Mediation and Peace*

Johannes Hörner
Yale University[†]

Massimo Morelli
Columbia University
and European University Institute[‡]

Francesco Squintani
Essex University
and Universita' degli Studi di Brescia[§]

December 2009 - PRELIMINARY AND INCOMPLETE

Abstract

This paper brings mechanism design to the study of conflict resolution in international relations. We determine when and how unmediated communication and mediation reduce the ex-ante probability of conflict, in a simple game where conflict is due to asymmetric information. Unmediated communication helps reducing the chance of conflict as it allows conflicting parties to reveal their types and establish type-dependent transfers to avoid conflict. Mediation improves upon unmediated communication when the intensity of conflict is high, or when the chance of power disparity among the players is higher. The mediator improves upon unmediated communication by not precisely reporting information to conflicting parties, and precisely, by not revealing to a player with probability one that the opponent is weak. Surprisingly, in our set up, mediators who can enforce settlements are no more effective in reducing the probability of conflict than mediators who can only make non-binding recommendations.

Keywords: Mediation, War and Peace, Imperfect Information, Communication Games, Optimal Mechanism.

JEL classification numbers: C7

*We are grateful for their comments to Alberto Bisin, Matt Jackson, Adam Meirowitz, Philippe Jehiel and Kris Ramsay, and to the audiences of the ESSET 2009, SAET 2009, Essex and EUI.

[†]Department of Economics.

[‡]Department of Political Science and Department of Economics

[§]Department of Economics

1 Introduction

The positive analysis of conflict and its potential sources has attracted the attention of game theorists for decades, in an increasingly fertile interaction with international relations scholars.[1] On the other hand, the normative analysis about which institutions or mechanisms we should use to reduce the possibility of conflict has not benefited by many interactions across the two disciplines yet. In particular, the powerful tools of mechanism design developed in economic theory have not yet been extensively applied to conflict resolution or to the minimization of the probability of future wars.[2] The literature on optimal auctions, optimal market design, organization theory and public good provision mechanisms have been very successful. Studying optimal mechanisms we can learn important lessons about what institution designers should consider most relevant in different situations, and this seems eminently relevant if we want to think of flexible institutions to help in the reduction or elimination of costly conflicts.

In this paper we set for ourselves a normative objective: we select one of the most studied sources of conflict, namely the presence of asymmetric information, and we ask what institutional mechanisms can be most useful to minimize the probability of war in different contexts. In particular, when the source of potential conflicts is information asymmetries, it is natural to assume that agents could benefit by communicating. So, we first ask when and how can cheap talk communication between the agents in conflict, without any third party, minimize the ex ante probability of war with respect to the benchmark situation without communication?[3] We then turn to the main topic of this paper: mediation. So, we ask *when and how* is it the case that communication through a mediator can strictly improve the ex-ante probability of peace with respect to unmediated communication?

We consider a very basic form of mediation: the mediator's objective is simply the minimization of the probability of war, an impartial objective, and the mediator does not

[1] See Jackson and Morelli (2009) for an updated survey of such a positive analysis work.

[2] As examples of the discussion in international relations on the importance of institutional design for conflict or international cooperation, see e.g. Koremenos et al (2001). For a discussion on the lacking applications of mechanism design to conflict, see e.g. Fey and Ramsay (2009).

[3] On the great relevance of allowing for pre-play communication in situations where bargaining breakdown is due to asymmetric information, see e.g. Baliga and Sjöström (2004).

possess any additional information beyond what she can learn from communicating with the parties in conflict.[4] Hence the mediator can only improve the chances of peace by controlling the flow of information between the parties.[5] Moreover, the mediator has no special powers, no independent budget for transfers or subsidies, no enforcement capability. Thus, the mediator's recommendations should be self-enforcing. This feature distinguishes our problem from the standard problems studied in the mechanism design literature.[6]

In our environment, the mediator's role is only that of helping communication, formulating proposals, and manipulating the information received (see Touval and Zartman (1985) for a discussion of these three principal mediator roles). In practice, this corresponds to the mediator's role in "collecting and judiciously communicating select confidential material" (Raiffa, 1982, 108-09). Obviously, the role for mediation that we identify cannot be performed by holding joint, face-to-face sessions with both parties, but requires private and separate caucuses, a practice that is often followed by mediators. In international relations, the practice of shuttle diplomacy has become popular since Henry Kissinger's efforts in the Middle East in the early 1970s and the Camp David negotiations mediated by Jimmy Carter, in which a third party conveys information back and forth between parties, providing suggestions for moving the conflict toward resolution (see, for example, Kydd, 2006).

We assume that mediator's proposals are self-enforcing. Indeed, countries are sovereign, and enforcement of contracts or agreements is often impossible.[7] In the terminology of Fisher (1995), our main focus is on "pure mediation", i.e. mediation only involving infor-

[4]As some scholars claim, "mediator impartiality is crucial for disputants' confidence in the mediator, which, in turn, is a necessary condition for his gaining acceptability, which, in turn, is essential for mediation success to come about" (see e.g. Young (1967) and the list of other scholars mentioned in Kleiboer (1996)). On the other hand, when a mediator possesses independent information that needs to be credibly transmitted about some crucial feature of the parties in conflict, some degree of bias may be optimal (see arguments in favor and against mediators' bias in Kydd (2003) and Rauchos (2006) respectively).

[5]Our notion of communication equilibrium, with and without a mediator, is related to the concepts introduced in the seminal contributions by Forges (1985) and Myerson (1986).

[6]Among the few papers studying self-enforcing mechanisms in mechanism design problems, see Matthews and Postlewaite, 1989; Cramton and Palfrey, 1995; Forges, 1999; Compte and Jehiel, 2006.

[7]Viewed another way, countries cannot commit not to initiate war if such an attack is a profitable deviation from an agreement. In this sense, even if the bargaining problem comes from asymmetric information, we also have a natural form of commitment problem built in. See Powell (2006) for a recent comprehensive discussion of the relative importance of asymmetric information and commitment problems in creating bargaining breakdown.

mation gathering and settlement proposal making, rather than "power mediation," which instead also involves mediator's power to reward, punish or enforce. But we will also conclude that, surprisingly, in the context of our model the availability of enforcement power would *not* reduce the ex-ante probability of conflict. The assumption that mediator's proposal are self-enforcing is formalized by requiring that, whenever a mediator recommends a peaceful settlement of the crisis, both parties must find the proposed settlement better for them than starting conflict (with its expected associated payoff lottery and costs) given what they learn from the mediator recommendation itself. Since war can be started unilaterally, this ex post individual rationality constraint is indispensable.

Having so far clarified our methodological choices and our general motivation, let us now describe the other basic features of our model and then offer a preview of our findings.

We consider the canonical conflict situation, in which two agents fight for a fixed amount of contestable resources. The exogenous cake to be either divided peacefully or contested in a costly conflict is a standard metaphor for many types of wars, for example related to territorial disputes or to the present and future sharing of the rents from the extraction of natural resources. Custody, partnership dissolution, labor management struggles, and all kinds of litigations and legal disputes could be considered equally relevant applications of the model, but we keep the international conflict example as the main one in the text, also for ease of terminology.

A player cannot observe the opponent's strength, resolve, or outside options. In particular, we assume that each player is strong (hawk) with some probability and weak (dove) with complementary probability. If the two players happen to be of the same type, war is a fair lottery. When they are not of equal strength, the stronger wins with higher probability. To complete this standard setting, we assume that all war lotteries are equally costly.[8] War takes place in our game of conflict unless both players opt for peace, i.e. war can be initiated unilaterally, and we assume that the players war declaration choice is simultaneous. The equilibrium that maximizes the ex-ante chances of peace is such that weak players opt

[8]It may be interesting to verify in future research whether allowing for different cost of symmetric and asymmetric wars can make any difference for the main results of the paper. We believe that the additional complexity would be more than the interesting results we could find in such an extension.

for peace, and strong players declare war unless the hawk type is sufficiently likely.

This simple setting has been the work-horse for models of war due to imperfect information, and it is also the setting common to the very few papers in the literature with an explicit mechanism design agenda. Specifically, Bester and Wärneryd (2006) study the case in which the mediator can enforce settlements, after collecting players' reports. Like us, Fey and Ramsay (2009) consider self-enforcing mechanisms. Unlike us, they do not characterize the optimal self-enforcing mechanism. They show that war can be fully avoided by the optimal mechanism if and only if the type distribution is independent across players and the players' payoffs depend only on their types, unlike in our game.

We first study unmediated communication, to then determine when and how mediation improves upon it. To make our case for negotiation tighter, we allow players to use public random devices *(sunspots)* to correlate their play, and hence increase the ex-ante chance of peace, in the unmediated communication game. Specifically, players send a cheap-talk message to each other first; then, for any pair of observed messages, they correlate their play either on war or on a peaceful cake division, depending on the realization of a public random device. In equilibrium, it must be the case that players do not want to unilaterally declare war, when the sunspot realization is associated to a peaceful cake division. When hawks are not too likely, so that war cannot be avoided in the basic conflict game, the optimal symmetric sunspot separating communication equilibrium is shown to improve on no-communication. Specifically, it allows players to reveal their type, and establish type-dependent cake divisions to avoid conflict. However, war cannot be fully avoided.

We then consider mediated communication. First, the mediator collects the players' messages privately. Then, he chooses message-dependent cake-division proposals, and correlates the players' war declaration choices optimally. Peace recommendations must be ex-post individually rational. It is clear that the mediator's optimal solution cannot be worse than the best equilibrium without the mediator. In fact, the mediator could always, trivially, make the messages she receives public, thereby mimicking the optimal unmediated communication equilibrium. Thus, the usefulness of mediation can be measured by looking at what regions of the parameters allow the mediator to induce a strict welfare improvement.

We show that an active mediator is particularly effective when the intensity and/or cost of conflict is high. On the other hand, when the intensity of conflict is expected to be low, the presence of the mediator brings about strict welfare gains only for a sufficiently high degree of expected power asymmetry. Interestingly, the intensity of conflict, the nature of the issue, and power asymmetry are considered among the most important variables that affect when is mediation most successful (see e.g. Bercovitch and Houston (2000) and Bercovitch et al. (1991)). Our findings resonate with well-documented stylized facts in the empirical literature on negotiation (Bercovich and Jackson 2001, Wall and Lynn, 1993), that show that parties are less likely to reach an agreement without a mediator when the intensity of conflict is high than when it is low. Further, Bercovitch et al (1991) show that mediation is useful mostly when the dispute are about resources, territory, or in any case divisible issues, which is our case. Regarding the role of mediation with different balances of power between the players, we predict maximum effectiveness with high power disparity and low intensity or with high intensity for any power disparity.[9]

In terms of mediator strategy or tactic, the model allows us to focus only on communication facilitation, settlement proposal formulation, separation of players, manipulation of their messages or obfuscation of parties' positions.[10] For low intensity and sufficiently large expected power asymmetry, the mediator's key strategy involves proposing equal split settlements even when he receives different messages. In this way the incentive to *exaggerate strength* in order to receive high payoff shares is reduced and the overall probability of peace increases. Instead, for high conflict intensity, the mediator can improve upon unmediated communication by offering unequal splits even when he observes both players reporting to be doves. In this way the incentives to *hide strength* by hawks are the ones kept in check.

In both cases the mediator's proposed settlements do not precisely reveal each player's report to the counterpart. Although it is widely believed that a successful mediator should establish credible reports to the conflicting parties, we find that a mediator that reports

[9]The empirical evidence on such variables is not conclusive, but mediation is often seen as the management of power imbalances. See e.g. Davis and Salem (1984).

[10]For an exhaustive discussion of all observed mediation techniques and types of mediation problems, see the survey by Wall and Lynn (1993).

precisely all the information transmitted would not act optimally. The mediator's optimal obfuscation strategy consists in not revealing with probability one that the opponent is weak, when this is the case. Specifically, in the case of high-conflict intensity, by not revealing to a self-declared dove when she is facing a dove, the mediator reduces the incentive for hawks to hide strength and then wage war if revealed that the opponent is weak. In the case of low intensity with high expected power asymmetry, by not revealing to a self-declared hawk when she is facing a dove, the mediator reduces the incentives for doves to exaggerate strength, in order to achieve a favorable settlement when it is revealed that the opponent is weak.

We finally conclude the analysis by showing that, surprisingly, an unbiased mediator who can enforce outcomes is exactly as effective as a mediator who can only propose self-enforcing agreements. This result is quite striking because there are well known environments where the mechanism designer's enforcement power changes the results (see, e.g., Cramton and Palfrey, 1995, or Compte and Jehiel, 2008). This result confirms our view that a mediator does not need enforcement power. "A mediated settlement that arises as a consequence of the use of leverage may not last very long because the agreement is based on compliance with the mediator and not on internalization of the agreement-changed attitudes and perceptions" (Kelman 1958).

The paper is organized as follows: section 2 introduces our basic model of conflict; section 3 studies unmediated communication, focusing on pure strategy equilibrium; section 4 characterizes optimal mediation, and displays the most important substantive results, in terms of *when and how* mediation strictly improves upon unmediated communication; section 5 extends the analysis to mixed strategies; section 6 compares our results with the case of mediators endowed with enforcement powers, and section 7 concludes. All proofs are in appendix.

2 The Game of Conflict

Let us consider a standard bilateral conflict problem, in which two parties want as much as possible of a given cake.[11] As it is standard, we normalize the value of such an exogenous cake to 1. If the two parties cannot agree to any peaceful self enforcing sharing rule and choose conflict, we assume that the destructive war would shrink the actual net value of the cake to $\theta < 1$.

War is modeled as a costly lottery, without the possibility of stalemate.[12] The probability of winning for each player depends on players' *types*: each player can be of type H or L, privately and independently drawn from the same distribution with probability q and $(1 - q)$ respectively. Such a private information characteristic can be thought of as related for example to resolve, military strength, leaders' stubbornness, outside options, etc. When the two fighting players are of the same type we assume that they have the same probability of winning, whereas when one player is stronger than the other one, she wins with probability $p > 1/2$. Hence a type H player who fights against an L type expects $p\theta$ from such a conflict. In the paper we will often refer to type H as a "hawk" and to a L type as a "dove" (with no reference to the hawk-dove game).

