

Mediation and Peace*

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

First version, May 2010;
this version, February, 2015

Abstract

This paper applies mechanism design to the study of international conflict resolution. Standard mechanisms in which an arbitrator can enforce her decisions are usually not feasible because disputants are sovereign entities. Nevertheless, we find that this limitation is inconsequential. Despite only being capable of making unenforceable recommendations, mediators can be equally effective as arbitrators. By using recommendation strategies that do not reveal that one player is weak to a strong opponent, a mediator can effectively circumvent the unenforceability constraint. This is because these strategies make the strong player agree to recommendations that yield the same payoff as arbitration in expectation. This result relies on the capability of mediators to collect confidential information from the disputants, before making their recommendations. Simple protocols of unmediated communication cannot achieve the same level of ex ante welfare, as they preclude confidentiality.

*We thank Sandeep Baliga, Alberto Bisin, Yeon-Koo Che, Matt Jackson, Philippe Jehiel, Philipp Kircher, Adam Meirowitz, Clara Ponsati, Kris Ramsay and Tomas Sjöström for their insightful comments, and Michael Rubin for excellent research assistance. The comments of audiences of ESSET 2009, the Wallis Conference 2010, and several seminar series are gratefully acknowledged.

[†]Yale University, Department of Economics.

[‡]Bocconi University and IGER.

[§]University of Warwick, Department of Economics.

1 Introduction

A fundamental objective of international relations is to understand what institutions and/or third party intervention strategies may be most effective for conflict resolution and prevention. This paper applies the powerful tool of mechanism design to the study of conflict in international relations, and derives provocative results comparing different conflict resolution institutions. The results we derive within our framework may also serve as a benchmark to stimulate future investigations of conflict resolution with mechanism design.

We consider conflicts that arise because of asymmetric information between disputants,¹ and compare the performance of institutions such as mediation, arbitration and unmediated peace talks, that are used to reduce information asymmetries and lead to peaceful negotiated resolution of disputes. Outside international relations, the fact that private information may make bargaining and negotiations fail has been invoked as an explanation for costly trials in the case of litigation, and strikes in the case of wage bargaining (see Kennan and Wilson, 1993, for an early review). In these types of ‘conflicts,’ it is possible to use the standard mechanisms identified by the revelation principle by Myerson (1979): If the parties agree to the arbitration, an arbitrator with enforcement power (the power of law or of the State) collects information from the disputants privately, and then makes binding recommendations. Further, attention can be restricted to equilibria in which the disputants’ reports to the mediator are truthful.

However, the key distinguishing feature of international crises and disputes is that the players involved are sovereign entities, and hence there is no legitimate or recognized third party to which they can credibly delegate decision and enforcement power (see e.g. Waltz, 1959).² For this reason, any mechanism design model aimed to deal with international relations

¹Such asymmetric information may be 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. 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).

²Beside being often impossible, enforcement may even be undesirable in international relations: “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).

has to dispense with the assumption that the third party can enforce her decisions and, instead, focus on self-enforcing mechanisms.³

Our formal analysis begins by studying the benchmark case of arbitration with enforcement power. Following Bester and Wärneryd (2006), we derive the optimal arbitration mechanism that minimizes the probability of conflict in a stylized conflict environment, in which two players dispute some surplus, and conflict leads to a surplus reduction.⁴ Each player can be either weak or strong, its type is private information, and a strong type's expected payoff if fighting a weak type is larger than the equal division of the surplus. In the optimal revelation mechanism, the arbitrator 'induces' war with some probability when both disputants report to be strong, so that weak players do not have an incentive to lie.⁵

When turning to study optimal mediation, we drop the enforcement assumption. A mediator's recommendations avoid conflict only when both disputants agree to them. When comparing optimal mediation and arbitration, we surprisingly find an arbitrator who can enforce her prescriptions is no more effective in preventing conflict than a mediator who can only propose self-enforcing agreements. This result is important for international relations. A key controversy is whether non-State institutions are effective only when backed by State power, or whether they may also play a role purely by exerting suasion or managing information.⁶ Here, we settle the score by finding that mediators can be so effective in managing information so as to make enforcement power entirely redundant.⁷

³Among the few papers studying self-enforcing mechanisms in contexts different from international relations, see Matthews and Postlewaite (1989), Banks and Calvert (1992), Cramton and Palfrey (1995), Forges (1999), Compte and Jehiel (2008), and Goltsman, Hörner, Pavlov and Squintani (2009).

⁴This is a standard metaphor for many types of wars, for example those related to territorial disputes or to the present and future sharing of the rents from the extraction of natural resources.

⁵The model is framed in the context of revelation mechanism, and it appears to assume that the arbitrator recommends war to disputants. This feature of our model should not be taken literally. In the real world, third-parties who are called to resolve disputes often quit in some circumstances; this usually results in conflict escalation by the contestants. We discuss the relevance of this assumption later.

⁶The first view is often put forth by scholars who belong to the "realist" tradition that follows Morgenthau (1948) and Waltz (1959). The opposite view is usually championed by those who set their work within the "liberalist" paradigm advocated by Keohane (1984) and Barnett and Finnemore (2004), among others.

⁷This contrasts with what usually prevails in mechanism design. Quite generally, optimal mechanisms do require commitment. Whether one considers standard bargaining with independent values or interdependent values, whether one is dealing with one or multiple buyers (Myerson 1985; Skreta, 2006), lack of commitment leads to predictions (*e.g.*, fixed price mechanisms) that differ from those obtained under commitment. When

This surprising result holds because the mediator can effectively circumvent the unenforceability constraint, by using sufficiently sophisticated recommendation strategies. Specifically, the lack of enforcement power makes it impossible for the mediator to make a strong player facing a weak type agree to the optimal arbitration settlement. But the mediator can circumvent this constraint by using a recommendation strategy that does not reveal to the strong player that the opponent is weak. Specifically, the mediator always recommends equal split of surplus when the players' types are the same. When the players' strength differ, the mediator recommends with some probability the settlement that makes a strong player indifferent between fighting a weak opponent or not. But with complementary probability, it recommends equal split of the surplus. The probability is chosen so that the strong player facing a low type achieves the same payoff as arbitration in expectation.⁸ Most importantly, these recommendations are self-enforcing. Evidently, the strong player has no reason to fight if given the lion's share of the surplus. When a strong player is proposed the equal split of surplus, it does not know her opponent's strength, in equilibrium. It would strictly prefer to fight if knowing that the opponent is weak, and strictly prefer not to fight in the opposite case. By the law of total probability, it can be verified that its equilibrium expected payoff for fighting exactly coincides with one half, as we explain later in details.

The optimal mediation mechanism that we derive requires that the mediator collects information from the disputants privately in closed door meetings, and keeps it confidential before organizing a summit in which making her proposals. Otherwise, the mediator would not be able to keep a strong player uncertain about its opponent's strength. The mediator's practice of privately meeting disputants is often called 'shuttle diplomacy' and has become popular since Henry Kissinger's efforts in the Middle East in the early 1970s and the Camp David negotiations mediated by Jimmy Carter (see, for example, Kydd, 2006, and Fey and Ramsay,

incomplete information is one-sided (as in Bester and Strausz, 2001), some general principles exist that would guide the analysis of the optimal mechanism with no commitment –though no interesting condition is known under which this delivers the same outcome as with commitment. To this day, there exist no general principles or methods that would allow an analysis under two-sided incomplete information and limited commitment.

⁸This is always possible because the strong type's arbitration payoff lies between equal share of surplus and its payoff of war against a weak player.

2010, for a detailed discussion of shuttle diplomacy).

A natural question is then whether it is possible to achieve the same ex ante welfare also with unmediated communication, or with mediators that act exclusively as communication facilitators, without conducting separate meetings with the disputants, or without keeping communications confidential. Our second main result establishes that this is not the case at least for simple protocols of unmediated communication, in which players meet once and communicate simultaneously, each choosing whether to report its type truthfully or to lie. In a subset of our model's parameter space, when the cost of conflict is high or asymmetric information is significant,⁹ optimal mediation with shuttle diplomacy yields a strictly lower chance of conflict than one or two rounds of direct communication among the disputants (of course, optimal mediation always weakly dominates direct communication among the disputants). This finding, which underlines the value of confidential mediation, may explain the wide-spread use of mediation for conflict resolution. According to the International Crisis Behavior (ICB) project, 30% of international crises for the entire 1918–2001 period were mediated, and the fraction rises to 46% for 1990–2001 (see Wilkenfeld et al., 2005).¹⁰

To gain some intuition for this results, recall that the optimal recommendation strategies by a mediator deters each player from lying about its type. For the parameters of interest, the punishment consists in quitting and inducing war among self-reported strong players. A weak disputant has a stronger incentive to reveal its weakness confidentially to a mediator, than directly to its opponent. In the first case, with some probability, the opponent will not learn that the player is weak in equilibrium, and will settle for an equal share of the surplus, instead of demanding more. So, the mediator's confidentiality facilitates truthtelling by the disputants, and allows to reduce the probability of war needed to penalize lying. As a result. optimal mediation yields strictly higher welfare than the optimal truthtelling

⁹The intensity of conflict and asymmetric information are considered among the most important variables explaining when mediation is most successful (see e.g. Bercovitch and Houston, 2000, Rauchhaus, 2006).

¹⁰Outside the realm of military conflict, communication facilitation with open communication is common: the WTO settlement procedure requires explicitly full transparency of all communications among disputants. See in particular items 5,6,10 in the Annex 2 of the WTO procedural rules.

equilibrium of our unmediated communication game. Of course, babbling equilibria are also strictly dominated by mediation, as no information is communicated. The result that semi-pooling mixed communication strategy equilibria are also strictly dominated is less intuitive and is presented in the appendix.¹¹

We conclude this introduction by briefly discussing the results derived here, with the aim of both making the boundaries of our work precise, and of stimulating future investigations of conflict resolution with the tools of mechanism design. In line with the mechanism design literature, we consider unbiased mediators who have no private information.¹² Further, our results hold when the mediator’s objective is the minimization of the *ex ante* probability of war. Hence, our mediator must be able to commit to quit and terminate the mediation in some circumstances, instead of seeking a peaceful agreement in all contingencies.¹³ Such commitments facilitate information disclosure by the contestants, and ultimately improve the *ex ante* chances of peaceful conflict resolution. In the online appendix, we provide evidence that mediators as well as disputants recognize the value of the commitment to quit. But we also argue that if commitment is hard to enforce, mediation may not be more effective than unmediated communication. 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. However, single states account for less than a third of the mediators in mediated conflicts (Wilkenfeld, 2005), so that our assumption can be seen as reasonable approximation for several mediated crises.

