anticipates that the opponent is likely a hawk. If the opponent plays  $x_j = 1/2$  regardless of her type, then player  $i$  can secure the payoff  $1/2$ , by also playing  $x_i = 1/2$ . Because the opponent is likely a hawk, even a hawk prefers to induce this known payoff, rather than triggering war by making a claim larger than  $1/2$ . (Of course, this is true *a fortiori* for a dove). In fact, the condition  $q \geq \gamma/(\gamma + 1)$  corresponds to the inequality  $1/2 \geq q\theta/2 + (1 - q)p\theta$ , where the right hand side is the expected war payoff of a hawk. In this region of high  $q$  one could say that the idea of peace by deterrence is supported.

When  $q < \gamma/(\gamma + 1)$ , hawks will trigger war, unless they expect the opponent to make demands below  $1/2$ . Because the game is symmetric, this implies that peace cannot be achieved when both players are hawks. But doves are less willing to trigger war, and would accept an unequal split to avoid fighting a hawk. Hence it is possible to achieve peace in hawk-dove player dyads, unless the chance  $q$  that the opponent is hawk is too small. The

---

<sup>5</sup>Our choice of this selection criterion as opposed to one that minimizes the probability of war inclusive of arming decisions is motivated by the observation that before the initial, and possibly long, process of militarization, the disputants are unlikely capable to form commitments on how to play in the bargaining game, so as to achieve the equilibrium that is optimal in the whole game. Instead, when the crisis erupts, they will attempt to coordinate on the equilibrium that minimizes the chances of a destructive war.

demands  $x_H \in [p\theta, 1 - (1 - p)\theta]$  and  $x_L = 1 - x_H$  induce equal split among doves, and make sure that neither a hawk nor a dove player wants to deviate and trigger war. In this intermediate range of values of  $q$  we have only wars between hawks.

When  $q < \gamma/(\gamma + 2)$ , even doves are unwilling to accept unequal splits: the optimal equilibrium is such that  $x_A = x_B = 1/2$ , and peace is achieved only by dove dyads and peace by deterrence is not a feature of the equilibrium. In summary, when treating the level of militarization  $q$  as an exogenous parameter, the probability of peace is increasing in  $q$  (weakly) for any type of pair, but this monotonic deterrence effect will be significantly altered when militarization is endogenized.

Having determined the optimal equilibrium of the crisis bargaining game, we can move one step back and calculate the equilibrium arming strategy  $q$ . Note that for intermediate values of  $q$  many demands by the hawk are consistent with equilibrium. We select the one least favorable to the hawk; i.e., the one with  $x_H = p\theta$ . This is the natural choice as it minimizes the incentive for militarization among the equilibria listed in Proposition 1 and thus is the selection that first minimize the risk of war for a given conflict and then among this set of equilibria which all induce the same probability of war for a given conflict this one minimizes the risk of war in the larger game. This selection also simplifies the proof of the main claim in this section, that unmediated cheap talk may increase militarization incentives. Our characterization will depend on whether the cost of arming,  $k$ , is larger than the critical threshold  $k^* \equiv [(1 - \theta)\gamma^2]/[2(\gamma + 1)]$ . Through out we focus the presentation on the case that  $k$  is smaller than  $\bar{k} \equiv [(1 - \theta)\gamma(\gamma + 1)]/[2(\gamma + 2)]$ , as this greatly simplifies the exposition without affecting the insights that come out of the analysis.<sup>6</sup>

**Proposition 2** *Assume that the cost of arming  $k$  is smaller than  $\bar{k}$ . Given the solution in Proposition 1 with  $x_H = p\theta$  when  $\gamma/(\gamma + 1) > q \geq \gamma/(\gamma + 2)$ , the equilibrium of the game with endogenous militarization that maximizes welfare is as follows.*

1. For  $k \in [0, k^*)$ , the disputants cannot improve on the grim outcome in which they

---

<sup>6</sup>We limit attention to the parameter space with  $k \leq \bar{k}$  because it is the most interesting region of parameters, where the most surprising results obtain, but the full characterization of equilibrium for the remaining range of high costs of militarization is available upon request. Importantly,  $k^* < \bar{k}$ .

*both arm with probability  $q = 1$ , and then fight;*

2. *For  $k \in [k^*, \bar{k}]$ , each disputant militarizes with probability  $q(k) = \gamma - 2k/(1 - \theta)$ , and only hawk dyads fight.*

For small costs of militarization, i.e., when  $k < k^*$ , the players cannot improve on the grim outcome: they arm and fight. Indeed, fighting is always an equilibrium of the crisis bargaining game. For small costs of militarization, players prefer to arm in anticipation of a fight, rather than saving the militarization cost and losing the war. In fact, the disputants are unable to commit to remain unarmed and peaceful, or even to randomize. Because the cost of arming is small, each disputant would have an incentive to deviate, arm and wage war whenever the opponent remains weak, arms or randomizes. For  $k \in [k^*, \bar{k}]$ , the disputants neither strictly prefer to remain peaceful, nor to militarize. Indeed, they randomize at the militarization stage, to then fight if and only if they both armed. Not surprisingly, the arming probability decreases as the cost of militarization increases, to reach  $q = \gamma/(\gamma + 2)$  for  $k = \bar{k}$ .<sup>7</sup>

**Unmediated Peace Talks** We now augment the game by adding an extra intermediate stage. We assume that after the militarization stage, but before playing the crisis bargaining game, direct bilateral peace talks take place. At this stage the two players simultaneously send unverifiable messages  $m_i \in \{l, h\}$ ,  $i = A, B$ , to each other. For simplicity, we assume that only one round of communication takes place. But to broaden the scope and capabilities of bilateral peace talks, we allow players to make use of a randomization device, whose realization is observed by both disputants. This modeling strategy is less esoteric than it might seem. Aside from capturing literal publicly observe random events that players could condition their behavior on, it is known that public randomization devices can be reconstructed as simultaneous cheap talk (see, e.g. Aumann and Hart, 2003). By allowing for public randomization devices we are ensuring that the logic of our result holds for a large class of models of communication (albeit in a reduced form way). So, in our model,

---

<sup>7</sup>The equilibrium characterization changes discontinuously for  $k > \bar{k}$ , because, once the dispute has arisen, players fight unless they are both doves.

the disputants can coordinate on different equilibria of the crisis bargaining game, as a function of both their messages, and of the realization of the public randomization device.

Our model then provides a simple framework to represent meetings in which the disputants conduct unmediated talks, share information and attempt to coordinate future play. Any equilibrium has the following form: with some probability, the meetings will be successful, and the disputants will agree on a peaceful resolution. With complementary probability, the peace talk meetings will fail, and lead to open conflict. In any equilibrium in which information is meaningfully revealed the probability that the meeting results in a peaceful resolution will depend on their type dyad, as we detail below in Proposition 3.

In order to show that unmediated peace talks may strictly reduce the probability of peace and welfare, it is sufficient to find equilibria of the augmented game that induce higher militarization and war probability than the equilibria described in Proposition 2 for some values of the parameters. Our equilibrium selection criterion will focus again on the equilibrium that maximizes the peace probability  $V$ , given any arming strategy  $q$ , in which disputants adopt a pure strategy when declaring their types.<sup>8</sup> The characterization of this equilibrium is in Proposition 3.

**Proposition 3** *For any given militarization strategy  $q$ , the optimal pure strategy equilibrium of the crisis bargaining game with bilateral peace talks is characterized by two parameters,  $p_H, p_M$  and has the following form. Each type truthfully reveals her type during peace talks. Then, hawk dyads  $(H, H)$  coordinate on the peaceful demands  $x_A = 1/2, x_B = 1/2$  with probability  $p_H$ , and fight with probability  $1 - p_H$ ; asymmetric dyads  $(H, L)$  coordinate on the peaceful demands  $(p\theta, 1 - p\theta)$  with probability  $p_M$ , and fight with probability  $1 - p_M$  — the case for  $(L, H)$  is symmetric; and dove dyads  $(L, L)$  achieve peace with probability one, making demands  $x_A = 1/2, x_B = 1/2$ .*

*Specifically, when  $q < \gamma/(\gamma + 2)$ ,  $p_H = 0$  and  $p_M \in (0, 1)$ ; whereas when  $\gamma/(\gamma + 2) \leq q < \gamma/(\gamma + 1)$ ,  $p_M = 1$  and  $p_H \in (0, 1)$ ; finally, when  $q \geq \gamma/(\gamma + 1)$ ,  $p_M = 1$  and  $p_H = 1$ .*

*Whenever  $q < \gamma/(\gamma + 1)$ , unmediated cheap talk strictly improves the peace chance, for any*

---

<sup>8</sup>Again, the motivation is that, when the crisis erupts, disputants will attempt to coordinate on the equilibrium that minimizes the chances of a destructive war.

given militarization strategy  $q$ .

The equilibrium in Proposition 3 is best illustrated by comparing it to the optimal equilibrium of the bargaining game without communication, reported in Proposition 1. In both cases disputants play a separating equilibrium (except for the trivial case in which  $q \geq \gamma/(\gamma + 1)$  where peace is always achieved). So, in both equilibria, players reveal their types by means of their choices. But in the crisis bargaining game without cheap talk, this information is revealed only after demands are made, whereas with unmediated cheap talk, the information is shared before the demands are made. Intuitively, sharing information before demands are made allows the players to use this information more efficiently in the crisis bargaining game. In fact, unmediated peace talks induce a strict improvement in the probability of peace for  $q < \gamma/(\gamma + 1)$ . Specifically, when  $\gamma/(\gamma + 2) \leq q < \gamma/(\gamma + 1)$  and both players are hawk, peace talks turn the sure event of war (under the game without talks) into one in which peace occurs with positive probability. Similarly, when  $q < \gamma/(\gamma + 2)$ , war takes place with probability one in asymmetric hawk dove dyads, in the equilibrium of Proposition 1, but with probability smaller than one when unmediated peace talks are introduced.

We now proceed to describe the militarization strategies in the augmented game with unmediated peace talks, given that the players coordinate on the equilibria reported in Proposition 3.

