Mediation and Peace*

Johannes Hörner[†] Massimo Morelli[‡] Francesco Squintani[§]

September 2010

Abstract

This paper brings mechanism design to conflict resolution. We determine when and how unmediated communication and mediation reduce the ex ante probability of conflict in a game of conflict with asymmetric information. Mediation improves upon unmediated communication when the intensity of conflict is high, or when asymmetric information is significant. 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. Arbitrators who can enforce settlements are no more effective than mediators who only make non-binding recommendations.

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

JEL classification numbers: C7

*We thank Sandeep Baliga, Alberto Bisin, Yeon-Koo Che, Matt Jackson, Philippe Jehiel, Adam Meirowitz, Clara Ponsati, Kris Ramsay Tomas Sjöström, and the audiences of the ESSET 2009, SAET 2009, conflict and mechanism design conferences at Princeton, Warwick, Madrid and UCL, and regular seminars at Columbia, Essex, Bristol, Siena and EUI. This paper also benefited from various conversations with Gianluca Esposito, Principal Administrator at the Council of Europe, and currently Consulting Counsel at the International Monetary Fund (IMF) (the views expressed by Gianluca Esposito do not necessarily reflect the views of the Council of Europe, of the IMF or of any of their respective bodies).

[†]Yale University, Department of Economics.

[‡]Columbia University, Department of Political Science and Department of Economics, and and European University Institute, Department of Economics.

[§]Essex University, 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] Yet the normative analysis about which institutions or mechanisms we should use to reduce the possibility of conflict has not benefited from many interactions across the two disciplines so far. 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] By studying optimal mechanisms, we can draw important lessons about which institution designers should consider most effective in different situations, and this seems eminently relevant if we want to think of flexible institutions to help for the reduction or elimination of costly conflicts. Our interest lies more particularly with mediation. This institution, under the conditions set up in this paper, is by construction one of the optimal mechanisms. Further, mediation has played an increasingly important role in international crisis resolution. According to the International Crisis Behavior (ICB) project, the most comprehensive empirical effort to date, 30% of international crises for the entire period 1918–2001 were mediated, and the fraction rises to 46% for the period 1990–2001 (see Wilkenfeld et al., 2005).

In this paper, we select one of the most studied sources of conflict, namely the presence of asymmetric information, and we examine what institutional mechanisms are most effective in minimizing the probability of war.[3] 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 investigate when and how unmediated cheap talk between the

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

[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]Blainey (1988) famously argued that wars begin when states disagree about their relative power and end when they agree again (see also, Brito and Intriligator, 1985, and Fearon, 1995). Wars may arise because of asymmetric information about military strength, but also about the value of outside options or about the contestants' political resolve, i.e. about the capability of the leaders and the peoples to sustain war. For example, it is known that Saddam Hussein grossly under-estimated the US administration political resolve, when invading Kuwait in 1990.

disputants can reduce the *ex ante* probability of war, relative to the benchmark without communication.[4] We then turn to the main topic of this paper, mediation: 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?[5] We conclude by evaluating the benefits of enforcement power in mediation. That is, we ask: How do arbitration and mediation differ in their capabilities to prevent conflict?

We assume that the mediator has no private information and is unbiased.[6] Further, the mediator is assumed fully committed to minimize the probability of war. Hence, our mediator must be willing to commit to deadlines and to break off talks when they come to a standstill, instead of seeking an agreement in all circumstances (see Watkins, 1998). Such commitments, in fact, facilitate information disclosure by the contestants, and ultimately improve the chances of peaceful conflict resolution.[7] Finally, we study mediators who have no independent budget for transfers or subsidies, and cannot impose peace to the contestants. To be sure, third-party states that mediate conflict, such as the United States, are neither unbiased nor powerless, but single states account for less than a third of the mediators in mediated conflicts (Wilkenfeld, 2005), so that we view our assumption not only as a useful theoretical benchmark, but also as a reasonable approximation for numerous instances of mediated crises.

Unlike most of the mechanism design literature, throughout most of the paper, we assume that the mediator's proposals must be self-enforcing.[8] Indeed, countries are sovereign,

[4] 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).

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

[6] 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 scholars mentioned in Kleiboer, 1996). On the other hand, when a mediator possesses independent information that needs to be credibly transmitted, some degree of bias may be optimal (see Kydd, 2003, and Rauchos, 2006).

[7] In the final discussion, we provide some anecdotal evidence supporting our assumption of full mediator's commitment, which is obviously also one of our normative prescriptions.

[8] Among the few papers studying self-enforcing mechanisms in mechanism design, see Matthews and Postlewaite (1989), Banks and Calvert (1992), Cramton and Palfrey (1995), Forges (1999), Compte and Jehiel (2008), and Goltsman et al. (2009).

and enforcement of contracts or agreements is often impossible.[9] In the terminology of Fisher (1995), our main focus is on "pure mediation," that is, on mediation only involving information gathering and settlement proposal making, rather than "power mediation," which instead also involves mediator's power to reward, punish or enforce. The assumption that the mediator's proposals 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's recommendation itself. Since war can be started unilaterally, this *ex post* individual rationality constraint is indispensable. But in order to describe the difference between arbitration and mediation, in the final section, we relax *ex post* individual rationality and introduce standard *ex interim* individual rationality constraints.

To achieve her objective, the mediators studied in this paper can facilitate communication, formulate proposals, and manipulate the information transmitted (see Touval and Zartman, 1985, for a discussion of these three roles; and Wall and Lynn, 1993, for an exhaustive discussion of all observed mediation techniques). Because we consider unmediated cheap talk as a benchmark, our mediators can only improve the chances of peace by controlling the flow of information between the parties.[10] In practice, this corresponds to the mediator's role in "collecting and judiciously communicating select confidential material" (Raiffa, 1982, 108–109). 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 ef-

[9]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.

[10]In the real world, mediators also often prevent conflict by facilitating communication or coordinating discussions among parties unwilling to communicate without a mediator. Such instances of mediation correspond to what we formally call unmediated communication. Our paper confirms the value of mediators as communication facilitators, by showing that communication often reduces the chance of conflict.

forts 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).

Having clarified our methodological choices and our general motivation, let us now describe the 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. Indeed, Bercovitch et al (1991) show that mediation is useful mostly when the disputes are about resources, territory, or in any case divisible issues. 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.

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.[11] 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 serves as a benchmark to assess the value of communication and mediation.

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

[11]It might be interesting to allow for different costs for symmetric and asymmetric wars, but the additional notational and computational costs appear a heavy price to pay.

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 avoided altogether in 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, and then determine when and how mediation improves upon it. Among others, communication enables players to coordinate their play, and this particular role can be abstracted away by assuming a public randomization device. More importantly, communication enables the players to convey their private information, if they wish. Specifically, we study the following communication protocol. First, players send a cheap-talk message to each other; second, for any pair of observed messages, they coordinate either on war or on a peaceful cake division, depending on the realization of a public randomization device. In equilibrium, it must be the case that players do not want to unilaterally declare war whenever they are supposed to coordinate on a peaceful cake division. When war cannot be avoided in the basic conflict game, the optimal 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.[12]

We then consider mediated communication. First, the mediator collects the players' messages privately. Then, she 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 unmedi-

[12] In a small parameter region, in which the cost of war is high, the players can improve on the separating equilibrium, by playing a mixed strategy equilibrium in the cheap talk game. Of course, mixed strategy equilibria are strictly dominated by mediation, as mixed strategies induce randomizations independent across players, instead of the optimally correlated randomizations chosen by the mediator (see, e.g. Aumann, 1974).

ated 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.[13] Our main results are as follows.

- *When does a mediator help*? A mediator helps in two distinct sets of circumstances. First, when the intensity (or cost) of conflict is high. Second, when the intensity of conflict is low, but the uncertainty regarding the disputants' strength is high. Interestingly, the intensity of conflict and asymmetric information are considered among the most important variables that affect when mediation is 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. Rauchhaus (2006) provides quantitative analysis showing that mediation is especially effective when it targets asymmetric information.

- *How does the mediator help?* 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. In the first case in which she is effective, i.e. when conflict intensity is high, the mediator can improve upon unmediated communication by offering unequal splits even when she observes both players reporting to be doves. This is equivalent to an *obfuscation* strategy by which the mediator does not reveal with probability one to a self-declared dove that she is facing a dove. With this strategy, the mediator reduces the incentive for hawks to *hide strength* and then wage war if revealed that the opponent is weak. In the second case, when the intensity is low but uncertainty

[13] The model by Banks and Calvert (1992) can be related to our construction. They also compare the solution of self-enforcing mediation, to what can be achieved without a mediator in an underlying two-by-two game. But their underlying game is very different from our game of conflict: They consider a coordination game with incomplete information. This makes their results difficult to relate to our results on mediation and conflict.

is high, the mediator's strategy involves proposing equal split settlements even when she receives different messages. Equivalently, the mediator does not always reveal to a self-declared hawk that she is facing a dove, and hence reduces the incentives for doves to *exaggerate strength* in order to achieve a favorable settlement when it is revealed that the opponent is weak. In both cases, the mediator's proposed settlements do not precisely reveal each player's report to her 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 precisely all the information transmitted would not act optimally. Specifically (and realistically), the mediator's optimal obfuscation strategy consists in not revealing with probability one that the opponent is weak, when this is the case.

- *Does enforcement power help?* We finally conclude the analysis by showing that an arbitrator who can enforce outcomes is no more effective in preventing conflict than a mediator who can only propose self-enforcing agreements. This result stands in stark contrast with well-known results in other environments (see, e.g., Cramton and Palfrey, 1995, Compte and Jehiel, 2008, or Goltsman et al., 2009). Our results confirms the view that a mediator does not necessarily 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. Section 4 characterizes optimal mediation, and establishes *when and how* mediation strictly improves upon unmediated communication. Section 5 compares mediation and arbitration. The final section offers some concluding comments. In particular, it discusses interim mechanism selection, mediator's commitment and contestants' renegotiation. All proofs are in appendices.

2 The Game of Conflict

Consider a bilateral conflict problem, in which two parties want as much as possible of a given cake.[14] As is standard, we normalize the value of such an exogenous cake to 1. If the two parties cannot agree to any peaceful sharing 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.[15] 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. This private characteristic can be thought of as related to resolve, military strength, leaders' stubbornness, outside options, etc. When the two fighting players are of the same type, 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).[16]

War can be initiated unilaterally, while "it takes two to tango," i.e., a peaceful agreement must be preferred by both players to war in order to work. More precisely, we can think that for any proposed split $(x, 1-x)$, $(x \in [0,1])$, there is a "war declaration game." In such a 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 to accept a peaceful split there are ways to implement such a split.[17] There always exists an equilibrium in which both players declare war in this game,

[14] 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.

[15] 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.

[16] To simplify the analysis, and keep the problem's dimensionality in check, we adopt a fully symmetric model. We believe that our results will hold approximately, for models that are close to symmetric.

[17] 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. 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

regardless of the split. In what follows, we focus on the equilibrium that maximizes the *ex ante* probability of peace, which will be denoted by V, the value.

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

The model has three parameters: θ, p, and q. Yet all results depend on only two statistics:[18]

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

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.

As a benchmark, we now calculate the splits $(x, 1-x)$ and equilibria in the consequent war-declaration game that maximize the *ex ante* probability of peace. First, note that for $q\theta/2 + (1-q)\,p\theta \geq 1/2$, or $\lambda \geq \gamma$, both doves and hawks choose peace in the peace-maximizing equilibrium of the game with $x = 1/2$. When $\lambda < \gamma$, the probability of peace V is maximized by setting x so that all doves play peace, together with the hawk type of one of the two players. This is achieved by setting $x \geq (1-q)\,p\theta + q\theta/2$ and $1 - x \geq (1-q)\,\theta/2 + q\,(1-p)\,\theta$, which is possible if and only if $(1-q)\,p\theta + q\,(1-p)\,\theta + \theta/2 \leq 1$, i.e., $\lambda \geq (\gamma - 1)\,/\,(\gamma + 3)$, which is always satisfied when $\gamma \leq 1$. When this condition fails, the probability of peace is maximized by setting $x = 1/2$ so that doves play peace, and hawks declare war. In sum, the optimal probability of peace absent communication or mediation is:

$$V = \begin{cases} (1-q)^2 = \frac{1}{(\lambda+1)^2} & \text{if } \lambda < \frac{\gamma-1}{\gamma+3}, \\ 1 - q = \frac{1}{\lambda+1} & \text{if } \frac{\gamma-1}{\gamma+3} \leq \lambda < \gamma, \\ 1 & \text{if } \lambda \geq \gamma. \end{cases}$$

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 and Sonin (2008).