¹¹Unlike us, Fey and Ramsay (2009) and (2010) do not find any advantage for shuttle diplomacy over unmediated communication. The reason is that private information in their model is about costs of war (private values), whereas in our model it is about the probability of winning (interdependent values).

¹²As 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 those 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 Rauchhaus, 2006).

¹³The optimal mediation mechanism has, among its properties, that quitting (or not making a proposal after receiving the private communication) needs to be interpreted as an implicit suggestion to go to war, which is a continuation equilibrium.

The paper is organized as follows. Section 2 provides the benchmark characterization of optimal arbitration, section 3 characterizes optimal mediation and shows that it achieves the same peace probability as optimal arbitration, section 4 studies unmediated communication. Section 5 concludes. All the proofs are in the appendix.

2 Arbitration

This section presents a simple model of conflict with asymmetric information, and lays out the arbitration program to minimize the chance of war.

Two players contest a pie of size normalized to one. War shrinks the value of the pie to $\theta < 1$. The expected payoffs in case of war depend on both players' private *types*. Each player can be of high (H) or low (L) type with probability q and $(1 - q)$, respectively. This private characteristic can be thought of as being related to political resolve, military strength, leaders' stubbornness, etc. When the two players are of the same type, the expected share of the (remaining) pie in case of war is $1/2$ for both. When a type H player fights against a low type, its expected share is $p > 1/2$, and hence its expected payoff is $p\theta > 1/2$.¹⁴

The model has three parameters: θ , p , and q . Yet, it turns out that a more parsimonious description of all results can be given in terms of only two statistics: $\lambda \equiv \frac{q}{1-q}$ and $\gamma \equiv \frac{p\theta - 1/2}{1/2 - \theta/2}$. The parameter $\lambda > 0$ is the high/low type odds ratio, and $\gamma \geq 0$ the ratio of benefits over cost of war for a high type: the numerator is the gain for waging war against a low type instead of accepting the equal split $(1/2, 1/2)$, and the denominator is the loss for waging war against a high type rather than accepting $(1/2, 1/2)$. Given that γ is increasing in θ , we will also interpret situations with low γ as situations of high intensity or cost of conflict. Note that when $\lambda \geq \gamma$, war can always be averted with the split $(1/2, 1/2)$ because the expected payoff of war for high types, $(1 - q)p\theta + q\theta/2$ is smaller than $1/2$. We shall henceforth assume $\lambda < \gamma$.

Under arbitration, a third party collects information privately from the disputants and makes binding decisions on how to resolve the dispute. Enforcement is hard or impossible

¹⁴If $p\theta < 1/2$, the problem is trivial, as war can always be averted with the equal split $(1/2, 1/2)$.

when the players are sovereign entities, but calculating the optimal arbitration mechanism will provide us with a useful benchmark. Invoking the version of the revelation principle proved by Myerson (1979), and proceeding as in Bester and Warneryd (2006), we can set up the arbitration game as a revelation mechanism, without loss of generality.

Hence, we stipulate that after being informed of its type, each player i chooses whether to participate in the arbitration or not. If both players agree to participate, each player i privately sends a report $m_i \in \{l, h\}$ to the arbitrator. Given reports $m = (m_1, m_2)$, the arbitrator prescribes a peaceful split $(x, 1 - x)$ with probability $p(m)$. With probability $1 - p(m)$ the arbitration fails, so that the players escalate the conflict and fight a war. Unlike the reports, the arbitrator's recommendation is public.

Again by the revelation principle, without loss of generality, we consider equilibria in which the players agree to participate in the arbitration and adopt truthful report strategies. Further, it can be shown that restricting attention to symmetric recommendations is without loss of generality, because the optimal arbitration program is linear.¹⁵ Symmetry 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 $p_L \equiv p(l, l)$, $p_M \equiv p(l, h) = p(h, l)$ and $p_H \equiv p(h, h)$. The optimal arbitration program determines b, p_L, p_M and p_H so as to minimize the war probability:

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

subject to the constraints that players are willing to reveal their type truthfully (the *interim* incentive compatibility constraints), and that the players are willing to participate to the arbitration regardless of their types (the *interim* participation constraints). Specifically, the

¹⁵Each 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, 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 averaging these two mechanisms. As the objective is linear, this symmetric mechanism is also optimal.

incentive-compatibility constraint for a low type is:

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

The left-hand side is a type L player's payoff when revealing its type. With probability $1 - q$, the opponent is also a low type: this leads to the equal split with probability p_L , and to war (with payoff $\theta/2$) with probability $1 - p_L$. With probability q , the opponent is a high type, so that the payoff is $1 - b$ with probability p_M , and $(1 - p)\theta$ with probability $1 - p_M$. The right-hand side is the player's payoff from exaggerating its strength, and reporting to be of type H . When the opponent is a low type, the payoff is b with probability p_M , and $\theta/2$ with probability $1 - p_M$. When the opponent is a high type, the payoff is $1/2$ with probability p_H , and $(1 - p)\theta$ with probability $1 - p_H$.

Similarly, the incentive compatibility constraint for a high type is:

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

The *interim* participation constraints (for the low type and the high type, respectively) are:

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

The left-hand sides are the low type and the high type's payoffs when accepting the arbitration and truthfully revealing their types. Evidently, they coincide with the left-hand sides of the incentive compatibility constraints. The right-hand sides are now the payoffs for refusing arbitration and triggering war. A type H player's payoff is $\theta/2$ with probability q , when the opponent's type is high, and the payoff $p\theta$ otherwise. A type L player's payoff is $(1 - p)\theta$ with

probability q and $\theta/2$ with probability $1 - q$.¹⁶ Relegating to the appendix the exact formulas, the main features of optimal arbitration are as follows.

Lemma 1 *The solution of the optimal arbitration program is such that:*

1. *The arbitrator's decisions make the players' types become common knowledge;*
2. *The low type's incentive compatibility constraint and the high type's participation constraint bind, the other constraints do not;*
3. *For $\lambda \leq \gamma/2$, low type dyads (L, L) do not fight ($p_L = 1$), asymmetric dyads (H, L) and (L, H) fight with positive probability ($0 < p_M < 1$), high type dyads (H, H) always fight ($p_H = 0$), and the arbitrator's prescriptions are self-enforcing ($b = p\theta$);*
4. *For $\gamma > \lambda > \gamma/2$, low type dyads and asymmetric dyads do not fight ($p_L = 1$ and $p_M = 1$), high type dyads fight with positive probability $p_H \in (0, 1)$, and the arbitrator's prescriptions are not self enforcing ($1/2 < b < p\theta$).*

It is easy to see that the players' types will be common knowledge after the arbitration is concluded: if the settlement is $(1/2, 1/2)$ each player knows that the opponent has the same type as its own, and they know that the opponent has the opposite type if an unequal split is proposed. Given that a low type must be discouraged from exaggerating strength, 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 quitting and instigating a conflict escalation when there are two self-proclaimed high types. When the players' type is likely high (λ large), it is enough to set $p_H < 1$ without instigating war in asymmetric dyads (so that $p_M = 1$). When λ is small, it is necessary to set $p_H = 0$ and $p_M < 1$, to deter low types from exaggerating strength.

Most importantly, when λ is small, the binding prescription by the arbitrator would also be self-enforcing. Because $b = p\theta$, a high type disputant would be willing to accept the split b

¹⁶Although these constraints are not linear because of the products $p_M b$, they can be turned into linear constraints by changing the variable b with $p_B = p_M b$ and the constraint $1/2 \leq b \leq 1$ with $p_B \leq p_M \leq 2p_B$.

even *ex post*, when knowing that the opponent is of low type. When the odds ratio λ is large, instead, the binding recommendation by the arbitrator is not self-enforcing, because $b < p\theta$. Upon realizing that the opponent is of low type, a type H disputant would like to not accept the arbitrator's decision and fight, so as to obtain the payoff $p\theta$.

3 Mediation

In the previous section, we have characterized the optimal solution for the case of third party intervention by an arbitrator endowed with the power to enforce its decisions. In this section, we consider mediation: The third party's decisions are not binding anymore.

The version of the revelation principle by Myerson (1982) guarantees that we can stipulate the following protocol without loss of generality, when representing mediation. After being informed of its 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 $(x, 1 - x)$ according to some cumulative distribution function $F(x|m)$, where the only recommendation leading to war in the support of $F(\cdot|m)$ is $x = 0$.¹⁷ Each of the contestants separately decides whether to accept the mediator's recommended split. Unless they both agree, war takes place (it is enough to have one player attacking to trigger war).

Again by the revelation principle, we restrict attention to equilibria in which the players reveal their type and accept the mediator's recommendation (unless they are meant to go to war). Hence, as in the case of optimal arbitration, the optimal mediation mechanism F minimizes the war probability subject to constraints determined by the structure of the equilibria we consider, formally stated in appendix. The equilibrium feature that players reveal their type truthfully translates into two (new) *interim* incentive compatibility constraints, one for each type of disputant. Whereas the feature that players accept the mediator recommendations defines *ex post* participation constraints, one for each split x recommended by F

¹⁷Clearly all recommendations leading to war induce the same payoffs, and hence can be subsumed by the recommendation $b = 0$.

with positive probability to each type of disputant. Here we note that they are at least as demanding as the incentive compatibility and participation constraints of arbitration.