**Proposition 4** *In the augmented game with unmediated bilateral peace talks, given the solution in Proposition 3, the disputants cannot improve on the grim outcome in which they both arm with probability  $q = 1$ , and then fight, for any  $k \in [0, \bar{k}]$ .*

*Hence, for  $k \in (k^*, \bar{k})$ , the expectation of unmediated peace talks induces more militarization, and lower peace chance, than in the equilibrium of the game without communication. For  $k \leq k^*$ , the equilibrium incentives are the same with and without unmediated peace talks.*

In sharp contrast with the benchmark case without communication, unmediated peace talks do not improve on the grim outcome in arming and fighting are certain, for the

relevant range  $[0, \bar{k}]$  of the cost parameter  $k$ . Hence, for any  $k \in (k^*, \bar{k})$ , unmediated peace talks may induce negative incentives for militarization and reduce the peace chance as well as overall welfare. The intuition for this seemingly perverse result is simple. Taking the probability that each nation is strong as given, cheap talk improves the chance for peaceful crisis resolutions, by reducing the probability of wasteful war when both nations are hawks (without any change in the hawks payoffs when they meet doves). This increases the payoff for nations to become strong, and thus in equilibrium nations arm with higher probability. Thus, although it is true that communication can improve the chances for peace when the militarization decisions are ignored, an analysis that captures how expectations of the bargaining process will unfold illustrates that this channel to improved negotiations creates particularly strong incentives for militarization and the latter effect dominates.

### 3 Third party intervention

We have shown that bilateral unmediated communication, while reducing the risk of conflict in ongoing crises, may yield more militarization and more destructive crises. We now show that there exist forms of third party interventions that do not suffer from these drawbacks. In fact, it is possible to avoid the militarization incentives mentioned above if the crises are expected to be dealt with by some kind of third party intervention, even if this third party has no enforcement power and still relies on the same information that is available with unmediated peace talks. That is, the third party has no special access to the private information of the two sides. This section establishes that there exist forms of mediation that change the communication incentives of players in a way that the crisis resolution objectives do not create the same militarization incentives created by the expectation of direct communication.

**Optimal Mediation, Given Militarization** We consider a type of procedural mediation that is often studied in mechanism design. We consider the case with a neutral or “honest broker” mediator who does not favor either of the disputants. Examples of such mediators might be Nobel Peace Prize winners like Martti Ahtisaari and Jimmy Carter, or

Nongovernmental Organizations like the International Crisis Group. The mediator has a mandate, or a preference, to minimize the probability of war when the international crisis erupts, after the militarization stage. That is, we assume our mediators have a narrow mandate to resolve the current conflict and do not take into account the incentives of disputants who anticipate their mediation techniques when they choose whether to militarize before negotiations. Indeed, it would not be realistic to presume that the mediator’s mandate included deterring strategic choices that took place before their intervention in the crisis. Rather, our mediator with her short-term mandate, takes the (symmetric) equilibrium militarization probability  $q$  as given, and tries to minimize the chance that this dispute ends in war. This assumption mirrors our selection in the benchmark bargaining game and its cheap-talk extension. There we focus on the best equilibria treating investment decisions as fixed. Our mediator has the same characteristics.

Also note that the mediator is not endowed with unrealistic levels of “commitment power” either. We assume that the mediator can commit to her proposal during the crisis, and will not renegotiate if one or both nations reject her recommendation. This means that we assume the mediator can quit and credibly refuse to broker any subsequent deals, leading to escalation of conflict and war. We also assume that the disputants cannot broker their own deals following the decision to reject the mediator’s offers. This mediation institution is close to what mediation scholars call procedural mediation (Bercovitch,1997).

We suppose that the mediator does not have access to any privileged or private information. She privately solicits unverifiable and costless messages from the disputants and then makes settlement recommendations. What  $A$  and  $B$  tell the mediator is entirely up to them. Upon hearing the settlement recommendation, the disputants,  $A$  and  $B$ , then bargain on the split of the pie according to the same crisis bargaining game described in the benchmark case above.

As in the previous section, we first derive the optimal solution of the bargaining game with such a mediator, taking the militarization probability,  $q$ , as exogenous and then we consider the induced equilibrium probability of arming.

**Proposition 5** *For any given militarization strategy  $q$ , the optimal equilibrium of a the*

*crisis bargaining game with an unbiased mediator can be characterized by three parameters,  $q_H, q_M, p_M$  and described as follows: Each player truthfully reveals her arms level to the mediator. Then, hawk dyads  $(H, H)$  coordinate on the peaceful demands  $(1/2, 1/2)$  with probability  $q_H$ , and fight with probability  $1 - q_H$ ; asymmetric dyads  $(H, L)$  coordinate on the peaceful demands  $(p\theta, 1 - p\theta)$  with probability  $p_M$ , on the demands  $(1/2, 1/2)$  with probability  $q_M$  and fight with probability  $1 - p_M - q_M$  — the case for  $(L, H)$  is symmetric; and dove dyads  $(L, L)$  achieve peace with demands  $(1/2, 1/2)$  with probability one.*

*Specifically, when  $q \leq \gamma/(\gamma + 2)$ ,  $q_H = q_M = 0$ , and  $p_M \in (0, 1)$ ; when  $\gamma/(\gamma + 2) \leq q < \gamma/(\gamma + 1)$ ,  $p_M + q_M = 1$ ,  $q_H \in (0, 1)$  and  $q_M \in (0, 1)$ ; and finally, when  $q \geq \gamma/(\gamma + 1)$ ,  $q_M = 1$ , and  $q_H = 1$ .*

*Whenever  $\gamma/(\gamma + 2) \leq q < \gamma/(\gamma + 1)$ , mediation improves the peace chance relative to unmediated cheap talk. For all values of  $q$ , mediation yields at least as large chance of peace as unmediated cheap talk.*

The proposition illustrates that the optimal form of mediation involves some degree of opaqueness, or obfuscation, in the mediator's choice of recommendations.<sup>9</sup> Specifically, mediators optimally obfuscate hawks' beliefs. In equilibrium, when a hawk receives recommendation  $(p\theta, 1 - p\theta)$ , she knows for sure that the opponent is a dove; whereas when receiving recommendation  $(1/2, 1/2)$ , the hawk is unsure about the opponent's type. Mediators, importantly, sometimes take actions that make hawks believe their opponents are also hawks. Specifically, when receiving the reports stating that one player is a hawk and the other is a dove, the mediator recommends the symmetric split  $(1/2, 1/2)$  with probability  $q_M$ , instead of the asymmetric splits  $(p\theta, 1 - p\theta)$ . This strategy is equivalent to not revealing to a hawk that the opponent is a dove. We shall later see how this optimal mediation strategy, which minimizes the difference between the hawk's and the dove's utility in the mediation game, leads to optimal incentives for militarization.

---

<sup>9</sup>A mediator who always truthfully relays the players' reports cannot achieve a peace chance higher than the peace chance achieved with unmediated direct peace talks.

**Mediation and Militarization** We now build on the above result to solve for the militarization equilibrium probability  $q$  given that the ensuing dispute is solved via mediation. The following result characterizes the equilibrium militarization and conflict probability. Importantly, it shows that mediators looking to reach peaceful settlements in the current crisis do not suffer from the drawbacks presented by unmediated bilateral peace talks. Not only can a mediator improve the peace chance given militarization strategies  $q$ . Such a mediator can also improve welfare in the whole game, which captures the “upstream” militarization decisions.

**Proposition 6** *In the crisis bargaining game with a peace seeking mediator the best equilibrium can be characterized as follows: For any militarization cost  $k$ , there is a unique symmetric equilibrium militarization probability  $q(k)$ , which strictly decreases in  $k$  for  $k \in [0, \bar{k}]$ , with  $q(0) = \gamma/(\gamma + 1)$ , and  $q(\bar{k}) = \gamma/(\gamma + 2)$ .<sup>10</sup>*

*Peace seeking mediators improve the chance of peace  $V$  and the welfare  $W$  with respect to unmediated peace talks, and to the benchmark double-auction without communication.*

Remarkably, mediation improves the chances of peaceful dispute resolution without creating negative distortions to the militarization decisions. The key to understanding how this is possible hinges on seeing how mediation differs from cheap-talk. Our mediator can hide information that is in the open with direct cheap talk. Specifically, it may hide information so as to make nations that arms more concerned about the strength of their opponent than those who choose not to arm. This makes a militarized nation willing to accept a less attractive settlement and lowers the overall expected payoff to arming (compared to cheap talk). As a consequence, nations are less willing to invest resources in militarization than when they expect the crisis to be settled by direct bilateral negotiations. As a result, the probability of being strong is lower in equilibrium, and the probability of war given any profile of militarization levels is no higher in the best mediated equilibrium than in the best cheap talk equilibrium (it is also no higher than in the benchmark without communication).

---

<sup>10</sup>The explicit formula for the mixing probability is cumbersome, and relegated to the appendix.

A more formal intuition can be grasped by noting that, given any dispute resolution institution, the equilibrium probability  $q$  of strategic militarization is given by

$$q = \frac{U(H, L) - U(L, L)}{U(H, L) - U(L, L) + U(L, H) - U(H, H)}, \quad (1)$$

where  $U(t_A, t_B)$  is player  $A$ 's expected payoff in the dispute, when the players' types are  $t_A$  and  $t_B$ . Hence, the equilibrium militarization strategy  $q$  increases in  $U(H, L) - U(L, L)$ , the gain for being a hawk instead of a dove, when facing a dove. Similarly,  $q$  decreases in  $U(L, H) - U(H, H)$ , the loss for being a hawk instead of a dove, when facing a hawk. Peace seeking mediation penalizes choosing to be a hawk more than bilateral and direct cheap talk, i.e., it makes  $U(H, L) - U(L, L)$  smaller and  $U(L, H) - U(H, H)$  larger. Hence, it makes the militarization strategy  $q$  smaller in equilibrium.

**The Institutional Optimality of Mediation** Recall that when a mediator is called in a dispute, her mandate is usually set within the boundaries of the dispute. The objective of preventing disputes and militarization is not realistically part of her commitment and to a large degree the militarization decision of the current parties to the dispute were made in the past. The success of the mediator's efforts are likely judged solely on the basis of how the current dispute is resolved.

But as we have seen because an institution (bilateral diplomacy or mediation) can effectively reduce the risk of war in the ongoing dispute, it may paradoxically incentivize arming, thus ultimately raise the risk of war. This concern mirrors the perverse effect of unmediated bilateral communication above.