War can be initiated unilaterally, while "it takes two to tango," i.e., a peaceful agreement must be preferred by both players to war. More precisely, there is a war declaration game for a given proposed split of the cake, $(x, 1 - x)$, $(x \in [0, 1])$. In this war declaration game, the two players simultaneously announce "peace" or "war", and if they both announce peace the settlement is accepted otherwise war takes place. We assume that when the two players choose a peaceful split $(x, 1 - x)$, there are ways to implement such a split.[13] We

[11]Depending on the context, of course, the interpretation of what the cake means ranges from territory or exploitation of natural resources to any measure of social surplus in a country or partnership.

[12]Allowing for the possibility of stalemate makes the problem inherently dynamic. A dynamic extension of our mediation model is definitely interesting, but beyond the scope of the present paper.

[13]If the cake is a resource that can be depleted in a short period and does not have spillovers on relative strength, then there is no commitment problem. A more complicated setting in which having a greater share of the cake affects relative strength as well would make this assumption more problematic (see e.g. Jackson and Morelli (2007)). If the cake sharing is instead to be interpreted as a durable agreement for example on the exploitation of a future stream of resources or gains from trade, then the commitment problem is non trivial. In this case the agreement could be about periodic tributes to be made in perpetuity, and there are ways to implement the agreement with sufficient use of dynamic incentives. See for example Schwartz

note that there exists always an equilibrium where both players declare war in this game, regardless of the split. In what follows, we focus on the equilibrium that maximizes the ex-ante probability of peace V.[14]

The model has three parameters: θ, p, and q. Yet all results depend on only two statistics:[15]

$$\lambda \equiv \frac{q}{1-q} \quad and \quad \gamma \equiv \frac{p\theta - 1/2}{1/2 - \theta/2}.$$

λ is the hawk/dove odds ratio, and γ represents 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. Given that γ is increasing in θ, we will also interpret situations with low γ as situations of high intensity or cost of conflict.

Throughout the paper, we assume anonymity: While we will allow splits to depend on player's reports, they cannot depend on the players' identities. So, absent communication or mediation, the only split compatible with anonymity or symmetry is the equal split $(1/2, 1/2)$. The equilibrium that maximizes the ex-ante probability of peace is such that low types choose peace in the war declaration game and high types choose peace if and only if

$$q\theta/2 + (1-q)p\theta \leq 1/2, \quad \text{i.e.,} \quad \lambda \geq \gamma. \tag{1}$$

Hence the peace-probability maximizing equilibrium induces probability of peace $V = 1$ if $\lambda \geq \gamma$, and $V = 1 - (1-q)^2$ otherwise.

3 Communication Game Without Mediation

Let us now consider the value of communication in our basic game of conflict. We consider a simple communication game without any mediator. We allow players to make use of random public devices (i.e., sunspots), to optimally coordinate their play. In the next

and Sonin (2008).

[14] If $p\theta < 1/2$, conflict can always be averted with the split $(1/2, 1/2)$; we shall therefore assume henceforth that $p\theta > 1/2$.

[15] This feature will allow us to give graphical illustrations of all the results.

section we will then allow the mediator to choose recommendations without necessarily revealing the messages she received, and the difference between the equilibria of the two game forms will be interpreted as a characterization of the benefits of mediation.

The communication game form without the mediator is as follows:

- Time 0: players learn their type privately.
- Time 1: simultaneous communication. Each player i simultaneously chooses (possibly randomly) a message $m_i \in \{h, l\}$. The messages (m_1, m_2) are publicly observed.
- Time 2: Depending on (m_1, m_2), a split $(x, 1-x)$ is selected. The realization of a public random device is observed.
- Time 3: The players simultaneously announce their war-peace declaration, on the basis of the split and on the sunspot realization.

We want to calculate the symmetric sunspot equilibrium with the smallest possible ex ante probability of war. We momentarily ignore mixed strategies by the players at the message stage, we will extend the analysis to consider this possibility in section 5. Evidently, there is always a pooling equilibrium where both types choose the same reporting strategy, and whose outcomes coincide with the equilibrium of the war-declaration game without communication. We now define a fully separating equilibrium in which each player truthfully reveals her type at time 1. By symmetry, the equal split $(1/2, 1/2)$ must follow any pair of equal messages. To characterize the rest of the equilibrium we need to introduce the following notation:

- Given messages (h, h), the players coordinate on peace with split $(1/2, 1/2)$ with probability p_H, and on war with probability $1 - p_H$.
- Given messages (h, l) or (l, h) the players coordinate on peace (with allocation $b > 1/2$ to the hawk and $1 - b$ to the dove) with probability p_M, and on war with probability $1 - p_M$.

- Given messages (l, l), the players coordinate on peace and equal split $(1/2, 1/2)$ with probability p_L, and on war with probability $1 - p_L$.

The players use the random sunspot realization to coordinate their play. When the sunspot realization is associated to peace, in equilibrium, it must be that they do not prefer to unilaterally deviate and declare war.

The optimal symmetric sunspot equilibrium in pure strategies is characterized by the following program:

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

subject to ex post individual rationality (IR) constraints and interim truth-telling, or incentive compatibility, constraints.

The ex-post IR constraints are:

$$b \geq p\theta, \ \ 1/2 \geq \theta/2, \ 1 - b \geq (1-p)\theta.$$

That is, a high type facing a low type must get a share b that makes going to war unprofitable against a low type. The second constraint, that a player would accept an equal split when the opponent's type is the same as her own, is clearly always satisfied. The third states that the low type's share against a high type cannot be so low that it is better for her to go to war after all.

The interim incentive compatibility constraints are denoted by IC_L^* and IC_H^*, for low and high types respectively. The "star" superscript refers to the fact that, when a player contemplates deviating at the message stage, she also anticipates and takes into account that she might prefer to declare war ex-post, although the sunspot realization suggests the players to coordinate on peace (see Matthews and Postlewaite, 1989). In particular, consider the right hand side of IC_L^*, below, which describes the payoff of a low type who deviates and sends the high message, i.e. who exaggerates strength. With probability $(1 - q)$ the low type exagerating strength will be facing another low type. In that case, she will agree to a peaceful settlement $b > 1/2$ if only if $b \geq \frac{\theta}{2}$, since the true strength of the two players is equal. Conversely, the misreporting low type will face a high type with probability q, and in that case she will accept the equal split if and only if $1/2 \geq (1-p)\theta$. The IC_L^* constraint is thus

$$(1-q)\left((1-p_L)\frac{\theta}{2}+p_L\frac{1}{2}\right)+q\left((1-p_M)(1-p)\theta+p_M(1-b)\right)\geq$$
$$(1-q)\left((1-p_M)\frac{\theta}{2}+p_M\max\{b,\frac{\theta}{2}\}\right)+q\left((1-p_H)(1-p)\theta+p_H\max\{\frac{1}{2},(1-p)\theta\}\right);$$

Similarly, the IC_H^* constraint is

$$(1-q)\left((1-p_M)p\theta+p_M b\right)+q\left((1-p_H)\frac{\theta}{2}+p_H\frac{1}{2}\right)\geq$$
$$(1-q)\left((1-p_L)p\theta+p_L\max\{\frac{1}{2},p\theta\}\right)+q\left((1-p_M)\frac{\theta}{2}+p_M\max\{1-b,\frac{\theta}{2}\}\right).$$

The right hand side of this constraint takes into account that if a high type deviates at the message stage, i.e. she hides strength, then she will face a high type with probability $(1-q)$ and a low type with probability q. In the first case, if offered the settlement $1/2$, she will opt for peace if and only if $1/2 \geq p\theta$; in the second case, if offered the share $(1-b)$, she will accept it if and only if $(1-b) \geq \frac{\theta}{2}$.

Solving the above optimization program, we obtain the following characterization.

Proposition 1 *The unique optimal sunspot symmetric separating equilibrium is characterized as follows.*

1. *Suppose that $\gamma \leq 1$.*

   (a) *When $\lambda < \gamma/(1+\gamma)$, both interim IC* constraints bind,*

   $$b > p\theta,\ p_H = 0,\ p_M = \frac{1}{(1+\gamma)(1-\lambda)},\quad \textit{and } V = \frac{1+\gamma+\lambda(1-\gamma)}{(1+\gamma)(1-\lambda)(1+\lambda)^2}.$$

   (b) *When $\lambda \in [\gamma/(1+\gamma), \min\{1/(1+\gamma), \gamma\}]$, both IC* constraints bind,*

   $$b > p\theta,\ p_M = 1,\quad p_H = 1-\frac{\gamma}{(1+\gamma)\lambda},\quad \textit{and } V = 1-\frac{\gamma\lambda}{(1+\gamma)(1+\lambda)^2}.$$

(c) *When* $\lambda \in [1/(1+\gamma), \gamma)$, *only the* IC_L^* *constraint binds,*

$$b = p\theta, \; p_M = 1, \; p_H = \frac{2\lambda - \gamma}{\lambda(2+\gamma)}, \quad \textit{and } V = \frac{2(1+\lambda)+\gamma}{2+\gamma+\lambda(2+\gamma)}.$$

2. *Suppose that* $\gamma > 1$.

   (a) *When,* $\lambda < \gamma/2$, *only the* IC_L^* *constraint binds,*

$$b = p\theta, \; p_H = 0, \; p_M = \frac{1}{1+\gamma-2\lambda}, \; \textit{and } V = \frac{1+\gamma}{(1+\gamma-2\lambda)(1+\lambda)^2}.$$

   (b) *When* $\lambda \in [\gamma/2, \gamma)$, *only the* IC_L^* *constraint binds,*

$$b = p\theta, \; p_M = 1, p_H = \frac{2\lambda - \gamma}{\lambda(\gamma+2)}, \; \textit{and } V = 1 - \frac{\gamma\lambda}{(2+\gamma)(1+\lambda)}.$$

The proof is in the appendix. Let us comment in detail on the properties of this separating equilibrium, starting from some general ones.

First, war is never optimal when both players report low strength: $p_L = 1$. Second, the truth-telling constraint for the low type, IC_L^*, is always binding. When the truth-telling constraint for the high type is not binding, $b = p\theta$; and when both truth-telling constraints are binding, the ex-post IR constraint $b \geq p\theta$ does not bind. Hence, b is either pinned down by the ex-post IR constraint $b \geq p\theta$, or by the joint interim truth-telling constraints. Third, given that the incentive to exaggerate strength is always present and needs to be discouraged, there needs to be positive probability of war following a high report. The most potent channel through which the low type's incentive to exaggerate strength can be kept in check, is by assigning a positive probability of war whenever there are two self-proclaimed high types. When λ is low (few high types) it is indeed optimal to set $p_H = 0$ and $p_M > 0$; whereas for higher values of λ, $p_H < 1$ and $p_M = 1$. Because war with hawks is more costly than war with doves, threatening war with a hawk is more effective to deter a dove from exaggerating strength, than threatening war with a dove. Further, when λ is sufficiently high, the likelihood of a hawk opponent is sufficiently high that prescribing war against a dove is not needed to deter a dove from exaggerating strength. But when λ is low, deterring misreporting by a dove requires having a positive probability of war when the opponent is a dove, in addition to having war for sure when both claim to be hawks.

The other properties of the characterization of Proposition 1 differ when $\gamma \geq 1$ and $\gamma < 1$.

Suppose first that $\gamma \geq 1$, so that the benefits from war are sufficiently high. Then the ex-post IR constraint always binds, and hence $b = p\theta$; and the interim high-type truth-telling IC_H^* constraint never binds. This is because the hawk hiding strength always prefers to wage war (both against hawks and against doves). When $b = p\theta$, in fact, the condition $\gamma \geq 1$ is equivalent to $1 - b \leq \theta/2$. As a result, the hawk obtains the payoff $p\theta$ against doves, regardless of her message, whereas against hawks she obtains $\theta/2$ for sure if hiding strength, and either $\theta/2$ (after a war recommendation) or $1/2$ (after settlement) when truthfully reporting.

Second, suppose that $\gamma < 1$. For $\lambda \leq 1/(1+\gamma)$, the high-type truth-telling constraint IC_H^* binds, and $b > p\theta$. To see why, suppose by contradiction that $b = p\theta$. For $\gamma < 1$, this would imply that $1 - b > \theta/2$. Consider a hawk pretending to be a dove. If she meets a dove, she can secure the payoff $p\theta$ by waging war. This is also the payoff for revealing being hawk and meeting a dove: She obtains $p\theta$ through war or through the split $b = p\theta$. If she meets a hawk, she gets $1 - b$ with probability p_M and $\theta/2$ with probability $1 - p_M$. By claiming to be a hawk, she gets $1/2$ with probability p_H and $\theta/2$ with probability p_M. But we know that p_M is larger than p_H, and because $1 - b > \theta/2$, this gives an incentive to pretend to be a dove (hiding strength) to secure peace more often than by revealing that she is a hawk, which contradicts IC_H^*. To make sure that both truth-telling constraints are satisfied, we must have $b > p\theta$, so as to reduce the payoff from hiding strength: This reduces the payoff from settling against a hawk when hiding strength and increases the payoff from settling against a dove when revealing to be hawk.

Next, note that p_H increases in λ, as in the case of $\gamma \geq 1$. Because the incentive to hide strength decreases as p_H increases relative to p_M, also the optimal b decreases in λ. When λ reaches the threshold $1/(1+\gamma)$, the offer b required for the high type truth-telling constraint to bind is exactly $p\theta$. Further increasing λ cannot induce a further decrease in b, because the ex-post IR constraint $b \geq p\theta$ becomes binding. So in the region where $\lambda \in [1/(1+\gamma), \gamma]$, the IC_H^* constraint does not bind and $b = p\theta$.

Figure 1 shows the probability of peace (our welfare measure here) induced by the optimal sunspot symmetric separating equilibrium. For $\gamma \geq 1$, we note that it is U-shaped in λ for $\lambda \leq \gamma/2$, and decreasing in λ when λ is between $\gamma/2$ and γ. To understand the forces leading to the U-shaped effect of λ in the lower region, note first that an increase in λ shifts probability mass from the LL dyad to the LH dyad and from the LH dyad to the HH dyad (the overall effect on the likelihood of the LH dyad is that it increases in λ if and only if $\lambda < 1$). Because $1 = p_L \geq p_M > p_H$, these shifts make the probability of peace initially decrease in λ. However, p_M strictly increases in λ for $\lambda \leq \gamma/2$, and eventually this makes the probability of peace increase in λ. Interestingly, despite the fact that p_H strictly increases in λ, for $\lambda > \gamma/2$, it still does not grow fast enough to compensate for the shift in probability mass towards the dyads with the higher probability of war. As a result, the probability of peace decreases in λ when λ is between $\gamma/2$ and γ.