The reason is twofold. First, requiring each type of each player to agree to participate in an arbitration before knowing which settlement will be proposed is less demanding than requiring them to agree to each settlement that can be proposed on path by a mediator. In principle, they may be willing to agree to some and not to others, but still be better off in expectation by participating in the arbitration. Second, when entertaining the possibility of not telling the truth in the mediation game, each type of player knows that it can also deviate and trigger war after the mediator has proposed a peaceful settlement, whereas this is impossible under arbitration. Because the constraints associated with mediation are weakly more demanding than the arbitration's constraints, the welfare achieved under optimal mediation cannot be larger than the welfare achieved under optimal arbitration.

Relegating to the appendix the exact formulas, we here report the main features of optimal mediation.

Lemma 2 *An optimal mediation mechanism is such that high type dyads are recommended the split $(1/2, 1/2)$ with probability q_H , and fight with probability $1 - q_H$; asymmetric dyads (H, L) are recommended $(p\theta, 1 - p\theta)$ with probability p_M and $(1/2, 1/2)$ with probability q_M , whereas they fight with probability $1 - p_M - q_M$, the case for (L, H) being symmetric; and low type dyads settle on the split $(1/2, 1/2)$ with probability q_L and on the splits $(p\theta, 1 - p\theta)$ and $(1 - p\theta, p\theta)$ with probability p_L each. Further:*

1. *Both the low-type incentive compatibility constraint and the high-type participation constraints bind, the other constraints do not;*
2. *For $\lambda \leq \gamma/2$, low types do not fight ($q_L + 2p_L = 1$), asymmetric dyads (H, L) fight with positive probability, and the mediator recommends only the split $(p\theta, 1 - p\theta)$, so that $q_M = 0$ and $0 < p_M < 1$, whereas high type dyads always fight ($q_H = 0$);*
3. *For $\gamma > \lambda > \gamma/2$, low types do not fight ($q_L + 2p_L = 1$); high type dyads fight with positive*

probability ($0 < q_H < 1$), asymmetric dyads (H, L) do not fight ($q_M + p_M = 1$) and the mediator recommends both splits $(p\theta, 1 - p\theta)$ and $(1/2, 1/2)$ with positive probability ($q_M > 0$ and $p_M > 0$);

4. For $\gamma < 1$, the mediator assigns the unequal splits $(p\theta, 1 - p\theta)$ and $(1 - p\theta, p\theta)$ to low type dyads with strictly positive probability ($p_L > 0$); whereas for $\gamma \geq 1$, she only recommends the equal split $(1/2, 1/2)$, so that $q_L = 1$.

For the same reasons as under arbitration, low types never fight among each other, and the low type incentive compatibility constraint binds. Again, when the chance of facing a high type is high, $\gamma > \lambda \geq \gamma/2$, only high type dyads fight, whereas when $\lambda < \gamma/2$ also asymmetric dyads fight. In fact, direct comparison of the results in Propositions 1 and 2 shows that the arbitration and mediation solutions coincide when $\lambda \leq \gamma/2$ and $\gamma \geq 1$.

But the optimal mediation mechanism is significantly more complex than optimal arbitration, when $\lambda > \gamma/2$. When not quitting and triggering war, the mediator does not always assign the split $(b, 1 - b)$ to (H, L) asymmetric dyads. With probability $q_M > 0$, she recommends $(1/2, 1/2)$ instead. This strategy is partially revealing: Unlike with optimal arbitration, players' types do not become common knowledge after recommendations are made. Specifically, here, a type H player does not learn the opponent's type precisely, when offered $1/2$.

This partially revealing recommendation strategy is adopted because, unlike an arbitrator, a mediator cannot force a type H disputant to accept a split $b < p\theta$ when facing a type L opponent. The high-type *ex post* participation constraint, $b \geq p\theta$, cannot be violated. The mediator *circumvents* this constraint by recommending the equal split $(1/2, 1/2)$ to asymmetric dyads with positive probability, thus lowering the *expected* payoff of a high type who faces a low type below $p\theta$. This circumvention is possible because a high type's payoff for fighting a high type is $\theta/2 < 1/2$. Hence, when offered the equal split, a high type is willing not to fight even if knowing that the opponent may be of low type (as long as q_M is not too high).

When $\gamma < 1$, there is another difference between optimal mediation and optimal arbitration. In this case, the mediator recommends the unequal splits $(b, 1 - b)$ and $(1 - b, b)$ with

probabilities $p_L > 0$ each to low type dyads. This strategy does not reveal the opponent's type to a self-reported type L disputant. It is used to keep in check the *double-deviation* of a high type who first pretends to be a low type and then wages war, deviating from the mediator's recommendation, if the recommended split reveals that the opponent's type is low —i.e., if the recommended split is $(1/2, 1/2)$. By not revealing to a self-reported low type player that it is facing a type L opponent, the mediator keeps in check the benefit for the double-deviation.

In spite of these differences between the arbitration and mediation optimal mechanisms, the most important result of this section is that the two mechanisms coincide in terms of peace probabilities and welfare.

Proposition 1 *An arbitrator who can enforce recommendations is exactly as effective in promoting peace as a mediator who can only propose self-enforcing agreements.*

We build the intuition for this unexpected result in stages. First, we establish that mediation and arbitration yield the same peace probability when $\gamma > \lambda \leq \gamma/2$. In fact, it is even the case that the two optimal mechanisms coincide when $\gamma \geq 1$ and $\lambda \leq \gamma/2$. When $\gamma < 1$ and $\lambda \leq \gamma/2$, the only difference concerns the values of the peaceful splits assigned to low type dyads (L, L) . But this difference is immaterial for peace probabilities and welfare, because low type dyads never fight in either mechanism.

Let us turn to the case in which $\gamma > \lambda > \gamma/2$. The only constraints binding in the arbitration solution are the low type's incentive compatibility constraint, and the high type's *interim* participation constraint. Analogously, the only binding constraints in the mediation solution are the low type's incentive compatibility constraint (which here includes the possibility of a double deviation), and the two high type's *ex post* participation constraints associated with the settlement proposals $p\theta$ and $1/2$. But the low type's incentive compatibility constraints of optimal mediation and arbitration are equivalent, because a low type never wages war after exaggerating strength, in the mediation solution.

The comparison of the high type's arbitration *interim* participation constraint and the two high type's mediation *ex post* participation constraints is less immediate. As pointed earlier,

the interim constraint is weaker than the ex post constraints. Specifically, the arbitration solution prescribes a settlement $b < p\theta$ to a type H player facing a type L opponent. The mediator cannot use this mechanism, as this recommendation would not satisfy the high type's *ex post* participation constraint. She circumvents this problem with a more complex mechanism, in which a high type facing a low type opponent is proposed the settlement $1/2$ with probability $q_M \in (0, 1)$ and $p\theta$ with complementary probability. Instead, high type dyads receive the same recommendations, with the same probabilities, in the optimal mediation and arbitration mechanism.

Further, because the two high type's *ex post* participation constraints bind, in the optimal mediation mechanism,

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

In words, a high type is indifferent between accepting the recommendation $1/2$ and waging war, thereby realizing payoff $p\theta$ if the opponent's type is low, and $\theta/2$ if it is high. By Bayes law, she assesses that the opponent's type is low with probability $\frac{(1-q)q_M}{(1-q)q_M + qq_H}$, and high with the complementary probability.

The crucial step in the intuition for the equivalence between mediation and arbitration for the case where $\lambda > \gamma/2$, is realizing that the mediator can choose q_M optimally so that the expected payoff of a high type meeting a low type under mediation, $(1 - q_M) \cdot p\theta + q_M \cdot 1/2$, is exactly equal to b , the split prescribed in optimal arbitration. This can always be done because b lies between $1/2$ and $p\theta$. Further, this choice of q_M also makes equality (2) always hold, by the law of total probability. In fact, using inequality (1) and $b = (1 - q_M) \cdot p\theta + q_M \cdot 1/2$, the binding arbitration high-type *interim* participation constraint can be written as:

$$\begin{aligned} & (1-q)(1-q_M)p\theta + (1-q)q_M \frac{1}{2} + q \left(p_H \frac{1}{2} + (1-p_H) \frac{\theta}{2} \right) \\ &= (1-q)(1-q_M)p\theta + (1-q)q_M \cdot p\theta + q \left(p_H \frac{\theta}{2} + (1-p_H) \frac{\theta}{2} \right), \end{aligned}$$

which is equivalent to the binding mediation high-type *ex post* participation constraint de-

scribed in equation (2), after simplification and rearrangement.¹⁸

Now, the welfare equivalence result of Proposition 1 becomes intuitive. Having established that the expected payoff of a type H player facing a type L player under mediation can be made equal to its payoff under arbitration, we add that also the payoffs of a type H player facing a high type can be equalized, as there is no issue of enforcement with the arbitration settlement $1/2$. So, the high type's *interim* mediation payoff can be made equal to its (optimal) arbitration payoff. Further, also the low type's *interim* mediation payoff can be equalized to its arbitration payoff, because the low type never entertains double deviations, as we pointed out earlier. In sum, both types achieve the same *interim* payoff under optimal mediation and arbitration. As a consequence, also the players' *ex ante* payoffs must coincide. And because the only payoff loss in the game is due to war, the two institutions must also be equivalent in terms of the *ex ante* chance of peace and welfare.

We conclude this section by considering the comparison between mediation and arbitration for the case in which the third party seeks a peaceful resolution in all instances, instead of being capable to commit to quit and trigger a conflict escalation for some reports by the players. As we earlier pointed out, the lack of commitment constrains the capability of third parties to resolve conflict. Within the standard framework of revelation games (generalized to allow for any arbitrary message space and allowing for mixed strategies), we show in the online appendix that arbitration without commitment cannot be more effective than mediation without commitment at minimizing the chance of conflict.¹⁹

¹⁸Intuitively, the arbitration *interim* and mediation *ex post* participation constraints are equivalent, when the mediator sets $q_H = p_H$ and q_M such that $b = (1 - q_M) \cdot p\theta + q_M \cdot 1/2$, because the net gains from violating these participation constraints are the same in the mediation and arbitration solutions. Under mediation, rearranging equation (2) as: $(1 - q)q_M[1/2 - p\theta] + qq_H[1/2 - \theta/2] = 0$, we see that a high type gains from fighting is $p\theta - 1/2$ when facing a low type, which occurs with probability $(1 - q)q_M$, and its loss for fighting is $1/2 - \theta/2$ with probability qq_H , when the opponent's type is high. Consider now arbitration: The high type faces a low type with probability $1 - q$, and is forced to accept $b = (1 - q_M)p\theta + q_M \cdot 1/2$, whereas it obtains $1/2$ with probability p_H and $\theta/2$ with probability $1 - p_H$ if facing a high type, which occurs with probability q . This is equivalent to saying that, with probability $(1 - q)(1 - q_M)$, the high type's payoff is $p\theta$, which is the same as the payoff from fighting; with probability $(1 - q)q_M$, its payoff is $1/2$ so that its fighting gain is $p\theta - 1/2$, and with probability qp_H its payoff is $1/2$ so that its fighting loss is $1/2 - \theta/2$, as under mediation.