We just showed that peace seeking mediation does not suffer from this drawback: it improves the chances of peaceful dispute resolution without creating negative militarization distortions. SO while the exact nature of the bargaining or mediation protocol will likely have affects on exact settlements, the probability of war, and the chance of peace, in general there should be a difference in militarization strategies across crises conditional on observed dispute settlement techniques. In fact, empirically, this is sometimes the case. For example, looking at the data from Svensson's (2007) study on the effects of direct negotiations and mediation on agreements in all intrastate armed conflicts between 1989 and 2003, one finds

that if you compare the size of government armies in crises with just bilateral negotiations to those with a mediator the first category averages 273,901 troops versus when a mediator is used it is only 124,277. A simple t-test rejects the hypothesis on no difference in these case at better than a .001 level.

Next, we will now show something much stronger, peace seeking mediation achieves the same welfare as a hypothetical optimal third party intervention mechanism, in which the third party has the broader mandate to commit to ‘punish’ nations for militarizing and entering a dispute in the first place. That is, a mediator with the mandate to consider total welfare, and balance long-term militarization incentives and short-run fighting incentives, can do no better than the short-sighted peace seeking mediator.

Such far-sighted institution is not likely available in the real world, and it serves only as benchmark to assess how much is lost by the narrow mandate. We only maintain the assumption that the optimal institution is budget balanced: That is, it cannot bribe or punish the players to force them to settle.

**Theorem 1** *Although peace seeking mediators may not take into account the incentives that they create for strategic militarization, they are an optimal institution in the militarization game, both in terms of arms reduction and peace chance maximization, among all budget-balanced institutions.*

An intuitive description of the proof of Theorem 1 may be obtained by considering equation (1) again. Recall from Proposition 5 that, under optimal mediation, (i) dove pairs never fight, (ii) the settlement  $(p\theta, 1 - p\theta)$  is chosen so as to keep the payoff of hawks meeting doves as low as possible, and (iii) hawk pairs are more likely to fight than any other type pair. Hence, she minimizes the incentives of a weak nation to pretend that it is strong. As an unintended consequence, she also minimizes the incentives for a weak nation to become strong. That is, peace seeking mediation minimizes  $U(H, L) - U(L, L)$  and maximizes  $U(L, H) - U(H, H)$ , among all budget-balanced institutions. Hence, she keeps the upstream incentives to arm in check, and minimizes strategic militarization  $q$  in equilibrium. In other words, such a mediator induces the same militarization incentives as

the hypothetical institution whose mandate includes deterring militarization which took place before the eruption of the crisis. Because, by construction, the mediator's strategy in Proposition 5 maximizes the chance of peace given any militarization strategy  $q$ , the fact that it also minimizes the equilibrium militarization probability implies that it maximizes the overall disputants' welfare, among all budget-balanced institutions.

## 4 Conclusion

This paper pushes for scholars of conflict to broaden their field of vision in thinking about institutions. How nations negotiate can affect more than whether the crisis at hand is resolved peacefully; the institution can have pronounced effects on the types of conflicts that are likely to emerge by shaping the incentive for nations to militarize. Our results show that not all forms of conflict resolution are equal. In particular, there are at least two dimensions upon which the conflict reducing effects of an institution should be considered. First is what we may call the short-run effect. The short-run effect is how a particular type of dispute resolution procedure—like a summit to enable bilateral peace talks or a mediated third party intervention—might increase the chance of peace today. From the short-run perspective direct peace talks look like they generate a significant benefit to all parties looking to avoid war. For every possible configuration of parameters the best cheap talk equilibrium has less conflict than the best equilibrium from the bargaining game without cheap talk.

In the bigger picture, however, there is a second effect we may call the long-run effect. The long-run effect is how the method of conflict resolution changes incentives to militarize and, therefore, shapes the types of conflict that arise. We show that which dispute resolution mechanism an actor expects to face can matter. Specifically when the cost of arming is low, but not too low, the equilibrium level of strategic arming generates a positive probability of peace when there is no direct communication, but war for sure when peace talks are expected. Mediation, on the other hand, improves the probability of peace in both the long and the short run. These changes in the probability of war result from how expected peace settlements change the expected payoffs to actors with different levels of

arms making some strategies in the short run game untenable when the players strengths are a function of their militarization strategies.

Our analysis depends on a number of assumptions, like that there are two possible levels of arms, that uncertainty is about the strength of players as opposed to their costs or preference for fighting, and that the mediator can commit to allowing war if her offer is rejected. The assumptions are important for our analysis and the construction of equilibria to this crisis analysis, but they are not crucial to our general point. We chose to analyze bilateral cheap talk communication and mediation because they are often studied in the conflict literature and represent seemingly similar ways to facilitate peace. The fact that they obtain peace in ways that are subtly distinct, but have decidedly different affects on upstream incentives suggests that this line of reasoning needs to be expanded.

Beyond broadening the focus of theoretical work on institutions to a study of short and long run incentives, this paper makes an important contribution to our understanding of communication and bargaining. Although the point must be developed within the context of a series of formal papers, the central conclusion transcends this approach and provides a logic about what makes third-party intervention potentially valuable. A series of paper going back to at least Kydd (2003) focus on the question of whether mediation can improve on direct communication. At the heart of this debate is whether there is value in a mediator with no independent private information and no ability to create external incentives. In the case of private values the answer is no (Fey and Ramsay 2010) but in the context of interdependent values the answer is yes (Horner Morelli Squintani, 2010). In this paper we have focused on how mediated and unmediated communication can differ as the latter need not created distortions to militarization whereas the latter does. The difference hinges on the ability of a mediator to selectively shape the beliefs players assign to their opponent being a hawk at the time that they must accept or reject an offer on the table. In settings with interdependent values this belief is important as it influences what a Hawk will accept. But in settings with private values the type of the other nation is not payoff relevant and thus the ability to influence this belief buys a mediator nothing. One take away of this is a better understanding of what makes a mediator effective – the ability to obfuscate by being

willing to sometimes let the participants fight when alternative settlements are possible – , and when such attributes can have the most value – when the uncertainty is about a common-values component. There is scope for improvement over direct communication by the disputants if a third-party has the stomach to sometimes stop short of brokering a settlement. This tendency, while seemingly counterproductive, can result in a type of obfuscation that ultimately discourages the very militarization that raises the risk of costly war-fighting in the first place.

## References

- [1] Allison, G.T., A. Carnesale, and J.S. Nye Jr (1985) “Hawks, Doves and Owls: A New Perspective on Avoiding Nuclear War,” *International Affairs* (Royal Institute of International Affairs), Vol. 61, No. 4, 581-89.
- [2] Baliga, S. and T. Sjöström (2004): “Arms Races and Negotiations,” *Review of Economic Studies*, **71(2)**, 351–69.
- [3] Bercovitch, J. (1997): “Mediation in International Conflict: An Overview of Theory, a Review of Practice,” in I.W. Zartmann and J.L. Rasmussen (eds) *Peacemaking in International Conflict: Methods and Techniques*, Washington D.C.: United States Institute of Peace Press.
- [4] Beardsley, K. (2008) “Agreement without Peace? International Mediation and Time Inconsistency,” *American Journal of Political Science*, Vol. 52, No 4, October, pp.723-740.
- [5] Bercovitch, J., J.T. Anagnoson, and D. Wille (1991): “Some conceptual issues and empirical trends in the study of successful mediation in international relations,” *Journal of Peace Research* **28**, 7–17.
- [6] Bercovitch, J. and A. Houston (2000): “Why Do They Do It? An Analysis of the Factors Influencing Mediation Behavior in International Conflicts,” *Journal of Conflict Resolution*, **44(2)**, 170–202.

- [7] Bercovich, J. and R. Jackson (2001): “Negotiation or Mediation? An Exploration of Factors Affecting the Choice of Conflict Management in International Conflict,” *Negotiation Journal*, **17**, 59–77.
- [8] Bester, H. and K. Wärneryd (2006): “Conflict and the Social Contract,” *Scandinavian Journal of Economics*, **108(2)**, 231–49.
- [9] Blainey, Geoffrey. (1973): *The Causes of War*, New York: the Free Press.
- [10] Brito, D.L. and M.D. Intriligator (1985): “Conflict, War, and Redistribution,” *American Political Science Review*, **79(4)**, 943–57.
- [11] Chapman, Terrence L. and Scott Wolfrord. (2010) “International Organization, Strategy, and Crisis Bargaining.” *International Organization*, **71(1)**, 227–242.
- [12] Chassang, S. and G. Padro i Miquel (2008) “Conflict and Deterrence under Strategic Risk,” NBER Working Paper 13964.
- [13] Collier, Peter and Anke Hoffler. (2006) “Military Expenditures in Post-Conflict Societies.” *Economics and Governance*, **7(1)**, 87–107.
- [14] Dixon, W. J. (1996): “Third Party Techniques for Preventing Conflict Escalation and Promoting Peaceful Settlements,” *International Organization*, **50:3**, 653–68.
- [15] Fang, Songying. (2010): “The Strategic Use of International Institutions in Dispute Settlement.” *Quarterly Journal of Political Science*, **63(2)**, 107–131.
- [16] Fearon, J.D. (1995): “Rationalist explanations for war,” *International Organization*, **49:3**, 379–414.
- [17] Fearon, James D. (1997) “Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs.” *The Journal of Conflict Resolution*, **41(1)**, 68–90.
- [18] Fey, M. and K. Ramsay (2009): “Mechanism Design Goes to War: Peaceful Outcomes with Interdependent and Correlated Types,” *Review of Economic Design*, **13(3)**, 233–250.