[18] This feature will allow us to give graphical illustrations of most results.

as is displayed on Figure 1.

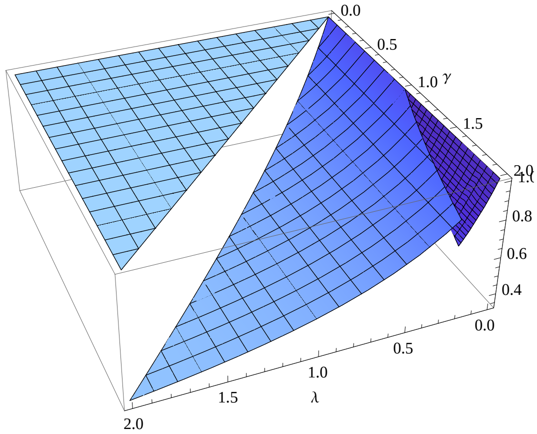


Figure 1: Probability of peace without communication

3 Communication Without Mediation

Communication Game We now consider the value of unmediated communication in our basic game of conflict. We augment our game to include communication prior to the war declaration game. Specifically, we consider the following communication protocol. After privately learning her type, each player i sends a message $m_i \in \{l, h\}$. The two messages are sent simultaneously. After observing each other's message, the players play a war-declaration game in which the split x may depend on the messages $m = (m_1, m_2)$. Their strategy in the war declaration game may depend also on the realization of a public randomization device. With probability p, the randomization device coordinates the players on both playing peace in the war declaration game, and with complementary probability, war takes place. The peace probability p may depend on the public message m.

Of course, in equilibrium, the players must be willing to follow the recommendation of the public randomization device in the war declaration game with splits $x(m)$ and obey the (possibly mixed) communication actions. We will determine the equilibrium with the smallest *ex ante* probability of war. That is, we calculate the optimal values of the split $x(\cdot)$ and the probabilities $p(\cdot)$ subject to the constraints that players are willing to use the equilibrium communication actions, and to follow the recommendations of the randomization device.

Before proceeding with the analysis, we briefly comment on the characteristics of our communication protocol. Relative to the benchmark without communication, we have introduced one round of binary cheap talk, and a public randomization device coordinating the play in the final war-declaration game. Following Aumann and Hart (2003), such a public randomization device can be replicated by an additional round of communication (using so-called jointly controlled lotteries). Hence our game can be reformulated as a two-round communication game without any extraneous randomization device. For the sake of tractability, we do not consider the possibility of further rounds of cheap talk.[19]

The restriction to binary messages is natural given the binary type space. As long as attention is restricted to pure-strategy equilibria, this restriction is without loss of generality, but given the lack of commitment, it is possible that more messages might help in mixed-strategy equilibria (though numerical simulations suggest it does not). On the other hand, the restriction to a fixed split $x(m)$, for every m, rather than the consideration of a lottery over splits, is without loss of generality.[20] Finally, we point out that, as described above, the game form encapsulates one shortcut. We have assumed that, with probability $1 - p(m)$, war takes place without the contestants being called to play the war declaration

[19] This might help, however. Aumann and Hart (2003) provide examples of games in which longer, indeed unbounded, communication protocols improve upon finite round communication.

[20] Note first that if some splits in a lottery induce war on the equilibrium path, i.e., after both players choose the equilibrium strategy at the communication stage, then such splits can be replaced with no loss by a war recommendation. After this change, we can replace without loss any lottery over peaceful recommendations with its certainty equivalent: A deterministic recommendation equal to the expected recommendation of the lottery, assigned with probability equal to the lottery's aggregate probability of peaceful recommendation. In fact, at the war declaration stage, the requirement that the players accept such a deterministic average split is less stringent than the requirement that they accept all splits in the support of the lottery. Further, lotteries over peaceful splits affect each player's equilibrium payoff at the communication stage only through their expectations. Finally, the payoff of a player who deviates at the communication stage turns out to be convex in the recommended split. Hence, the deviation payoff is lower when replacing a lottery with its certainty equivalent, thus making the equilibrium requirement less stringent. The reason for the convexity of the deviation payoff is that, off the equilibrium path, the player optimally chooses whether to declare war or accept the split conditionally on the split realized with the lottery. Hence, the expected value of the lottery off the equilibrium path is the maximum between the realized split and the war payoff, a convex function.

game. Because war can be started unilaterally, both players declaring war is always an equilibrium of the war declaration game.[21]

Pure-strategy Equilibria. We momentarily ignore mixed strategies by the players at the message stage. Those will be considered in the next subsection. Evidently, there is always a pooling equilibrium in which both types choose the same reporting strategy, whose outcomes coincide with the equilibrium of the war-declaration game without communication. We now consider separating equilibrium, i.e., equilibrium in which each player truthfully reveals her type.

Let us consider here only equilibria with splits $x(m)$ and probabilities $p(m)$ that are symmetric across players. Such symmetry restriction entails that $x(h,h) = x(l,l) = 1/2$, and that we only need to find another split value, i.e., $b \equiv x(h,l) = 1 - x(l,h)$, given that the message space contains only two elements. We shall later see that this restriction is without loss of generality, because the separating equilibrium which minimizes the *ex ante* probability of peace is calculated by solving a linear program.[22] To shorten notation further, we let $p_L \equiv p(l,l), p_H \equiv p(h,h)$ and $p_M \equiv p(h,l) = p(l,h)$.

Armed with these definitions, the optimal separating equilibrium is characterized by the following program. Maximize the peace probability

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

subject to the following *ex post* individual rationality (IR) constraints and *ex interim* incentive compatibility constraints (IC_L^*, IC_H^*).[23] First, reporting truthfully must be optimal.

[21]While for some splits x, such an equilibrium may be weakly dominated, we could expand the war declaration game by including a small first-strike advantage, and make the war equilibrium always strict.

[22]We will see that each player's constraints are linear in the maximization arguments. Thus, the constraint set is convex. Hence, suppose that an asymmetric mechanism maximizes the probability of peace. Because the set-up is symmetric across players, the anti-symmetric mechanism, obtained by interchanging the players' identities, is also optimal. But then, the constraint set being convex, it contains also the symmetric mechanism obtained by mixing the above optimal mechanisms. As the objective function is linear in the maximization argument, such symmetric mixed mechanism is also optimal.

[23]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*, even when players are supposed to coordinate on peace. This explains the maxima on the right-hand side of the two constraints.

For the dove, this constraint (IC_L^*) states that

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

The left-hand side is the dove's equilibrium payoff. With probability $1-q$, the opponent is also a dove, in which case the equal split $1/2$ occurs with probability p_L and the payoff from war, $\theta/2$, is collected with probability $(1-p_L)$. With probability q, the opponent is hawk. With probability p_M, this leads to the split $1-b$, and with probability $1-p_M$ to the payoff from war $(1-p)\,\theta$. The right-hand side is the expected payoff from exaggerating strength. When the opponent is dove, the split b is recommended with probability p_M. In principle, the player may deviate from the recommendation, and collect the war payoff $\theta/2$, hence the payoff is $\max\{b,\theta/2\}$. Further, war is recommended with probability $1-p_M$. When the opponent is hawk, the split $1/2$ is recommended with probability p_H, and war with probability $1-p_H$. Similarly, for the hawk, the constraint (IC_H^*)

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

must hold, where the left-hand side is the equilibrium payoff and the right-hand side is the expected payoff from "hiding strength."[24]

Second, players must find it optimal to accept the split. Given that, in a separating equilibrium, messages reveal types, this requires that

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

That is, a hawk facing a self-proclaimed dove must get a share b that makes war unprofitable against a dove. Similarly, the dove's share against a hawk cannot be so low that it is better

[24]Although the constraints (IC_L^*) and (IC_H^*) are not linear because of the maxima and of the products p_Mb, they can be turned into linear constraints as follows. First, one replaces each constraint with four constraints in which the left-hand sides equal the left-hand side of the original constraint with one of the four pairs of the arguments of the two maxima, in lieu of the maxima. Second, one changes the variable b with $p_B=p_Mb$ and the constraint $1/2\leq b\leq 1$ with $p_B\leq p_M\leq 2p_B$.

for her to go to war. The constraint that a player would accept an equal split when the opponent's type is the same as her own, $1/2 > \theta/2$, is always satisfied.

Solving this program yields the following characterization. We here omit the precise equilibrium formula, presented in the Appendix, as it is quite burdensome.

Proposition 1 *There is a unique best separating equilibrium in the communication game without mediation. This equilibrium displays the following characteristics, for $\lambda < \gamma$:*

- *The ex ante probability of peace is strictly greater than in the absence of communication.*
- *Dove dyads do not fight: $p_L = 1$.*
- *Hawk dyads fight with positive probability, $p_H < 1$, and the dove's incentive compatibility constraint IC_L^* binds.*
- *If $\gamma \geq 1$ and/or $\lambda \geq (1+\gamma)^{-1}$, then the hawk's incentive compatibility constraint IC_H^* does not bind and $b = p\theta$; and further:*
  - *if $\lambda < \gamma/2$, then hawk dyads fight with probability one, $p_H = 0$, and asymmetric dyads fight with positive probability, $p_M \in (0,1)$;*
  - *if $\lambda \geq \gamma/2$ (which covers also the case $\lambda \geq (1+\gamma)^{-1}$), then hawk dyads fight with positive probability, $p_H \in (0,1)$, and asymmetric dyads do not fight, $p_M = 0$.*
- *If $\gamma < 1$ and $\lambda < (1+\gamma)^{-1}$, then IC_H^* binds and $b > p\theta$; and further $p_H = 0$ and $p_M \in (0,1)$ for $\lambda < \gamma/(1+\gamma)$, whereas $p_H \in (0,1)$ and $p_M = 1$ otherwise.*

We now elaborate on the characterization described above.

First, the separating equilibrium always improves upon the war declaration game without communication. While intuitive, this result is far from obvious: While at least one equilibrium of the communication game must be at least as good as the optimal war declaration game equilibrium, it is not an obvious implication that the separating equilibrium would strictly improve upon all equilibria without communication. Second, war is never optimal when both players report low strength: $p_L = 1$; intuitively, there is no need to punish self-reported doves by means of war, as they receive lower splits on average than if reporting to be hawks.

Third, the truth-telling constraint for the low type, IC_L^*, is always binding. Given that the incentive to exaggerate strength is always present and must 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 $p_M = 1$. 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 is sufficiently high that prescribing war against a dove is not needed to deter a dove to exaggerate strength. But when λ is low, deterring misreporting by a dove requires having a positive probability of war when exactly one of the players claims to be a hawk, in addition to having war for sure when both claim to be.

Fourth, when the truth-telling constraint for the high type, IC_H^*, is not binding, then $b = p\theta$; and when both truth-telling constraints are binding, then 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 *ex interim* truth-telling constraints. Intuitively, both (IC_H^*) and the constraint $b \geq p\theta$ need b sufficiently large to be satisfied. On the other hand, keeping in check the (binding) constraint (IC_L^*) requires keeping b as low as possible. Hence b will be such that either IC_H^* binds, or $b = p\theta$.

The other properties of the characterization of Proposition 1 are best described by distinguishing the cases $\gamma \geq 1$ 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 *ex 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 doves). When $b = p\theta$, 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 $\theta/2$ 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 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 $1 - p_H$. 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 both the payoff from settling against a hawk when hiding strength and 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, we can reduce b as λ increases. 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 2 shows the probability of peace in the best separating equilibrium. For $\gamma \geq 1$, 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.