¹⁹The theoretical mechanism design literature does not provide us with a result on the generality of our result beyond the framework we consider. Bester and Strausz (2001) shows that there is no loss by considering standard revelation games, when the mechanism designer can impose transfers to a single agent, but cannot

The result is intuitive. With commitment, the arbitrator’s capability to impose peaceful settlement is an advantage at conflict resolution over mediation. But without commitment, the arbitrator’s enforcement power constrains her capability to resolve conflict more than a mediator’s capability to make non-binding recommendations. Indeed, we find that arbitration without commitment is ineffective in our framework. There does not exist an arbitration mechanism to which the players would agree to participate.

4 Unmediated Peace Talks

We now consider unmediated peace talks. By the revelation principle, unmediated communication cannot improve the peace chance and welfare over mediation. We seek to establish if and when mediation yields a higher welfare. To make the comparison simpler, we study unmediated peace talks with the same game form and equilibria as in the previous section. The only difference is that, now, disputants communicate directly, instead of through a mediator.

Here, we stipulate that there is a peace conference, in which both players $i = A, B$ simultaneously send unverifiable messages $m_i \in \{l, h\}$ to each other. As well as allowing disputants to communicate, the peace conference also provides a public correlation device. With probability $p(m)$, the conference is successful, and leads to a settlement proposal $x(m)$. If both players agree to the proposal, war is avoided. With probability $1 - p(m)$, the peace conference fails, and this leads to open conflict. Hence, we adopt the same game form as before, with the only difference that, here, players communicate directly. We also note that, following Aumann and Hart (2003), the public correlation device in our game can be replicated by an additional round of communication among the players (using so-called jointly controlled lotteries). Hence our game can be reformulated as a two-round communication game without public correlation of play. Again to simplify comparison with the analysis of mediation, we restrict attention to pure-strategy equilibria in which players report truthfully their type, i.e.,

precommit, as long one considers mixed strategy as well as truthful equilibria. But Bester and Strausz (2000) show that with two agents, there can be an improvement if the message space is larger than the agents’ type spaces. There are two agents, here, but our result is proved for any arbitrary message space.

to separating equilibria. Further we focus on equilibria with $x(m)$ and $p(m)$ symmetric across players, where $x(h, h) = x(l, l) = 1/2$, and where we let $\tilde{b} \equiv x(h, l) = 1 - x(l, h)$, $\tilde{p}_L \equiv p(l, l)$, $\tilde{p}_H \equiv p(h, h)$ and $\tilde{p}_M \equiv p(h, l) = p(l, h)$.

The program that calculates the optimal separating equilibrium with unmediated communication determines the values of $\tilde{b}, \tilde{p}_L, \tilde{p}_M, \tilde{p}_H$ that minimize the chance of war, subject to the constraints that both types communicate truthfully (incentive compatibility constraints with double deviations), and agree to each proposal $x(m)$ (*ex post* participation constraints). The detailed program formulation is reported in the Appendix. Solving this program yields the following equilibrium characterization.

Proposition 2 *In the optimal separating equilibrium of the unmediated communication game,*

1. *If $\gamma \geq 1$ and/or $\lambda \geq (1 + \gamma)^{-1}$, then the only binding constraints are the low type's incentive compatibility and the high-type's participation constraints (so that, $\tilde{b} = p\theta$). If $\lambda \leq \gamma/2$, then the equilibrium coincides with the optimal mediation mechanism; whereas if $\gamma/2 < \lambda < \gamma$, then high type dyads settle on $(1/2, 1/2)$ with probability $\tilde{p}_H < q_H$, low type and asymmetric dyads do not fight ($\tilde{p}_M = \tilde{p}_L = 1$), and the chance of peace is strictly lower than under optimal mediation.*
2. *If $\gamma < 1$ and $\lambda < (1 + \gamma)^{-1}$, then the only binding constraints are the incentive compatibility constraints, and $\tilde{b} > p\theta$, and the chance of peace is strictly lower than under optimal mediation (either because $\tilde{p}_M = p_M + q_M = 1$ and $\tilde{p}_H < q_H$, or because $\tilde{p}_M < p_M$ and $q_M = \tilde{p}_H = q_H = 0$).*

Mediation strictly improves welfare upon unmediated communication both when $\gamma < 1$, so that war is expected to be very costly or intense, and when $\gamma/2 < \lambda < \gamma$, so that high types are neither too likely nor too unlikely. In both cases, a mediator's optimal recommendation strategy does not always reveal to a player the opponent's type. Specifically, when $\gamma/2 < \lambda < \gamma$, the optimal strategy does not always reveal the opponent's type to a type H player. When $\gamma < 1$, instead, it does not always reveal the opponent's type to a type L player. Evidently,

these strategies are not feasible when players reveal their type to each other directly, and our unmediated communication set-up yields a lower chance of peace than mediation.

When $\gamma \geq 1$ and $\gamma/2 < \lambda < \gamma$, the infeasibility of the optimal mediation mechanism requires a higher war probability in high type dyads in the optimal separating equilibrium of the unmediated communication game. It is impossible to keep the payoff of a self-reported high type low by recommending the equal split $(1/2, 1/2)$ when it faces a low type. Hence, the incentive for a low type to exaggerate strength needs to be kept under check by increasing the chance of war in self-reported high dyads (i.e., by setting $\tilde{p}_H < q_H = p_H$).

When $\gamma < 1$, the infeasibility of the optimal mediation mechanism in our unmediation communication set up has interesting and unexpected implications. In the optimal separating equilibrium of the unmediated communication game, the double deviation of misreporting one's high type, and then wage war if the opponent reveals to be weak is kept in check by increasing the self-reported high type's share \tilde{b} against a low type above $p\theta$. As a result, the high-type participation constraint does not bind, whereas both types' incentive compatibility constraint bind, which is unusual in standard mechanism design solutions.

We conclude this section briefly discussing the robustness of Proposition 2. First, we extend the comparison between optimal mediation and the optimal separating equilibrium, to the other equilibria of our unmediated communication game (the formal comparison is presented in appendix). It is intuitive that the optimal mediation yields a strictly higher peace probability than the pooling equilibrium in which no information is disclosed in the communication stage, as asymmetric information is the source of conflict, here. And it turns out that also semi-pooling equilibria, in which disputants mix at the communication stage cannot achieve the peace probability of optimal mediation. As is known (Forges, 1986), mixed strategies profiles induce independent random outcomes, and often cannot replicate the correlated randomness determined by communication through a mediator.

Second, we consider the generality of the unmediated communication protocol we study. The pure communication strategy equilibrium outcomes do not change when considering more

rounds of communication, or richer message spaces (and also the focus on symmetric equilibria is without loss of generality). But it is possible that more messages and/or more rounds might help with mixed-strategy equilibria.^{20,21} The online appendix explores the robustness of our results when allowing for richer communication protocols. The suboptimality of unmediated communication relative to mediation persists in our robustness exercises.²²

A more fundamental caveat about Proposition 2 concerns our maintained assumption that mediators are able to commit to quit and induce a conflict escalation, after receiving the disputants reports m . This capability of commitment allows them to maximize the *ex ante* chances of peace. Hence, our mediator must be able to commit to not seek a peaceful agreement in all contingencies. In the online appendix we provide evidence that mediators as well as disputants often recognize the value of this commitment, especially when disputants claim to be strong. But if mediators cannot make these commitments, however, our results argue that they are not more effective than unmediated communication.

5 Concluding Remarks

By applying mechanism design techniques to the study of international conflict resolution, this paper derives novel results comparing arbitration, mediation and unmediated communication. These results may also serve as a benchmark to stimulate future investigations of conflict resolution with mechanism design. Given the sovereignty of disputants, it is often not feasible in international relations to rely on arbitration by a third party with the power to enforce settlements. Nonetheless, we show that a mediator without enforcement power can be

²⁰Aumann and Hart (2003) provide examples of games in which longer, indeed unbounded, communication protocols improve upon finite round communication.

²¹Instead, the restriction to a single peaceful split $x(m)$, for every m , rather than the consideration of a lottery over peaceful splits, can be shown to be without loss of generality even if allowing for mixed communication strategies.

²²In particular, numerical simulations suggest that two rounds of cheap talk with binary messages does not improve on one round. An important *caveat*, however, is that there is no theoretical support for either restricting attention to two (or indeed, finitely many) messages per round, given that incomplete information is two-sided; nor does it follow that more rounds do not help: well-known examples illustrate that the equilibrium payoff set might increase discontinuously once a countable infinity of rounds of cheap talk is permitted.

as effective as an arbitrator in minimizing the chance of war. The mediator can circumvent the constraint that her recommendations be self-enforcing, by using recommendation strategies that do not always reveal to a disputant that the opponent is weak. This mechanism relies on the mediator's capability to gather information from the disputants privately, under a confidentiality agreement. Simple forms of unmediated communication, or of third-party involvement with full transparency (e.g., communication facilitation) cannot always achieve the same outcome, because they rule out confidentiality. Specifically, mediation outperforms unmediated communication in our model, when the expected intensity or cost of conflict are high, and/or when the likelihood that a disputant is strong is intermediate.