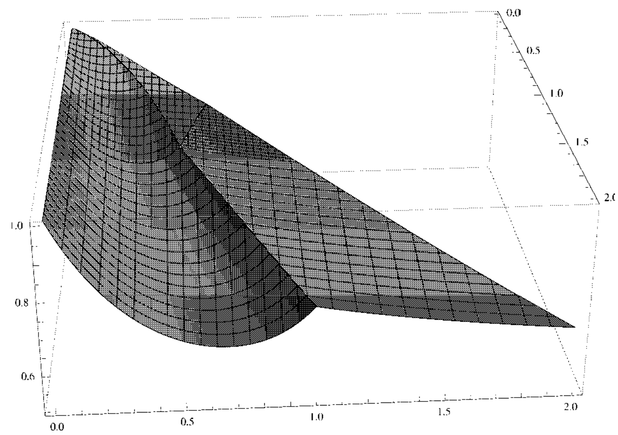

Figure 1: Probability of peace in the separating equilibrium (λ from bottom to front, γ from left to right)

In sum, we derive the following key substantive results.

Result 1 *When does unmediated communication improve the chances of peace?*

- *When the proportion of hawks is not too large. Otherwise, the high chance that the opponent is a hawk is a sufficient deterrent for the players to avoid conflict.*

Result 2 *How does unmediated communication improve the chances of peace?*

- *It allows to reveal whether a player is a hawk or a dove, and establish appropriate splits to reduce the probability of war when one player is a hawk.*
- *Specifically, when it is revealed that a hawk meets a dove, the players coordinate with positive probability on an unequal split favorable to the hawk, whereas when two hawks meet, the players coordinate on the equal split with positive probability.*

4 Optimal Mediation

In the previous section, we have characterized the equilibrium outcome in the case in which players send public messages to each other. Now we want to characterize the effect of adding an active mediator who collects the players' private message, and makes optimal recommendations, with respect to unmediated communication.

The game form is the same as above, with the difference that messages at time 1 are not publicly observed. They are individually reported to a mediator, who then chooses the split and the correlation of play according to the following symmetric mechanism.[16]

- Each player i reports his type to the mediator.
- Given reports (H, H), the mediator recommends the peaceful split $(1/2, 1/2)$ with probability q_H, the splits $(\beta, 1-\beta)$ and $(1-\beta, \beta)$ with probability p_H each, and war with probability $1 - 2p_H - q_H$. It is assumed that $\beta > 1/2$.
- Given reports (H, L), the mediator recommends the peaceful split $(1/2, 1/2)$ with probability q_M, the split $(\beta, 1-\beta)$ with probability p_M^+, the split $(1-\beta, \beta)$ with probability p_M^-, and war with probability $1 - p_M^+ - p_M^- - q_M$, with $p_M^- < p_M^+$.
- Given reports (L, L), the mediator recommends the peaceful split $(1/2, 1/2)$ with probability q_L, the splits $(\beta, 1-\beta)$ and $(1-\beta, \beta)$ with probability p_L each, and war with probability $1 - 2p_L - q_L$.

[16]As we shall discuss later, more complicated symmetric mechanisms cannot do better, see footnote 18.

Hence the optimal mediator solution solves the program

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

subject to the six ex-post IR constraints

- $\beta(qp_H+(1-q)p_M^+) \geq qp_H\theta/2+(1-q)p_M^+p\theta$,
- $(1-\beta)(qp_H+(1-q)p_M^-) \geq qp_H\theta/2+(1-q)p_M^-p\theta$,
- $(qq_H+(1-q)q_M)\cdot 1/2 \geq qq_H\theta/2+(1-q)q_Mp\theta$,
- $(qp_M^-+(1-q)p_L)\beta \geq qp_M^-(1-p)\theta+(1-q)p_L\frac{\theta}{2}$,
- $(qp_M^++(1-q)p_L)(1-\beta) \geq qp_M^+(1-p)\theta+(1-q)p_L\frac{\theta}{2}$,
- $(qq_M+(1-q)q_L)\cdot 1/2 \geq qq_M(1-p)\theta+(1-q)q_L\frac{\theta}{2}$,

as well as the high-type interim IC* constraint

$$\begin{aligned}
&q(p_H\beta+p_H(1-\beta)+q_H/2+(1-2p_H-q_H)\theta/2)\\
&+(1-q)(p_M^+\beta+p_M^-(1-\beta)+q_M/2+(1-p_M^+-p_M^--q_M)p\theta) \geq\\
&\max\{(qp_M^++(1-q)p_L)(1-\beta), qp_M^+\theta/2+(1-q)p_Lp\theta\}\\
&+\max\{(qp_M^-+(1-q)p_L)\beta, qp_M^-\theta/2+(1-q)p_Lp\theta\}\\
&+\max\{(qq_M+(1-q)q_L)\cdot 1/2, qq_M\theta/2+(1-q)q_Lp\theta\}\\
&+q(1-p_M^+-p_M^--q_M)\theta/2+(1-q)(1-2p_L-q_L)p\theta,
\end{aligned}$$

and for the low-type interim IC* constraint

$$\begin{aligned}
&q(p_M^-\beta+p_M^+(1-\beta)+q_M/2+(1-p_M^+-p_M^--q_M)(1-p)\theta)\\
&+(1-q)(p_L\beta+p_L(1-\beta)+q_L/2+(1-2p_L-q_L)\frac{\theta}{2}) \geq\\
&\max\{(qp_H+(1-q)p_M^-)(1-\beta), qp_H(1-p)\theta+(1-q)p_M^-\frac{\theta}{2}\}\\
&+\max\{(qp_H+(1-q)p_M^+)\beta, qp_H(1-p)\theta+(1-q)p_M^+\frac{\theta}{2}\}\\
&+\max\{(qq_H+(1-q)q_M)\cdot 1/2, qq_H(1-p)\theta+(1-q)q_M\frac{\theta}{2}\}\\
&+q(1-2p_H-q_H)(1-p)\theta+q(1-p_M^+-p_M^--q_M)\theta/2,
\end{aligned}$$

Solving this program, we obtain the following result.[17]

Proposition 2 *An optimal solution to the mediator's problem is such that:*

- *For* $\lambda \leq \gamma/2$*,*

$$q_L + 2p_L = 1, p_H = q_H = q_M = p_M^- = 0, \beta = p\theta, p_M^+ = \frac{1}{1+\gamma-2\lambda}.$$

Further, for $\gamma \geq 1$*,*

$$p_L \leq \frac{2\lambda}{(\gamma - 2\lambda + 1)(\gamma - 1)},$$

while for $\gamma < 1$*,*

$$p_L \geq \frac{(1-\gamma)\lambda}{2\gamma^2} \frac{(\lambda - \gamma)(\gamma + 2)}{(\lambda - \gamma - 1)}.$$

The ex-ante peace probability is

$$V = \frac{\gamma + 1}{(1+\gamma-2\lambda)(1+\lambda)^2}.$$

- *For* $\lambda \geq \gamma/2$*,*

$$q_L + 2p_L = 1, p_M^+ + q_M = 1, p_H = p_M^- = 0, \beta = p\theta, q_H = \frac{2\lambda - \gamma}{\lambda(\gamma+1-\lambda)}, q_M = \frac{2\lambda - \gamma}{\gamma(\gamma+1-\lambda)},$$

and $q_L \geq \frac{\lambda(2\lambda-\gamma)}{\gamma^2(\gamma-\lambda+1)}$*. Further, for* $\gamma \geq 1$*,*

$$p_L \leq 2\frac{(\gamma - \lambda)(\gamma + 2)\lambda}{(\gamma - \lambda + 1)\gamma(\gamma - 1)},$$

whereas for $\gamma < 1$*,*

$$p_L \geq \frac{(1-\gamma)\lambda}{2\gamma^2} \frac{(\lambda - \gamma)(\gamma + 2)}{(\lambda - \gamma - 1)}.$$

The ex-ante peace probability is

$$V = \frac{\gamma + 1}{(\gamma - \lambda + 1)(\lambda + 1)}.$$

[17] Interestingly, we find that the solution of our mediation program is unique only for $\lambda \leq \gamma/2$. Here, we only report the simplest solution, i.e. the one where $p_M^- = 0$ and $p_H = 0$, so that the unequal split is only recommended when a hawk meets a dove, and it is always favorable to the hawk. We show in the appendix that, for $\gamma/2 \leq \lambda \leq \gamma$, there exists also solutions with $p_M^- > 0$ and $p_H > 0$. Of course, all solutions yield the same ex-ante probability of peace.

In order to interpret the features of the optimal mediation solution, and to ease comparison with the discussion following Proposition 1, let us distinguish the cases $\gamma \geq 1$ and $\gamma < 1$.

Suppose that $\gamma \geq 1$. If $\lambda > \gamma/2$, then $q_H > 0$: the mediator sometimes recommends the equal split $(1/2, 1/2)$ when one player reports to be a hawk, and the other claims to be a dove. In this way, the ex-post IR constraint of the high type who is recommended the equal split becomes binding. We remark that this ex-post constraint was slack in the unmediated communication equilibrium. By making a slack constraint binding, the mediator increases the probability of peace. Indeed, the mediator lowers the gain from pretending to be a hawk, by making exagerating strength less profitable against doves. When $\lambda \leq \gamma/2$ instead, $q_H = 0$ and the mediator does not improve upon unmediated communication. In this case, in fact, in both the optimal mediated and unmediated communication solution, war needs to occur with probability one in dyads of hawks, to avoid that doves misreport their type. But then, the above-mentioned slack constraint is not relevant for either the mediated or the unmediated communication solution, and the mediator cannot improve upon unmediated communication.

In contrast with the case of $\gamma \geq 1$, the mediator always yields a strict welfare improvement when $\gamma < 1$. When $\lambda > 1/(1+\gamma)$, so that $b = p\theta$ in the perfectly separating equilibrium, it is also the case that $\lambda > \gamma/2$ (note that $1/(1+\gamma) > \gamma/2$), and hence the mediator helps for the same reasons as when $\gamma \geq 1$. When $\lambda < 1/(1+\gamma)$, the mediator makes sure that the interim high-type IC* constraint is satisfied with $\beta = p\theta$. In fact, the mediator offers $1-\beta$ with positive probability when both players report to be doves. By doing so, the mediator makes sure that a hawk hiding strength will wage war when proposed $1-\beta$. This eliminates the incentive to hide strength in order to seek peace against hawks that we observed in the unmediated communication equilibrium. Hence, the expected payoff of hiding strength is lower, and the interim high-type IC* constraint is satisfied with $\beta = p\theta$. Note that the ex-post IR constraint $b \geq p\theta$ was slack in the unmediated communication equilibrium. By making the ex-post IR constraint binding, the mediator can improve the objective function, i.e. increase the probability of peace.

Figure 2 shows the probability of peace induced by the optimal mediation solution compared to the probability of peace induced by the optimal sunspot symmetric separating equilibrium.

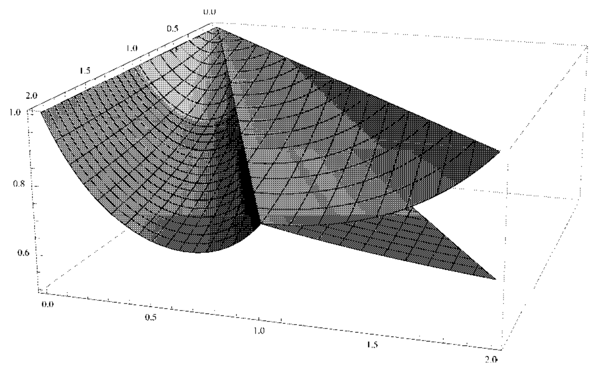

Figure 2: Probability of peace in the mediated vs. unmediated case (λ from bottom to front, γ from left to right)

In sum, we derive the following key substantive results.

Result 3 *When does mediation improve on unmediated communication?*

- *When the intensity and/or cost of conflict is high (γ low), mediation always brings about strict welfare improvements with respect to unmediated cheap-talk.*
- *When conflict is not expected to be very costly or intense (high γ), on the other hand, mediation provides a large improvement in welfare if and only if the proportion of hawks is intermediate, i.e., for high expected power asymmetry.*

Result 4 *How does mediation improve on unmediated communication?*

- *When the proportion of hawks is intermediate (high expected power asymmetry), the mediator lowers the reward for a dove from mimicking a hawk, by not always giving the lion's share to a declared hawk facing a dove (or, equivalently by not always revealing a self-reported hawk that she is facing a dove).*

- *When the probability of facing a hawk is low and conflict is expected to be costly, the mediator's strategy is instead to offer with some probability unequal split to two parties reporting low type (or, equivalently the mediator does not always reveal a dove that she is facing a dove). This lowers the incentive to hide strength and seek peace with a hawk.*

5 Potential Benefits of Mixed Strategies

We have conducted the comparison between cheap talk and active mediation without admitting the possibility for players to mix at the reporting stage. Evidently, mixing cannot improve the chances of peace in the game with a mediator, because any mixing on the final outcomes of the war-declaration game can be reproduced by mediator's (possibly correlated) mixing. However, there may be a potential for improvement in the unmediated communication game. As a robustness check, we now consider this possibility. For brevity, we relegate the maximization programs for the mixed strategy equilibria to the appendix. Here, we show that, while there is no mixed-strategy equilibrium of the unmediated communication game where the hawk randomizes between sending the high and low report, there is a mixed strategy equilibrium where the dove randomizes, which yields a higher ex-ante peace probability, in a restricted parameter region, when $\gamma < 1$.

Proposition 3 *Allowing players to play mixed strategies in the unmediated communication game, the optimal equilibrium is such that the high type H always reports message h and the low type L reports message l with probability σ, where $\sigma < 1$ if and only if $\gamma < 1$ and*

$$\frac{\gamma}{1+\gamma} > \lambda > \max\left\{ \frac{-1-\gamma\left(5+6\gamma\right)+\sqrt{\left(1+3\gamma\right)\left(1+\gamma\left(11+8\gamma\left(3+2\gamma\right)\right)\right)}}{2\left(1+\gamma\right)\left(1+3\gamma\right)}, \right.$$
$$\left. \frac{-1-\gamma\left(8+3\gamma\right)+\sqrt{1+\gamma\left(16+\gamma\left(54+\gamma\left(48+25\gamma\right)\right)\right)}}{2\left(\gamma^2-1\right)} \right\}.$$

For $\lambda < 2\gamma^2/\left(1+3\gamma\right),$

$$p_M = \frac{2\gamma-\lambda+\gamma\lambda}{2(1+\gamma)(\gamma-\lambda)},\ p_H = 0,\ \sigma = 1+\frac{\lambda}{2}(1-1/\gamma),\ b = (1+\gamma(1-\theta))/2$$

$$and\ V = \frac{\lambda(\gamma^2(4+3\lambda) - \lambda - 2\gamma\lambda(3+2\lambda))}{4\gamma(\gamma-\lambda)(1+\lambda)^2}.$$

For $\lambda > 2\gamma^2/\left(1+3\gamma\right)$,

$$p_M = 1,\ p_H = 0,\ b = p\theta,\ \sigma = \frac{(1+\gamma)(1+\lambda)}{(1+2\gamma)},\ and\ V = \frac{\gamma^2}{(1+2\gamma)^2}.$$

Overall this result shows that the role of mixing in the unmediated communication game is rather limited. It may improve on pure-strategy communication only in a small subset of the parameter region where both interim IC* constraints bind. Specifically, it turns out that mixing by the low type, may relax the incentive of the high type to hide strength.