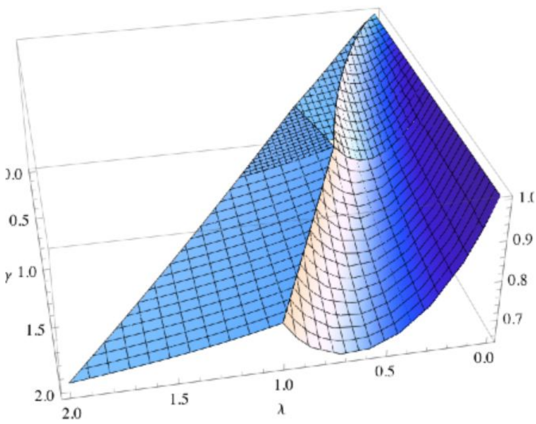


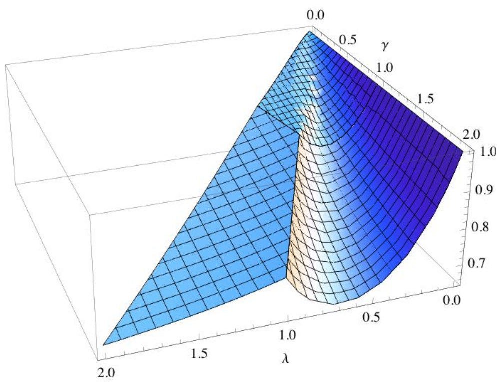
Figure 2: Probability of peace in the separating equilibrium

Mixed-strategy Equilibria. This subsection considers mixed strategy equilibria. Mixing can help, though its role in the unmediated communication game is relatively limited (compare Figure 2 and right panel of Figure 3). The following result states that, while there is no mixed-strategy equilibrium in which the hawk randomizes between sending the high and low report, there exists a mixed-strategy equilibrium in which the dove randomizes.[25] Furthermore, in some parameter region, depicted in the right panel of Figure 3, such a mixed strategy equilibrium yields a higher *ex ante* peace probability than the separating

[25] In this subsection, and this subsection only, attention is restricted to symmetric equilibria. That is, we did not establish whether asymmetric mixed-strategy equilibria may yield a higher welfare.

equilibrium. The specific definition of the region in which mixing improves upon the separating equilibrium is rather cumbersome, as is the explicit description of the mixed-strategy equilibrium, and so it is relegated to the Appendix. But it is interesting to note that mixing may improve only in a small subset of the parameter region in which both *ex interim* IC* constraints bind: mixing by the dove may relax the incentive of the hawk to hide strength. We summarize our findings as follows.

Proposition 2 *Allowing for mixed strategies in the unmediated communication game, the best equilibrium is such that the hawk always sends message h and the dove sends l with probability σ, where $\sigma < 1$ in the parameter region shown in the right panel of Figure 3.*



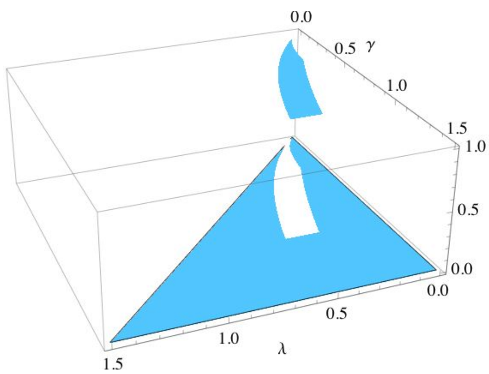


Figure 3: Welfare in the best (pure or mixed) equilibrium, and region where mixing occurs

4 Mediation

In the previous section, we have characterized the optimal equilibrium in the case in which players send public messages. In this section, we consider an active mediator who collects the players' private messages and makes optimal recommendations.

We modify the game form to account for such a mediator. In the first stage, messages are no longer public. They are separately reported to a mediator, who then proposes the split and randomly correlates the play in the consequent war declaration game. More precisely, the version of the revelation principle proved in Myerson (1982) guarantees that the following game form entails no loss of generality:

- After being informed of her type, each player i privately sends a report $m_i \in \{l, h\}$ to the mediator.

- Given reports $m = (m_1, m_2)$, the mediator recommends a split $(b, 1-b)$ according to some cumulative distribution function $F(b|m)$, where we denote by $1 - F(1|m)$ the probability with which the mediator recommends war.[26] Unlike the reports, the mediator's recommendation is public.

- War takes place if recommended. Otherwise, the contestants play the war declaration game with the recommended split.

Again by the revelation principle, we may restrict attention to distributions F such that truthful type revelation and obedience to the mediator's recommendation are part of the equilibrium. As before, this imposes both *ex interim* incentive compatibility constraints and *ex post* individual rationality constraints, which we now describe. To simplify notation, we restrict attention to mechanisms that are symmetric across players, where $F(\cdot|m_1, m_2) = 1 - F(\cdot|m_2, m_1)$ for all (m_1, m_2), and to discrete distributions F. We shall later see that these restrictions entail no loss of generality.

Let $\Pr[m_{-i}, b, m_i]$ denote the equilibrium joint probability that the players send messages (m_i, m_{-i}) and that the mediator offers $(b, 1-b)$, and set $\Pr[b, m_i] \equiv \Pr[h, b, m_i] + \Pr[l, b, m_i]$. When player i is a hawk, she reports $m_i = h$ in equilibrium, and *ex post* individual rationality requires that

$$b \Pr[b, h] \geq \Pr[l, b, h] p\theta + \Pr[h, b, h]\theta/2, \text{ for all } b \in [0, 1], \tag{1}$$

which ensures that, if recommended the split b, i.e., for all b such that $\Pr[b, h] > 0$, the hawk prefers accepting the split to starting a war. Similarly, when i is a dove, *ex post*

[26]The term "war recommendation" is part of the game theoretic jargon we use to describe the formal model. It should not be taken literally. In the real world, mediators do not literally recommend war, but they may walk away from the table, and this usually results in conflict escalation by the contestants.

individual rationality dictates that

$$b\Pr[b,l] \geq \Pr[h,b,l](1-p)\theta + \Pr[l,b,l]\theta/2, \text{ for all } b \in [0,1]\,. \tag{2}$$

Ex interim incentive compatibility requires that, when player i is a hawk, she truthfully reports $m_i = h$. The associated constraint (IC_H^*) dictates that

$$q(1-F(1|hh))\theta/2 + (1-q)(1-F(1|hl))p\theta + \int_0^1 bdF(b|h) \geq q(1-F(1|lh))\theta/2 +$$
$$(1-q)(1-F(1|ll))p\theta + \int_0^1 \max\{b, \Pr[l|b,l]p\theta + \Pr[h|b,l]\theta/2\}dF(b|l), \tag{3}$$

where $\Pr[m_{-i}|b,m_i] = \Pr[m_{-i},b,m_i]/\Pr[b,m_i]$ whenever $\Pr[b,m_i] > 0$, and $F(\cdot|m_i) \equiv qF(\cdot|m_i,h) + (1-q)F(\cdot|m_i,l)$, for m_i and m_{-i} taking values l and h. Note that, as in the optimal separating equilibrium program, player i might behave opportunistically after deviating, as reflected by the maxima on the right-hand side.

Similarly, to ensure truth-telling by player i when a dove, the following constraint (IC_L^*) must be satisfied:

$$q(1-F(1|lh))(1-p)\theta + (1-q)(1-F(1|ll))\theta/2 + \int_0^1 (1-b)dF(b|l) \geq q(1-F(1|hh))(1-p)\theta +$$
$$(1-q)(1-F(1|lh))\theta/2 + \int_0^1 \max\{1-b, \Pr[l|b,h]\theta/2 + \Pr[h|b,h](1-p)\theta\}dF(b|h). \tag{4}$$

In the best equilibrium, the mediator seeks to minimize the probability of war, i.e.,

$$(1-q)^2(1-F(1|hh)) + 2q(1-q)(1-F(1|lh)) + q^2(1-F(1|ll)).$$

Because recommendations need to be self-enforcing, there is *a priori* no reason to restrict the mediator in the number of splits to which he assigns positive probability. In fact, recommendations convey information about the most likely opponents' revealed types, and it might be in the mediator's best interest to scramble such information by means of multiple recommendations. Nevertheless, Proposition 3 below shows that relatively simple mechanisms reach the maximal probability of peace among all possible mechanisms, including asymmetric ones. These simple mechanisms can be described as follows. Given reports (h,h), the mediator recommends the peaceful split $(1/2, 1/2)$ with probability q_H, and war with probability $1 - q_H$. Given reports (h, l), the mediator recommends the peaceful split $(1/2, 1/2)$ with probability q_M, the split $(b, 1-b)$ with probability p_M, and war with probability $1 - p_M - q_M$, for some $b \geq 1/2$. Given reports (l, l), the mediator recommends the peaceful split $(1/2, 1/2)$ with probability q_L, the splits $(b, 1-b)$ and $(1-b, b)$ with probability p_L each, and war with probability $1 - 2p_L - q_L$.

Again, we relegate the explicit formulas of the solution to the Appendix, and restrict ourselves here to the description of its main features.

Proposition 3 *A solution to the mediator's problem is such that, for all $\lambda < \gamma$:*

- *Doves do not fight: $q_L + 2p_L = 1$.*
- *The low-type incentive compatibility constraint IC_L^* binds, whereas the high-type incentive compatibility constraint IC_H^* does not, and $b = p\theta$.*
- *For $\gamma \geq 1$ and $\lambda > \gamma/2$, hawk dyads fight with positive probability, $q_H \in (0, 1)$, mixed dyads do not fight ($p_M + 2q_M = 1$), and mediation strictly improves upon cheap talk.*
- *For $\gamma \geq 1$ and $\lambda \leq \gamma/2$, the solution exactly reproduces the separating equilibrium of the cheap talk game (specifically, $q_L = 1$, $q_M = 0$, $p_M \in (0, 1)$ and $q_H = 0$), and mediation yields the same welfare as cheap talk.*
- *For $\gamma < 1$, the probability p_L of unequal splits among dove dyads is bounded above zero, and mediation strictly improves upon cheap talk.*

We now comment on the solution and we make some comparisons with the optimal separating equilibrium characterized in Proposition 1.

Suppose that $\gamma > 1$. If $\lambda > \gamma/2$, then $q_M > 0$: 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 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 exaggerating strength less profitable against doves. When $\lambda \leq \gamma/2$ instead, $q_H = q_M = 0$ and the mediator does not improve upon unmediated communication. In this case, in fact, in both the mediated and the best (unmediated) separating equilibrium, 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 program, 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 IC_H^* constraint is satisfied with $b = p\theta$. In fact, the mediator offers $1-b$ 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 $1-b$. This eliminates the incentive to hide strength in order to seek peace against hawks that we observed in the unmediated equilibrium. Hence, the expected payoff of hiding strength is lower, and the IC_H^* constraint is satisfied with $b = p\theta$. Note that the *ex post* individual rationality constraint $b \geq p\theta$ was slack in the unmediated equilibrium. By making this rationality constraint binding, the mediator can improve the objective function, i.e. increase the probability of peace.

Figure 4 shows the probability of peace in the mediated game compared to the probability of peace induced by the best separating equilibrium.

We can now precisely answer the first set of questions presented in the introduction:

- When does mediation improve on unmediated communication?

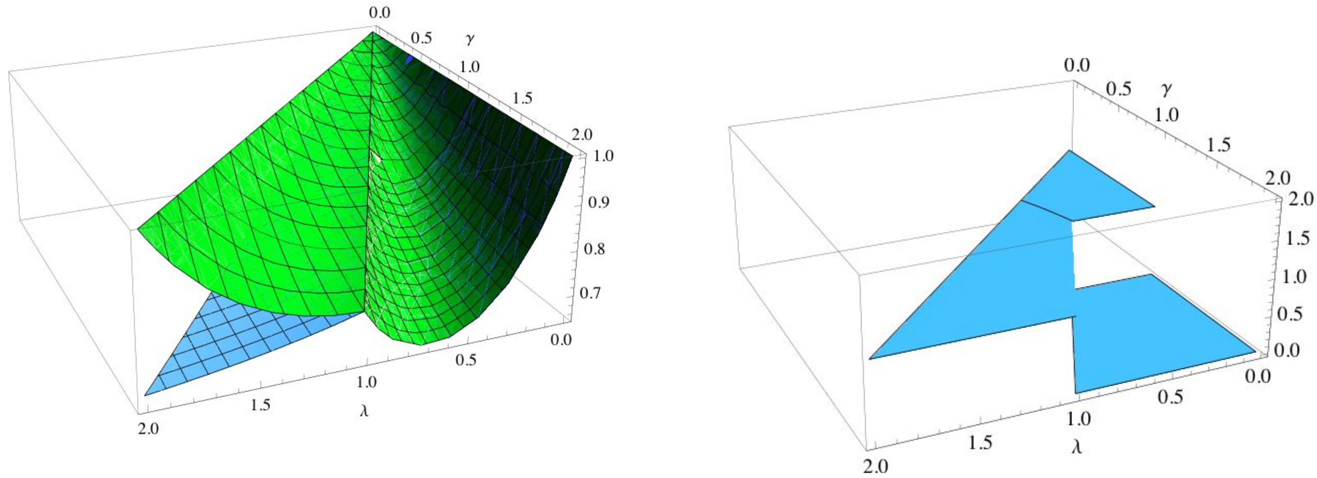


Figure 4: Probability of peace in the mediated vs. unmediated case, and region where mediation dominates

- When the intensity and/or cost of conflict is high (low γ), 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), mediation provides an improvement in welfare if and only if the proportion of hawks is intermediate, i.e., for high expected power asymmetry.