- [19] Fey, M. and K. Ramsay (2010): “When Is Shuttle Diplomacy Worth the Commute?: Information Sharing through Mediation”, *World Politics*, 62(4): 529-560.
- [20] Fisher, R.J. (1995): “Pacific, Impartial Third Party Intervention in International Conflict: a Review and Analysis.” In *Beyond Confrontation: Learning Conflict Resolution in the Post Cold War Era*. Ed. Vasquez, Johnson, Jaffe and Stamato, Ann Arbor: University of Michigan Press, 39–59.
- [21] Frazier, Daniel V. and W. J. Dixon. (2006) “Third Party Intermediaries and Negotiated Settlements, 1945-2000,” *International Interactions*, **32:4**, 385-408.
- [22] Garfinkel, Michelle R. (1990) “Arming as a Strategic Investment in a Cooperative Equilibrium,” *American Economic Review*, Vol. **80:1**, 50-68.
- [23] Garfinkel, Michelle R. and Stergios Skaperdas (1996) *The Political Economy of Conflict and Appropriation*, editors, Cambridge: Cambridge University Press.
- [24] Gilligan, Michael, Leslie Johns, and B. Peter Rosendorf. (2010) “Strengthening International Courts and the Early Settlement of Disputes.” *Journal of Conflict Resolution*, **54(1)**, 5–38.
- [25] Horner, J., M. Morelli and F. Squintani (2010): “Mediation and Peace”, mimeo, Yale University.
- [26] Intriligator, Michael D. (1964) “Some Simple Models of Arms Races,” *General Systems*, 9, 143-147.
- [27] Jackson, M.O. and M. Morelli (2011): “Reasons for War: an Updated Survey,” *Handbook on the Political Economy of War* edited by Chris Coyne, Elgar Publishing.
- [28] Jackson, M.O., and M. Morelli (2009): “Strategic Militarization, Deterrence and Wars,” *Quarterly Journal of Political Science*, **Vol. 4-4**, 279-313;
- [29] Jervis, Robert (1978): “Cooperation under the Security Dilemma,” *World Politics*, 30(2), 167-214.

- [30] Kelman, H.C. (1958): “Compliance, Identification, and Internalization: Three processes of attitude change,” *Journal of Conflict Resolution* **2**, 51–60.
- [31] Kleiboer, M. (1996): “Understanding Success and Failure of International Mediation,” *Journal of Conflict Resolution*, **40(2)**, 360–89.
- [32] Kydd, Andrew (2003): “Which Side Are You On? Bias, Credibility, and Mediation,” *American Journal of Political Science*, **47(4)**, 597–611.
- [33] Kydd, Andrew (2000): “Arms races and arms control: Modeling the hawk perspective,” *American Journal of Political Science*, 44 (2), 228-44.
- [34] Kydd (2006)
- [35] Matthews, Steven A., and Andrew Postlewaite. “Pre-play communication in two-person sealed-bid double auctions.” *Journal of Economic Theory*, Vol. 48, no. 1: pp. 238-263.
- [36] Meirowitz, A.H. and K. Ramsay (2009): “Investing in Failure: Equilibrium Constraints on Prior Beliefs in Bargaining Problems,” mimeo: Princeton University.
- [37] Meirowitz, A.H. and A.E. Sartori (2008): “Strategic Uncertainty as a Cause of War,” *Quarterly Journal of Political Science*, **3**, 327–352.
- [38] Myerson, R.B. (1979): “Incentive Compatibility and the Bargaining Problem,” *Econometrica*, **47(1)**, 61–73.
- [39] Myerson, R.B. (1982): “Optimal coordination mechanisms in generalized principal-agent problems,” *Journal of Mathematical Economics*, 10(1): 67–81.
- [40] Nash, John. (1953) “Two-person bargaining games.” *Econometrica*, Vol. 21, No. 1: pp. 128-140.
- [41] Powell, R. (1993), “Guns, Butter and Anarchy,” *American Political Science Review*, 87, 115-32;

- [42] Powell, R. (2006) "War as a Commitment Problem," *International Organization*, 60, 169-203.
- [43] Raiffa, H. (1982): *The Art and Science of Negotiation*, Harvard University Press.
- [44] Ramsay, Kristopher W. (2011). "Cheap Talk Diplomacy, Voluntary Negotiations, and Variable Bargaining Power," *International Studies Quarterly*, Vol. 55, No. 4, December, pp. 1003-1023.
- [45] Rauchhaus, R. W. (2006): "Asymmetric Information, Mediation and Conflict Management," *World Politics*, 58, 207-241.
- [46] Sartori, Anne E. (2002), "The Might of the Pen: A Reputational Theory of Communication in International Disputes," *International Organization*, Vol. 56, No. 1, Winter, pp. 123-151.
- [47] Schelling, Thomas C. (1966): *Arms and Influence*, New Haven and London, Yale University Press.
- [48] Smith, A. (1998) "Fighting Battles, Winning Wars," *The Journal of Conflict Resolution*, Vol. 42, No. 3, Opening up the Black Box of War: Politics and the Conduct of War, June, pp. 301-320.
- [49] Smith, A. (1998) "International Crises and Domestic Politics," *American Political Science Review*, Vol. 92, No. 3, September, pp. 623-638.
- [50] Savun, B. (2008) "Information, Bias, and Mediation Success", *International Studies Quarterly*, Vol. 52, No. 1. pp. 25-47.
- [51] Trager, R. (2011) "Multi-Dimensional Diplomacy", *International Organization*, Vol. 65, No. 3. Summer, pp. 469-506.
- [52] Touval, S., and I.W. Zartman (1985): "Introduction: Mediation in theory," In *International mediation in theory and practice*, edited by S. Touval and I.W. Zartman, 7-17. Boulder, CO: Westview.

- [53] United Nations General Assembly Security Council (2000): Report A/55/305-S/2000/809.
- [54] Wagner, R. H. (2000): “Bargaining and War,” *American Journal of Political Science*, 44(3), 469-84.
- [55] Wall, J.A. and A. Lynn (1993): “Mediation: A Current Review,” *Journal of Conflict Resolution*, **37(1)**, 160–94.
- [56] Watkins, M. (1998): “Building Momentum in Negotiations: Time-Related Costs and Action-Forcing Events,” *Negotiation Journal*, **14(3)**, 241–256.
- [57] Wilkenfeld, J., Young, K.J., Quinn, D.M. and V. Asal (2005): *Mediating International Crises*, Oxon, U.K.: Routledge.
- [58] Wittman, Donald. (2009) “Bargaining in teh Shadow of War: When is a Peaceful Resolution Most Likely?” *American Journal of Political Science*, **55(4)**, 588–602.
- [59] Young, O.R. (1967): *The intermediaries: Third parties in international crises*, Princeton, N.J.: Princeton University Press.

## 5 Appendix

**Proof of Proposition 1.** First, note that for  $q\theta/2 + (1 - q)p\theta \leq 1/2$ , or  $q \geq \gamma/(\gamma + 1)$ , both doves and hawks can achieve peace by coordinating on the claims  $x_A = x_B = 1/2$ . When  $q < \gamma/(\gamma + 1)$ , it is impossible that hawk dyads achieve peace. But it is possible to achieve peace for all other type dyads, as long as the hawk's claim is compatible with an opponent dove's demand. This is achieved by setting  $x_L + x_H = 1$ , so that peace can be achieved in equilibrium. Also, a hawk must prefer to demand  $x_H$  rather than triggering war against a dove by making a higher demand (if meeting a hawk, the demand will result in war anyway). Hence, we need that  $x_H \geq p\theta$ . Further, a dove must prefer to post her demand  $x_L$  rather than triggering war with a hawk, but collecting a higher share of the pie with a dove, by making the demand  $1 - x_L$ . This requirement translates in the following inequality:  $(1 - q)/2 + qx_L \geq (1 - q)(1 - x_L) + q(1 - p)\theta$ . Bringing together these conditions, we obtain the condition that  $q \geq \gamma/(\gamma + 2)$ . When this condition fails, it is impossible to achieve peace for mixed hawk-dove dyads. As a result, only dove dyads will achieve peace, by making compatible claims  $x_L = 1/2$ . ■

**Proof of Proposition 2.** We calculate the equilibrium militarization strategy  $q$  given the solution of the crisis bargaining game in Proposition 1.

We first search for completely mixed strategies, i.e., we impose the indifference condition

$$I_L(q) = I_H(q) - k;$$

where  $I_L(q)$  and  $I_H(q)$  are the interim payoffs of doves and hawks respectively.

On the basis of Proposition 1, there are two cases to consider.

Case 1,  $q \leq \gamma/(\gamma + 2)$  Here, peace is only achieved by dove dyads, and so:

$$I_L(q) = q(1 - p)\theta + (1 - q)/2$$

$$I_H(q) = q\theta/2 + (1 - q)p\theta.$$

Solving the indifference condition, we obtain

$$k(q) = \frac{1}{2}q(1 - \theta) + p\theta - \frac{1}{2},$$

which is an increasing expression in  $q$ , such that  $k(0) = p\theta - 1/2$ .

Case 2.  $\gamma/(\gamma + 2) < q \leq \gamma/(\gamma + 1)$ . Here, only hawk dyads result in war. Because  $x_H = p\theta$  and  $x_L = 1 - p\theta$ , we obtain:

$$I_L(q) = q(1 - p\theta) + (1 - q)/2$$

$$I_H(q) = q\theta/2 + (1 - q)p\theta.$$

Solving the indifference condition yields:

$$k(\lambda) = (1 - \theta) \frac{\gamma - \lambda(1 - \gamma)}{2(\lambda + 1)},$$

which is a decreasing function in  $\lambda$  such that  $k^+(\gamma/2) = (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)} < k^-(\gamma/2) = \frac{(\gamma+3)\gamma}{2(\gamma+2)}(1 - \theta)$  and  $k(\gamma) = (1 - \theta) \frac{\gamma^2}{2(\gamma+1)}$ . Inverting the expression  $k(\lambda)$ , and changing variable to  $q$ , we obtain:

$$q(k) = \gamma - \frac{2k}{1 - \theta}$$

Hence, we obtain that for  $k \in [0, (1 - \theta) \frac{\gamma^2}{2(\gamma+1)})$ , there is no completely mixed strategy equilibrium  $q$ .

For  $k \in [(1 - \theta) \frac{\gamma^2}{2(\gamma+1)}, (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}]$ , the unique completely mixed strategy equilibrium  $q(k) = \gamma - \frac{2k}{1 - \theta}$  is decreasing in  $k$ , and such that  $\lambda \left( (1 - \theta) \frac{\gamma^2}{2(\gamma+1)} \right) = \gamma$  and  $\lambda \left( (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)} \right) = \gamma/2$ .

For  $k \in ((1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}, (1 - \theta) \frac{\gamma}{2})$ , there is no completely mixed strategy equilibrium  $q$ .

Considering that, for  $k \geq (1 - \theta) \frac{\gamma}{2}$ , there is a pure strategy equilibrium, in which  $q = 0$ , we can disregard case 1, which yields higher militarization probability and lower welfare.