References

- [1] Aumann, R.J. and S. Hart (2003): "Long Cheap Talk," *Econometrica*, **71(6)**, 1619–60.
- [2] Banks, J.S. and R.L. Calvert (1992): "A Battle-of-the-Sexes Game with Incomplete Information," *Games and Economic Behavior*, **4**, 347–372.
- [3] Barnett, M.N. and M. Finnemore (2004): *Rules for the world: International Organizations in Global Politics*, Cornell University Press, Ithaca: NY.
- [4] 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.
- [5] Bester, H. and R. Strausz (2000): "Imperfect Commitment and the Revelation Principle: the Multi-Agent Case" *Economics Letters*, **69**, 165-171.
- [6] Bester, H. and R. Strausz (2001): "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case", *Econometrica* **69**, 1077-1098.
- [7] Bester, H. and K. Wärneryd (2006): "Conflict and the Social Contract," *Scandinavian Journal of Economics*, **108(2)**, 231–49.
- [8] Blainey, G. (1988): *The Causes of War*. Third ed. New York: The Free Press.
- [9] Brito, D.L. and M.D. Intriligator (1985): "Conflict, War, and Redistribution," *American Political Science Review*, **79(4)**, 943–957.
- [10] Compte, O. and P. Jehiel (2009): "Veto Constraint in Mechanism Design: Inefficiency with Correlated Types," *American Economic Journal: Microeconomics*, **1**, 182–206.
- [11] Cramton P.C. and T.R. Palfrey (1995): "Ratifiable mechanisms: learning from disagreement," *Games Economic Behavior*, **10**, 255–283.

- [12] Fearon, J.D. (1995): “Rationalist Explanations for War,” *International Organization*, **49:3**, 379–414.
- [13] Fey, M. and C. Ramsay (2009): “Mechanism Design Goes to War: Peaceful Outcomes with Interdependent and Correlated Types,” *Review of Economic Design*, **13(3)**, 233–250.
- [14] Fey, M. and K. Ramsay (2010): “When Is Shuttle Diplomacy Worth the Commute? Information Sharing through Mediation,” *World Politics*, **62(4)**: 529-60.
- [15] Forges, F. (1986): “An Approach to Communication Equilibria,” *Econometrica*, **54(6)**: 1375-1385
- [16] 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.
- [17] Goltsman, M., J. Hörner, G. Pavlov and F. Squintani (2009): “Mediation, Arbitration and Negotiation,” *Journal of Economic Theory*, **144(4)**: 1397-1420.
- [18] Kelman, H.C. (1958): “Compliance, identification, and internalization: Three processes of attitude change,” *Journal of Conflict Resolution* **2**, 51–60.
- [19] Kennan, J. and R. Wilson (1993): “Bargaining with Private Information”, *Journal of Economic Literature*, **31(1)**, 45–104
- [20] Keohane, R. O. (1984): *After Hegemony*, Princeton, NJ: Princeton University Press.
- [21] Kleiboer, M. (1996): “Understanding Success and Failure of International Mediation,” *Journal of Conflict Resolution*, **40(2)**, 360–89.
- [22] Kydd, A. (2003): “Which Side Are You On? Bias, Credibility, and Mediation,” *American Journal of Political Science*, **47(4)**, 597–611.
- [23] Kydd, A. (2006): “When Can Mediators Build Trust,” *American Political Science Review*, **100**, 449–462.
- [24] Matthews, S.A. and A. Postlewaite (1989): “Pre-play communication in two-person sealed-bid double auctions,” *Journal Economic Theory* **48**, 238–263.
- [25] Morgenthau, H. (1948), *Politics Among Nations: The Struggle for Power and Peace*, New York NY: Alfred A. Knopf.
- [26] Myerson, R.B. (1979): “Incentive Compatibility and the Bargaining Problem,” *Econometrica*, **47(1)**, 61–73.
- [27] Myerson, R. B. (1982): “Optimal coordination mechanisms in generalized principal-agent problems,” *Journal of Mathematical Economics*, **10(1)**, 67–81.

- [28] Myerson, R.B. (1985). “Analysis of Two Bargaining Games with Incomplete Information,” in A. Roth (ed.), *Game Theoretic Models of Bargaining*, Cambridge University Press, 115–147.
- [29] Rauchhaus, R. W. (2006): “Asymmetric Information, Mediation and Conflict Management,” *World Politics*, **58**, 207–241.
- [30] Skreta, V. (2006). “Sequentially optimal mechanisms,” *The Review of Economic Studies*, **73**, 1085–1111.
- [31] Waltz, K. (1959): *Man, the State and War: A Theoretical Analysis*. New York: Columbia University Press.
- [32] Wilkenfeld, J., Young, K.J., Quinn, D.M. and V. Asal (2005): *Mediating International Crises*, Oxon, U.K.: Routledge.
- [33] World Trade Organization Legal Texts: “Annex 2: Understanding on the Rules and Procedures Governing the Settlement of Disputes”, <http://www.wto.org/english/docs>
- [34] Young, O.R. (1967): *The Intermediaries: Third parties in International Crises*, Princeton, N.J.: Princeton University Press.

Appendix

Proof of Lemma 1. We prove that the solution of the mediator’s program with enforcement power is such that only the type *L interim* incentive compatibility (*LIC*) constraint and the type *H interim* participation (*HP*) constraint bind. Further, for $\lambda \leq \gamma/2$, $p_M = \frac{1}{\gamma-2\lambda+1}$, $p_H = 0$, and $V = \frac{(\gamma+1)}{(\gamma-2\lambda+1)(\lambda+1)^2}$, whereas for $\lambda \geq \gamma/2$, $p_M = 1$, $p_H = \frac{2\lambda-\gamma}{(\gamma-\lambda+1)\lambda}$, and $V = \frac{\gamma+1}{(\gamma-\lambda+1)(\lambda+1)}$.

We first solve the relaxed program: $\min_{b,p_L,p_M,p_H} W(b,p_L,p_M,p_H)$ subject to the *LIC* and *HP* constraints and to $p_L \leq 1$, $p_M \leq 1$ and $p_H \geq 0$.

We note that $p_L = 1$ in the solution because p_L appears only in the right-hand side (*RHS*) of the *LIC* constraint, which increases in p_L . Then, we note that the *LIC* constraint must bind in the solution, or else one could increase p_H thus reducing the value of the objective function W , without violating the *HP* constraint. Then, we note that the *HP* constraint must bind in the solution, or else one could decrease b and make the *LIC* constraint slack.

Solving for b and p_H as a function of p_M in the system defined by the binding *LIC* and *HP* constraints, and plugging back the resulting expressions in the objective function, we obtain $W = -p_M \frac{\gamma+1}{(\lambda+1)(\gamma+1-\lambda)} + K$, where K is an inconsequential constant. Hence, the probability of war W is minimized by setting $p_M = 1$ whenever possible. Substituting $p_M = 1$, in the system defined by the binding *LIC* and *HP* constraints, we obtain $p_H = \frac{2\lambda-\gamma}{(\gamma-\lambda+1)\lambda}$ and $b = \frac{1}{2} \frac{q-\theta-4p^2\theta^2-2p\theta-3q\theta+4p\theta^2+2q\theta^2+4p^2q\theta^2+4pq\theta-6pq\theta^2+1}{q+\theta-2p\theta-2q\theta+2pq\theta}$. The quantity p_H is strictly positive for

$\lambda \geq \gamma/2$ and always smaller than one. At the same time $1/2 < b < p\theta$ whenever $\lambda \geq \gamma/2$. When $\lambda \leq \gamma/2$, setting $p_M = 1$ makes p_H negative. Hence, we set $p_H = 0$ and obtain $p_M = \frac{1}{\gamma-2\lambda+1}$ and $b = p\theta$ from the system defined by the binding *LIC* and *HP* constraints, and we verify that the quantity p_M is strictly positive and smaller than one when $\lambda \leq \gamma/2$.

The proof is concluded by showing that the solution of the relaxed program calculated above does not violate the type *H interim* incentive compatibility and type *L interim* participation constraints in the complete optimal arbitration program. Indeed, for $\lambda \geq \gamma/2$, we verify that the slacks of these constraints are, respectively $\frac{(1-\theta)(\gamma-\lambda)(\gamma+1)}{2(\gamma-\lambda+1)} > 0$, and $\frac{(\gamma+1)(1-\theta)}{2(\gamma-\lambda+1)} > 0$. Similarly, for $\lambda \leq \gamma/2$, the slacks are $\frac{(1-\theta)(\gamma-\lambda)(\gamma+1)}{2(\gamma-2\lambda+1)(\lambda+1)} > 0$, and $\frac{(\gamma+1)(1-\theta)}{2(\gamma+1-2\lambda)(\lambda+1)} > 0$. \square

The Mediation Program We restrict attention to symmetric mechanisms, 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 this entails 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 $\Pr[b, m_i] \equiv \Pr[h, b, m_i] + \Pr[l, b, m_i]$.

The *H* type and *L* type *ex post* participation constraints are, respectively:

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

Likewise, the *H* and *L* type *interim* incentive-compatibility constraints with double deviation are, respectively:

$$\begin{aligned} qF(0|h, h)\theta/2 + (1 - q)F(0|h, l)p\theta + \int_0^1 b dF(b|h) &\geq qF(0|l, h)\theta/2 + \\ (1 - q)F(0|l, l)p\theta + \int_0^1 \max\{b, \Pr[l|b, l]p\theta + \Pr[h|b, l]\theta/2\} dF(b|l), \text{ and} \\ qF(0|lh)(1 - p)\theta + (1 - q)F(0|ll)\theta/2 + \int_0^1 (1 - b) dF(b|l) &\geq qF(0|hh)(1 - p)\theta + \\ (1 - q)F(0|lh)\theta/2 + \int_0^1 \max\{1 - b, \Pr[l|b, h]\theta/2 + \Pr[h|b, h](1 - p)\theta\} dF(b|h). \end{aligned}$$

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 .

The mediator seeks to minimize the probability of war, $W(F) = (1 - q)^2 F(0|hh) + 2q(1 - q)F(0|lh) + q^2 F(0|ll)$ subject to all the above constraints.

Proofs of Lemma 2 and of Proposition 1. These results follow from this Lemma.