• 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 to a self-reported hawk that she is facing a dove). This lowers the incentive to exaggerate strength and achieves a favorable peace settlement with a dove.
- Instead, when the probability of facing a hawk is low and conflict is expected to be costly, the mediator's strategy is to offer with some probability unequal split to two parties reporting low type (or, equivalently the mediator does not

always reveal to a dove that she is facing a dove). This lowers the incentive to hide strength and seek peace with a hawk.

Mediation also improves upon the best mixed-strategy equilibrium (whenever the latter is non-degenerate, which requires, among others, $\gamma < 1$, i.e., a high cost of conflict). How this can be the case should not be surprising for the reader familiar with the literature on correlated equilibrium (see, e.g. Aumann, 1974). 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. The mixing by the dove may improve welfare upon the pure-strategy equilibrium, but at the cost of inducing war with positive probability within dove dyads. This does not occur with a mediator, who induces war only when at least one of the players is a hawk.

To conclude, note one sharp difference between mediated and unmediated communication: the *ex ante* probability of peace is *decreasing* in λ, for $\lambda \in [\gamma/2, \gamma)$, without mediation, whereas in the same range the *ex ante* probability of peace is *increasing* in λ with mediation. This difference could have an important impact on an important debate in international relations, namely the debate on "deterrence:" the higher is λ (or q), the more "deterred" a country should be from initiating a war, due to a higher likelihood of facing a hawk. Hence, a possible interpretation of this difference is that an international system with a level of deterrence higher than another is "good" for peace if every bilateral crisis is dealt with using mediators, whereas it is "bad" if direct communication is the most common way in which countries try to avoid wars.

5 The Role of Enforcement

Even though the cause of war is asymmetric information, the analysis of the optimal mediation problem involves a significant enforcement problem. Countries are sovereign, and enforcement of contracts or agreements is often impossible. Because war can be started unilaterally, we have incorporated *ex post* IR and *ex interim* IC* constraints in the formulation of the optimal mediation program. 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.

So far, mediation has reduced to optimal information elicitation from, and transmission to, the conflicting parties. One might also wonder whether the mediator could further reduce the *ex ante* probability of war if she was an arbitrator, i.e. if she were endowed with enforcement power. For this, it is enough to compare our findings with those of Bester and Wärneryd (2006). Rather than imposing *ex post* IR constraints and *ex interim* IC* constraints, they impose *ex interim* IR and IC constraints. Conflicting parties must be willing to participate in 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. Hence, they abstract away from enforcement, and their model is more suitable to describe arbitration than the pure mediation that considered here.

Formally, invoking the version of the revelation principle proved by Myerson (1979), the Bester-Wärneryd problem can be summarized as follows. The parties truthfully report their types L, H to the arbitrator. The arbitrator recommends peaceful settlement with probability $p(m)$ after report m. Because recommendations are enforced by an external agency, they can restrict attention to a single peaceful recommendation $x(m)$, for each report pair m.[27] Symmetry is without loss of generality because the arbitrator's program is linear, and entails that the settlement is $(1/2, 1/2)$ if the players report the same type, that the split is $(b, 1-b)$ if the reports are (h, l), and $(1-b, b)$ if they are (l, h), for some $b \in [1/2, 1]$. Let $p_L = p(l, l)$, $p_M = p(l, h) = p(h, l)$ and $p_H = p(h, h)$. The arbitrator

[27]In fact, both participation and revelation decisions are taken *before* knowing the arbitrator's recommendation, and hence the players' payoffs depend only on the expected recommendation, and not on the realized one. Hence, as in footnote 20, any lottery over peaceful recommendations can be replaced without loss with its certainty equivalent.

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

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

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

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

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

In general, the solution of the program with an arbitrator (with enforcement power) provides an upper bound to the solution of the program with a mediator (without enforcement power), as described in Section 4. Surprisingly, the solution of the latter program yields the same welfare as the solution of the former program.[28] Specifically, for $\lambda \leq \gamma/2$, the mechanisms with and without enforcement coincide. When $\lambda > \gamma/2$, the simplest optimal mechanism with enforcement is such that $b < p\theta$, which is not self-enforcing. But the optimal mechanism without enforcement obfuscates the players' reports, and this obfuscation succeeds in fully circumventing the enforcement problem.

Proposition 4 *An arbitrator who can enforce recommendation is exactly as effective in promoting peace as a mediator who can only propose self-enforcing agreements.*

[28]This result facilitates the proof of Proposition 3. It is enough to establish that the simple mechanism characterized there, and described in closed form in the Appendix, satisfies the more stringent constraints of the mediator's program. Because this mechanism achieves the same welfare as the solution to the arbitration problem, it must be optimal, *a fortiori*, in the mediator's program.

The intuition is as follows. First, note that the dove's IC constraint and hawk's *ex 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 dove's IC* constraint and the two *ex post* hawk's IR constraints. Recall that, in our solution, the hawk is always indifferent between war and peace if recommended a settlement. Further, the dove's IC* constraint in the mediator's problem with self-enforcing recommendations is identical to the dove's IC constraint in the arbitrator's program, because a dove never wages war after exaggerating strength in the solution of mediator's problem with self-enforcing recommendation.

Further, the hawk's *ex interim* IR constraint integrates the two binding hawk's *ex post* IR constraints in the arbitrator's problem. While requiring a constraint to hold in expectation is generally a weaker requirement than having the two constraints, it turns out that the induced welfare is the same. This is easiest to see when $\lambda \leq \gamma/2$, as in this case the only settlement ever granted to a hawk is b, when the opponent is dove. For any mechanism with this property, the *ex interim* IR and the *ex post* IR constraints trivially coincide. Let us now consider the case $\lambda > \gamma/2$. In this case, the optimal truthful arbitration mechanism prescribes a settlement $b < p\theta$ that is not *ex post* IR for a hawk meeting a dove, as well as prescribing a settlement with slack, equal to $1/2$, to same type dyads. The mediator cannot reproduce this mechanism. But it circumvents the problem with the obfuscation strategy whereby the hawk is made exactly indifferent between war and peace when recommended either the split $1/2$ or the split $b = p\theta$. Hence, it optimally rebalances the *ex post* IR constraints so as to achieve the same welfare as the arbitrator.

We can now answer the last question that we posed in the introduction.

- *Does enforcement power help? How do mediation and arbitration differ in terms of conflict resolution?*
  - In our war-declaration game, there is no difference in terms of optimal *ex ante* probability of peace between the two institutions.

– Either the two optimal mechanism coincide, for λ low relative to γ, or the mediator's optimal obfuscation strategy circumvents her lack of enforcement power.

6 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 draw a number of lessons.

First of all, we have shown *when* 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 little impact on the probability of winning; and even when neither θ nor p is low, mediation can still be strictly better than direct communication when the *ex ante* chance of power asymmetry is high (intermediate q). A world with higher level of deterrence (higher λ in that intermediate range) turns out to be better for peace only if mediation is used. While most of our analysis has considered mediators without enforcement powers, it turns out that an arbitrator who can enforce outcomes is exactly as effective as a mediator who can only propose self-enforcing agreements.

Second, we have shown *how* mediation improves upon unmediated communication. This is achieved by not reporting to a player with probability one that her opponent is weak. Specifically, when the *ex ante* chance of power asymmetry is high, the mediator is mostly effective in her 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 reduces the probability of war between hawks, and hence the *ex ante* probability of war. When the expected intensity or cost of conflict are high, regardless of the expected degree of uncertainty, the mediator reduces the temptation to hide strength by a strong player. The mediator's strategy is to lower the reward from mimicking a dove by giving sometimes an unequal split to two parties reporting being a 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.

This is obviously a very limited theory of mediation. We have mentioned some of the limitations of our mediator in the introduction. Let us emphasize some that might lead to fruitful future research, while delineating circumstances in international relations under which none of such issues is likely to be a major concern.

Our analysis takes as given how the process starts, and how it ends. Because we consider the games with and without a mediator separately, our analysis did not address the incentives of each disputant to seek the assistance of a mediator in the first place, given that such a call for mediation is likely to convey information. Another related issue that we have not explored here involves the *ex ante* incentives of players to engage in strategic militarization (see, for example, Meirowitz and Sartori, 2008). We have also taken our mediator as having commitment power, and while we do not assume that disputants have commitment power, we have ignored their incentives to re-negotiate.

In order for mediation to occur, both disputants must consent to the involvement of a mediator as a third party. Individual interests, rather than “shared values” are the main driving force behind acceptance of mediation (Princen, 1992). Whether or not a party gains from accepting mediation depends on how such a decision will be perceived by the other party, and what such a party can guarantee in the absence of a mediator. These features of the problem naturally lead to multiplicity of equilibria. Suppose in fact that we were to augment our mediation game to include a stage in which, immediately after being informed of their types, the contestants simultaneously and independently choose whether to accept the mediator or resort to unmediated cheap talk. Further assume that mediation will take place if and only if both players agree. This game admits both equilibria in which mediation always takes place and some in which it never occurs.

As an example of the latter, consider an equilibrium in which all players' types reject the mediator, and such that if a player deviates and asks for mediation, the opponent will maintain her prior beliefs about her opponent. In this case, both players are always indifferent between accepting mediation or not, as this is irrelevant given the opponent's equilibrium veto. For this reason, the postulated off-path beliefs can be viewed as reasonable. An equilibrium in which the optimal mediation that we characterized takes place is the one in which both players agree to mediation, and if either deviates, players maintain their prior beliefs, babble in the communication stage, and declare war. This off-path behavior can be described as each player threatening the other "not to listen and declare war, unless a mediator is called in." Such a harsh punishment may not always be realistic, but, by construction, it delivers the highest possible welfare in the game. Deviations are deterred, because the outcome of rejecting the mediator is equivalent to triggering war, and deterrence is equivalent to an *ex interim* individual rationality condition that is weaker than the *ex post* individual rationality conditions we already imposed to solve our program.

To conclude on the matter of the choice of the mechanism by informed parties, note that scholars often put forward other motives for desiring mediation. Bercovitch (1992, 1997), for instance, argues that disputants might view mediation as an expression of their commitment to peaceful conflict resolution, and seek it out of a desire to improve their relationships with each other. Ultimately, whether and how the acceptance of a mediator affects perceptions and incentives is an empirical question. Using the ICB data, Wilkenfeld et al. (2005) argue that symmetric crises are more likely to be mediated than asymmetric ones, suggesting that mediation initiative might fail when power disparity is extreme. Note, however, that our model only considers (*ex ante*) symmetric crises, and there seems to be no empirical evidence that would correlate the occurrence (as opposed to the success) of mediation with the level of uncertainty.

Turning to the question of strategic militarization, we believe that our model can provide a simple benchmark to explore the *ex ante* incentives for countries to arm when different conflict resolution institutions prevail, and in particular when focusing on the optimal one (the mediator we study in this paper). Specifically, our results may be related with the results by a recent paper by Meirowitz and Ramsay (2009). 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 *ex interim* IC constraints (Theorem 2). Because we show that in our simple game, optimal arbitration and optimal mediation coincide, and both can be characterized via *ex interim* IC and IR constraints, it would be interesting to assess the implications of their results in the contest of the optimal crisis-resolution mechanisms that we characterize in this paper.