The benefits of mediation over mixing in an unmediated communication game are evident. By randomizing over recommendations, the mediator can reproduce any distribution induced by mixing. In unmediated communication, however, because players must mix independently of each other, they cannot generate the optimal correlated distribution chosen by the mediator. In practice, the low type mixing between reporting high and low may improve welfare upon the pure-stategy equilibrium, but at the cost of inducing war with positive probability between dove dyads. This event is optimally ruled out by the mediator, who induces war only when at least one of the players is a hawk.

6 The Role of Enforcement

To analyze the role of mediation, we have chosen to use a canonical model where war is a costly lottery and may take place due to asymmetric information about the players' strength or resolve. Even though the cause of war is asymmetric information, the analysis of the optimal mediation problem involves also a significant enforcement problem. Indeed, countries are sovereign, and enforcement of contracts or agreements is often impossible. Since war can be started unilaterally, we have incorporated ex post IR and interim IC* constraints in the formulation of the optimal mediation program. Therefore in our model, the residual ex-ante chance of war that results in the optimal mediation solution, can be thought as being due to a combination of asymmetric information and enforcement problems.

The only role we have so far attributed to mediation is to optimally manage information elicited by the conflicting parties. One might also wonder whether the mediator could further reduce the ex-ante probability of war if it were endowed with enforcement power. The answer to this question can be obtained by simply comparing our findings with those in Bester and Wärneryd (2006): Rather than imposing ex-post IR constraints and interim IC* constraints like we do, they impose interim IR constraints and interim IC constraints. In practice, they only require that conflicting parties are willing to begin the mediation process, and to reveal their information to the mediator. But mediator's recommendations are enforceable by external actors, such as the international community or the mediator itself, and hence they abstract from the enforcement problem that we introduce.

Formally, invoking the revelation principle (Myerson 1979), and restricting attention to symmetric mechanisms, the Bester-Wärneryd problem is expressed as follows. The parties report their types L, H to the mediator. The mediator recommends peaceful settlement with probability p_L after the reports (l, l), with probability p_H after the reports (h, h), and with probability p_M after the reports (l, h) and (h, l). Symmetry entails that the settlement is $(1/2, 1/2)$ if the two players report to be of the same type. If the reports are (H, L), settlements are $(b, 1-b)$, and if they are (L, H) settlements are $(1-b, b)$. The mediator chooses b, p_L, p_M and p_H so as to solve the program

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

subject to interim individual rationality (for the high type and the low-type respectively)

$$(1-q)(p_M b + (1-p_M) p\theta) + q\left(p_H \frac{1}{2} + (1-p_H)\frac{\theta}{2}\right) \geq (1-q) p\theta + q\frac{\theta}{2},$$

$$(1-q)\left(p_L \frac{1}{2} + (1-p_L)\frac{\theta}{2}\right) + q(p_M(1-b) + (1-p_M)(1-p)\theta) \geq (1-q)\frac{\theta}{2} + q(1-p)\theta,$$

and to the interim incentive compatibility constraints (for the high type and the low-type respectively)

$$(1-q)((1-p_M)p\theta + p_M b) + q\left((1-p_H)\frac{\theta}{2} + p_H\frac{1}{2}\right) \geq$$
$$(1-q)\left((1-p_L)p\theta + p_L\frac{1}{2}\right) + q\left((1-p_M)\frac{\theta}{2} + p_M(1-b)\right),$$

$$(1-q)\left((1-p_L)\frac{\theta}{2}+p_L\frac{1}{2}\right)+q\left((1-p_M)(1-p)\theta+p_M(1-b)\right)\geq$$
$$(1-q)\left((1-p_M)\frac{\theta}{2}+p_M b\right)+q\left((1-p_H)(1-p)\theta+p_H\frac{1}{2}\right).$$

Surprisingly, we show that the solution of our mediation program, where the mediator's recommendations are self-enforcing, yields the same welfare as the solution of Bester and Wärneryd's program, in which the mediator can enforce outcomes.

Proposition 4 *In the optimal symmetric solution of the mediator's program with enforcement power,*

For $\lambda \leq \gamma/2$,

$$p_M=\frac{1}{\gamma-2\lambda+1},\ p_H=0,\ \textit{and}\ V=\frac{(\gamma+1)}{(\gamma-2\lambda+1)(\lambda+1)^2}.$$

For $\lambda \geq \gamma/2$,

$$p_M=1,\ p_H=\frac{2\lambda-\gamma}{(\gamma-\lambda+1)\lambda},\ \textit{and}\ V=\frac{\gamma+1}{(\gamma-\lambda+1)(\lambda+1)}.$$

The ex-ante probability of peace in the optimal symmetric solution of the mediator's program with self-enforcing recommendations is the same as the probability of peace in the optimal symmetric solution of the mediator's program with enforcement powers.

This striking and surprising result can be intuitively explained as follows. First note, that the low-type IC constraint and high type interim IR constraint are the only ones binding in the solution of the mediator's program with enforcement power. Conversely, the only binding constraints in the mediator's program with self-enforcing recommendations are the low-type IC* constraint and the two ex-post high-type IR constraints. Recall in fact, that in our solution, the high type is always indifferent between war and peace if recommendeded a settlement. Further, the low-type IC* constraint in the mediator's problem with self-enforcing recommendation is identical to the low-type IC constraint in the mediator's program with enforcement power, because a dove never wages war after exaggerating strength in the solution of mediator's problem with self-enforcing recommendation. Finally, the high-type interim IR constraint integrates the two binding high-type ex-post IR constraints in the mediator's problem with self-enforcing recommendation. While requiring a constraint to hold in expectation only is generally a weaker requirement than having the two constraints, it turns out that in our stylized model of conflict, the induced welfare is the same.

This Proposition immediately implies the following substantive result.[18]

Result 5 *An unbiased mediator who can enforce outcomes is exactly as effective as a mediator who can only propose self-enforcing agreements.*

7 Concluding Remarks

This paper brings mechanism design to the study of conflict resolution in international relations. We have determined when and how unmediated communication and mediation reduce the ex-ante probability of conflict, in a simple game where conflict is due to asymmetric information. From the analysis of this paper we have drawn a number of lessons.

First of all, we have shown under which conditions mediation improves upon unmediated communication. Mediation is particularly useful when the intensity of conflict and/or cost of war is high (low θ); when power asymmetry has large impact on the probability of winning (large p); and then, even when neither θ is low, nor p is large, mediation can still be useful when the ex-ante chance of power asymmetry is higher (intermediate q).

The mediator improves upon unmediated communication by not reporting to a player with probability one that the opponent has revealed that she is weak. Specifically, when the ex-ante chance of power asymmetry is high, the mediator is mostly effective over unmediated communication in its ability to keep in check the temptation to exaggerate strength by a dove. The mediator lowers the reward from mimicking a hawk by not always giving the lion's share to a hawk facing a dove. This allows to reduce the probability of war between hawks, and hence the ex-ante probability of war.

[18] As a by-product, this result implies that the mechanism we are considering to solve the mediator's problem with self-enforcing agreements is fully general. Indeed, it can be shown that, first, the constraint set for this problem is more restrictive than the constraint set of the mediator's program with enforcement power. Second, the revelation principle applies in the latter program.

Instead, when the expected intensity or cost of conflict are high, regardless of the expected degree of power asymmetry, the mediator is mostly effective in improving upon unmediated communication in the task of reducing the temptation to hide strength by a strong player. The mediator's optimal strategy in this case is to lower the reward from mimicking a dove by giving sometimes an unequal split to two parties reporting low type. This allows to lower the split proposed to avoid war between a hawk and a dove. In turn, this allows to reduce the probability of war of the hawk-dove dyad.

While the core of our analysis has considered the optimal strategy of mediators who are not endowed with enforcement powers, we have concluded the analysis by showing that, surprisingly, an unbiased mediator who can enforce outcomes is exactly as effective as a mediator who can only propose self-enforcing agreements. This result is quite striking because there are well known games in which enforcement power matters: Interim IR and IC constraints are not equivalent to ex-post IR and interim IC* constraints in the optimal mechanism, but rather they are usually less restrictive.

Our analysis can be extended in several directions. The most promising one is the following. Our results, together with the results by a recent paper by Meirowitz and Ramsay (2009), may be helpful in drawing conclusions on the issue of strategic militarization (see also Meirowitz and Sartori (2008)). Suppose in fact that, prior to enter a crisis, the two players must costly and secretly invest in their military might. Meirowitz and Ramsay (2009) characterize the equilibrium investment strategies in relation to any general crisis-resolution mechanism, and hence any bargaining or communication protocol, that satisfy interim IC constraints (Theorem 2). It would be very interesting to assess the implications of their results in the contest of the optimal crisis-resolution mechanism that we characterize in this paper.

References

[1] Baliga, Sandeep and Tomas Sjöström (2004): "Arms Races and Negotiations," *Review of Economic Studies,* **Vol. 71(2),** 351-69.

[2] Bercovitch, Jacob (1989): "International dispute mediation: A comparative empirical analysis." In *Mediation research,* edited by K. Kressel and D. G. Pruitt, 53-67. San Francisco: Jossey-Bass.

[3] 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* **Vol. 28**, 7-17.

[4] Bercovitch, Jacob and Allison Houston (2000): "Why Do They Do It? An Analysis of the Factors Influencing Mediation Behavior in International Conflicts." *Journal of Conflict Resolution,* **Vol. 44(2)**, 170-202.

[5] Bercovich, Jacob and R. Jackson (2001): "Negotiation or Mediation? An Exploration of Factors Affecting the Choice of Conflict Management in International Conflict", *Negotiation Journal,* **Vol 17**, 59-77.

[6] Bester, H. and K. Wärneryd (2006): "Conflict and the Social Contract," *Scandinavian Journal of Economics,* **Vol. 108(2)**, 231-49.

[7] Blume, A., and J. Sobel (1995): "Communication-Proof Equilibria in Cheap Talk Games," *Journal of Economic Theory,* **Vol. 65(2)**, 359-82.

[8] Brito, D.L. and M.D. Intriligator (1985) " Conflict, War, and Redistribution, " *American Political Science Review,* Vol. 79, No. 4, 943-957.

[9] Compte, Olivier, and Philippe Jehiel (2009): "Veto Constraint in Mechanism Design: Inefficiency with Correlated Types," *American Economic Journal: Microeconomics,* **Vol 1,** 182–206.

[10] Cramton Peter C., Palfrey Thomas R. (1995): "Ratifiable mechanisms: learning from disagreement," Games Economic Behavior, **Vol 10,** 255–283.

[11] Davis, A., and R. Salem (1984): "Procedures for guiding the divorce mediation process," *Mediation Quarterly,* **Vol. 6**, 17-25.

[12] Fearon, J.D. (1995) "Rationalist explanations for war," *International Organization,* 49:3, 379-414.

[13] Fey, M. and C. Ramsay (2009): "Mechanism Design Goes to War: Peaceful Outcomes with Interdependent and Correlated Types," *Review of Economic Design,* **Vol. 13(3)**,

[14] 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. PP.39-59.

[15] Forges, Francoise (1986), "An approach to communication equilibria," *Econometrica,* **Vol 54**, 1375-1385.

[16] Forges, Francoise (1999): "Ex Post Individually Rational Trading Mechanisms," in *Current Trends in Economics: Theory and Applications*, ed. Charalambos Aliprantis, Nicholas Yannelis and Ahmet Alkan, Springer, 157-76.

[17] Jackson, M.O. and M. Morelli (2007): "Political Bias and War," *American Economic Review,* Vol. **97:4**, 1353-1373.

[18] Jackson, Matthew O. and Massimo Morelli (2009): "Reasons for War: an Updated Survey." *Handbook on the Political Economy of War* (forthcoming), edited by Chris Coyne, Elgar Publishing.

[19] Kelman, H.C. (1958): "Compliance, identification, and internalization: Three processes of attitude change." *Journal of Conflict Resolution* /**Vol. 2**, 51-60.

[20] Kleiboer, Marieke (1996): "Understanding Success and Failure of International Mediation," *Journal of Conflict Resolution,* **Vol. 40(2)**, 360-89.

[21] Koremenos, Barbara, Charles Lipson and Duncan Snidal (2001): "The Rational Design of International Institutions." *International Organization,* **Vol. 55(4)**, 761-99.

[22] Kydd, Andrew (2003), "Which Side Are You On? Bias, Credibility, and Mediation." *American Journal of Political Science*, **Vol. 47(4)**, 597-611.

[23] Kydd, Andrew (2006): "When Can Mediators Build Trust", *American Political Science Review,* **Vol 100**, 449-462.

[24] Matthews Steven A., Postlewaite Andrew (1989): "Pre-play communication in two-person sealed-bid double auctions", *Journal Economic Theory* **Vol 48,** 238–263

[25] Meirowitz, Adam H. and Kris Ramsay (2009) "Investing in Failure: Equilibrium Constraints on Prior Beliefs in Bargaining Problems" mimeo: Princeton University.

[26] Meirowitz, Adam H. and Ann E. Sartori (2007): "Secrecy and War: The Origins of Private Information," *Quarterly Journal of Political Science,* **Vol.**

[27] Myerson, Roger B. (1979): "Incentive Compatibility and the Bargaining Problem," *Econometrica*, **Vol. 47(1)**, 61-73.

[28] Myerson, Roger (1986): "Multi-Stage Games with Communication," *Econometrica*, **Vol. 54(2)**, 323-

[29] Myerson, Roger B. and Mark A. Satterthwaite (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, **Vol. 29(2)**, 265-81.