Lemma 3 *The optimal mediation program solution $(b, p_L, q_L, p_M, q_M, q_H)$ is such that: For $\lambda \leq \gamma/2$, further, $q_L + 2p_L = 1$, $b = p\theta$, $q_H = q_M = 0$, $p_M = \frac{1}{1+\gamma-2\lambda}$, $p_L \leq \frac{2\lambda}{(\gamma-2\lambda+1)(\gamma-1)}$ if $\gamma \geq 1$, and $p_L \geq \frac{(1-\gamma)\lambda(\lambda-\gamma)(\gamma+2)}{2\gamma^2(\lambda-\gamma-1)}$ if $\gamma < 1$, and the peace probability is $V = \frac{\gamma+1}{(1+\gamma-2\lambda)(1+\lambda)^2}$; For $\gamma/2 < \lambda < \gamma$, instead, $q_L + 2p_L = 1$, $p_M + q_M = 1$, $b = p\theta$, $q_H = \frac{2\lambda-\gamma}{\lambda(\gamma+1-\lambda)}$, $q_M = \frac{2\lambda-\gamma}{\gamma(\gamma+1-\lambda)}$, $q_L \geq \frac{\lambda(2\lambda-\gamma)}{\gamma^2(\gamma-\lambda+1)}$, $p_L \leq 2\frac{(\gamma-\lambda)(\gamma+2)\lambda}{(\gamma-\lambda+1)\gamma(\gamma-1)}$ if $\gamma \geq 1$, and $p_L \geq \frac{(1-\gamma)\lambda(\lambda-\gamma)(\gamma+2)}{2\gamma^2(\lambda-\gamma-1)}$ if $\gamma < 1$, and the peace probability is $V = \frac{\gamma+1}{(\gamma-\lambda+1)(\lambda+1)}$; For all γ and λ , only the type L incentive compatibility and the type H participation constraints bind.*

Proof. Consider the mechanisms subject to the mediation constraints reported above. The *ex post* participation constraints are stronger than the following (type H and type L , respectively) *interim* participation constraints $\int_0^1 b dF(b|h) \geq \Pr[l, h]p\theta + \Pr[h, h]\theta/2$, and $\int_0^1 b dF(b|l) \geq \Pr[h, l](1-p)\theta + \Pr[l, l]\theta/2$, for all $b \in [0, 1]$, and that the *interim* incentive compatibility constraints with double deviation are stronger than the *interim* incentive compatibility constraints obtained by substituting the maxima in the associated inequalities with their first argument (the *interim* payoff induced by accepting peace recommendations later in the game).

By the revelation principle of Myerson (1979), the optimal *ex ante* probability of peace within the class of mechanisms which satisfy these *interim* incentive compatibility and participation constraints cannot be larger than the probability of peace calculated in Lemma 1. Because these constraints are weaker than the mediation constraints, a mediation mechanism cannot yield a higher peace probability than the one calculated in Lemma 1. Hence, to prove Lemma 3, it is enough to show that the stated expressions for the choice variables $(b, p_L, q_L, p_M, q_M, q_H)$ satisfy the mediation constraints and achieve the same peace probability as in Lemma 1.

Specialized to the mechanisms described by $(b, p_L, q_L, p_M, q_M, q_H)$, the *ex ante* peace probability takes the following form: $V(b, p_L, q_L, p_M, q_M, q_H) = q^2(1 - q_H) + 2q(1 - q)(1 - p_M - q_M) + (1 - q)^2(1 - 2p_L - q_L)$, and the mediation constraints take the following forms. The type H and type L *ex post* participation (HP and LP) constraints are respectively:

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

the H and type L incentive compatibility constraints with double deviations (HIC^* and LIC^*)

are:

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

Plugging in the stated expressions for $b, p_L, q_L, p_M, q_M, q_H$ into equation defining V above shows that our mediation mechanism gives the same peace chance as the optimal arbitration mechanism of Lemma 1.

Importantly, we also verify the equilibrium interim utility of high and low types are the same with the optimal arbitration and mediation mechanisms, because:

$$p_M \cdot p\theta + q_M \cdot 1/2 = \frac{1}{2} \frac{q - \theta - 4p^2\theta^2 - 2p\theta - 3q\theta + 4p\theta^2 + 2q\theta^2 + 4p^2q\theta^2 + 4pq\theta - 6pq\theta^2 + 1}{q + \theta - 2p\theta - 2q\theta + 2pq\theta},$$

where the right hand side is the expression for $b \in (1/2, p\theta)$ for the case of arbitration.

The verification that our mediation mechanism satisfies the constraints above is long and not particularly informative, so it is left for the the online appendix.

This concludes the proof of Lemma 3. Lemma 2 and Proposition 1 follow. \square

The Unmediation Communication Program Simplifying notation, the optimal separating equilibrium with unmediated communication is calculated by the following program. Minimize the war probability

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

subject to the type H and L truthtelling (HIC^* and LIC^*) constraints:

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

and the (HP and LP) constraints that the high type (respectively, the low type) accept the peaceful proposals: $b \geq p\theta$, $1 - b \geq (1 - p)\theta$.

Proof of Proposition 2 The proof follows from the following Lemma.

Lemma 4 *The best separating equilibrium is characterized as follows.*

Suppose that $\gamma \leq 1$. When $\lambda < \frac{\gamma}{1+\gamma}$, both truthtelling constraints bind, $b > p\theta$, $p_H = 0$, $p_M = \frac{1}{(1+\gamma)(1-\lambda)}$, and $V = \frac{1+\gamma+\lambda(1-\gamma)}{(1+\gamma)(1-\lambda)(1+\lambda)^2}$. When $\lambda \in [\frac{\gamma}{1+\gamma}, \min\{\frac{1}{1+\gamma}, \gamma\}]$, both truthtelling constraints bind, $b > p\theta$, $p_M = 1$, $p_H = 1 - \frac{\gamma}{(1+\gamma)\lambda}$, and $V = 1 - \frac{\gamma\lambda}{(1+\gamma)(1+\lambda)^2}$. When $\lambda \in [\frac{1}{1+\gamma}, \gamma)$, the type L truthtelling constraint and the type H proposal acceptance constraint bind, $b = p\theta$, $p_M = 1$, $p_H = \frac{2\lambda-\gamma}{\lambda(2+\gamma)}$, and $V = \frac{2(1+\lambda)+\gamma}{2+\gamma+\lambda(2+\gamma)}$.

Suppose that $\gamma > 1$. When $\lambda < \gamma/2$, the type L truthtelling constraint and the type H proposal acceptance constraint bind, $b = p\theta$, $p_H = 0$, $p_M = \frac{1}{1+\gamma-2\lambda}$, and $V = \frac{1+\gamma}{(1+\gamma-2\lambda)(1+\lambda)^2}$. When $\lambda \in [\gamma/2, \gamma)$, the type L truthtelling constraint and the type H proposal acceptance constraint bind, $b = p\theta$, $p_M = 1$, $p_H = \frac{2\lambda-\gamma}{\lambda(\gamma+2)}$, and $V = 1 - \frac{\gamma\lambda}{(2+\gamma)(1+\lambda)}$.

The proof of lemma 4 proceeds in two parts.

Part 1 ($\gamma \geq 1$). We set up the following relaxed problem:

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

subject to the type L relaxed truthtelling constraint (LIC):

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

to the type H proposal acceptance constraint $b \geq p\theta$ and to the relaxed probability constraints $p_L \leq 1, p_M \leq 1, 0 \leq p_H$.

Step 1. We want to show that $p_L = 1$. We first note that setting $p_L = 1$ maximizes the LHS of the LIC constraint and does not affect its RHS. It is immediate to see that the HP constraint is not affected either.

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

Step 3. We want to show that the *HP* constraint binds. Suppose it does not. Then $b > p\theta$, and it is possible to reduce b without violating the *HP* constraint. But this makes the *LIC* constraint slack, because $-b$ appears in its LHS and b in the RHS, thus contradicting step 2.

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 binding *LIC* and *HP* constraints define the function $p_M(p_H) = \frac{(1-\lambda p_H(\gamma+2))}{(\gamma-2\lambda+1)}$, 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 the war chance of W requires minimization p_H . Setting $p_H = 0$ and substituting it in the above expression for $p_M(p_H)$ yields $p_M = \frac{1}{1+\gamma-2\lambda}$. Because $\lambda \leq \gamma/2$, it follows that $p_M \leq 1$, as required. By substitution, we verify that the peace chance is $V = \frac{1+\gamma}{(1+\gamma-2\lambda)(1+\lambda)^2}$.

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 decreases in p_H in the above expression $p_M(p_H)$, the solution requires setting $p_M = 1$ and $p_H = \frac{2\lambda-\gamma}{\lambda(\gamma+2)}$. When $\lambda \geq \gamma/2$, $p_H \geq 0$ and hence the solution is admissible. By substitution, we verify that the peace chance is $V = 1 - \frac{\gamma\lambda}{(2+\gamma)(1+\lambda)}$.

Step 6. We want to show that the solution constructed above satisfies all the constraints of the unrelaxed program that defines the optimal separating equilibrium of the unmediated communication game. The *LP* 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 *LIC** and *LIC* constraints coincide. The *HP* high-type constraint $1-b = 1-p\theta \leq \theta/2$ 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 $1-b \leq \theta/2$. As a result, the *HIC** constraint becomes: $(1-q)((1-p_M)p\theta + p_M b) + q((1-p_H)\frac{\theta}{2} + p_H\frac{1}{2}) \geq (1-q)((1-p_L)p\theta + p_L p\theta) + q((1-p_M)\frac{\theta}{2} + p_M\frac{\theta}{2})$ which is satisfied because $b = p\theta$. The probability constraints are obviously satisfied.

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

Case 1. Let us consider the following relaxed problem:

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

subject to the type H and L relaxed truth-telling constraints (HIC and LIC), respectively:

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

which embed the assumption (to be verified afterwards) that $1-b \geq \theta/2$, and to the relaxed probability constraints: $p_L \leq 1, p_M \leq 1, 0 \leq p_H$.

Step 1. As in step 1 of part 1, we conclude that $p_L = 1$.