Turning to the issue of commitment by the mediator, we interpret the outcome of war as a failure of the disputants to agree with the conditions set by the mediator. Our analysis suggests that the mediator's success relies on its ability to employ so-called action-forcing events. The importance of using such events is stressed by Watkins (1998). Mediators are well aware of the importance of being able to break off talks with no intention of resuming them. For the technique to be effective, a deadline must be credible. According to Avi Gil, one of the key architects of the Oslo peace process, "A deadline is a great but risky tool. Great because without a deadline it's difficult to end negotiations. [The parties] tend to play more and more, because they have time. Risky because if you do not meet the deadline, either the process breaks down, or deadlines lose their meaning" (Watkins, 1998). Among the many cases in which this technique was used, see for instance Curran and Sebenius (2003)'s account of how a deadline was employed by former Senator George Mitchell in the Northern Ireland negotiations. Committing to such deadlines might be somewhat easier for professional mediators whose reputation is at stake, but they have been also used both by unofficial and official individuals, including Pope John Paul II and former U.S. President Jimmy Carter.[29] Meanwhile, institutions like the United Nations increasingly sets time limits to their involvement upfront (see, for instance, the U.N. General Assembly report, 2000).

[29]See Bebchik (2002) on how Clinton and Ross attempted to impress upon Arafat the urgency of accepting the proposal being offered for a final settlement, calling it a "damn good deal" that would not be within his grasp indefinitely.

Nevertheless, it is a fact that, no matter their moral authority, mediators may sometimes struggle to walk away when the deadline expires (as Pope John Paul II found out in the Beagle Channel dispute, for instance). This fact suggests the formulation and solution of a model of mediation that introduces limited mediator's commitment, a seemingly fairly involved theoretical task. One of the main issues is the identification of a game form and mechanism space that do not entail loss of generality. It is known, in fact, that the revelation principle may fail in the presence of limited commitment (see, for example, Bester and Strausz, 2000).

Finally, while we have required recommendations to be self-enforcing, there is a sense in which they need not be renegotiation-proof, as they might be Pareto-dominated for the players. For instance, when there is common belief that both players are hawks, they would be better off settling for an equal split rather than going to war, although doing so is part of the solution. Yet renegotiation-proofness does not seem to be a first and foremost concern of real world mediators. It is not overly realistic to think that, after the mediator quits, contestants who struggled to find an agreement in the presence of the mediator, will autonomously sit down at the negotiation table again, in search for a Pareto improving agreement. Indeed, while the literature on the causes of conflict underlines that contestants may not be able to individually commit to peaceful conflict resolutions, it may well be the case that they can jointly or even individually commit to belligerent resolutions, when such commitments are *ex ante* valuable. Audience costs, for instance, are recognized to provide an important channel that makes war threats credible (see, for instance, Tomz, 2007).

Nevertheless, renegotiation-proofness is an important issue that deserves a careful analysis. A technical difficulty lies in the fact that there is no widely accepted notion of renegotiation-proofness for games with incomplete information. The definitions that apply to our game usually yield non-existence, given the restriction imposed by *ex post* individual rationality and incentive compatibility. Forges (1990), for instance, defines an equilibrium to be renegotiation-proof if it is the case that, for every further (exogenous) proposal that players can simultaneously accept or reject after the mediator's recommendation, players would not unanimously prefer the exogenous proposal. Unfortunately, it can be shown that this requirement is impossible to satisfy in our problem.[30] Coming up with a satisfactory definition appears like an important issue for future research. Note, however, that any such definition would exacerbate the incentives for the mediator to obfuscate the information that he owns, so as to prevent the possibility that players that are supposed to go to war commonly believe that some specific, peaceful split is better for both of them.

[30] At the renegotiation stage, on the equilibrium path, the war-declaration stage can be viewed as a static Bayesian game. Let q_i denote the probability that i is a dove at this stage. If $q_1 < 1$ and $q_2 < 1$, then doves reach this stage with positive probability, and in that event, both players would prefer the exogenous proposal "peace and an equal split" to the recommendation of war. Hence there must be peace in such a subform. If $q_1 = 1$ but $q_2 < 1$, then player 2's dove and player 1 would agree on the proposal "peace and split $p\theta$;" and if $q_1 = q_2 = 1$, they would agree on the proposal "peace and equal split." Hence, the only candidate for a renegotiation-proof equilibrium involves the mediator always suggesting peace, which cannot satisfy incentive compatibility and *ex post* rationality.

References

[1] Aumann, R.J. (1974): "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics* 1, 67–96.

[2] Aumann, R.J. and S. Hart (2003): "Long Cheap Talk," *Econometrica*, **71(6)**, 1619–60.

[3] Baliga, S. and T. Sjöström (2004): "Arms Races and Negotiations," *Review of Economic Studies,* **71(2)**, 351–69.

[4] Banks, J.S. and R.L. Calvert (1992): "A Battle-of-the-Sexes Game with Incomplete Information," *Games and Economic Behavior*, **4**, 347–372.

[5] Bebchik, B. (2002): "The Philosophy and Methodology of Ambassador Dennis Ross as an International Mediator," *International Negotiation*, **7**, 115–131.

[6] Bercovitch, J. (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.

[7] Bercovitch, J. (1992): "The Structure and Diversity of Mediation in International Relations," in. J. Bercovitch and J.Z. Rubin (eds) *Mediation in International Relations: Multiple Approaches to Conflict Management*, New York: St. Martin's Press.

[8] Bercovitch, J. (1997): "Mediation in International Conflict: An Overview of Theory, a Review of Practice," in I.W. Zartmann and J.L. Rasmussen (eds) *Peacemaking in International Conflict: Methods and Techniques*, Washington D.C.: United States Institute of Peace Press.

[9] Bercovitch, J., J.T. Anagnoson, and D. Wille (1991): "Some conceptual issues and empirical trends in the study of successful mediation in international relations," *Journal of Peace Research* **28**, 7–17.

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

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

[12] Bester, H. and R. Strausz (2000): "Imperfect Commitment and the Revelation Principle: the Multi-Agent Case,"*Economics Letters* **69,** 165-171.

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

[14] Blainey, G. (1988): *The Causes of War.* Third ed. New York: The Free Press.

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

[16] Brito, D.L. and M.D. Intriligator (1985): " Conflict, War, and Redistribution, " *American Political Science Review*, **79(4)**, 943–957.

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

[18] Cramton P.C. and T.R. Palfrey (1995): "Ratifiable mechanisms: learning from disagreement," Games Economic Behavior, **10**, 255–283.

[19] Curran, D. and J.K. Sebenius (2003): "The Mediator as Coalition Builder: George Mitchell in Northern Ireland," *International Negotiation*, **8(1)**, 111–147.

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

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

[22] Fey, M. and C. Ramsay (2009): "Mechanism Design Goes to War: Peaceful Outcomes with Interdependent and Correlated Types," *Review of Economic Design*, **13(3)**, 233–250.

[23] Fisher, R.J. (1995): "Pacific, Impartial Third Party Intervention in International Conflict: a Review and Analysis." In *Beyond Confrontation: Learning Conflict Resolution in the Post Cold War Era.* Ed. Vasquez, Johnson, Jaffe and Stamato, Ann Arbor: University of Michigan Press, 39–59.

[24] Forges, F. (1986): "An approach to communication equilibria," *Econometrica,* **54**, 1375–1385.

[25] Forges, F. (1990): "Some Thoughts on Efficiency and Information," in *Frontiers of Game Theory*, Binmore, Kirman and Tani (eds.), 133–149.

[26] Forges, F. (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.

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

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

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

[30] Kleiboer, M. (1996): "Understanding Success and Failure of International Mediation," *Journal of Conflict Resolution*, **40(2)**, 360–89.

[31] Koremenos, B., Lipson C. and D. Snidal (2001): "The Rational Design of International Institutions," *International Organization*, **55(4)**, 761–99.

[32] Kydd, A. (2003): "Which Side Are You On? Bias, Credibility, and Mediation," *American Journal of Political Science*, **47(4)**, 597–611.

[33] Kydd, A. (2006): "When Can Mediators Build Trust,"*American Political Science Review,* **100**, 449–462.

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

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

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

[37] Myerson, R.B. (1979): "Incentive Compatibility and the Bargaining Problem," *Econometrica*, **47(1)**, 61–73.

[38] Myerson, R.B. (1986): "Multi-Stage Games with Communication," *Econometrica*, **54(2)**, 323–358.

[39] Myerson, R.B. and M.A. Satterthwaite (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, **29(2)**, 265–81.

[40] Powell, R. (2006): "War as a Commitment Problem," *International Organization*, **60**, 169–203.

[41] Princen, T. (1992): *Intermediaries in International Conflict*, Princeton, N.J.: Princeton University Press.

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

[43] Rauchhaus, R. W. (2006): "Asymmetric Information, Mediation and Conflict Management," *World Politics*, **58**, 207–241.

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

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

[46] Tomz, M. (2007): "Domestic Audience Costs in International Relations: An Experimental Approach," *International Organization*, **61**, 821–40.

[47] 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.

[48] United Nations General Assembly Security Council (2000): Report A/55/305-S/2000/809.

[49] Wall, J.A. and A. Lynn (1993): "Mediation: A Current Review," *Journal of Conflict Resolution*, **37(1)**, 160–94.

[50] Watkins, M. (1998): "Building Momentum in Negotiations: Time-Related Costs and Action-Forcing Events," *Negotiation Journal*, **14(3)**, 241–256.

[51] Wilkenfeld, J., Young, K.J., Quinn, D.M. and V. Asal (2005): *Mediating International Crises*, Oxon, U.K.: Routledge.

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

NOT FOR PUBLICATION: Mediation and Peace

Material Submitted for a Web-Appendix

Appendix A - Unmediated Communication

Proof of Proposition 1 All the statements in the proposition, but the comparison with no-communication, follow from the following characterization lemma:

Lemma 1 *The best separating equilibrium is characterized as follows.*

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

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

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

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

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

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

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

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

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

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

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

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

The proof of lemma 1 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

$$p_M = \frac{(1-\lambda p_H(\gamma+2))}{(\gamma-2\lambda+1)}, \tag{5}$$

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 minimization p_H. Setting $p_H = 0$ and solving for p_M in (5) 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 (5), the solution requires setting $p_M = 1$ and, from (5), $p_H = \frac{2\lambda-\gamma}{\lambda(\gamma+2)}$. 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:

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

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:

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

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

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:

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

$$p_M = \frac{(1 - \lambda p_H(1+\gamma))}{(\gamma+1)(1-\lambda)}. \tag{6}$$

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:

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

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 satisfied. 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 IC* constraint coincides with the (binding) relaxed high-type 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 (6) 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 satisfied. 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.

This concludes the proof of the characterization lemma. One can then verify by inspection that the above full characterization determines all the characteristics highlighted in Proposition 1, but the comparison with no communication, which we now determine.

For $\gamma > 1$, $\lambda < \gamma/2$ and $\lambda < \frac{\gamma-1}{\gamma+3}$, the separating equilibrium optimal value $\frac{1+\gamma}{(1+\gamma-2\lambda)(1+\lambda)^2}$ is evidently larger than the optimal no-communication value $\frac{1}{(1+\lambda)^2}$.

Suppose that $\gamma > 1$, $\lambda < \gamma/2$, and $\lambda > \frac{\gamma-1}{\gamma+3}$. The separating equilibrium optimal value and the no-communication values are, respectively, $\frac{1+\gamma}{(1+\gamma-2\lambda)(1+\lambda)^2}$ and $\frac{1}{\lambda+1}$. The difference is:

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

and this quantity is positive if and only if $\lambda < \frac{\gamma-1}{2}$, which is always true for $\gamma > 1$ and $\lambda > \frac{\gamma-1}{\gamma+3}$, as $\frac{\gamma-1}{2} < \frac{\gamma-1}{\gamma+3}$ requires that $\gamma < 1$.

Suppose that $\gamma > 1$ and $\lambda > \gamma/2$, and $\lambda > \frac{\gamma-1}{\gamma+3}$. The separating equilibrium optimal value is $1 - \frac{\gamma\lambda}{(2+\gamma)(1+\lambda)}$. Taking the difference with the no-communication value,

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

Suppose that $\gamma < 1$ and $\lambda < \gamma/(1+\gamma)$. The separating equilibrium optimal value is $\frac{1+\gamma+\lambda(1-\gamma)}{(1+\gamma)(1-\lambda)(1+\lambda)^2}$. Hence,

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

because $\lambda < \gamma < 1$, the above is positive if $\lambda - \gamma + \lambda\gamma + 1 > 0$, i.e. $\lambda > \frac{\gamma-1}{\gamma+1}$ which always holds.