Now consider pure-strategies. Suppose that  $q = 0$ , then  $I_L(q) = 1/2$  and  $I_H(q) = p\theta$ . Hence,  $q = 0$  is an equilibrium if and only if  $k \geq p\theta - 1/2 = (1 - \theta) \frac{\gamma}{2}$ . In the remaining region,  $k < (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}$ , the ‘grim’ equilibrium, in which disputants arm and then fight is the unique equilibrium. ■

**Proof of Proposition 3.** In order to prove this result, we first reformulate the problem by substituting the last stage of the game, the double auction game, with the following, simpler,

model. Given messages  $m = (m_A, m_B)$ , and the outcome of the public randomization device, nature selects a split proposal  $x$ , and the players simultaneously choose whether to agree to the pie division  $(x, 1 - x)$ .

It is evident that any outcome of this simpler model is also an equilibrium of our model. Suppose, in fact, that the players agree to the split division  $(x, 1 - x)$  in the simpler model. Then, they can achieve the outcome  $(x, 1 - x)$  in the double auction game, by making demands  $x_A = x$  and  $x_B = 1 - x$ .

Focusing on fully-separating equilibria, one can also establish the converse result, that any equilibrium of our model (with the double auction game) is also an outcome of this simpler model. In fact, when types fully separate at the cheap talk stage, each player's type is common knowledge in the double auction stage. Hence, in equilibrium, the players know each other's demands. Suppose that, in equilibrium, player  $A$  demands  $x$ . Player  $B$ 's best response is either to demand  $1 - x$ , or to trigger war with a higher demand. Hence, either peace takes place with the split  $(x, 1 - x)$ , or war occurs. Player  $B$ 's choice of whether to trigger war or make the demand  $1 - x$  follows exactly the same calculation that she would make in the simpler model we introduced above, if nature selected the split proposal  $(x, 1 - x)$ . Because the same argument applies to player  $A$ , we have shown that any fully-separating equilibrium of our model (with the double auction game) is also an outcome of this simpler model.

Note, now, that pure-strategy equilibria of our model belong to two different categories: pooling equilibria and fully-separating equilibria. Of course, the pooling equilibria coincide with the equilibria of the game of conflict without peace talks. Hence, the solution of our problem can be achieved by comparing the equilibria described in Proposition 1, with the optimal (fully-separating) outcomes of the simpler game introduced in this proof, which are fully characterized in Lemma 1 in Horner Morelli and Squintani (2010). These outcomes are the one reported in the statement of Proposition 3. They induce a higher peace chance than the equilibria described in Proposition 1, because it is the case that  $p_H > 0$ , when  $\gamma/2 \leq \lambda < \gamma$ , and that  $p_M > 0$ , when  $\lambda < \gamma/2$ . ■

**Proof of Proposition 4 .** In the notation of the separating equilibrium, for any given militarization probability, the interim payoffs are:

$$I_L(q) = q(p_M(1-b) + (1-p_M)(1-p)\theta) + (1-q)/2$$

$$I_H(q) = q(p_H/2 + (1-p_H)\theta/2) + (1-q)(p_M b + (1-p_M)p\theta).$$

We search for completely mixed strategies, i.e., we impose that  $I_L(q) = I_H(q) - k$ . There are two cases.

Case 1.  $\lambda \leq \gamma/2$ . Substituting the solution described in Proposition 3 into the above expressions, we obtain:

$$I_L(q) = q \left[ \frac{1}{1+\gamma-2\lambda}(1-p\theta) + \left(1 - \frac{1}{1+\gamma-2\lambda}\right) (1-p)\theta \right] + (1-q)/2$$

$$I_H(q) = q\theta/2 + (1-q)p\theta.$$

Solving the indifference condition, and reparametrizing to get rid of  $p, q$ , we obtain the  $k$  which makes the players indifferent for  $\lambda$  and  $\gamma$  and  $\theta$  fixed:

$$k(\lambda) = (1-\theta) \frac{(\gamma - \lambda + \lambda\gamma - 2\lambda^2)(\gamma + 1)}{2(\gamma - 2\lambda + 1)(\lambda + 1)}$$

Because  $2\lambda \leq \gamma$ , this is always positive.

We first differentiate  $k(\lambda)$ ,

$$\frac{\partial k(\lambda)}{\partial \lambda} = \frac{1}{2}(\gamma - 2\lambda + 1)^{-2}(\lambda + 1)^{-2}(\gamma - 4\lambda - 1)(\gamma + 1)(1 - \theta)$$

$$\propto \gamma - 4\lambda - 1$$

The expression is positive for  $\lambda < (\gamma - 1)/4$  and negative for  $\lambda > (\gamma - 1)/4$ , on the range  $\lambda \in [0, \gamma/2]$ . Then, we calculate the extremes of the range:  $k(0) = (1 - \theta) \frac{\gamma}{2}$  and  $k(\gamma/2) = (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}$ . This concludes that the function  $k(\lambda)$  equals  $(1 - \theta) \frac{\gamma}{2}$  at  $\lambda = 0$  then increase until  $\lambda = (\gamma - 1)/4$  and then decrease until  $\lambda = \gamma/2$  reaching  $(1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}$ . Noting that  $(1 - \theta) \frac{\gamma}{2} > (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)} > 0$ , we determine the following conclusions:

For  $[k \in (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}, (1 - \theta) \frac{\gamma}{2})$ , there exists a unique equilibrium  $\lambda(k)$ , the function  $\lambda$  is strictly decreasing in  $k$ , it starts at  $\lambda = \gamma/2$  for  $k = (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}$  and reaches  $\lambda(q) =$

0 for  $k = (1 - \theta) \frac{\gamma}{2}$ . The explicit equilibrium solution is cumbersome, and its omission inconsequential.

For  $k \in [0, (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)})$ , there does not exist any equilibrium such that  $\lambda \leq \gamma/2$ .

Case 2.  $\lambda \in [\gamma/2, \gamma)$ . The interim payoffs are:

$$I_L(q) = q(1 - p\theta) + (1 - q)/2$$

$$I_H(q) = q \left[ \frac{2\lambda - \gamma}{\lambda(\gamma + 2)}(1/2) + \left(1 - \frac{2\lambda - \gamma}{\lambda(\gamma + 2)}\right) \theta/2 \right] + (1 - q)p\theta.$$

Hence, the indifference condition yields:

$$k = (1 - \theta) \frac{(\gamma + 1)\gamma}{2(\gamma + 2)},$$

which is constant in  $\lambda$ . So, for  $k = (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}$ , all  $\lambda \in [\gamma/2, \gamma)$  are an equilibrium, and there is no completely mixed equilibrium for  $k < (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}$ .

Now consider pure-strategies. Suppose that  $q = 0$ , then  $I_L(q) = 1/2$  and  $I_H(q) = p\theta$ . Hence,  $q = 0$  is an equilibrium if and only if  $k \geq p\theta - 1/2 = (1 - \theta) \frac{\gamma}{2}$ . In the remaining region,  $k < (1 - \theta) \frac{(\gamma+1)\gamma}{2(\gamma+2)}$ , the ‘grim’ equilibrium, in which disputants arm and then fight is the unique equilibrium. ■

**Proof of Proposition 5.** The proof of this result consists in showing that the optimal strategies of a Myerson mediator coincide with the optimal strategies of a mediator characterized in Horner, Morelli, and Squintani (2010), Lemma 4.

The basic set up is the same in the two papers: There are two players, whose strength can be high with probability  $q$  and of low with probability  $1 - q$ ; if fighting, the payoffs are  $(\theta/2, \theta/2)$  when players have the same strength, and  $p\theta, (1 - p)\theta$  if the first player is stronger. Unlike here, when modeling mediation in HMS, we directly appeal to the revelation principle (Myerson, 1982) and restrict attention, without loss of generality, to direct revelation mechanisms. In a direct revelation mechanism the players privately report their “types”, in this case their realized level of militarization, to the mediator. Then, the mediator makes a recommendation to the players, possibly randomizing across recommendations. In this context, such a recommendation may also be to go to “war”. Furthermore, without loss of

generality, we restricted attention to truthful equilibria of the direct and obedient revelation mechanism, in which the mediators strategies are such that players reveal their types truthfully to the mediator, and the mediator's recommendation are obeyed by the players. By the revelation principle we know that the ex-ante peace probability induced by *any* mediation scheme, within the class of games described above, can also be achieved as the truthful equilibrium of the game induced by the direct and obedient revelation mechanism. For future reference, we here report the mediation program derived in HMS.

$$\min_{b, p_H, p_M, q_M, p_L, q_L} (1 - 2p_L - q_L)(1 - q)^2 + (1 - p_M - q_M)2q(1 - q) + (1 - q_H)q^2$$

subject to the high-type *ex post* IR constraint

$$bp_M \geq p_M p \theta, \quad (qq_H + (1 - q)q_M) \cdot 1/2 \geq qq_H \theta/2 + (1 - q)q_M p \theta,$$

to the low type *ex post* IR constraint

$$p_L b \geq p_L \theta/2, \quad (qp_M + (1 - q)p_L)(1 - b) \geq qp_M(1 - p)\theta + (1 - q)p_L \theta/2, \\ (qq_M + (1 - q)q_L) \cdot 1/2 \geq qq_M(1 - p)\theta + (1 - q)q_L \theta/2,$$

to the high-type *ex interim* IC\* constraint

$$q(q_H/2 + (1 - q_H)\theta/2) + (1 - q)(p_M b + q_M/2 + (1 - p_M - q_M)p\theta) \geq \\ \max\{(qp_M + (1 - q)p_L)(1 - b), qp_M \theta/2 + (1 - q)p_L p \theta\} + \max\{(1 - q)p_L b, (1 - q)p_L p \theta\} \\ + \max\{(qq_M + (1 - q)q_L) \cdot 1/2, qq_M \theta/2 + (1 - q)q_L p \theta\} \\ + q(1 - p_M - q_M)\theta/2 + (1 - q)(1 - 2p_L - q_L)p\theta,$$

and to the low-type *ex interim* IC\* constraint

$$q(p_M(1 - b) + q_M/2 + (1 - p_M - q_M)(1 - p)\theta) \\ + (1 - q)(p_L b + p_L(1 - b) + q_L/2 + (1 - 2p_L - q_L)\frac{\theta}{2}) \geq \\ \max\{(1 - q)p_M b, (1 - q)p_M \frac{\theta}{2}\} + \max\{(qq_H + (1 - q)q_M) \cdot 1/2, qq_H(1 - p)\theta + (1 - q)q_M \frac{\theta}{2}\} \\ + q(1 - q_H)(1 - p)\theta + q(1 - p_M - q_M)\theta/2,$$

To see that such an upper bound can be reached by Myerson mediation in our model, it is sufficient to note that any truthful equilibrium achieved by the direct and obedient

revelation mechanism is also an equilibrium of our Nash demand game augmented with a Myerson mediator.