Step 2. We want to show that the LIC constraint binds. Indeed, if it does not, we can increase p_H without violating either relaxed truth-telling constraint (note that the LHS of the HIC constraint increases in p_H).

Step 3. We want to show that the HIC constraint binds. Suppose not. We can then reduce b because the LHS of the HIC constraint increases in b and the RHS decreases in b . This makes the LIC constraint slack, without changing p_M and p_H . But in light of step 2, this cannot minimize the objective function W . Hence, the HIC 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 program. The binding LIC and HIC constraints define the functions: $p_M(p_H) = \frac{(1-\lambda p_H(1+\gamma))}{(\gamma+1)(1-\lambda)}$ and $b(p_H) = \frac{2\lambda+\gamma-\theta\lambda-\theta\gamma-2\lambda p_H+\theta\lambda p_H-3\lambda\gamma p_H+2\theta\lambda\gamma p_H-\lambda^2 p_H-\lambda\gamma^2 p_H-\lambda^2\gamma p_H+\theta\lambda\gamma^2 p_H+1}{2(1-\lambda p_H-\lambda\gamma p_H)(\lambda+1)}$. 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 $p_M(p_H)$ and $b(p_H)$ when $p_H = 0$ we obtain: $b = -\frac{1}{2\lambda+2}(-2\lambda-\gamma+\theta\lambda+\theta\gamma-1)$ and $p_M = \frac{1}{(\gamma+1)(1-\lambda)}$. Because $1 \geq \gamma \geq \lambda$, it is the case that $p_M \geq 0$, 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 verify that the probability of war is $V = \frac{1+\gamma+\lambda(1-\gamma)}{(1+\gamma)(1-\lambda)(1+\lambda)^2}$.

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 unrelaxed program. Again, the LIC^* and LIC constraints coincide. We need to show that the HP constraint $b \geq p\theta$ is satisfied. In fact, simplification yields: $b - p\theta = \frac{1}{2}(\lambda+1)^{-1}(1-\gamma)(1-\theta)\lambda > 0$. Then, we note that $1-b-\theta/2 = \frac{1}{2}(\lambda+1)^{-1}(1-\gamma)(1-\theta)\lambda \geq 0$. As a result, the HIC^* and HIC constraints coincide, and, further, the LP constraint $1-b \geq (1-p)\theta$ is satisfied, because $\theta/2 > (1-p)\theta$.

Step 6. We want to show that for $\lambda \in [\gamma/(1+\gamma), \min\{1/(1+\gamma), \gamma\}]$, the values $p_M = 1$, $p_H = 1 - \frac{\gamma}{(1+\gamma)\lambda}$ solve the relaxed program. When $\lambda > \gamma/(1+\gamma)$, setting $p_H = 0$ violates the constraint $p_M = 1$. Further, p_M decreases in p_H in the above formula $p_M(p_H)$. Hence

minimization of p_H , which induces minimization of the objective function W , requires setting $p_M = 1$. Solving for b and p_H , we obtain: $p_H = \frac{\lambda - \gamma + \lambda\gamma}{(\gamma + 1)\lambda} = 1 - \frac{\gamma}{(1 + \gamma)\lambda}$. 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 $\lambda \in [\gamma/(1 + \gamma), \min\{1/(1 + \gamma), \gamma\}]$, setting $p_M = 1, p_H = 1 - \frac{\gamma}{(1 + \gamma)\lambda}$ solves the unrelaxed program. Again, the LIC^* and LIC constraints coincide. We verify that the HP constraint $b \geq p\theta$ is satisfied, because $b - p\theta = \frac{(\lambda + \lambda\gamma - 1)(\theta - 1)\gamma}{2(\gamma + 1)(\lambda + 1)}$ and this quantity is positive if and only if $\lambda \leq \frac{1}{\gamma + 1}$. We then inspect $1 - b - \theta/2 = \frac{(1 - \theta)(\lambda - \gamma + \lambda\gamma - \gamma^2 + 1)}{2(\gamma + 1)(\lambda + 1)}$ and note that $\lambda - \gamma + \lambda\gamma - \gamma^2 + 1 \geq 0$ 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 $\lambda \geq \frac{\gamma}{\gamma + 1}$. Because $1 - b \geq \theta/2$, the HIC^* and HIC constraints and the LP constraint is satisfied, as $\theta/2 > (1 - p)\theta$.

Case 2. When $\lambda \in [1/(1 + \gamma), \gamma)$, our proof approach is to verify that the solution of the relaxed program considered in Part 1 of this proof (the case where $\gamma \geq 1$), $p_L = 1, p_M = 1, p_H = \frac{2\lambda - \gamma}{\lambda(2 + \gamma)}$, solves the unmediated communication program. Because this verification is not very informative, we relegate it to the online Appendix.

This concludes the proof of the Lemma 4, and hence of Proposition 2. \square

Proposition 3 *All pooling equilibria of our unmediated communication game yield a strictly lower peace probability than the optimal separating equilibrium, and hence than optimal mediation, for $\lambda < \gamma$.*

Proof. In any pooling equilibrium, the splits $(x(m), 1 - x(m))$ are independent of m . When $\lambda < \gamma$, the probability of peace is maximized by setting x so that low type players do not fight, as well as the high type of one of the two players (say player 1). This is achieved by setting $x \geq p\theta$, so that the high type of player 1 agrees to x , and $1 - x \geq (1 - q)\theta/2 + q(1 - p)\theta$, so that the low type of player 2 agrees to x . These two inequalities are both satisfied for some x if and only if $(1 - q)\theta/2 + q(1 - p)\theta + p\theta \leq 1$, i.e., $\lambda \geq \frac{1}{2}(\gamma - 1)$, 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 low types agree to x and high types do not. In sum, the optimal probability of peace in a pooling equilibrium is: $V = (1 - q)^2 = \frac{1}{(\lambda + 1)^2}$ if $\lambda < \frac{1}{2}(\gamma - 1)$, and $V = 1 - q = \frac{1}{\lambda + 1}$ if $\frac{1}{2}(\gamma - 1) \leq \lambda < \gamma$. Direct comparison shows that it is lower than the probability of peace in the optimal separating equilibrium calculated in lemma 4. \square

Proposition 4 *There does not exist any mixed strategy equilibrium of our unmediated communication game that yields the same peace probability as optimal mediation for $\gamma < 1$ or*

$$\gamma/2 < \lambda < \gamma.^{23}$$

Proof. We first consider the case in which $\gamma < 1$, and the optimal mediation mechanism is such that $p_L > 0$, $q_L + 2p_L = 1$ and $0 < q_H < 1$. This mechanism cannot be reproduced with any mixed strategy equilibrium of the unmediated communication game. In fact, $q_H < 1$ implies that there is one message, say h , that is played with strictly positive probability $\sigma(h|H) > 0$ by high type players, and such that war is associated with positive probability $1 - p(h, h) > 0$ to the pair of messages (h, h) . The fact that $p_L > 0$ implies that the low type L cannot play a pure strategy, so that it must be the case that $\sigma(h|L) > 0$. But this fact, together with $1 - p(h, h) > 0$, contradicts $q_L + 2p_L = 1$: with some probability, a low type dyad (L, L) needs to fight, in the hypothesized mixed strategy equilibrium.

Turning to the case where $\gamma > 1$ and $\gamma/2 < \lambda < \gamma$, and the optimal mediation mechanism is such that $q_L + 2p_L = 1$, $0 < q_H < 1$, $0 < q_M < 1$ and $p_M + q_M = 1$. This mechanism cannot be reproduced with any mixed strategy equilibrium of the unmediated communication game. In fact, $0 < q_M < 1$ and $0 < p_M < 1$ imply that the high type H cannot play a pure strategy, and that the pair of messages (l, h) yields the split $(p\theta, 1 - p\theta)$, whereas the pair (l, l) yields the split $(1/2, 1/2)$. At the same time, $p_M + q_M = 1$, $q_L + 2p_L = 1$ and $0 < q_H < 1$ imply that war is only associated with the pair of messages (h, h) , with probability $1 - q_H$. With probability q_H , the pair of messages (h, h) results in the split $(1/2, 1/2)$.

Consider a high type's payoff for sending messages h and l respectively: $U_H(h) = q[q_H \cdot 1/2 + (1 - q_H)\theta/2]\sigma(h|H) + p\theta(1 - \sigma(h|H)) + (1 - q)p\theta$, and $U_H(l) = q[(1 - p\theta)\sigma(h|H) + (1 - \sigma(h|H)) \cdot 1/2] + (1 - q) \cdot 1/2$. We want to show the contradiction $U_H(h) > U_H(l)$ for all q_H and $\sigma(h|H)$. Evidently, $U_H(h) - U_H(l)$ increases in q_H , so we set $q_H = 0$, and note that $U_H(h) - U_H(l) = q\sigma(h|H)[\theta/2 - (1 - p\theta)] + [1 - q\sigma(h|H)][p\theta - 1/2] > 0$, because $p\theta > 1/2$ and because $\gamma > 1$ implies that $\theta/2 > 1 - p\theta$. \square

²³For expositional simplicity, this result is proved in the standard framework of revelation games, where it is assumed that the message space coincides with the type space. Inspection of the proof however, reveals that it holds also if allowing an arbitrary finite message space.

Mediation and Peace

Johannes Hörner, Massimo Morelli and Francesco Squintani

Online Appendix not Submitted for Publication

Proofs Omitted from the Manuscript

Completion of the Proof of Lemma 3. We need to verify that the mechanism with values for $b, p_L, q_L, p_M, q_M, q_H$ provided in Lemma 3 satisfies the mediation constraints. We distinguish the two cases.