Suppose that $\gamma < 1$ and $\lambda < \gamma/(1+\gamma)$. The separating equilibrium optimal value is $1 - \frac{\gamma\lambda}{(1+\gamma)(1+\lambda)^2}$. Hence,

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

Suppose that $\gamma < 1$ and $\lambda \in [1/(1+\gamma), \gamma)$, so that the separating equilibrium optimal value is $\frac{2(1+\lambda)+\gamma}{2+\gamma+\lambda(2+\gamma)}$ and

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

This concludes the proof of Proposition 1.

Proof of Proposition 2. The Proposition follows from this Lemma.

Lemma 2 *Allowing players to play mixed strategies in the unmediated communication game, the optimal equilibrium is such that the hawk always sends message h and the dove sends 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 $\lambda < 2\gamma^2/\left(1+3\gamma\right)$,

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

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

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

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

Proof. We proceed in three parts.

Part 1. (The low type mixes).

The choice variables are b, σ, p_L, p_M, and p_H. We have 19 constraints, i.e. one IC for the low type which is binding, four IC for the high type to get rid of the maximum in the constraint, two *ex post* constraints for high type, four *ex post* constraints for low type, and eight probability constraints. First we rearrange the IC constraint for low type and express b in terms of the other variables. Substituting b into objective function and constraints, we get rid of b and IC constraint for low type. After simplifying the constraints, we are left with the following constraints, referred to as constraints Ci, $i = 1, \ldots, 9$. (We omit the constraints that all probabilities must be in $[0,1]$.)

1. $ICH1 : (1+\gamma)p_M(1+\lambda-2\sigma)-(1+\gamma)pH(1+\lambda-\sigma)+p_L\sigma$;
2. $ICH2 : -pH+p_M+\frac{(pH+p_L-2p_M)\sigma}{1+\lambda}$;
3. $ICH3 : (1+\lambda)(-\gamma+\lambda)pH+(-1+\gamma-2\lambda)\sigma(pH-p_M)+(pH+p_L-2p_M)\sigma^2$;
4. $ICH4 : (1+\lambda)(-\gamma+\lambda)p_H+((-1+\gamma-2\lambda)p_H+\gamma(p_L+\lambda p_L-p_M)+p_M+2\lambda p_M)\sigma+(pH+p_L-2p_M)\sigma^2$;
5. $EXH1 : p_M+p_L\sigma+pH(-1-(2+\gamma)\lambda+\sigma)-p_M(\gamma-2\lambda+2\sigma)$;

6. $EXH2 : \lambda + \gamma(-1+\sigma)$;
7. $EXL1 : pH(1+\lambda-\sigma)(1+(2+\gamma)\lambda-\sigma)+\sigma(p_M(2+(3+\gamma)\lambda-2\sigma)+p_L(-1-\lambda+\sigma))$;
8. $EXL3 : p_M(2+(3+\gamma)\lambda-2\sigma)+p_L\sigma+pH(-1-(2+\gamma)\lambda+\sigma)$;
9. $EXL4 : 1-\frac{(1+\gamma)\lambda}{-1-\lambda+\sigma}$.

- case 1: C5 binds

  This section covers the case that only C5 binds. We do not assume C5 binds *ex ante*. We set up the following relaxed problem:

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

  subject to the following constraints:

  1. $p_L \leq 1$,
  2. $0 \leq p_M \leq 1$,
  3. $p_H \geq 0$,
  4. $0 \leq \sigma \leq 1$,
  5. $C5 \geq 0 \Leftrightarrow p_L\sigma \geq (1+(2+\gamma)\lambda-\sigma)p_H + (\gamma-2\lambda+2\sigma-1)p_M$.

  - Case 1.1: Parameter Region is $1/2<\lambda \leq \frac{1}{2}(-1+\sqrt{5})$ and $\frac{1-\lambda}{\lambda} < \gamma < 2\lambda$, or $\lambda > \frac{1}{2}(-1+\sqrt{5})$ and $\lambda < \gamma < 2\lambda$.

    1. We want to show that $p_L = 1$. Suppose $p_L < 1$. We can set $p_L = 1$ and increase p_H to make sure C5 is satisfied. By doing so, no constraint will be violated and the objective function is strictly decreased.
    2. We want to show that C5 binds. Suppose it does not. We can increase p_H without violating other constraints and decrease the objective function.
    3. Suppose $(\gamma-2\lambda+2\sigma-1) > 0$. Then $\frac{MCp_M}{MCp_H} = \frac{2\sigma+(\gamma-2\lambda-1)}{1+\lambda-\sigma+(1+\gamma)\lambda} < \frac{2\sigma}{1+\lambda-\sigma} = \frac{MUp_M}{MUp_H}$, since $(\gamma-2\lambda-1) < 0$ and $(1+\gamma)\lambda > 0$. Therefore, we want p_M to be as large as possible and p_H to be as small as possible, i.e. $p_M = 1$ or $p_H = 0$.
       If $\sigma \leq \gamma-2\lambda+2\sigma-1$, $p_H = 0$ and $p_M = \frac{\sigma}{\gamma-2\lambda+2\sigma-1}$.
       If $\sigma \geq \gamma-2\lambda+2\sigma-1$, $p_M = 1$ and $p_H = \frac{1-\sigma+2\lambda-\gamma}{1-\sigma+2\lambda+\gamma\lambda}$.
    4. Suppose $(\gamma-2\lambda+2\sigma-1) \leq 0$. We have $p_L\sigma + (-\gamma+2\lambda-2\sigma+1)p_M \geq (1+(2+\gamma)\lambda-\gamma)p_H$. Then $p_L = 1$, $p_M = 1$ and $p_H = \frac{1-\sigma+2\lambda-\gamma}{1-\sigma+2\lambda+\gamma\lambda}$.
    5. To sum up, we conclude that:

(a) If $0 \leq \sigma \leq 1+2\lambda-\gamma$, then $p_L = 1$, $p_M = 1$, $p_H = \frac{1-\sigma+2\lambda-\gamma}{1-\sigma+2\lambda+\gamma\lambda}$, and $V = \frac{\gamma(1+\lambda-\sigma)^2}{(1+\lambda)(1+(2+\gamma)\lambda-\sigma)}$.

(b) If $1 \geq \sigma \geq 1+2\lambda-\gamma$, then $p_L = 1$, $p_H = 0$, $p_M = \frac{\sigma}{\gamma-2\lambda+2\sigma-1}$, and $V = -\frac{(1+\lambda-\sigma)((-1+\gamma-2\lambda)(1+\lambda)+(1+\gamma)\sigma)}{(1+\lambda)^2(1-\gamma+2\lambda-2\sigma)}$.

Under the parameter region we specify above, we know that $1+2\lambda-\gamma \geq 1$. Since $\sigma \leq 1$, only case (a) is possible. And V is minimized when $\sigma = 1$.

6. The solution to the relaxed problem is $p_L = 1$, $p_M = 1$, $p_H = \frac{2\lambda-\gamma}{2\lambda+\gamma\lambda}$, $\sigma = 1$, and $V = \frac{\gamma\lambda}{2+\gamma+2\lambda+\gamma\lambda}$. Substituting these into the original problem, we can show that all the constraints are satisfied. Therefore, this is also the solution to the original problem.

– Case 1.2: Parameter Region is $0<\lambda \leq \frac{1}{2}$ and $\gamma > 1$, or $\lambda > \frac{1}{2}$ and $\gamma > 2\lambda$.

1. We want to show that $p_L = 1$. Suppose not. We can increase p_L and p_H and decrease the objective function without violating the other constraints.
2. It is easy to show that C5 binds. Suppose not. We can increase p_H and decrease the objective function without violating the other constraints.
3. Suppose $(\gamma - 2\lambda + 2\sigma - 1) > 0$.
If $\frac{MCp_M}{MCp_H} = \frac{2\sigma+(\gamma-2\lambda-1)}{1+\lambda-\sigma+(1+\gamma)\lambda} \leq \frac{2\sigma}{1+\lambda-\sigma} = \frac{MUp_M}{MUp_H}$, then we want p_M to be as large as possible and p_H to be as small as possible, i.e. $p_M = 1$ or $p_H = 0$. If $\sigma \leq \gamma - 2\lambda + 2\sigma - 1$, $p_H = 0$ and $p_M = \frac{\sigma}{\gamma-2\lambda+2\sigma-1}$. If $\sigma \geq \gamma - 2\lambda + 2\sigma - 1$, $p_M = 1$ and $p_H = \frac{1-\sigma+2\lambda-\gamma}{1-\sigma+2\lambda+\gamma\lambda}$.
If $\frac{MCp_M}{MCp_H} = \frac{2\sigma+(\gamma-2\lambda-1)}{1+\lambda-\sigma+(1+\gamma)\lambda} > \frac{2\sigma}{1+\lambda-\sigma} = \frac{MUp_M}{MUp_H}$, we want p_M to be as small as possible and p_H to be as large as possible, i.e. $p_M = 0$ and $p_H = \frac{\sigma}{1+(2+\gamma)\lambda-\sigma}$.
4. Suppose $(\gamma - 2\lambda + 2\sigma - 1) \leq 0$. We have $p_L\sigma + (-\gamma + 2\lambda - 2\sigma + 1)p_M \geq (1 + (2+\gamma)\lambda - \gamma)p_H$. Then $p_L = 1$, $p_M = 1$ and $p_H = \frac{1-\sigma+2\lambda-\gamma}{1-\sigma+2\lambda+\gamma\lambda}$.
5. To sum up, we conclude that:
   (a) If $0 \leq \sigma \leq \frac{1+2\lambda-\gamma}{2}$, we have $p_L = 1$, $p_M = 1$, $p_H = \frac{1-\sigma+2\lambda-\gamma}{1-\sigma+2\lambda+\gamma\lambda}$, and $V = \frac{\gamma(1+\lambda-\sigma)^2}{(1+\lambda)(1+(2+\gamma)\lambda-\sigma)}$.
   (b) If $\frac{(-1+\gamma-2\lambda)(1+\lambda)}{-1+\gamma+2\gamma\lambda} > \sigma \geq \frac{1+2\lambda-\gamma}{2}$, $p_L = 1$, $p_M = 0$, $p_H = \frac{\sigma}{1+(2+\gamma)\lambda-\sigma}$. And $V = \frac{(1+\lambda-\sigma)(1-\sigma+\lambda(3+2\lambda+\gamma(1+\lambda+\sigma)))}{(1+\lambda)^2(1+(2+\gamma)\lambda-\sigma)}$.
   (c) If $\max(\frac{1+2\lambda-\gamma}{2}, \frac{(-1+\gamma-2\lambda)(1+\lambda)}{-1+\gamma+2\gamma\lambda}) \leq \sigma \leq 1+2\lambda-\gamma$, then $p_L = 1$, $p_M = 1$, $p_H = \frac{1-\sigma+2\lambda-\gamma}{1-\sigma+2\lambda+\gamma\lambda}$, and $V = \frac{\gamma(1+\lambda-\sigma)^2}{(1+\lambda)(1+(2+\gamma)\lambda-\sigma)}$.
   (d) If $\max(1+2\lambda-\gamma, \frac{(-1+\gamma-2\lambda)(1+\lambda)}{-1+\gamma+2\gamma\lambda}) \leq \sigma \leq 1$, $p_L = 1$, $p_H = 0$, $p_M = \frac{\sigma}{\gamma-2\lambda+2\sigma-1}$, and $V = -\frac{(1+\lambda-\sigma)((-1+\gamma-2\lambda)(1+\lambda)+(1+\gamma)\sigma)}{(1+\lambda)^2(1-\gamma+2\lambda-2\sigma)}$.
6. Under the parameter region we specify above, all the cases specified above are possible. After comparing all the minimized values, we find that case (d) achieves the minimized V when $\sigma = 1$.