[30] Powell, Robert (2006): "War as a Commitment Problem," *International Organization*, **Vol. 60**, 169-203.

[31] Rauchhaus, Robert W. 2005. "Asymmetric Information, Mediation and Conflict Management.". published in 2006 somewhere.

[32] Raiffa, Howard (1982): *The Art and Science of Negotiation,* Harvard University Press.

[33] Schelling, Thomas C. (1960): *The Strategy of Conflict.* Cambridge, MA: Harvard University Press.

[34] Schwarz, M. and K. Sonin (2008): "A Theory of Brinkmanship, Conflicts, and Commitments," *Journal of Law, Economics, and Organization.* Vol. **24(1)**: 161-183.

[35] 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.

[36] Wall, James A. and Ann Lynn (1993): "Mediation: A Current Review," *Journal of Conflict Resolution*, **Vol. 37(1)**, 160-94.

[37] Young, O.R. (1967): *The intermediaries: Third parties in international crises.* Princeton, NJ: Princeton University Press.

Appendix

Proof of Proposition 1

The proof proceeds in two parts.

Part 1 ($\gamma \geq 1$).

We set up the following relaxed problem:

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

subject to the high-type ex-post IR constraints:

$$b \geq p\theta$$

to the probability constraints:

$$p_L \leq 1, p_M \leq 1, 0 \leq p_H$$

and ex-ante low-type IC* constraint:

$$(1-q)\left((1-p_L)\frac{\theta}{2} + p_L\frac{1}{2}\right) + q\left((1-p_M)(1-p)\theta + p_M(1-b)\right) \geq$$
$$(1-q)\left((1-p_M)\frac{\theta}{2} + p_M b\right) + q\left((1-p_H)(1-p)\theta + p_H\frac{1}{2}\right)$$

Step 1. We want to show that $p_L = 1$. We first note that setting $p_L = 1$ maximizes the LHS of the relaxed low-type IC* constraint and does not affect the RHS. It is immediate to see that the high-type ex-post constraint is not affected either.

Step 2 We want to show that the relaxed low-type IC* constraint binds. Suppose it does not. It is possible to increase p_H thus decreasing the objective function without violating the constraint (note that there is no constraint that $p_H < 1$ in the relaxed problem).

Step 3. We want to show that the high-type ex-post constraint binds. Suppose it does not. Then $b > p\theta$, and it is possible to reduce b without violating the ex-post constraint.

But this makes the low–type relaxed IC* constraint slack, because $-b$ appears in the LHS and b in the RHS. Because step 2 concluded that the low–type relaxed IC* constraint cannot be slack in the solution, we have proved that the ex-post constraint cannot be slack.

Step 4. We want to show that for $\lambda \leq \gamma/2$: $p_H = 0, p_M = \frac{1}{1+\gamma-2\lambda}$ in the relaxed program. The low–type relaxed IC* constraint and ex-post constraint define the function

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

substituting this function into the objective function

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

duly simplified in light of step 1, we obtain the following expression:

$$W = p_H \frac{(2\lambda + \gamma + 3)\,\lambda}{(\gamma - 2\lambda + 1)(\lambda + 1)} + \frac{2\gamma - 3\lambda + \lambda\gamma - 2\lambda^2}{(\gamma - 2\lambda + 1)(\lambda + 1)},$$

where we note that, because $\gamma \geq 2\lambda$, the coefficient of p_H is positive and the whole expression is positive. Hence, minimization of W requires minization p_H. Setting $p_H = 0$ and solving for p_M in (2) yields

$$p_M = \frac{1}{1 + \gamma - 2\lambda}.$$

Because $\lambda \leq \gamma/2$, it follows that $p_M \leq 1$, as required. We note that the probability of war equals:

$$C = \frac{(2\gamma - 3\lambda + \lambda\gamma - 2\lambda^2)\,\lambda}{(\gamma - 2\lambda + 1)(\lambda + 1)^2}.$$

Step 5. We want to show that for $\lambda \geq \gamma/2$, $p_M = 1, p_H = \frac{2\lambda-\gamma}{\lambda(\gamma+2)}$ in the relaxed problem. In light of the previous step, the solution $p_H = 0$ yields $p_M > 1$ and is not admissible when $\lambda > \gamma/2$. Because p_M decrease in p_H in (2), the solution requires setting $p_M = 1$ and, from (2), $p_H = \frac{2\lambda-\gamma}{\lambda(\gamma+2)}$. When $\lambda \geq \gamma/2$, $p_H \geq 0$ and hence the solution is admissible. We note that the probability of war equals:

$$C = \frac{\gamma\lambda}{(\gamma + 2)(\lambda + 1)}.$$

Step 6. We want to show that the solution constructed satisfies all the program constraints. The low-type ex-post constraint $1-b \geq (1-p)\theta$ is trivially satisfied, when $b = p\theta$. Because $b > \theta/2$ and $1/2 > (1-p)\theta$, the low-type ex-ante IC* constraint coincides with the low-type ex-ante relaxed IC* constraint. The condition $1 - b = 1 - p\theta \leq \theta/2$ yields $2 - 2p\theta \leq \theta$, i.e. $1 - \theta \leq 2p\theta - 1$, i.e. $\gamma = \frac{2p\theta-1}{1-\theta} \geq 1$. Hence, for $\gamma \geq 1$, we conclude that $1 - b \leq \theta/2$. So, after simplification, the ex-ante high-type IC* constraint becomes:

$$\begin{aligned}
&(1-q)\,p\theta + q\left((1-p_H)\frac{\theta}{2} + p_H\frac{1}{2}\right)\\
&= (1-q)\left((1-p_M)p\theta + p_M b\right) + q\left((1-p_H)\frac{\theta}{2} + p_H\frac{1}{2}\right) \geq\\
&(1-q)\left((1-p_L)p\theta + p_L p\theta\right) + q\left((1-p_M)\frac{\theta}{2} + p_M\frac{\theta}{2}\right)\\
&= (1-q)\,p\theta + q\theta/2,
\end{aligned}$$

which is satisfied (with slack when $\lambda \geq \gamma/2$). The probability constraints are obviously satisfied.

Part 2 ($\gamma < 1$). We allow for two cases:

Case 1. I will temporarily consider the following relaxed problem:

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

subject to the low-type and high-type relaxed IC* constraints:

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

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

which embed the assumption (to be verified ex-post) that $1-b \geq \theta/2$, and to the probability constraints:

$$p_L \leq 1, p_M \leq 1, 0 \leq p_H$$

Step 1. As in the previous case, we conclude that $p_L = 1$.

Step 2. We want to show that the low-type relaxed IC* constraint binds. Indeed, if it does not, we can increase p_H without violating neither relaxed IC* constraints (note that the LHS of the high-type relaxed IC* constraint increases in p_H).

Step 3. We want to show that the high-type relaxed IC* constraint binds. Suppose not. We can then reduce b because the LHS of the high-type relaxed IC* constraint increases in b and the RHS decreases in b. This makes the low-type relaxed IC* constraint slack, without changing p_M and p_H. But in light of step 2, this cannot minimize the objective function. Hence, the high-type relaxed IC* constraint must bind.

Step 4. We want to show that for $\lambda < \gamma/(1+\gamma)$, $p_H = 0$ and $p_M = \frac{1}{(1+\gamma)(1-\lambda)}$ solve the relaxed problem. The binding relaxed ex-ante IC* constraints define the function: $[p_M, b]\,(p_H)$, after substituting λ for q and γ for p, we obtain:

$$\begin{aligned} b &= \frac{2\lambda + \gamma - \theta\lambda - \theta\gamma - 2\lambda p_H + \theta\lambda p_H - 3\lambda\gamma p_H + 2\theta\lambda\gamma p_H - \lambda^2 p_H - \lambda\gamma^2 p_H - \lambda^2\gamma p_H + \theta\lambda\gamma^2 p_H + 1}{2\,(1 - \lambda p_H - \lambda\gamma p_H)\,(\lambda + 1)} \\ p_M &= \frac{(1 - \lambda p_H(1+\gamma))}{(\gamma+1)\,(1-\lambda)}. \end{aligned} \tag{3}$$

Substituting p_M into the objective function

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

duly simplified in light of step 1, we obtain:

$$W = p_H \frac{\lambda}{1-\lambda} + \frac{2\gamma - \lambda - \lambda\gamma - \lambda^2 - \lambda^2\gamma}{(\gamma+1)\,(\lambda+1)\,(1-\lambda)},$$

because the coefficient of p_H is positive, this quantity is minimized by setting $p_H = 0$. Then, solving for p_M and b when $p_H = 0$ we obtain:

$$\begin{aligned} b &= -\frac{1}{2\lambda + 2}\,(-2\lambda - \gamma + \theta\lambda + \theta\gamma - 1) \\ p_M &= \frac{1}{(\gamma+1)\,(1-\lambda)} \end{aligned}$$

we know that $1 \geq \gamma \geq \lambda$, so $p_M \geq 0$, but the condition $p_M \leq 1$ yields $\frac{1}{(\gamma+1)(1-\lambda)} - 1 \leq 0$, i.e. $\lambda \leq \frac{\gamma}{\gamma+1}$, as stated. We note that the probability of war equals:

$$C = \frac{(\lambda - 2\gamma + \lambda\gamma + \lambda^2 + \lambda^2\gamma)\,\lambda}{(\gamma+1)\,(\lambda+1)\,(\lambda-1)}.$$

Step 5. We want to show that for $\lambda < \gamma/(1+\gamma)$, $p_H = 0$ and $p_M = \frac{1}{(1+\gamma)(1-\lambda)}$ solve the original problem. Again, the low-type ex-ante IC* constraint coincides with the relaxed low-type ex-ante IC* constraint. We need to show that the ex-post constraint $b \geq p\theta$ is satified. In fact, simplification yields:

$$b - p\theta = \frac{1}{2}\,(\lambda+1)^{-1}\,(1-\gamma)\,(1-\theta)\,\lambda > 0.$$

Finally we show that the high-type ex-ante IC* constraint coincides with the (binding) relaxed high-type ex-ante IC* constraint, i.e. that $1 - b \geq \theta/2$. Note in fact, that this implies that the ex-post constraint $1 - b \geq (1-p)\,\theta$ is satisfied, because $\theta/2 > (1-p)\,\theta$. Indeed, after simplification, we obtain:

$$1 - b - \theta/2 = \frac{1}{2}\,(\lambda+1)^{-1}\,(1-\gamma)\,(1-\theta)\,\lambda \geq 0.$$

Step 6. We want to show that for $\lambda \in [\gamma/(1+\gamma), \min\{1/(1+\gamma), \gamma\}]$, $p_M = 1, p_H = 1 - \frac{\gamma}{(1+\gamma)\lambda}$ solves the relaxed problem. When $\lambda > \gamma/(1+\gamma)$, setting $p_H = 0$ violates the constraint $p_M = 1$. Further, the expression (3) reveals that p_M decreases in p_H. Hence minimization of p_H, which induces minimization of W, requires setting $p_M = 1$. Solving for b and p_H, we obtain:

$$b = -\frac{(-\lambda - 3\gamma + 2\theta\gamma - \lambda\gamma - \gamma^2 + \theta\gamma^2 - 1)}{2\lambda + 2\gamma + 2\lambda\gamma + 2}$$
$$p_H = \frac{\lambda - \gamma + \lambda\gamma}{(\gamma+1)\,\lambda} = 1 - \frac{\gamma}{(1+\gamma)\lambda}.$$

The condition that $p_H \geq 0$ requires that $\lambda \geq \frac{\gamma}{\gamma+1}$ as stated.

Step 7. We want to show that for $\lambda \in [\gamma/(1+\gamma), \min\{1/(1+\gamma), \gamma\}]$, $p_M = 1, p_H = 1 - \frac{\gamma}{(1+\gamma)\lambda}$ solves the original problem. Again, the low-type ex-ante IC* constraint coincides with the relaxed low-type ex-ante IC* constraint. We need to show that the ex-post constraint $b \geq p\theta$ is satified. In fact, simplification yields:

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

and this quantity is positive if and only if $\lambda \leq \frac{1}{\gamma+1}$. Finally we show that the high-type ex-ante IC* constraint coincides with the (binding) relaxed high-type ex-ante IC* constraint, i.e. that $1-b \geq \theta/2$. Note in fact, that this implies that the ex-post constraint $1-b \geq (1-p)\theta$ is satisfied, because $\theta/2 > (1-p)\theta$. Indeed, after simplification, we obtain:

$$1-b-\theta/2 = \frac{1}{2}(\gamma+1)^{-1}(\lambda+1)^{-1}(1-\theta)\left(\lambda-\gamma+\lambda\gamma-\gamma^2+1\right)$$

and $\lambda-\gamma+\lambda\gamma-\gamma^2+1 \geq 0$ if and only if $\lambda \geq \frac{1}{\gamma+1}(\gamma+\gamma^2-1)$ but because $\frac{1}{\gamma+1}(\gamma+\gamma^2-1) < \frac{\gamma}{\gamma+1}$, this condition is less stringent than $\lambda \geq \frac{\gamma}{\gamma+1}$.

Case 2. We want to show that for $\lambda \in [1/(1+\gamma), \gamma)$, $p_M = 1, p_H = \frac{2\lambda-\gamma}{\lambda(2+\gamma)}$ solve the original problem. Consider now the same relaxed problem that we considered in the proof for the case of $\gamma \geq 1$. We know from the analysis for the case $\gamma \geq 1$, that this relaxed problem is solved by $p_H = 0, p_M = \frac{1}{1+\gamma-2\lambda}, b = p\theta$ for $\lambda < \gamma/2$ and by $p_M = 1, p_H = \frac{2\lambda-\gamma}{\lambda(\gamma+2)}, b = p\theta$ for $\lambda \in [\gamma/2, \gamma)$. We now note that

$$\frac{1}{\gamma+1} - \gamma/2 = \frac{1}{2}(\gamma+1)^{-1}(1-\gamma)(\gamma+2)$$

and this quantity is positive when $\gamma \leq 1$. Hence the possibility that $\lambda < \gamma/2$ is ruled out: On the domain $1/(1+\gamma) \leq \lambda \leq \gamma \leq 1$, the solution to the relaxed problem is $p_M = 1, p_H = \frac{2\lambda-\gamma}{\lambda(\gamma+2)}$, with $b = p\theta$. We now need to show that this is also the solution of the original problem. Again, the low-type ex-ante IC* constraint coincides with the relaxed low-type ex-ante IC* constraint. Consider the ex-ante high-type IC* constraint. The condition $1-b = 1-p\theta \geq \theta/2$ yields $\gamma = \frac{2p\theta-1}{1-\theta} \leq 1$. Hence, for $\gamma \leq 1$, we conclude that $1-b \geq \theta/2$, and hence that $1-b \geq (1-p)\theta$. So the ex-ante high-type IC* constraint becomes:

$$(1-q)\left((1-p_M)p\theta + p_M p\theta\right) + q\left((1-p_H)\frac{\theta}{2} + p_H\frac{1}{2}\right) - (1-q)p\theta - q\left((1-p_M)\frac{\theta}{2} + p_M(1-p\theta)\right) \geq 0$$

and indeed, after simplification, the LHS equals:

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

a positive quantity as long as $\lambda+\lambda\gamma-1$, i.e., $\lambda>\frac{1}{\gamma+1}$, which is exactly the condition under which we operate.