Suppose in fact that the mediator's recommendation is for a peaceful split  $(x, 1 - x)$  of the pie, accepted by the players in the equilibrium of the game induced by the direct revelation mechanism. Then, there is also an equilibrium of the Nash demand game in which the players demand precisely  $x_A = x$  and  $x_B = 1 - x$ , thus resolving the dispute peacefully. And we have seen before that the crisis bargaining game always has war equilibria, to reproduce the outcome of a "war recommendation" in the game induced by the direct revelation mechanism. Hence, the direct revelation mechanism yields an upper bound for the peace probability that mediation can achieve in the Nash demand game augmented with a Myerson mediator.

This concludes that the optimal strategies of a Myerson mediator coincide with the optimal strategies of a mediator characterized in Horner, Morelli, and Squintani (2010), Lemma 4 which states that, for any fixed  $q$ , the specific values of the control variables are:

- For  $q \leq \gamma/(\gamma + 2)$ ,  $b = p\theta$ ,  $q_L = 1$ ,  $q_H = q_M = 0$ ,  $p_M = \frac{1-q}{(\gamma+1)(1-q)-2q}$ .
- For  $\frac{\gamma}{\gamma+2} \leq q < \frac{\gamma}{\gamma+1}$ ,  $b = p\theta$ ,  $q_L = 1$ ,  $p_M + q_M = 1$ ,  $q_H = \frac{1-q}{q} \frac{2q-\gamma(1-q)}{(\gamma+1)(1-q)-q}$ ,  $q_M = \frac{1}{\gamma} \frac{2q-\gamma(1-q)}{(\gamma+1)(1-q)-q}$ .
- For  $q \geq \frac{\gamma}{\gamma+1}$ ,  $b = p\theta$ ,  $q_L = 1$ ,  $q_M = 1$ , and  $q_H = 1$ .  $\square$

**Proof of Proposition 6.** Consider the Myerson mediation game. The expected equilibrium payoffs of a hawk and dove before they report their type to the mediator are, respectively:

$$I_L(q) = q(p_M(1 - p\theta) + q_M/2 + (1 - p_M - q_M)(1 - p)\theta) + (1 - q)/2; \quad (2)$$

$$I_H(q) = q(q_H/2 + (1 - q_H)\theta/2) + (1 - q)(p_M p\theta + q_M/2 + (1 - p_M - q_M)p\theta). \quad (3)$$

Hence, at the beginning of the game, the payoff for militarizing is  $I_H(q) - k$ , whereas the payoff for remaining dove is  $I_L(q)$ . An equilibrium with full militarization,  $q = 1$ , exists

when  $I_H(1) - k \geq I_L(1)$ ; likewise, an equilibrium with no militarization,  $q = 0$ , exists when  $I_L(0) \geq I_H(0) - k$ ; whereas an equilibrium with mixed militarization strategy  $q$  exists when  $I_L(q) = I_H(q) - k$ .

When  $k > p\theta - 1/2$ , the unique symmetric equilibrium is  $q = 0$ . In fact, for all  $q$ ,

$$\begin{aligned}
& I_L(q) - I_H(q) + k \\
= & q(p_M(1 - p\theta) + q_M/2 + (1 - p_M - q_M)(1 - p)\theta) + (1 - q)/2 \\
& - (q(q_H/2 + (1 - q_H)\theta/2) + (1 - q)(p_M p\theta + q_M/2 + (1 - p_M - q_M)p\theta)) + k \\
> & q(p_M(1 - p\theta) + q_M/2 + (1 - p_M - q_M)(1 - p)\theta) + (1 - q)/2 \\
& - (q(q_H/2 + (1 - q_H)\theta/2) + (1 - q)(p_M p\theta + q_M/2 + (1 - p_M - q_M)p\theta)) + p\theta - 1/2 \\
= & \frac{1 - \theta}{2(1 + \lambda)} (2\lambda p_M - \lambda q_H - \lambda + 2\lambda q_M + \gamma q_M + \lambda \gamma q_M) \\
= & \begin{cases} \frac{1 - \theta}{2(1 + \lambda)} \left( 2\lambda \frac{1}{1 + \gamma - 2\lambda} - \lambda \right) & \text{if } \lambda < \gamma/2 \\ \frac{1 - \theta}{2(1 + \lambda)} \left( 2\lambda - \lambda \frac{2\lambda - \gamma}{\lambda(\gamma + 1 - \lambda)} - \lambda + \gamma \frac{2\lambda - \gamma}{\gamma(\gamma + 1 - \lambda)} + \lambda \gamma \frac{2\lambda - \gamma}{\gamma(\gamma + 1 - \lambda)} \right) & \text{if } \gamma/2 \leq \lambda < \gamma \\ \frac{1 - \theta}{2} \gamma & \text{if } \lambda \geq \gamma. \end{cases} \\
= & \begin{cases} \frac{1 - \theta}{2(1 + \lambda)} \frac{(2\lambda - \gamma + 1)}{(\gamma - 2\lambda + 1)} \lambda & \text{if } \lambda < \gamma/2 \\ \frac{1 - \theta}{2(1 + \lambda)} \frac{\lambda + 1}{\gamma - \lambda + 1} \lambda & \text{if } \gamma/2 \leq \lambda < \gamma \\ \frac{1 - \theta}{2} \gamma & \text{if } \lambda \geq \gamma. \end{cases}
\end{aligned}$$

These quantities are all positive.

So, suppose that  $k \leq p\theta - 1/2 = (1 - \theta)\gamma/2$ . We now search for mixed strategy equilibria.

The indifference condition is:

$$I_L(q) + k = I_H(q)$$

Case 1.  $\lambda \leq \gamma/2$ . Substituting the mediator's solution into the expressions (2) and (3), we obtain:

$$\begin{aligned}
I_L(q) &= q \left[ \frac{1}{1 + \gamma - 2\lambda} (1 - p\theta) + \left( 1 - \frac{1}{1 + \gamma - 2\lambda} \right) (1 - p)\theta \right] + (1 - q)/2 \\
I_H(q) &= q\theta/2 + (1 - q)p\theta.
\end{aligned}$$

Solving the indifference condition, and reparametrizing to get rid of  $p, q$ , we obtain the  $k$  which makes the players indifferent for  $\lambda$  and  $\gamma$  and  $\theta$  fixed:

$$k(\lambda) = (1 - \theta) \frac{(\gamma - \lambda + \lambda\gamma - 2\lambda^2)(\gamma + 1)}{2(\gamma - 2\lambda + 1)(\lambda + 1)}$$

Because  $2\lambda \leq \gamma$ , this is always positive.

We first differentiate  $k(\lambda)$ ,

$$\begin{aligned} \frac{\partial k(\lambda)}{\partial \lambda} &= \frac{1}{2}(\gamma - 2\lambda + 1)^{-2}(\lambda + 1)^{-2}(\gamma - 4\lambda - 1)(\gamma + 1)(1 - \theta) \\ &\propto \gamma - 4\lambda - 1 \end{aligned}$$

The expression is positive for  $\lambda < (\gamma - 1)/4$  and negative for  $\lambda > (\gamma - 1)/4$ , on the range  $\lambda \in [0, \gamma/2]$ . Then, we calculate the extremes of the range:

$$k(0) = (1 - \theta) \frac{\gamma}{2} \text{ and } k(\gamma/2) = (1 - \theta) \frac{(\gamma + 1)\gamma}{2(\gamma + 2)}.$$

This concludes that the function  $k(\lambda)$  equals  $(1 - \theta) \frac{\gamma}{2}$  at  $\lambda = 0$  to then increase until  $\lambda = (\gamma - 1)/4$  and then decrease until  $\lambda = \gamma/2$  reaching  $(1 - \theta) \frac{(\gamma + 1)\gamma}{2(\gamma + 2)}$ . Noting that  $(1 - \theta) \frac{\gamma}{2} > (1 - \theta) \frac{(\gamma + 1)\gamma}{2(\gamma + 2)} > 0$ , we determine the following conclusions:

- For  $k \in (1 - \theta) \frac{(\gamma + 1)\gamma}{2(\gamma + 2)}, (1 - \theta) \frac{\gamma}{2}$ , there exists a unique equilibrium  $\lambda(k)$ , the function  $\lambda$  is strictly decreasing in  $k$ , it starts at  $\lambda = \gamma/2$  for  $k = (1 - \theta) \frac{(\gamma + 1)\gamma}{2(\gamma + 2)}$  and reaches  $\lambda(q) = 0$  for  $k = (1 - \theta) \frac{\gamma}{2}$ .
- For  $k \in [0, (1 - \theta) \frac{(\gamma + 1)\gamma}{2(\gamma + 2)})$ , there does not exist any equilibrium such that  $\lambda \leq \gamma/2$ .

The explicit equilibrium solution is cumbersome, and its omission inconsequential.

Case 2. For  $\gamma \geq \lambda \geq \gamma/2$ , substituting the mediator's solution in the expressions

$$I_L(q) = q \left[ \left( 1 - \frac{2\lambda - \gamma}{\gamma(\gamma + 1 - \lambda)} \right) (1 - p\theta) + \frac{2\lambda - \gamma}{2\gamma(\gamma + 1 - \lambda)} \right] + (1 - q)/2$$

$$\begin{aligned} I_H(q) &= q \left[ \frac{2\lambda - \gamma}{2\lambda(\gamma + 1 - \lambda)} + \left( 1 - \frac{2\lambda - \gamma}{\lambda(\gamma + 1 - \lambda)} \right) \theta/2 \right] \\ &\quad + (1 - q) \left[ \left( 1 - \frac{2\lambda - \gamma}{\gamma(\gamma + 1 - \lambda)} \right) p\theta + \frac{2\lambda - \gamma}{2\gamma(\gamma + 1 - \lambda)} \right]. \end{aligned}$$

Again, solving for the indifference condition and reparametrizing, we obtain

$$k(\lambda) = (1 - \theta) \frac{(\gamma - \lambda)(\gamma + 1)}{2(\gamma - \lambda + 1)},$$

or

$$q = (1 + \gamma) \frac{2k - \gamma(1 - \theta)}{2k(2 + \gamma) - (1 + \gamma)^2(1 - \theta)}$$

this is again always positive.