7. The solution to the relaxed problem is $p_L = 1$, $p_M = \frac{1}{1+\gamma-2\lambda}$, $p_H = 0$, $\sigma = 1$, and $V = \frac{\lambda(\gamma(2+\lambda)-\lambda(3+2\lambda))}{(1+\gamma-2\lambda)(1+\lambda)^2}$. Substituting these into the original problem, we can show that all the constraints are satisfied. Therefore, this is also the solution to the original problem.

- case 2: C1 binds

This section covers the case that C1 binds and C5 might bind. We do not assume C1 binds *ex ante*.

We set up the following relaxed problem:

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

subject to the following constraints:

1. $p_L \leq 1$,
2. $0 \leq p_M \leq 1$,
3. $p_H \geq 0$,
4. $0 \leq \sigma \leq 1$,
5. $C1 \geq 0 \Leftrightarrow p_L\sigma \geq (1+\gamma)(1+\lambda-\sigma)p_H + (1+\gamma)(-1-\lambda+2\sigma)p_M$.
6. $C5 \geq 0 \Leftrightarrow p_L\sigma \geq (1+(2+\gamma)\lambda-\sigma)p_H + (\gamma-2\lambda+2\sigma-1)p_M$.

- Case 2.1:

  Parameter Region is $0<\lambda \leq \frac{1}{2}$ and $\lambda < \gamma \leq \frac{\lambda}{1-\lambda}$, or $1/2<\lambda \leq \frac{1}{2}(-1+\sqrt{5})$ and $\lambda < \gamma < \frac{1-\lambda}{\lambda}$.

  1. We want to show that $p_L = 1$. Suppose not. We can increase p_L and p_H and decrease the objective function without violating other constraints.
  2. It's easy to show that either C1 or C5 binds. Suppose both are not binding. We can increase p_H without violating other constraints. Here, we first consider the case where C1 binds.
  3. Suppose $2\sigma-\lambda-1 \geq 0$. Then $\frac{MCp_M}{MCp_H} = \frac{2\sigma-\lambda-1}{1+\lambda-\sigma} < \frac{2\sigma}{1+\lambda-\sigma} = \frac{MUp_M}{MUp_H}$. Therefore, $p_M = 1$ or $p_H = 0$. If $\sigma \geq (1+\gamma)(-1-\lambda+2\sigma)$, we have $p_M = 1$ and $p_H = \frac{\sigma+(1+\gamma)(1+\lambda-2\sigma)}{(1+\gamma)(1+\lambda-\sigma)} \geq 0$. If $\sigma \leq (1+\gamma)(-1-\lambda+2\sigma)$, we have $p_H = 0$, $p_M = \frac{\sigma}{(1+\gamma)(-1-\lambda+2\sigma)} \leq 1$.
  4. Suppose $2\sigma-\lambda-1 < 0$, we have $p_L = 1$, $p_M = 1$, and $p_H = \frac{\sigma+(1+\gamma)(1+\lambda-2\sigma)}{(1+\gamma)(1+\lambda-\sigma)}$.
  5. To sum up ,we can show that:
     (a) If $0 \leq \sigma \leq \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$, then $p_L = 1$, $p_M = 1$ and $p_H = \frac{\sigma+(1+\gamma)(1+\lambda-2\sigma)}{(1+\gamma)(1+\lambda-\sigma)}$.

(b) If $1 \geq \sigma \geq \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$, then $p_L = 1$, $p_H = 0$, and $p_M = \frac{\sigma}{(1+\gamma)(-1-\lambda+2\sigma)}$.

6. Since in the parameter region specified above $\frac{(1+\gamma)(1+\lambda)}{1+2\gamma} \geq 1$, we know that only case (a) is possible. Hence, $p_L = 1$, $p_M = 1$, $p_H = \frac{\sigma+(1+\gamma)(1+\lambda-2\sigma)}{(1+\gamma)(1+\lambda-\sigma)}$, and $V = \frac{\gamma(1+\lambda-\sigma)\sigma}{(1+\gamma)(1+\lambda)^2}$. Notice that V is a quadratic function of σ which is maximized at $\sigma = \frac{1+\lambda}{2}$.
7. Substituting p_L, p_M, and p_L into C5, we have the following constraint on σ:

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

which is equivalent to

$$\sigma_1 \leq \sigma \leq \sigma_2,$$

where

$$\sigma_1 = \frac{1}{2}(2 + \gamma + 3\lambda + 2\gamma\lambda - \sqrt{2\gamma\lambda(3+4\lambda) + \lambda(4+5\lambda) + (\gamma + 2\gamma\lambda)^2}),$$

$$\sigma_2 = \frac{1}{2}(2 + \gamma + 3\lambda + 2\gamma\lambda + \sqrt{2\gamma\lambda(3+4\lambda) + \lambda(4+5\lambda) + (\gamma + 2\gamma\lambda)^2}).$$

Since $\frac{1+\lambda}{2} < \sigma_1 \leq 1 < \sigma_2$, we know that V is minimized at $\sigma = 1$.
8. Next we consider the case that C5 is binding. Using the same method, we get the minimized value which is larger than the V specified above. Hence, the solution to the relaxed problem is $p_L = 1$, $p_M = 1$, $p_H = 1 - \frac{\gamma}{\lambda+\gamma\lambda}$, $\sigma = 1$, and $V = \frac{\gamma\lambda}{(1+\gamma)(1+\lambda)^2}$. Substituting the solution to the original problem, we show that all the constraints are satisfied. Hence, this is also the solution to the original problem.

– Case 2.2: Parameter Region is $0<\lambda \leq \frac{1}{2}$ and $\frac{\lambda}{1-\lambda} \leq \gamma \leq 1$.

1. p_L=1. Suppose not. We can increase p_L and p_H without violating other constraints.
2. It is easy to show that either C1 or C5 binds. Suppose both are not binding. We can increase p_H without violating other constraints. Here, we first consider the case where C1 binds.
3. Suppose $2\sigma - \lambda - 1 \geq 0$, $\frac{MCp_M}{MCp_H} = \frac{2\sigma-\lambda-1}{1+\lambda-\sigma} < \frac{2\sigma}{1+\lambda-\sigma} = \frac{MUp_M}{MUp_H}$. Therefore, $p_M = 1$ or $p_H = 0$. If $\sigma \geq (1+\gamma)(-1-\lambda+2\sigma)$, we have $p_M = 1$ and $p_H = \frac{\sigma+(1+\gamma)(1+\lambda-2\sigma)}{(1+\gamma)(1+\lambda-\sigma)} \geq 0$. If $\sigma \leq (1+\gamma)(-1-\lambda+2\sigma)$, we have $p_H = 0$, $p_M = \frac{\sigma}{(1+\gamma)(-1-\lambda+2\sigma)} \leq 1$
4. Suppose $2\sigma - \lambda - 1 < 0$, we have $p_L = 1$, $p_M = 1$, and $p_H = \frac{\sigma+(1+\gamma)(1+\lambda-2\sigma)}{(1+\gamma)(1+\lambda-\sigma)}$.
5. To sum up ,we can show that:

   (a) If $0 \leq \sigma \leq \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$, then $p_L = 1$, $p_M = 1$ and $p_H = \frac{\sigma+(1+\gamma)(1+\lambda-2\sigma)}{(1+\gamma)(1+\lambda-\sigma)}$.

(b) If $1 \geq \sigma \geq \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$, then $p_L = 1$, $p_H = 0$, and $p_M = \frac{\sigma}{(1+\gamma)(-1-\lambda+2\sigma)}$.

6. Since in the parameter region specified above $\frac{(1+\gamma)(1+\lambda)}{1+2\gamma} \leq 1$, we know that both case (a) and (b) are possible.

(a) If $0 \leq \sigma \leq \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$, $p_L = 1$, $p_M = 1$, $p_H = \frac{\sigma+(1+\gamma)(1+\lambda-2\sigma)}{(1+\gamma)(1+\lambda-\sigma)}$, and $V = \frac{\gamma(1+\lambda-\sigma)\sigma}{(1+\gamma)(1+\lambda)^2}$.

Substituting p_L, p_M, and p_L into C5, we have the following constraint:

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

which is equivalent to

$$\sigma_1 \leq \sigma \leq \sigma_2,$$

where

$$\sigma_1 = \frac{1}{2}(2+\gamma+3\lambda+2\gamma\lambda-\sqrt{2\gamma\lambda(3+4\lambda)+\lambda(4+5\lambda)+(\gamma+2\gamma\lambda)^2}),$$

$$\sigma_2 = \frac{1}{2}(2+\gamma+3\lambda+2\gamma\lambda+\sqrt{2\gamma\lambda(3+4\lambda)+\lambda(4+5\lambda)+(\gamma+2\gamma\lambda)^2}).$$

Taking into account all constraints on σ, we have the following problem:

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

such that

$$0 \leq \sigma \leq \frac{(1+\gamma)(1+\lambda)}{1+2\gamma} = \sigma_3,$$

$$\sigma_1 \leq \sigma \leq \sigma_2.$$

We can show that if $\gamma < \frac{1}{4}(3\lambda-\sqrt{8\lambda+9\lambda^2})$, or $\gamma > \frac{1}{4}(3\lambda+\sqrt{8\lambda+9\lambda^2})$, then $\sigma_3 < \sigma_1$ and the feasible region of σ is empty. If $\frac{1}{4}(3\lambda-\sqrt{8\lambda+9\lambda^2}) \leq \gamma \leq \frac{1}{4}(3\lambda+\sqrt{8\lambda+9\lambda^2})$, then $\frac{1+\lambda}{2} < \sigma_1 \leq \sigma \leq \sigma_3 \leq 1$. Since V is a quadratic function of σ, it is obvious that V is minimized at $\sigma = \sigma_3 = \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$, and $V = \frac{\gamma^2}{(1+2\gamma)^2}$.

Here we rearrange the parameter region. We show that if $0 \leq \gamma \leq 1$ and $\frac{2\gamma^2}{1+3\gamma} \leq \lambda \leq \frac{\gamma}{1+\gamma}$, then $p_L = 1$, $p_M = 1$, $p_H = 0$, $\sigma = \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$, and $V = \frac{\gamma^2}{(1+2\gamma)^2}$.

(b) If $1 \geq \sigma \geq \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$, we have $p_L = 1$, $p_H = 0$, $p_M = \frac{\sigma}{(1+\gamma)(-1-\lambda+2\sigma)}$, and

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

Substituting p_L, p_M, and p_L into C5, we have the following constraint:

$$\frac{(-\lambda+\gamma(2+\lambda-2\sigma))\sigma}{(1+\gamma)(1+\lambda-2\sigma)} \leq 0,$$

which is equivalent to

$$0 \leq \sigma \leq \frac{1+\lambda}{2} \ or \ \frac{1}{2}(2+\lambda-\frac{\lambda}{\gamma}) \leq \sigma.$$

Taking into account all constraints on σ, we have the following reduced problem:

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

such that

$$1 \geq \sigma \geq \frac{(1+\gamma)(1+\lambda)}{1+2\gamma},$$

$$0 \leq \sigma \leq \frac{1+\lambda}{2} \ or \ \frac{1}{2}(2+\lambda-\frac{\lambda}{\gamma}) \leq \sigma.$$

If $0 \leq \lambda \leq \frac{2\gamma^2}{1+3\gamma}$, then $1 \geq \sigma \geq \frac{1}{2}(2+\lambda-\frac{\lambda}{\gamma})$. If $\frac{\gamma}{1+\gamma} \geq \lambda > \frac{2\gamma^2}{1+3\gamma}$, then $1 \geq \sigma \geq \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$. Since the curve of V is inverse U-shaped, we know that the minimal can be achieved at $\sigma = 1$, $\sigma = \frac{1}{2}(2+\lambda-\frac{\lambda}{\gamma})$, or $\sigma = \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$. When $\sigma = 1$, $V_1 = \frac{\lambda(\lambda+\lambda^2+\gamma(-1+\lambda+\lambda^2))}{(1+\gamma)(-1+\lambda)(1+\lambda)^2}$. When $\sigma = \frac{1}{2}(2+\lambda-\frac{\lambda}{\gamma})$, $V_2 = \frac{\lambda(\lambda+2\gamma\lambda(3+2\lambda)-\gamma^2(4+3\lambda))}{4\gamma(1+\lambda)^2(-\gamma+\lambda)}$. When $\sigma = \frac{(1+\gamma)(1+\lambda)}{1+2\gamma}$, $V_3 = \frac{\gamma^2}{(1+2\gamma)^2}$.
To sum up, we have following three cases:

i. V_1 is chosen when $0 < \gamma < \gamma^*$ and $0 \leq \lambda \leq \frac{-1-2\gamma}{2(1+\gamma)} + \frac{1}{2}\sqrt{\frac{1+11\gamma+24\gamma^2+16\gamma^3}{(1+\gamma)^2(1+3\gamma)}}$ or $\gamma^* < \gamma < 1$ and $0 \leq \lambda \leq \frac{-1-8\gamma-3\gamma^2}{2(-1+\gamma^2)} - \frac{1}{2}\sqrt{\frac{1+16\gamma+54\gamma^2+48\gamma^3+25\gamma^4}{(-1+\gamma^2)^2}}$.

ii. V_2 is chosen when $\gamma^* < \gamma < 1$ and $\frac{-1-8\gamma-3\gamma^2}{2(-1+\gamma^2)} - \frac{1}{2}\sqrt{\frac{1+16\gamma+54\gamma^2+48\gamma^3+25\gamma^4}{(-1+\gamma^2)^2}} < \lambda < \frac{2\gamma^2}{1+3\gamma}$.

iii. V_3 is chosen when $0 < \gamma < \gamma^*$ and $\frac{-1-2\gamma}{2(1+\gamma)} + \frac{1}{2}\sqrt{\frac{1+11\gamma+24\gamma^2+16\gamma^3}{(1+\gamma)^2(1+3\gamma)}} < \lambda < \frac{\gamma}{1+\gamma}$, or $\gamma^* < \gamma < 1$ and $\frac{\gamma}{1+\gamma} \geq \lambda > \frac{2\gamma^2}{1+3\gamma}$.