Proof of Proposition 3 We proceed in three parts.

Part 1. (The low type mixes). Suppose that the low type mix between the low message (with probability σ) and the high message, whereas the high type only sends the high message. Let $\chi:=\frac{q}{1-\pi}$ be the posterior of facing a high type after the high message, where $\pi:=(1-q)\sigma$ is the probability of low message. The optimal semi-pooling equilibrium solves the following program:

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

subject to the indifference condition for the low type:

$$\begin{aligned}&\pi((1-p_L)\frac{\theta}{2}+p_L\frac{1}{2})+(1-\pi)((1-p_M)(\chi(1-p)\theta+(1-\chi)\frac{\theta}{2})+p_M(1-b))\\ =\ &\pi((1-p_M)\frac{\theta}{2}+p_Mb)+(1-\pi)((1-p_H)(\chi(1-p)\theta+(1-\chi)\frac{\theta}{2})+p_H\frac{1}{2})\end{aligned}$$

to the ex-ante IC* constraint for the high type

$$\begin{aligned}&\pi((1-p_M)p\theta+p_Mb)+(1-\pi)((1-p_H)(\chi\frac{\theta}{2}+(1-\chi)p\theta)+p_H\frac{1}{2})\geq\\ &\pi((1-p_L)p\theta+p_L\max\{p\theta,\frac{1}{2}\})\\ &+(1-\pi)((1-p_M)(\chi\frac{\theta}{2}+(1-\chi)p\theta)+p_M\max\{1-b,\chi\frac{\theta}{2}+(1-\chi)p\theta\})\end{aligned}$$

to the high-type ex-post constraints:

$$b\geq p\theta,\ 1/2\geq\chi\theta/2+(1-\chi)p\theta$$

to the low-type ex-post constraints:

$$1-b\geq\chi(1-p)\theta+(1-\chi)\theta/2,\ 1/2\geq\theta/2,\ b\geq\theta/2,\ \frac{1}{2}\geq\chi(1-p)\theta+(1-\chi)\frac{\theta}{2}$$

and to the probability constraints:

$$0 \leq p_L \leq 1, 0 \leq p_M \leq 1, 0 \leq p_H \leq 1, 0 \leq \sigma \leq 1.$$

We note that for $\sigma = 1$, the constraint set is weakly more restrictive than the perfectly-separating equilibrium program.

In order to solve the above program we distinguish two cases.

Case 1 ($\gamma \geq 1$ or $\gamma < 1$ and $\lambda \in [1/(1+\gamma), \gamma)$).

Consider the following relaxed problem:

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

subject to the ex-ante indifference condition for the low-type

$$\begin{aligned} &\pi((1 - p_L)\frac{\theta}{2} + p_L\frac{1}{2}) + (1 - \pi)((1 - p_M)(\chi(1 - p)\theta + (1 - \chi)\frac{\theta}{2}) + p_M(1 - b)) \\ = \ &\pi((1 - p_M)\frac{\theta}{2} + p_M b) + (1 - \pi)((1 - p_H)(\chi(1 - p)\theta + (1 - \chi)\frac{\theta}{2}) + p_H\frac{1}{2}) \end{aligned}$$

to the high-type ex-post constraints:

$$b \geq p\theta, \ \ 1/2 \geq \chi\theta/2 + (1 - \chi)p\theta$$

and to the probability constraints:

$$p_L \leq 1, p_M \leq 1, 0 \leq p_H, \sigma \leq 1.$$

Step 1. We want to show that $p_L = 1$. Consider the low-type relaxed ex-ante IC* constraint (duly simplified)

$$\begin{aligned} &\pi((1 - p_L)\frac{\theta}{2} + p_L\frac{1}{2}) + (1 - \pi)((1 - p_M)(\chi(1 - p)\theta + (1 - \chi)\frac{\theta}{2}) + p_M(1 - b)) \\ = \ &\pi((1 - p_M)\frac{\theta}{2} + p_M b) + (1 - \pi)((1 - p_H)(\chi(1 - p)\theta + (1 - \chi)\frac{\theta}{2}) + p_H\frac{1}{2}) \end{aligned}$$

The right-hand side increases in p_L and the left-hand side increases in p_H. Clearly, increasing both p_L and p_H decreases the objective function, without violating the high-type ex-post constraints. Noting that in the relaxed problem, there is no constraint that p_H, we conclude that $p_L = 1$.

Step 2. We want to show that the first high-type ex-post constraint binds. Suppose not, i.e. $b > p\theta$. Then b can be reduced and p_H can be increased maintaining the ex-ante indifference condition for the low-type, thus decreasing the objective function.

Step 3. We want to show that the second high-type ex-post constraint can be substituted with the constraint $\sigma \geq \frac{\gamma-\lambda}{\gamma}$. In fact, after substitution the constraint is:

$$1/2 \geq \frac{q}{1-(1-q)\sigma}\theta/2 + \left(1 - \frac{q}{1-(1-q)\sigma}\right) p\theta,$$

the right-hand side is decreasing in σ, because $p\theta > 1/2 > \theta/2$, and the value $\frac{\gamma-\lambda}{\gamma}$ is obtained subtituting p, θ, q with γ, λ.

We now proceed to find the optimal (p_H, p_M) as a function of σ, and we will later establish the optimal σ.

Step 4. We want to show that whenever $\sigma \geq 2\lambda - \gamma + 1$, the optimal solution is $p_H = 0$ and $p_M = \frac{\sigma}{2\sigma-2\lambda+\gamma-1}$, and further, the objective function C is strictly decreasing in σ. We solve for p_M as a function of p_H and σ in the low-type relaxed ex-ante IC* constraint (after plugging in $b = p\theta$), to obtain: $p_M = \frac{\sigma - p_H + \sigma p_H - 2\lambda p_H - \lambda\gamma p_H}{2\sigma-2\lambda+\gamma-1}$. We then plug back the expression in the objective function and obtain:

$$C = \frac{\left(1-\sigma+3\lambda-\gamma+\sigma\gamma-\lambda\gamma+2\sigma\lambda\gamma+2\lambda^2\right)\left(\lambda-\sigma+1\right)}{\left(2\sigma-2\lambda+\gamma-1\right)\left(\lambda+1\right)^2} p_H + K_1(\sigma),$$

where $K_1(\sigma)$ is a function independent of p_H. Because $\sigma \in \left[\frac{\gamma-\lambda}{\gamma}, 1\right]$, the term $1-\sigma+3\lambda-\gamma+\sigma\gamma-\lambda\gamma+2\sigma\lambda\gamma+2\lambda^2$ is strictly positive, because $\sigma > 2\lambda-\gamma+1$, the term $2\sigma-2\lambda+\gamma-1$ is positive. Hence, the coefficient of p_H is positive, and C is minized by minizing p_H. Setting $p_H = 0$ and solving back p_M we obtain

$$p_M = \frac{\sigma}{2\sigma-2\lambda+\gamma-1},$$

which is smaller than one if and only $\sigma \geq 2\lambda - \gamma + 1$. Substituting in the objective function, and differentiating it with respect to σ, we obtain:

$$\frac{\partial^2 C}{\partial\sigma\partial\sigma} = -2\frac{(\sigma - 2\lambda + \gamma - 1)(\gamma + 1)\sigma}{(2\sigma - 2\lambda + \gamma - 1)^2 (\lambda + 1)^2}$$

which is strictly negative when $\sigma \geq 2\lambda - \gamma + 1$. Hence, the objective function is strictly decreasing in σ.

Step 5. We want to show that whenever $\sigma \leq 2\lambda - \gamma + 1$, the optimal solution is $p_H = \frac{2\lambda - \sigma - \gamma + 1}{2\lambda - \sigma + \lambda\gamma + 1}$ and $p_M = 1$; and the objective function C is strictly decreasing in σ. Again from the low-type relaxed ex-ante IC* constraint, we obtain an expression for p_H as a function of σ and p_M, we plug this expression in the objective function to obtain:

$$C = \frac{(1 - \sigma + 3\lambda - \gamma + \sigma\gamma - \lambda\gamma + 2\sigma\lambda\gamma + 2\lambda^2)(\lambda - \sigma + 1)}{(\sigma - 2\lambda - \lambda\gamma - 1)(\lambda + 1)^2} p_M + K_2(\sigma).$$

Because $\sigma \leq 2\lambda - \gamma + 1$, the coefficient of p_M is negative, and C is minimized by setting $p_M = 1$. Setting $p_M = 1$ and solving back p_H we obtain:

$$p_H = \frac{2\lambda - \sigma - \gamma + 1}{2\lambda - \sigma + \lambda\gamma + 1}$$

which is positive when $\sigma \leq 2\lambda - \gamma + 1$. Substituting in the objective function, and differentiating it with respect to σ, we obtain:

$$\frac{\partial C}{\partial\sigma} = -\frac{(3\lambda - \sigma + 2\lambda\gamma + 1)(\lambda - \sigma + 1)}{(2\lambda - \sigma + \lambda\gamma + 1)^2 (\lambda + 1)}\gamma$$

which is strictly negative when $\sigma \leq 2\lambda - \gamma + 1$. Hence the objective function is strictly decreasing in σ.

We have concluded that the solution of the relaxed problem requires $\sigma = 1$. Hence, the pure-strategy equilibrium is the solution of the program where the low type mixes.

Case 2 ($\gamma < 1$ and $\lambda \leq 1/(1+\gamma)$). We have verified this result with a semi-numerical analysis that is available upon request. A FULLY ANALYTICAL PROOF IS TO BE COMPLETED.

Part 2 (The high type mixes). Suppose that the high type mix between the high message (with probability ρ) and the low message. The low type only sends the low message. Let $\zeta := \frac{q(1-\rho)}{1-q\rho}$ be the posterior of facing a high type after the low message. Let $\pi := 1 - q\rho$ be the probability of low message. The optimal equilibrium is found by solving the following program.

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

subject the ex-ante IC* constraint for the for the low type:

$$\begin{aligned}
&\pi((1-p_L)(\zeta(1-p)\theta + (1-\zeta)\theta/2) + p_L\frac{1}{2}) + (1-\pi)((1-p_M)(1-p)\theta + p_M(1-b)) \\
\geq\ &\pi((1-p_M)(\zeta(1-p)\theta + (1-\zeta)\frac{\theta}{2}) + p_M \max\{b, (\zeta(1-p)\theta + (1-\zeta)\frac{\theta}{2})\}) \\
&+(1-\pi)((1-p_H)(1-p)\theta + p_H \max\{\frac{1}{2}, (1-p)\theta\})
\end{aligned}$$

to the indifference condition for the high type

$$\begin{aligned}
&\pi((1-p_M)(\zeta\frac{\theta}{2} + (1-\zeta)p\theta) + p_M b) + (1-\pi)((1-p_H)\frac{\theta}{2} + p_H\frac{1}{2}) = \\
&\pi((1-p_L)(\zeta\frac{\theta}{2} + (1-\zeta)p\theta) + p_L\frac{1}{2}) + (1-\pi)((1-p_M)\frac{\theta}{2} + p_M(1-b))
\end{aligned}$$

to the high-type ex-post constraints:

$$b \geq \zeta\theta/2 + (1-\zeta)p\theta,\ 1/2 \geq \theta/2,\ 1/2 \geq \zeta\frac{\theta}{2} + (1-\zeta)p\theta,\ 1-b \geq \frac{\theta}{2}$$

to the low-type ex-post constraints:

$$1-b \geq (1-p)\,\theta,\ 1/2 \geq \zeta\,(1-p)\,\theta + (1-\zeta)\,\theta/2$$

and to the probability constraints:

$$0 \leq p_L \leq 1, 0 \leq p_M \leq 1, 0 \leq p_H \leq 1, 0 \leq \sigma \leq 1.$$

But is immediate to note that the constraint set is empty. Indeed, the third high-type ex-post constraint is equivalent to:

$$-\frac{1}{2}(1-\theta)\frac{\gamma - \lambda(1-\rho)}{1+\lambda(1-\rho)} \geq 0,$$

which cannot be the case for $\gamma > \lambda$.