Differentiating it, we obtain:

$$\frac{\partial k(\lambda)}{\partial \lambda} = -\frac{1}{2}(\lambda - \gamma - 1)^{-2}(\gamma + 1)(1 - \theta) < 0.$$

Calculating it at the two extremes  $\lambda = \gamma/2$  and  $\lambda = \gamma$ , we obtain:

$$k(\gamma/2) = (1 - \theta) \frac{(\gamma + 1)\gamma}{(\gamma + 2)2}, \text{ and } k(\gamma) = 0.$$

This concludes that, for  $k \in \left[0, (1 - \theta) \frac{(\gamma + 1)\gamma}{(\gamma + 2)2}\right]$ , there is a unique mixed strategy equilibrium  $\lambda(k)$ , with  $\gamma/2 \leq \lambda \leq \gamma$ , the function  $\lambda(k)$  is strictly decreasing, it starts at  $\lambda = \gamma$  for  $k = 0$ , and reaches  $\lambda = \gamma/2$  for  $k = (1 - \theta) \frac{(\gamma + 1)\gamma}{(\gamma + 2)2}$ .

Wrapping up the two cases, we conclude that there is a unique mixed strategy equilibrium  $\lambda^*(k)$ . It is strictly decreasing in  $k$  for  $k \in \left[0, (1 - \theta) \frac{\gamma}{2}\right]$ , with  $\lambda(0) = \gamma$  and  $\lambda\left((1 - \theta) \frac{\gamma}{2}\right) = 0$ , for  $k > (1 - \theta) \frac{\gamma}{2}$ ,  $\lambda(k) = 0$ .

Turning to check for pure-strategy equilibria, we first suppose that  $q = \lambda = 0$ , then  $I_L(q) = 1/2$  and  $I_H(q) = p\theta$ . Hence, for  $k \leq p\theta - 1/2$ ,  $\lambda = 0$  is not an equilibrium. Then, suppose that  $q = 1$ . The mediator's solution is to assign  $q_L = q_M = q_H = 1$ , so that the split  $1/2$  is always assigned regardless of the reports. Hence, the interim payoffs are  $I_L(q) = 1/2 = I_H(q)$ . So, becoming hawk with probability one is never an equilibrium. This result completes the proof of Proposition 6.  $\square$

**Proof of Theorem 1.** To prove this result, we first show a useful Lemma that compares the symmetric equilibrium militarization probability in the *arbitration* dispute resolution solution of HMS, with the *mediation* HMS dispute resolution solution. The arbitration

dispute resolution in HMS is formulated as mediation, with the only difference that players must commit to agree to the third party recommendation, before such recommendations are made. In other terms, the arbitrator is capable of enforcing its recommendation, whereas the mediator's recommendations need to be self enforcing.

Formally, the HMS arbitration program is formalized as follows:

$$\min_{b, p_L, p_M, p_H} (1-q)^2 (1-p_L) + 2q(1-q)(1-p_M) + q^2(1-p_H)$$

subject to *ex interim* individual rationality (for the hawk and dove, respectively)

$$(1-q)(p_M b + (1-p_M)p\theta) + q(p_H/2 + (1-p_H)\theta/2) \geq (1-q)p\theta + q\theta/2,$$

$$(1-q)(p_L/2 + (1-p_L)\theta/2) + q(p_M(1-b) + (1-p_M)(1-p)\theta) \geq (1-q)\theta/2 + q(1-p)\theta,$$

and to the *ex interim* incentive compatibility constraints (for the hawk and dove, respectively)

$$(1-q)((1-p_M)p\theta + p_M b) + q((1-p_H)\theta/2 + p_H/2) \geq$$

$$(1-q)((1-p_L)p\theta + p_L/2) + q((1-p_M)\theta/2 + p_M(1-b)),$$

$$(1-q)((1-p_L)\theta/2 + p_L/2) + q((1-p_M)(1-p)\theta + p_M(1-b)) \geq$$

$$(1-q)((1-p_M)\theta/2 + p_M b) + q((1-p_H)(1-p)\theta + p_H/2).$$

**Lemma 1** *The arbitration dispute resolution solution of HMS yields the same symmetric equilibrium militarization probability as the mediation HMS dispute resolution solution.*

**Proof** For any  $\gamma$  (including  $\gamma < 1$ ), HMS show that the arbitration solution is:

1. For  $\lambda \leq \gamma/2$ ,  $b = \frac{1}{2}(\gamma(1-\theta) + 1)$ ,  $p_L = 1$ ,  $p_M = \frac{1}{\gamma-2\lambda+1}$ ,  $p_H = 0$ ;
2. For  $\gamma/2 < \lambda \leq \gamma$ ,  $b = -\frac{3\lambda-3\gamma-2\theta\lambda+2\theta\gamma+\lambda\gamma-\theta\lambda\gamma-\gamma^2+\theta\gamma^2-1}{-2\lambda+2\gamma+2}$ ,  $p_L = 1$ ,  $p_M = 1$ ,  $p_H = \frac{2\lambda-\gamma}{(\gamma-\lambda+1)\lambda}$ .

The militarization strategy  $q$  in the arbitration game is given by the indifference condition:

$$\begin{aligned} I_L(q) &= (1-q)((1-p_M)p\theta + p_M b) + q((1-p_H)\theta/2 + p_H/2) \\ &= (1-q)((1-p_L)\theta/2 + p_L/2) + q((1-p_M)(1-p)\theta + p_M(1-b)) - k = I_H(q) - k \end{aligned} \quad (4)$$

Substituting the above solutions in the indifference condition, we find the expressions:

$$k(\lambda) = (1-\theta) \frac{(\gamma - \lambda + \lambda\gamma - 2\lambda^2)(\gamma + 1)}{2(\gamma - 2\lambda + 1)(\lambda + 1)}, \text{ for } \lambda \leq \gamma/2, \quad (5)$$

$$k(\lambda) = (1-\theta) \frac{(\gamma - \lambda)(\gamma + 1)}{2(\gamma - \lambda + 1)}, \text{ for } \gamma/2 < \lambda \leq \gamma, \quad (6)$$

which correspond to the solutions for the militarization game with the HMS optimal mediation solution (for any  $\gamma$ , including  $\gamma < 1$ ).  $\square$

As a consequence of the above Lemma, we can prove the Theorem simply by showing that the HMS arbitration solution achieves the same outcome as our hypothetical institution that includes militarization deterrence in its objectives. In fact, we prove a stronger result. We show that the HMS arbitration solution achieves the same welfare as a hypothetical institution that not only aims to keep militarization in check, but is even capable of enforcing its recommendations. Such an institution is represented by the following program. Let the dove and hawk interim expected utilities be, respectively,

$$I_L = q(p_M(1-p\theta) + q_M/2 + (1-p_M - q_M)(1-p)\theta) + (1-q)/2$$

$$I_H = q(q_H/2 + (1-q_H)\theta/2) + (1-q)(p_M p\theta + q_M/2 + (1-p_M - q_M)p\theta).$$

The optimal institution chooses  $q, b, p_L, p_M$  and  $p_H$  so as to solve the program

$$\min_{q, b, p_L, p_M, p_H} (1-q)^2(1-p_L + p_L\theta) + 2q(1-q)(1-p_M + p_M\theta - k) + q^2(1-p_H + p_H\theta - 2k)$$

subject to the ex-ante obedience constraints:

$$q(1-q)[I_H - k - I_L] = 0, q[I_H - k - I_L] \geq 0, (1-q)[I_H - k - I_L] \leq 0$$

to the *ex interim* individual rationality (for the hawk and dove, respectively)

$$I_H \geq (1-q)p\theta + q\theta/2,$$

$$I_L \geq (1-q)\theta/2 + q(1-p)\theta,$$

and to the *ex interim* incentive compatibility constraints (for the hawk and dove, respectively)

$$I_H \geq (1 - q) ((1 - p_L)p\theta + p_L/2) + q ((1 - p_M)\theta/2 + p_M(1 - b)),$$

$$I_L \geq (1 - q) ((1 - p_M)\theta/2 + p_M b) + q ((1 - p_H)(1 - p)\theta + p_H/2).$$

In order to proceed with the proof, we distinguish two parts. We first show that HMS arbitrators achieve the same peace chance as the optimal institution define above. Then, we show that they achieve the same militarization probability minimization as the optimal institution. The analysis holds for any  $\gamma$ , including the case  $\gamma < 1$ .

To tackle the first problem, we set up the following relaxed problem which describes necessary constraints satisfied by the hypothetical optimal institution defined above. We choose  $\{b, p_L, p_M, p_H, q\}$  so as to minimize the war probability

$$W = (1 - q)^2 (1 - p_L) + 2q(1 - q)(1 - p_M) + q^2(1 - p_H)$$

subject to hawk *ex interim* individual rationality constraint

$$(1 - q)(p_M b + (1 - p_M)p\theta) + q(p_H/2 + (1 - p_H)\theta/2) \geq (1 - q)p\theta + q\theta/2,$$

to the dove *ex interim* incentive compatibility constraint

$$\begin{aligned} & (1 - q) ((1 - p_L)\theta/2 + p_L/2) + q ((1 - p_M)(1 - p)\theta + p_M(1 - b)) \geq \\ & (1 - q) ((1 - p_M)\theta/2 + p_M b) + q ((1 - p_H)(1 - p)\theta + p_H/2). \end{aligned}$$

and to the militarization indifference condition (4).