Since the solution given in case (b) is superior to that in case (a), the above solution is the final solution.

7. Next we consider the case that C5 is binding. Using the same method, we get the minimized value which is not smaller than the value for V specified above. Substituting the solution to the original problem, we show that all the constraints are satisfied. Hence, this is also the solution to the original problem.

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:

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

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

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

to the indifference condition for the high type

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

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:

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

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:

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

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,

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

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.

Appendix B – Mediation

For reasons of clarity, the proof of Proposition 3 is postponed to after the proof of Proposition 4

Proof of Proposition 4. The proof follows from this Lemma.

Lemma 3 *The solution of the mediator's program with enforcement power is such that: For $\lambda \leq \gamma/2$,*

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

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

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

Proof. 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 *ex 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 *ex 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 *ex interim* incentive compatibility constraint, which is increasing in p_L. Second, note that the low-type *ex 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*ex interim* individual rationality constraint. Third, note that the high-type *ex interim* individual rationality constraint must be binding in the relaxed program's solution, or else one could decrease b and make the low-type *ex interim* incentive compatibility slack.

Solving for b and p_H as a function of p_M in the system defined by the low-type *ex interim* incentive compatibility and high-type *ex 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 *ex interim* incentive compatibility and high-type *ex 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 of Lemma 3 and hence of Proposition 4 is concluded by showing that this solution does not violate the high-type *ex interim* incentive compatibility and low-type *ex interim* individual rationality constraints in the complete program.

Indeed, for $\lambda \geq \gamma/2$, we verify that the slacks of these constraints are, respectively

$$\frac{1}{2}(\gamma - \lambda + 1)^{-1}(1 - \theta)(\gamma - \lambda)(\gamma + 1) > 0,$$
$$\text{and } \frac{1}{2}(\gamma - \lambda + 1)^{-1}(\gamma + 1)(1 - \theta) > 0.$$

Similarly, for $\lambda \leq \gamma/2$, the slacks are

$$\frac{1}{2}(\gamma - 2\lambda + 1)^{-1}(\lambda + 1)^{-1}(1 - \theta)(\gamma - \lambda)(\gamma + 1) > 0,$$
$$\text{and } \frac{1}{2}(\gamma + 1 - 2\lambda)^{-1}(\lambda + 1)^{-1}(\gamma + 1)(1 - \theta) > 0.$$

Proof of Proposition 3. The characterization follows from this Lemma.

Lemma 4 *A solution to the mediator's problem is such that:*

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

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

Further,

$$p_L \le \frac{2\lambda}{(\gamma - 2\lambda + 1)(\gamma - 1)}, \text{ if } \gamma \ge 1, p_L \ge \frac{(1-\gamma)\lambda}{2\gamma^2}\frac{(\lambda - \gamma)(\gamma + 2)}{(\lambda - \gamma - 1)} \text{ if } \gamma < 1;$$

The ex ante peace probability is

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

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

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

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

$$p_L \le 2\frac{(\gamma - \lambda)(\gamma + 2)\lambda}{(\gamma - \lambda + 1)\gamma(\gamma - 1)} \text{ if } \gamma \ge 1, p_L \ge \frac{(1-\gamma)\lambda}{2\gamma^2}\frac{(\lambda - \gamma)(\gamma + 2)}{(\lambda - \gamma - 1)} \text{ if } \gamma < 1;$$

The ex ante peace probability is

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

Proof. Consider the general mechanisms subject to the *ex post* IR and *ex interim* IC* constraints (1)-(4). It is straightforward to observe that the *ex post* IR constraints constraints are stronger than the following (high-type and low-type, respectively) *ex interim* IR constraints

$$\int_0^1 bdF(b|h) \ge \Pr[l,h]p\theta + \Pr[h,h]\theta/2,$$
$$\int_0^1 bdF(b|l) \ge \Pr[h,l](1-p)\theta + \Pr[l,l]\theta/2, \text{ for all } b \in [0,1]$$

and that the *ex interim* IC* constraint are stronger than the *ex interim* IC constraint obtained by substituting the maxima with their first argument (the interim payoff induced by accepting peace recommendations later in the game).

By the revelation principle by Myerson (1979), the optimal *ex ante* probability of peace within the class of mechanisms which satisfy these *ex interim* IC and IR constraints cannot be larger than the *ex ante* probability of peace identified in Lemma 3 in Appendix D. Because the *ex interim* IC and IR constraints are weaker than the *ex interim* IC* and *ex post* IR constraints, it follows that any mechanism subject to the constraints (1)–(4) cannot yield a higher *ex ante* probability of peace than the one identified in Lemma 3.

Hence, to prove the result, it is enough to show that the formulas for the choice variables $(b, p_L, q_L, p_M, q_M, q_H)$ satisfy the constraints (1)-(4) and achieve the same *ex* *ante* probability of peace as in Lemma 3. Specialize to the mechanisms described by $(b, p_L, q_L, p_M, q_M, q_H)$, the *ex post* IR constraints take the following form, for the high type:

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

and for the low type:

$$\begin{aligned}&p_L b \geq p_L\theta/2,\ (qp_M + (1-q)p_L)(1-b) \geq qp_M(1-p)\theta + (1-q)p_L\theta/2,\\&(qq_M + (1-q)q_L) \cdot 1/2 \geq qq_M(1-p)\theta + (1-q)q_L\theta/2,\end{aligned}$$

whereas the high-type *ex interim* IC* constraint is

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

and the low-type *ex interim* IC* constraint is

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

It is straightforward to verify that the values provided in Lemma 4 are such that the *ex ante* IC* constraint in which the low type does not wage war after misreporting is binding. Also, plugging in our two sets of values for the choice variables gives the same welfare as in Proposition 3. We are left with showing that all other constraints are satisfied. We distinguish the two cases.

Step 1. Suppose that $\lambda < \gamma/2$, so that $q_M = q_H = 0$. After simplification, the low-type IC* constraint becomes

$$\begin{aligned}&q(p_M(1-p\theta) + (1-p_M)(1-p)\theta) + (1-q) \cdot 1/2 \geq\\&(1-q)p_M p\theta + q(1-p)\theta + q(1-p_M)\theta/2,\end{aligned}$$

which is binding for $p_M = \frac{1}{1+\gamma-2\lambda}$. Consider the high-type IC* constraint

$$\begin{aligned}&q\theta/2 + (1-q)(p_M b + (1-p_M)p\theta) \geq \max\{(qp_M + (1-q)p_L)(1-b), qp_M\theta/2 + (1-q)p_L p\theta\}\\&+\max\{(1-q)p_L b, (1-q)p_L p\theta\} + \max\{(1-q)q_L \cdot 1/2, (1-q)q_L p\theta\} + q(1-p_M)\theta/2,\end{aligned}$$

Note that

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

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

$$(1-q)\,p_L b = (1-q)\,p_L p\theta$$

and that

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

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

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

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

Similarly, we find that the two high-type *ex post* constraints

$$p_M b \geq p_M p\theta, \text{ and } (qq_H + (1-q)q_M)\cdot 1/2 \geq qq_H\theta/2 + (1-q)q_M p\theta$$

are satisfied — the second one because both sides equal zero.

We need to show that the low-type *ex post* constraints are satisfied. Indeed:

$$p_L p\theta > p_L\theta/2, \;(1-q)q_L \cdot 1/2 > (1-q)q_L\theta/2,$$

whereas

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

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 $\gamma < 1$, $p_L \geq 0 \geq \frac{2\lambda}{(\gamma-2\lambda+1)(\gamma-1)}$.

Finally the probability constraints are satisfied. In fact, $0 \leq p_M \leq 1$ requires only that $1 \leq 1+\gamma-2\lambda$, i.e., that $\lambda \leq \gamma/2$.

Step 2. Suppose that $\lambda \geq \gamma/2$. Consider the low-type constraint, first. After simplifying maxima, as the low type always accepts the split if exaggerating strength, the low-type IC* constraint is satisfied as an equality when plugging in the expressions $p_M + q_M = 1$, $b = p\theta$, $q_H = \frac{2\lambda-\gamma}{\lambda(\gamma+1-\lambda)}$, $q_M = \frac{2\lambda-\gamma}{\gamma(\gamma+1-\lambda)}$.

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

$$(qp_M + (1-q)p_L)\cdot(1-b) \leq qp_M\theta/2 + (1-q)p_L p\theta$$

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

$$(1-q)p_L b = (1-q)p_L p\theta$$

and that

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

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 verify that the consequentially simplified high-type IC* constraint is satisfied with equality, when substituting in the expressions $p_M + q_M = 1$, $b = p\theta$, $q_H = \frac{2\lambda-\gamma}{\lambda(\gamma+1-\lambda)}$, $q_M = \frac{2\lambda-\gamma}{\gamma(\gamma+1-\lambda)}$.

We then verify that the two high-type *ex post* constraints

$$p_M b \geq p_M p\theta, \text{ and } (qq_H + (1-q)q_M) \cdot 1/2 \geq qq_H\theta/2 + (1-q)q_M p\theta$$

are satisfied with equality when substituting in the expressions for $b = p\theta$, $q_H = \frac{2\lambda-\gamma}{\lambda(\gamma+1-\lambda)}$, $q_M = \frac{2\lambda-\gamma}{\gamma(\gamma+1-\lambda)}$.

Finally, show that the low-type *ex post* constraints are satisfied. In fact

$$p_L p\theta > p_L\theta/2, \text{ and } (qq_M + (1-q)q_L) \cdot 1/2 > qq_M(1-p)\theta + (1-q)q_L\theta/2,$$

whereas

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

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 $\gamma \geq 1$, $p_L \leq 2\frac{(\gamma-\lambda)(\gamma+2)\lambda}{(\gamma-\lambda+1)\gamma(\gamma-1)}$ and if $\gamma < 1$, $p_L \geq 0 \geq 2\frac{(\gamma-\lambda)(\gamma+2)\lambda}{(\gamma-\lambda+1)\gamma(\gamma-1)}$.

Finally the probability constraints are satisfied. In fact, because $\gamma + 1 - \lambda > 0$, $2\lambda - \gamma - \lambda(\gamma+1-\lambda) = (\lambda+1)\,(\lambda-\gamma) < 0$, and $2\lambda - \gamma - \gamma(\gamma+1-\lambda) = (\gamma+2)\,(\lambda-\gamma)$, the conditions $0 \leq q_H \leq 1$ and $0 \leq q_M \leq 1$ require only that $2\lambda - \gamma \geq 0$.

Having proved that the claimed solution satisfies all constraints, the proof of Lemma 4, and hence Proposition 3 is now concluded.