Part 3 (Both types mix). Suppose that the low type mixes between the low message (with probability σ) and the high message. The high type mixes between the high message (with probability ρ) and the low message. Let $\chi := \frac{q\rho}{1-\pi}$ be the posterior of facing a high type after the high message. Let $\pi := (1-q)\sigma + q(1-\rho)$ be the probability of a low message. Let $\zeta := \frac{q(1-\rho)}{\pi}$ be the posterior of facing a high type after the low message. The optimal equilibrium solves the following program:

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

subject the ex-ante IC* constraint for the for the low type:

$$\pi((1-p_L)(\zeta(1-p)\theta + (1-\zeta)\theta/2) + p_L\frac{1}{2}) + (1-\pi)((1-p_M)(\chi(1-p)\theta + (1-\chi)\frac{\theta}{2}) + p_M(1-b))$$
$$= \pi((1-p_M)(\zeta(1-p)\theta + (1-\zeta)\theta/2) + p_M b) + (1-\pi)((1-p_H)(\chi(1-p)\theta + (1-\chi)\frac{\theta}{2}) + p_H\frac{1}{2})$$

to the indifference condition for the high type

$$\pi((1-p_M)(\zeta\frac{\theta}{2} + (1-\zeta)p\theta) + p_M b) + (1-\pi)((1-p_H)(\chi\frac{\theta}{2} + (1-\chi)p\theta) + p_H\frac{1}{2}) =$$
$$\pi((1-p_L)(\zeta\frac{\theta}{2} + (1-\zeta)p\theta) + p_L\frac{1}{2}) + (1-\pi)((1-p_M)(\chi\frac{\theta}{2} + (1-\chi)p\theta) + p_M(1-b))$$

to the high-type ex-post constraints:

$$b \geq \zeta\theta/2 + (1-\zeta)p\theta,\ 1/2 \geq \chi\theta/2 + (1-\chi)p\theta,\ \frac{1}{2} \geq \zeta\frac{\theta}{2} + (1-\zeta)p\theta,\ 1-b \geq \chi\frac{\theta}{2} + (1-\chi)p\theta$$

to the low-type ex-post constraints:

$$1-b \geq \chi(1-p)\theta + (1-\chi)\theta/2,\ 1/2 \geq \zeta(1-p)\theta + (1-\zeta)\theta/2,\ b \geq \zeta(1-p)\theta + (1-\zeta)\frac{\theta}{2}$$
$$\frac{1}{2} \geq \chi(1-p)\theta + (1-\chi)\frac{\theta}{2}$$

and to the probability constraints:

$$0 \leq p_L \leq 1, 0 \leq p_M \leq 1, 0 \leq p_H \leq 1, 0 \leq \sigma \leq 1, 0 \leq \rho \leq 1.$$

But is immediate to note that the constraint set is empty. Indeed, second and fourth high-type ex-post constraints are equivalent to:

$$X := \frac{1}{2}(1-\theta)\frac{(\rho+\sigma)\lambda - \gamma}{\rho\lambda + 1 - \sigma} \geq 0,\ Z := \frac{1}{2}(1-\theta)\frac{\rho\lambda - \lambda + \sigma\gamma}{\rho\lambda - \lambda - \sigma} \geq 0.$$

Evidently, $X \geq 0$ requires $\lambda \geq \frac{\gamma}{\rho+\sigma}$, which, in light of $\gamma > \lambda$, requires $\rho + \sigma > 1$. Consider Z, note that it increases in λ. When λ takes its upper value γ,

$$Z = \frac{1}{2}(1-\theta)\frac{(1-\sigma-\rho)\gamma}{\sigma+\gamma(1-\rho)}$$

which is positive if and only if $\sigma + \rho \leq 1$. This concludes that whenever $\gamma > \lambda$, either $X < 0$ or $Z < 0$ or both.

Proof of Proposition 2 We first find the solution of the relaxed program:

$$\begin{aligned}\min_{\beta, p_L, q_L, p_M^+, p_M^-, q_M, p_H, q_H} \quad & (1-q)^2(1-2p_L-q_L) + 2q(1-q)(1-q_M-p_M^+-p_M^-) \\ & +q^2(1-2p_H-q_H)\end{aligned}$$

subject to the high-type ex-post constraints:

1. $\left(qp_H + (1-q)p_M^+\right)\beta \geq qp_H\theta/2 + (1-q)p_M^+ p\theta$
2. $\left(qp_H + (1-q)p_M^-\right)(1-\beta) \geq qp_H\theta/2 + (1-q)p_M^- p\theta$
3. $\left(qq_H + (1-q)q_M\right)/2 \geq qq_H\theta/2 + (1-q)q_M p\theta$

to ex-ante IC* constraints where the low type does not wage war after mis-reporting (after simplification)

$$\begin{aligned}&q(p_M^-\beta + p_M^+(1-\beta) + q_M/2 + (1-p_M^+-p_M^--q_M)(1-p)\theta) \\ &+(1-q)((2p_L+q_L)/2 + (1-2p_L-q_L)\theta/2) \geq \\ &(1-q)(p_M^-(1-\beta) + p_M^+\beta + q_M/2 + (1-p_M^+-p_M^--q_M)\theta/2) \\ &+q((2p_H+q_H)/2 + (1-2p_H-q_H)(1-p)\theta)\end{aligned}$$

and to the probability constraints:

$$q_L + 2p_L \leq 1, \; p_M^+ + p_M^- + q_M \leq 1, \; 0 \leq p_H, \; 0 \leq q_H.$$

Then, we will show that its solution (with innocuous additional constraints) solves also the complete problem.

Step 1. We want to show that $q_L + 2p_L = 1$. We first note that setting $q_L + 2p_L = 1$ maximizes the LHS of the relaxed low-type IC* constraint and does not affect the RHS. It is immediate to see that the high-type ex-post constraints are not affected. Because $q_L + 2p_L = 1$ minimizes the first-term of the objective function, we conclude that it must be part of the solution.

Step 2 We want to show that the relaxed low-type IC* constraint binds. Suppose it does not. It is then possible to increase q_H (thus increasing $2p_H + q_H$) thus decreasing the objective function without violating the constraint (note that there is no constraint that $2p_H + q_H \leq 1$ in the relaxed problem). Indeed, $1/2 > (1-p)\,\theta$ and $2p_H + q_H$ appears only on the RHS of the relaxed low-type IC* constraint:

$$
\begin{aligned}
&q(p_M^-\beta + p_M^+(1-\beta) + q_M/2 + (1 - p_M^+ - p_M^- - q_M)(1-p)\theta) + (1-q)/2 > \\
&q(2p_H + q_H)/2 + (1-q)(p_M^-(1-\beta) + p_M^+\beta + q_M/2) \\
&+q(1 - 2p_H - q_H)(1-p)\theta + (1-q)(1 - p_M^+ - p_M^- - q_M)\theta/2.
\end{aligned}
$$

Because $1/2 > \theta/2$, increasing q_H softens the third high-type ex-post constraints, and leaves unchanged the other two. Hence, we have shown that the relaxed low-type IC* constraint must bind in the solution.

Step 3. We want to show that the first high-type ex-post constraint binds. Suppose it does not. Then, it is possible to reduce β without violating the constraint, nor the other high-type ex-post constraints (the second one becomes slacker, and the third high-type is unchanged). The objective function is unchanged. As long as $p_M^+ > p_M^-$, reducing β makes the makes the low–type relaxed IC* constraint slack, because the effect on the constraint is:

$$
q(p_M^- d\beta - p_M^+ d\beta) - (1-q)(-p_M^- d\beta + p_M^+ d\beta) > 0.
$$

Because step 2 concluded that the low–type relaxed IC* constraint cannot be slack in the solution, we have proved that the first high-type ex-post constraint cannot be slack.

Step 4. We want to show that the second high-type ex-post constraint binds. Suppose not, then it is possible to increase p_M^- and decrease q_M so that $dp_M^- = -dq_M > 0$ without violating the constraint, nor the other high-type ex-post constraints (the first one is unchanged, and the third one becomes slacker). Further the objective function is unchanged, and the effect on the low-type ex-ante IC* constraint is:

$$\begin{aligned} & q(dp_M^-\beta + dq_M/2) - (1-q)(dp_M^-(1-\beta) + dq_M/2) \\ \propto\ & q(\beta - 1/2) - (1-q)((1-\beta) - 1/2) = \beta - 1/2 > 0. \end{aligned}$$

Because step 2 concluded that the low–type relaxed IC* constraint cannot be slack in the solution, we have proved that the second high-type ex-post constraint cannot be slack.

Step 5. We want to show that the third high-type ex-post constraint binds. Suppose not, then it is possible to reduce p_M^+ and increase q_M so that $dq_M = -dp_M^+ > 0$, without violating the ex-post high type constraints (the first one becomes slacker, and the second one is unchanged) The objective function is unchanged. The effect on the relaxed low-type IC* constraint is:

$$\begin{aligned} & q(dp_M^+(1-\beta) + dq_M/2 + (-dp_M^+ - dq_M)(1-p)\theta) \\ & -(1-q)(dp_M^+\beta + dq_M/2 + (-dp_M^+ - dq_M)\theta/2) \\ = & \left[q(-(1-\beta) + 1/2) - (1-q)(-\beta + 1/2)\right] dq_M = (\beta - 1/2)\, dq_M > 0 \end{aligned}$$

so that the constraint becomes slack. Because step 2 concluded that the low–type relaxed IC* constraint constraint cannot be slack in the solution, we have proved that the third high-type ex-post constraint cannot be slack.

Step 6. We want to show that for $\lambda \leq \gamma/2$: $p_H = q_H = p_M^- = q_M = 0$, and $p_M^+ = \frac{1}{1+\gamma-2\lambda}$ in the relaxed program. We have concluded that all constraints in the relaxed problem bind. We now solve for the system that includes all such binding constraints to find the expressions for $\left(\beta, q_M, p_M^+, p_M^-\right)$ as function of (p_H, q_H), and we substitute the resulting expressions in the objective function:

$$W = 2(1-q)(1 - q_M - p_M^+ - p_M^-) + q(1 - q_H - p_H), \tag{4}$$

duly simplified in light of step 1. We obtain an following expression of W as a function of p_H and q_H only. Substituting the variables p and q for γ and λ, we obtain

$$W = \frac{\lambda(1+\gamma)}{(\lambda+1)(\gamma-2\lambda+1)}q_H + 2\frac{\lambda(1+\gamma)}{(\lambda+1)(\gamma-2\lambda+1)}p_H + \frac{(2\gamma-3\lambda+\lambda\gamma-2\lambda^2)}{(\lambda+1)(\gamma-2\lambda+1)}.$$

We note that for $\gamma \geq 2\lambda$, it is also the case that $\gamma+1 \geq 2\lambda$, and hence the coefficients of q_H and p_H are positive. Further, for $\gamma \geq 2\lambda$, it is also the case that $2\gamma - 3\lambda + \lambda\gamma - 2\lambda^2 \geq 0$, and hence the probability of war is positive. Hence the function W is minimized by setting $p_H = q_H = 0$. We now need to check that, by doing so, we do not violate the remaining probability constraints.

Consider the second high-type ex-post constraint, when $p_H = 0$, it takes the form: $p_M^-(1-\beta) = p_M^- p\theta$. This allows for the solutions $\beta = 1 - p\theta$ and $p_M^- = 0$. The first one is not admissible because $\beta \geq 1/2$, hence we conclude that $p_M^- = 0$. Similarly the third high-type ex-post constraint, when $q_H = 0$, takes the form $(1-q)q_M/2 = (1-q)q_M p\theta$. This only allows the solution $q_M = 0$. Solving for β and p_M^+ in the system including the first high-type ex-post constraint, and the ex-ante low-type relaxed IC* constraint, we obtain $\beta = p\theta$, $p_M^+ = \frac{1}{\gamma+1-2\lambda}$. When $\gamma \geq 2\lambda$, $p_M^+ \leq 1$, hence the solution is admissible.

Step 7. For $\lambda < \gamma/2$, we want to show that the solution constructed satisfies all the program constraints. We first consider the low-type IC* constraint, which is satisfied because, noting that $p_M^- = 0 = p_H = q_M = q_H$, we only need to show that

$$\left(qp_H + (1-q)p_M^+\right)p\theta = \left(qp_H + (1-q)p_M^+\right)\beta \geq qp_H(1-p)\theta + (1-q)p_M^+\theta/2.$$

Consider the high-type IC* constraint: Note that

$$\left(qp_M^+ + (1-q)p_L\right)(1-\beta) \leq qp_M^+\theta/2 + (1-q)p_L p\theta$$

as long as either $\gamma > 1$ or $p_L \geq \frac{(1-\gamma)\lambda}{2\gamma}p_M^+ = \frac{(1-\gamma)\lambda}{2\gamma^2}\frac{(\lambda-\gamma)(\gamma+2)}{(\lambda-\gamma-1)}$ for $\gamma < 1$, that

$$\left(qp_M^- + (1-q)p_L\right)\beta = \left(qp_M^- + (1-q)p_L\right)\beta = qp_M^-\theta/2 + (1-q)p_L p\theta,$$

and that

$$(qq_M + (1-q)q_L)\cdot 1/2 \leq qq_M\theta/2 + (1-q)q_L p\theta.$$

Then we substitute in the high-type IC* constraint (duly simplified):

$$q\theta/2+(1-q)(p_M^+\beta+(1-p_M^+)p\theta)\geq q\theta/2+(1-q)\,p\theta$$

which is clearly satisfied because $\beta=p\theta$.

Finally, we need to show that the low-type ex-post constraints are satisfied. Indeed:

$$\begin{aligned}\left(qp_M^-+(1-q)p_L\right)p\theta &> qp_M^-(1-p)\theta+(1-q)p_L\frac{\theta}{2}\\ (qq_H+(1-q)q_M)\cdot 1/2 &> qq_H(1-p)\theta+(1-q)q_M\frac{\theta}{2}\end{aligned}$$

whereas

$$\left(qp_M^++(1-q)p_L\right)(1-p\theta)\geq qp_M^+(1-p)\theta+(1-q)p_L\theta/2$$

as long as $p_L\,(\gamma-1)=p_L\frac{(\theta+2p\theta-2)}{(1-\theta)}\leq 2\frac{q}{(1-q)}p_M=2\lambda p_M$. So that if $\gamma\geq 1$, $p_L\leq\frac{2\lambda}{(\gamma-2\lambda+1)(\gamma-1)}$ and if $\gamma<1$, $p_L\geq 0\geq\frac{2\lambda}{(\gamma-2\lambda+1)(\gamma-1)}$.

Step 8. We want to show that for $\gamma/2\leq\lambda\leq\gamma$, setting $p_H=p_M^-=0$, $\beta=p\theta$, $p_M^++q_M=1$ and $q_H=\frac{2\lambda-\gamma}{\lambda(1-\lambda+\gamma)}$, $q_M=\frac{2\lambda-\gamma}{\gamma(1-\lambda+\gamma)}$ maximizes the relaxed program. We have concluded that all constraints in the relaxed problem bind. We now let $Q=p_M^-+p_M^++q_M$, solve system that includes all such binding constraints to find the expressions for $\left(\beta,q_M,p_M^+,p_M^-,q_H\right)$ as function of (p_H,Q), and we substitute the resulting expressions in the objective function (4), to obtain the following expression:

$$W=-\frac{\gamma+1}{(1-\lambda+\gamma)\,(\lambda+1)}Q+\frac{(1-\lambda+2\gamma-\lambda^2+\lambda\gamma)}{(1-\lambda+\gamma)\,(\lambda+1)}.$$

Because $\lambda<\gamma$, we note that W is positive for $Q\in[0,1]$ and that the coefficient of Q is negative, so that minimizing W requires maximizing Q, which cannot exceed one, by the probability constraint $p_M^-+p_M^++q_M\leq 1$. We now show that it is possible to set $p_H=0$ and obtain $Q=1$ without violating any of the relaxed program constraints. In fact, when $p_H=0$, the first high-type ex-post constraint implies that $\beta=p\theta$ and the second high type ex-post constraint implies that $p_M^-=0$. Solving the system including the low-type ex-ante IC* constraint and the third high-type ex-post contraint, together with the restriction $p_M^++q_M=1$ yields $q_H=\frac{2\lambda-\gamma}{\lambda(1-\lambda+\gamma)}$, $q_M=\frac{2\lambda-\gamma}{\gamma(1-\lambda+\gamma)}$ and $p_M^+=1-q_M$. Because $\gamma/2\leq\lambda\leq\gamma$, we obtain that $0 \leq \frac{2\lambda-\gamma}{\lambda(1-\lambda+\gamma)}$ and $\frac{2\lambda-\gamma}{\gamma(1-\lambda+\gamma)} \leq 1$, hence the probability constraints are not violated.