To solve the relaxed problem, we first solve  $b$  in the militarization indifference condition, and substitute it in the hawk *ex interim* individual rationality constraint, and in the hawk *ex interim* incentive compatibility constraint. Rearranging, they now take the forms:

$$\begin{aligned} H &= k + \frac{1}{2}\theta - kq - p\theta - \frac{1}{2}q\theta + pq\theta + \frac{1}{2}p_L - qp_L + qp_M - \frac{1}{2}\theta p_L + q\theta p_L \\ &\quad - q\theta p_M + \frac{1}{2}q^2 p_H + \frac{1}{2}q^2 p_L - q^2 p_M - \frac{1}{2}q^2 \theta p_H - \frac{1}{2}q^2 \theta p_L + q^2 \theta p_M \geq 0 \\ L &= p\theta - \frac{1}{2}\theta - k + \frac{1}{2}\theta p_M + \frac{1}{2}q\theta p_H - p\theta p_M - \frac{1}{2}q\theta p_M - pq\theta p_H + pq\theta p_M \geq 0. \end{aligned}$$

Note now that  $W$  evidently decreases in  $p_L$ , that  $L$  is independent of  $p_L$ , and that  $\partial H/\partial p_L = \frac{1}{2}(q-1)^2(1-\theta) > 0$ . Because setting  $p_L = 1$  makes  $W$  as small as possible without violating the constraints  $H$  and  $L$ , it has to be part of the solution.

Substituting  $p_L = 1$  in  $W$ ,  $H$ , and  $L$ , we obtain:

$$\begin{aligned} H &= k - q - kq - p\theta + \frac{1}{2}q\theta + pq\theta + qp_M - q\theta p_M + \frac{1}{2}q^2 - \frac{1}{2}q^2\theta \\ &\quad + \frac{1}{2}q^2 p_H - q^2 p_M - \frac{1}{2}q^2\theta p_H + q^2\theta p_M + \frac{1}{2}, \\ L &= p\theta - \frac{1}{2}\theta - k + \frac{1}{2}\theta p_M + \frac{1}{2}q\theta p_H - p\theta p_M - \frac{1}{2}q\theta p_M - pq\theta p_H + pq\theta p_M, \\ W &= 2q(1-q)(1-p_M) + q^2(1-p_H). \end{aligned}$$

Now, we observe that  $\partial L/\partial p_H = -\theta q(p-1/2) < 0$  that  $\partial L/\partial p_M = -\theta(p-1/2)(1-q) < 0$ , and that  $L = -k$  when  $p_M = 1$  and  $p_H = 1$ . Because  $W$  decreases in both  $p_M$  and  $p_H$ , this concludes that the dove incentive compatibility constraint must bind.

We now solve for  $p_M$  in the constraint  $L = 0$  and substitute it in the expressions for  $W$  and  $H$ . We obtain

$$\begin{aligned} W &= q^2 p_H + K_1(p, \theta, q, k) \\ H &= -\frac{1}{2}q^2(1-\theta)p_H + K_2(p, \theta, q, k), \end{aligned}$$

where the explicit formulas of  $K_1$  and  $K_2$  are inessential. Because  $W$  increases in  $p_H$  and  $H$  decreases in  $p_H$ , this concludes that the constraint  $H = 0$  must bind, unless  $p_H = 0$  (which does not matter, as it is part of the HMS arbitration solution).

Solving  $p_H$  in the hawk ex interim individual rationality constraint, and substituting the solution in the objective  $W$ , we obtain:

$$W = \frac{1}{\theta-1} (2kq - 2k + 2p\theta + q\theta - 2pq\theta - 1).$$

This function increases in  $q$  for  $k \leq (1-\theta)\gamma/2 = p\theta - 1/2$ , because:

$$\frac{\partial W}{\partial q} = \frac{2p\theta - \theta - 2k}{1-\theta} \geq \frac{2p\theta - \theta - (2p\theta - 1)}{1-\theta} = 1 > 0.$$

Hence, the minimization of  $W$  under the constraints that  $H = 0$ ,  $L = 0$ ,  $0 \leq p_M \leq 1$ ,  $0 \leq p_H \leq 1$  and  $0 \leq q \leq 1$  is equivalent to the minimization of  $q$  subject to the constraints that  $H = 0$ ,  $L = 0$ ,  $0 \leq p_M \leq 1$ ,  $0 \leq p_H \leq 1$  and  $0 \leq q \leq 1$ .

Note now that setting  $q = 0$  together with  $H = 0$  and  $L = 0$  yields  $p_H = \frac{1}{\theta}(2p - 1)(2p\theta - 2k - 1) \rightarrow +\infty$ , because  $\theta(2p - 1)(2p\theta - 2k - 1) \geq 0$  when  $k \leq p\theta - 1/2$ . Hence, the solution must have an interior  $q$ .

We are now ready to show that the minimal value of  $q$  subject to the constraints that  $H = 0$ ,  $L = 0$ ,  $0 \leq p_M \leq 1$ , and  $0 \leq p_H \leq 1$  is exactly the equilibrium value of  $q$  in the militarization game, assuming that disputes are solved with the HMS optimal arbitration solution. In order to do so, we take the following approach. We first reparametrize all expressions in  $\lambda = q/(1 - q)$  and  $\gamma = (2p\theta - 1)/(1 - q)$ . Then, we prove that, for every  $\lambda$ , the minimal value of  $k$  subject to the constraints that  $H = 0$ ,  $L = 0$ ,  $0 \leq p_M \leq 1$ , and  $0 \leq p_H \leq 1$  coincides with the expressions (5) and (6) obtained when solving the militarization indifference condition (4) after plugging in the HMS optimal arbitration solution. Because the expressions for  $k(\lambda)$  in (5) and (6) are strictly decreasing in  $\lambda$ , this concludes that the inverse function  $k^{-1}(k) = \lambda$  identifies the minimal  $q$  subject to the constraints that  $H = 0$ ,  $L = 0$ ,  $0 \leq p_M \leq 1$ , and  $0 \leq p_H \leq 1$ , via the increasing relation  $q/(1 - q) = \lambda$ .

So, reparametrizing the expressions for  $H$  and  $L$ , setting both of them equal to zero, and solving for  $k$  and  $p_H$  as a function of  $p_M$ , we obtain

$$p_H = \frac{(2\lambda p_M - p_M - \gamma p_M + 1)}{(\gamma - \lambda + 1)\lambda} \quad (7)$$

$$k(\lambda) = -\frac{1}{2}(1 - \theta) \frac{(\lambda p_M - \lambda\gamma - \gamma + \lambda^2)(\gamma + 1)}{(\gamma + 1 - \lambda)(\lambda + 1)}. \quad (8)$$

Because  $k$  decreases in  $p_M$  for all  $\lambda < \gamma$ , we want to set  $p_M$  as large as possible. When  $p_M = 1$ ,  $p_H = \frac{(2\lambda - \gamma)}{(\gamma - \lambda + 1)\lambda}$ , which is positive if and only if  $\lambda > \gamma/2$ .

Likewise, solving for  $k$  and  $p_M$  as a function of  $p_H$ , we obtain

$$k(\lambda) = \frac{1}{2}(1 - \theta) \frac{(\gamma - \lambda + \lambda\gamma - 2\lambda^2 + \lambda^2 p_H)(\gamma + 1)}{(\gamma - 2\lambda + 1)(\lambda + 1)}. \quad (9)$$

For  $\lambda < \gamma/2$ , it is the case that  $\gamma - 2\lambda + 1 > 0$ , and hence this expression increases in  $p_H$ . We thus set  $p_H = 0$  and it is easy to verify that  $p_M = \frac{1}{\gamma - 2\lambda + 1} \in (0, 1)$ .

Because we have recovered the HMS optimal arbitration solution, the part of the proof concerning peace chance is concluded.

We now turn to show that HMS arbitrators are the optimal mechanism in terms of militarization probability minimization, and achieve the same militarization probability as the optimal institution defined at the beginning of this proof. Our proof approach will be to show that the HMS optimal arbitration solution is the solution of the following relaxed problem:

$$\min_{p_L, p_M, p_H, b} q \text{ s.t. } H = 0, L = 0, I_H - k = I_L \quad (10)$$

In order to do so, we first reparametrize all constraints in  $\lambda = q/(1-q)$  and  $\gamma = (2p\theta - 1)/(1-q)$ . Then, we prove that, for every  $\lambda$ , the minimal value of  $k(\lambda) = I_H - I_L$  subject to the constraints  $0 \leq p_L \leq 1$ ,  $0 \leq p_M \leq 1$ ,  $0 \leq p_H \leq 1$ ,  $H = 0$  and  $L = 0$  coincides with the expressions (5) and (6). Because these expressions strictly decrease in  $\lambda$ , the same reasoning as in the previous part of the proof then concludes that the problem (10) is solved by the HMS optimal arbitration solution.

So, we first differentiate  $k(\lambda) = I_H - I_L$  with respect to  $p_L$ , and obtain the negative derivative  $-\frac{1}{2}(1-q)(1-\theta)$ . Because we know that  $\partial H/\partial p_L > 0$  and  $\partial L/\partial p_L = 0$ , minimization of  $k(\lambda)$  requires setting  $p_L = 1$ . Then, we note that  $\partial k(\lambda)/\partial b > 0$ , whereas  $\partial H/\partial b > 0$ , and hence the constraint  $H = 0$  must bind. Then, we see that  $\partial k(\lambda)/\partial p_M = b - p\theta - q(1-\theta) < 0$ , because  $H = 0$  implies that  $b - p\theta = \frac{q\theta p_H - qp_H + 2p\theta p_M - 2pq\theta p_M}{2(1-q)p_M} - p\theta = -\frac{1}{2}(1-\theta)\frac{qp_H}{(1-q)p_M} \leq 0$ . This, together with  $\partial L/\partial p_M = -\theta(p-1/2)(1-q) < 0$  implies that  $L$  binds.

Solving for  $p_H$  and  $b$  in the constraints  $H = 0$ ,  $L = 0$ , after imposing  $p_L = 1$ , and substituting the results in  $k(\lambda)$ , we again obtain the expressions (7) and (8), so that we conclude that for  $\lambda > \gamma/2$ , the solution is  $p_M = 1$ ,  $p_H = \frac{(2\lambda-\gamma)}{(\gamma-\lambda+1)\lambda}$ . Likewise, solving for  $p_H$  and  $b$  in the constraints  $H = 0$ ,  $L = 0$ , after imposing  $p_L = 1$ , and substituting the results in  $k(\lambda)$ , we obtain expression (9), and conclude that the solution is  $p_H = 0$  and  $p_M = \frac{1}{\gamma-2\lambda+1} \in (0, 1)$ , for  $\lambda < \gamma/2$ .

Because we have recovered the HMS optimal arbitration solution, also the part of the proof concerning militarization probability is concluded.  $\square$