Step 9. For $\lambda \geq \gamma/2$, we want to show that the solution constructed (with innocuous additional constraints) satisfies all the program constraints. We first consider the low-type IC* constraint: We only need to show that, off the equilibrium path, the low-type settles:

$$\left(qp_H + (1-q)p_M^+\right)\beta \geq qp_H(1-p)\theta + (1-q)p_M^+\theta/2$$
$$(qq_H + (1-q)q_M)\cdot 1/2 \geq qq_H(1-p)\theta + (1-q)q_M\theta/2$$

and indeed:

$$p\theta > \theta/2$$
$$1/2 > \frac{qq_H}{qq_H+(1-q)q_M}(1-p)\theta + \frac{(1-q)q_M}{qq_H+(1-q)q_M})\theta/2.$$

Then we consider the high–type IC* constraint. We proceed in two steps. We first determine the off-path behavior of the high type and show that

$$\left(qp_M^+ + (1-q)p_L\right)\cdot(1-\beta) \leq qp_M^+\theta/2 + (1-q)p_Lp\theta$$

as long as either $\gamma > 1$ or $p_L \geq \frac{(1-\gamma)\lambda}{2\gamma}p_M^+ = \frac{(1-\gamma)\lambda}{2\gamma^2}\frac{(\lambda-\gamma)(\gamma+2)}{(\lambda-\gamma-1)}$ for $\gamma < 1$, that

$$\left(qp_M^- + (1-q)p_L\right)p\theta = qp_M^-\theta/2 + (1-q)p_Lp\theta$$

and that

$$(qq_M + (1-q)q_L)\,1/2 \leq qq_M\theta/2 + (1-q)q_Lp\theta$$

as long as $q_L \geq \frac{1-\theta}{2p\theta-1}\frac{q}{1-q}q_M$, i.e. $q_L \geq \frac{\lambda}{\gamma}q_M = \frac{\lambda(2\lambda-\gamma)}{\gamma^2(\gamma-\lambda+1)}$.

Then we substitute in the high–type IC* constraint (duly simplified):

$$\begin{aligned}
&q(q_H/2 + (1-q_H)\theta/2) + (1-q)(p_M^+\beta + q_M/2 + (1-p_M^+ - q_M)p\theta) \geq \\
&(qp_M^+ + (1-q)p_L)\{q_{l,1-\beta}\theta/2 + (1-q_{l,1-\beta})p\theta\} \\
&+(qp_M^- + (1-q)p_L)\{q_{l,\beta}\theta/2 + (1-q_{l,\beta})p\theta\} \\
&+(qq_M + (1-q)q_L)\{q_{l,1/2}\theta/2 + (1-q_{l,1/2})p\theta\} \\
&+q(1-p_M^+ - p_M^- - q_M)\theta/2 + (1-q)(1-2p_L-q_L)p\theta
\end{aligned}$$

$$
\begin{aligned}
&= qp_M^-\theta/2 + (1-q)p_L p\theta + qp_M^+\theta/2 + (1-q)p_L p\theta + qq_M\theta/2 + (1-q)q_L p\theta \\
&\quad + q(1 - p_M^+ - p_M^- - q_M)\theta/2 + (1-q)(1-2p_L - q_L)p\theta \\
&= q\theta/2 + (1-q)\,p\theta.
\end{aligned}
$$

Substituting in the expressions for q_M, q_H and $p_M^+ + q_M = 1$, we verify that the high-type IC* constraint is satisfied with equality.

Finally, we need to show that the low-type ex-post constraints are satisfied. Indeed:

$$
\begin{aligned}
\left(qp_M^- + (1-q)p_L\right)p\theta &> qp_M^-(1-p)\theta + (1-q)p_L\frac{\theta}{2} \\
\left(qq_M + (1-q)q_L\right)\cdot 1/2 &> qq_M(1-p)\theta + (1-q)q_L\frac{\theta}{2}
\end{aligned}
$$

whereas

$$
\left(qp_M^+ + (1-q)p_L\right)(1-p\theta) \geq qp_M^+(1-p)\theta + (1-q)p_L\frac{\theta}{2}
$$

as long as $p_L\,(\gamma-1) = p_L\frac{(\theta+2p\theta-2)}{(1-\theta)} \leq 2\frac{q}{(1-q)}p_M = 2\lambda p_M$. So that if $\gamma \geq 1$, $p_L \leq 2\frac{(\gamma-\lambda)(\gamma+2)\lambda}{(\gamma-\lambda+1)\gamma(\gamma-1)}$ and if $\gamma < 1$, $p_L \geq 0 \geq 2\frac{(\gamma-\lambda)(\gamma+2)\lambda}{(\gamma-\lambda+1)\gamma(\gamma-1)}$.

The proof is concluded because, obviously, the probability constraints are satisfied.

Additional Step. We want to show that for $\gamma/2 \leq \lambda \leq \gamma$, the solution $p_H = p_M^- = 0$, $\beta = p\theta$, $p_M^+ + q_M = 1$ and $q_H = \frac{2\lambda-\gamma}{\lambda(1-\lambda+\gamma)}, q_M = \frac{2\lambda-\gamma}{\gamma(1-\lambda+\gamma)}$ is not always the unique solution of the mediator's problem. We have previously concluded that, for $\gamma/2 \leq \lambda \leq \gamma$, any admissible solution to the following system is a solution to the relaxed problem (where we change notation, so that $p_M^- = p_N$ and $p_M^+ = p_M$):

$$
\begin{cases}
1 = p_M + p_N + q_M \\
\left(qq_H + (1-q)q_M\right)/2 = qq_H\theta/2 + (1-q)q_M p\theta \\
\left(qp_H + (1-q)p_N\right)(1-\beta) = qp_H\theta/2 + (1-q)p_N p\theta \\
\left(qp_H + (1-q)p_M\right)\beta = qp_H\theta/2 + (1-q)p_M p\theta \\
q(p_N\beta + p_M(1-\beta) + q_M/2 + (1-p_M-p_N-q_M)(1-p)\theta) + (1-q)/2 \\
= (1-q)(p_N(1-\beta) + p_M\beta + q_M/2 + (1-p_M-p_N-q_M)\theta/2) \\
+q((2p_H+q_H)/2 + (1-2p_H-q_H)(1-p)\theta)
\end{cases}
$$

Further, we know that the system is under-identified, and allows a continuum of solutions as functions of p_H (or say, p_N). We have focused on the solution where $p_H = 0$ and hence $p_N = 0$. But suppose that $p_N > 0$, and see if other solutions are admissible. We take an example with $p = 0.8, q = 0.5, \theta = 0.8$, so that $\gamma/2 = 0.7 \leq \lambda = 1 \leq \gamma = 1.4$, and we are within the required bounds. Picking $p_N = 0.001$, we solve the above system, and obtain the solution: $p = 0.8$, $q = 0.5$, $\theta = 0.8$, $\beta = 0.598\,83$, $\lambda = 1.0$, $\gamma = 1.4$, $p_H = 0.203\,75$, $q_H = 2.107\,0 \times 10^{-2}$, $p_M = 0.983\,95$, $p_N = 0.001$, $q_M = 1.505\,0 \times 10^{-2}$. Because all probabilities are positive, $\beta \in [1/2, p\theta]$ and $q_H + p_H < 1$, the solution is admissible. So, $\beta = 0.598\,83$, $p_H = 0.203\,75$, $q_H = 2.107\,0 \times 10^{-2}$, $p_M = 0.983\,95$, $p_N = 0.001$, $q_M = 1.505\,0 \times 10^{-2}$ is a solution to the relaxed problem for the parameters $(p = 0.8, q = 0.5, \theta = 0.8)$.

Now, we check that it is a solution to the mediator's problem. We first consider the low-type IC* constraint: We only need to show that, off the equilibrium path, the low-type settles. Indeed:

$$\begin{gathered}(qp_H + (1-q)p_N)(1-\beta) - (qp_H(1-p)\theta + (1-q)p_N\theta/2) = 2.457\,0 \times 10^{-2} > 0 \\ (qp_H + (1-q)p_M)\beta - (qp_H(1-p)\theta + (1-q)p_M\theta/2) = 0.142\,53 > 0 \\ (qq_H + (1-q)q_M) \cdot 1/2 - (qq_H(1-p)\theta + (1-q)q_M\theta/2) = 4.334\,4 \times 10^{-3} > 0\end{gathered}$$

Then we consider the high–type IC* constraint. We proceed in two steps. We first determine the off-path behavior of the high type and show that

$$(qp_M + (1-q)p_L) \cdot (1-\beta) \leq qp_M\theta/2 + (1-q)p_Lp\theta$$

as long as $p_L \geq 4.820\,3 \times 10^{-3}$, that

$$(qp_N + (1-q)p_L)\beta \leq qp_N\theta/2 + (1-q)p_Lp\theta$$

as long as $p_L \geq 4.829\,5 \times 10^{-3}$ and that

$$(qq_M + (1-q)q_L)\,1/2 \leq qq_M\theta/2 + (1-q)q_Lp\theta$$

as long as $q_L \geq 0.010\,75$.

Then we substitute in the high–type IC* constraint, and obtain that it is satisfied with equality.

Finally, we need to show that the low-type ex-post constraints are satisfied. Indeed:

$$\left(qp_M^- + (1-q)p_L\right)\beta > qp_M^-(1-p)\theta + (1-q)p_L\frac{\theta}{2}$$

for all $p_L \in [0,1]$,

$$(qq_M + (1-q)q_L) \cdot 1/2 > qq_M(1-p)\theta + (1-q)q_L\frac{\theta}{2}$$

for all $q_L \in [0,1]$, and

$$(qp_M + (1-q)p_L)(1-\beta) > qp_M(1-p)\theta + (1-q)p_L\frac{\theta}{2}$$

for all $p_L \in [0,1]$.

This concludes that the solution $\beta = 0.598\,83$, $p_H = 0.203\,75$, $q_H = 2.107\,0 \times 10^{-2}$, $p_M = 0.983\,95$, $p_N = 0.001$, $q_M = 1.505\,0 \times 10^{-2}$ indeed solves the mediator's problem for the parameters $(p = 0.8, q = 0.5, \theta = 0.8)$.

Proof of Proposition 4 We first solve the following relaxed program:

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

subject to high-type interim individual rationality:

$$(1-q)(p_M b + (1-p_M)p\theta) + q\left(p_H\frac{1}{2} + (1-p_H)\frac{\theta}{2}\right) \geq (1-q)p\theta + q\frac{\theta}{2}$$

to low-type interim incentive compatibility:

$$\begin{aligned}(1-q)\left((1-p_L)\frac{\theta}{2} + p_L\frac{1}{2}\right) + q\left((1-p_M)(1-p)\theta + p_M(1-b)\right) \geq \\ (1-q)\left((1-p_M)\frac{\theta}{2} + p_M b\right) + q\left((1-p_H)(1-p)\theta + p_H\frac{1}{2}\right),\end{aligned}$$

and to

$$p_L \leq 1, p_M \leq 1 \text{ and } p_H \geq 0.$$

First, note that $p_L = 1$ in the solution because p_L appears in the constraints only in the right-hand side of the low-type interim incentive compatibility constraint, which is increasing in p_L. Second, note that the low-type interim incentive compatibility must be binding in the relaxed program's solution, or else one could increase p_H thus reducing the value of the objective function, without violating the high-type interim individual rationality constraint. Third, note that the high-type interim individual rationality constraint must be binding in the relaxed program's solution, or else one could decrease b and make the low-type interim incentive compatibility slack.

Solving for b and p_H as a function of p_M in the system defined by the low-type interim incentive compatibility and high-type interim individual rationality constraints, and plugging back the resulting expressions in the objective function, we obtain

$$C = -p_M \frac{\gamma + 1}{(\lambda + 1)(\gamma + 1 - \lambda)} + K,$$

where K is an inconsequential constant. Hence, the probability of conflict is minimized by setting $p_M = 1$ whenever possible. Substituting $p_M = 1$, in the expression for p_H earlier derived, we obtain $p_H = \frac{2\lambda - \gamma}{(\gamma - \lambda + 1)\lambda}$, which is strictly positive for $\lambda \geq \gamma/2$ and always smaller than one.

Solving for b and p_M as a function of p_H in the system defined by the low-type interim incentive compatibility and high-type interim individual rationality constraints, and plugging back the resulting expressions in the objective function, we obtain

$$C = \frac{(\gamma + 1)\lambda}{(\gamma - 2\lambda + 1)(\lambda + 1)} p_H + K,$$

where K is another inconsequential constant. The coefficient of p_H is positive for $\lambda \leq \gamma/2$, hence the probability of conflict is minimized by setting $p_H = 0$, which entails $p_M = \frac{1}{\gamma - 2\lambda + 1}$, a quantity positive and smaller than one when $\lambda \leq \gamma/2$.

The proof is concluded by showing that this solution does not violate the high-type interim incentive compatibility and low-type interim individual rationality constraints in the complete program. Indeed, for $\lambda \geq \gamma/2$, we verify that the slacks of these constraints are, respectively $\frac{1}{2}(\gamma - \lambda + 1)^{-1}(1 - \theta)(\gamma - \lambda)(\gamma + 1) > 0$ and $\frac{1}{2}(\gamma - \lambda + 1)^{-1}(\gamma + 1)(1 - \theta) > 0$. Similarly, for $\lambda \leq \gamma/2$, the slacks are $\frac{1}{2}(\gamma - 2\lambda + 1)^{-1}(\lambda + 1)^{-1}(1 - \theta)(\gamma - \lambda)(\gamma + 1) > 0$ and $\frac{1}{2}(\gamma + 1 - 2\lambda)^{-1}(\lambda + 1)^{-1}(\gamma + 1)(1 - \theta) > 0$.