Step 1. Suppose that $\lambda < \gamma/2$, so that $q_M = q_H = 0$. After simplification, the LIC^* constraint becomes $q(p_M(1-p\theta) + (1-p_M)(1-p)\theta) + (1-q) \cdot 1/2 \geq (1-q)p_M p\theta + q(1-p)\theta + q(1-p_M)\theta/2$, which binds for $p_M = \frac{1}{1+\gamma-2\lambda}$. Consider the HIC^* constraint $q\theta/2 + (1-q)(p_M b + (1-p_M)p\theta) \geq \max\{(qp_M + (1-q)p_L)(1-b), qp_M\theta/2 + (1-q)p_L p\theta\} + \max\{(1-q)p_L b, (1-q)p_L p\theta\} + \max\{(1-q)q_L \cdot 1/2, (1-q)q_L p\theta\} + q(1-p_M)\theta/2$. 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 $p_L \geq \frac{(1-\gamma)\lambda}{2\gamma} p_M = \frac{(1-\gamma)\lambda(\lambda-\gamma)(\gamma+2)}{2\gamma^2(\lambda-\gamma-1)}$ 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 HIC^* 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 satisfied because $b = p\theta$.

Similarly, we find that the two HP constraints $p_M b \geq p_M p\theta$, 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 LP 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 LIC^* constraint, first. After simplifying maxima, as the low type always accepts the split if exaggerating strength, the LIC^* 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 HIC^* constraint. We proceed in two steps. We first consider the high type's off-path behavior after lying, and show that $(qp_M + (1 - q)p_L) \cdot (1 - b) \leq qp_M\theta/2 + (1 - q)p_Lp\theta$ as long as either $\gamma > 1$ or $p_L \geq \frac{(1-\gamma)\lambda}{2\gamma}p_M^+ = \frac{(1-\gamma)\lambda(\lambda-\gamma)(\gamma+2)}{2\gamma^2(\lambda-\gamma-1)}$ for $\gamma < 1$, that $(1 - q)p_Lb = (1 - q)p_Lp\theta$ and that $(qq_M + (1 - q)q_L)1/2 \leq qq_M\theta/2 + (1 - q)q_Lp\theta$ as long as $q_L \geq \frac{1-\theta}{2p\theta-1}\frac{q}{1-q}q_M$, i.e., $q_L \geq \frac{\lambda}{\gamma}q_M = \frac{\lambda(2\lambda-\gamma)}{\gamma^2(\gamma-\lambda+1)}$.

Then we verify that the consequentially simplified HIC^* 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 HP constraints $p_Mb \geq p_Mp\theta$, and $(qq_H + (1 - q)q_M) \cdot 1/2 \geq qq_H\theta/2 + (1 - q)q_Mp\theta$ are satisfied with equality when substituting in the expressions for $b = p\theta$, $q_H = \frac{2\lambda-\gamma}{\lambda(\gamma+1-\lambda)}$, $q_M = \frac{2\lambda-\gamma}{\gamma(\gamma+1-\lambda)}$.

We then show that the LP are satisfied. In fact $p_Lp\theta > p_L\theta/2$, 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 $\gamma < 1$, $p_L \geq 0 \geq 2\frac{(\gamma-\lambda)(\gamma+2)\lambda}{(\gamma-\lambda+1)\gamma(\gamma-1)}$.

We then see that 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 $2\lambda - \gamma \geq 0$. \square

Case 2 of Part 2 of the Proof of Lemma 4. We want to show that, when $\lambda \in [1/(1 + \gamma), \gamma)$, $p_L = 1$, $p_M = 1$, $p_H = \frac{2\lambda-\gamma}{\lambda(2+\gamma)}$, solve the unmediated communication program.

Consider the same relaxed program 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 program 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 of the relaxed program 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 unmediated communication program. Again, the LIC^* and LIC constraints coincide. Consider the HIC^* 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 HIC^* constraint becomes: $(1 - q)((1 - p_M)p\theta + p_Mp\theta) + q((1 - p_H)\frac{\theta}{2} + p_H\frac{1}{2}) - (1 - q)p\theta - q((1 - p_M)\frac{\theta}{2} + p_M(1 - p\theta)) \geq 0$ and indeed, after simplification, the LHS equals: $\frac{(\lambda+\lambda\gamma-1)(1-\theta)\gamma}{2(\gamma+2)(\lambda+1)}$, 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. \square

Evidence on Commitment

One of the main assumptions of our model is the mediator’s commitment to lead to a conflict escalation if this is prescribed by the optimal solution. In the real world, this translates into the mediator’s commitment to quit and terminate the mediation, at least with some probability. In this section we provide two types of evidence that mediators in the real world are capable of keeping this commitment, and even that disputants prefer having a mediator with a credible reputation of being able to keep this commitment. The first type of evidence is from a data set that has settlement attempts as units of observation, namely the Issue Correlates of War dataset (ICOW), by Hensel et al (2008). The second type of evidence will be from an illustrative case study.

In fact, mediators often make clear to the disputants under which circumstances they will quit. Such contingent plans of action often include deadlines. 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.²⁴ Meanwhile, institutions like the United Nations increasingly set time limits to their involvement upfront (see, for instance, the U.N. General Assembly report, 2000).

Evidence from settlement attempts data

The ICOW data set considers each settlement attempt as a distinct observation.²⁵ The key source of variation is whether mediators resume the mediation process through a new settlement attempt following failure or whether they withdraw for good. The theory assumes the mediator’s commitment to withdraw from mediation, expecting that this condition is essential to successful mediation. Mediation attempts in the presence of commitment to quit

²⁴Bebchik (2002) describes 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.

²⁵We use ICOW instead of other datasets because it’s level of analysis is appropriate for our purposes. The International Crisis Bargaining dataset (ICB), Brecher and Wilkenfeld (1997 and 2000) for example, would not be useful because it collapses all settlement attempts for each crisis into one observation.

should be expected to be more successful, and mediators with such a reputation more sought after to serve as a mediator compared to mediators that cannot commit to quit. We probe the ICOW data for information that is consistent with these two hypotheses.

To ascertain whether these patterns exist, we construct rough indicators of mediation termination and differentiate cases according to the circumstances under which they terminated. The settlement attempt may end with 1) an agreement fully implemented; 2) an agreement reached but not ratified by at least one party; 3) an agreement reached and ratified but not complied with by at least one party; 4) no agreement reached and/or escalation of violent hostilities.²⁶ We say a mediator has “quit” if the mediator does not initiate a new mediation attempt involving the same two primary disputants within at least 10 years of the end of the mediation failure. We consider this a “responsive quit” if the settlement attempt ended under any conditions 2–4. A non trivial proportion of failed mediation attempts satisfy this criteria and a significant portion of the mediators in the sample will have quit at least once. Mediators who have established a reputation for actually following through on commitments to withdraw from the process in response to disputant transgressions should be considered more effective mediators and therefore should be more likely to be invited as a mediator in future settlement attempts.

Variable	frequency	Rel. Frequency (total)	Rel. Frequency (failed)
quit1	38	28.4% (134 total)	41.8% (91 total)
quit2	39	29.1% (134 total)	41.5% (94 total)

Table 1 describes how frequently mediators quit in the sample of 134 mediation attempts present in the ICOW dataset. The relative frequency of quitting in the panel varies slightly based on the definition of quitting used, which vary by how strict the criteria are for coding the result as a responsive quit, but overall the data suggest that approximately 30% of the mediation attempts resulted in the mediator quitting by our definition. Of failed mediation attempts, mediators quit responsively in approximately 42% of cases. It is clear that mediators do quit in a significant portion of cases in which quitting is called for.

Variable	frequency	Rel. Frequency (total)
Quit1	116	86.6% (134 total)
Quit2	116	86.6% (134 total)

²⁶We use two alternative definitions of agreement, leading to quit1 and quit2 variables, but they lead to very similar results.

Table 2 demonstrates that an overwhelming majority of mediators in the sample quit the process (by our definition) at least once in their career. 116 of the 134 mediation attempts in the data were mediated by a third party that had quit at least once previously in the panel. This suggests mediators that demonstrate they will withdraw in response to belligerent transgressions are those asked again to mediate crises. The imprecision of the measurement and the limited variation does not allow us to claim that this is airtight evidence that the mediator's capability to credibly threaten to quit when appropriate is a key factor in determining the mediator selection, but the evidence is at least consistent with the general perception that the ability to commit we assume in our model is an important characteristic of real world mediation.

Illustrative Case Study: Kofi Annan in Kenya and Syria

When the contested December 2007 Kenyan Presidential election resulted in widespread ethnic violence, threatening to escalate to full-scale civil war, the international community rushed to establish peace-making efforts. After weeks of multiple uncoordinated mediation attempts that failed to bring the disputants closer to a negotiated settlement, the African Union appointed former UN Secretary-General Kofi Annan to head the Panel of Eminent African Personalities, a coalition of political actors to bring about a solution to the escalating violence. The most important pillar of Annan's mediation strategy was the consolidation of the disparate mediation efforts and interested third parties behind a single process. He demanded, and received, assurances that the international community back the Panel's efforts and refrain from pedaling alternative or competing mediation attempts. Throughout the negotiations, Annan constantly devised signals to communicate the lack of alternative negotiating channels to the disputants in order to ensure their engagement with the Panel's efforts and proposals.

While working hard to overcome the challenges associated with bridging the bargaining gap separating Kibaki and Odinga, Annan also made a point early on to communicate to the disputants that the Panel would not remain involved indefinitely. As the pace of progress slowed, Annan reiterated that he 'would not be available forever and that an alternative had to be found' (Lindenmayer and Kaye, 2009). When the two sides began stalling for time, refusing to make meaningful concessions on the political issue, Annan announced a suspension of the talks. In an interview with the Centre for Humanitarian Dialogue, Griffiths (2008), Annan described the motivation for his decision:

"...Look, this is getting nowhere, so I've decided I'm suspending talks." ... "I had done my duty and believed the leaders should do theirs."

Annan clearly used the suspension of the mediation process, and the threat of its termination, to coerce the parties into giving up more than they hoped at the bargaining table.

He even went so far to strengthen his signal by suggesting his successor, Cyril Ramaphosa of South Africa. Kibaki and the PNU vehemently rejected this proposed replacement, claiming Ramaphosa to be partial to Odinga’s claim. But nevertheless, the signal seems to have been perceived as credible, as the disputants increased the urgency with which they engaged with the Panel’s peace-making process (Lindenmayer and Kaye 2009).

Annan’s commitment to withdraw from mediation in the event the parties refused to abide by his negotiating plan was a calculated complement to his demand to consolidate the international community behind one mediation process. These two tactics served the strategy to remove the disputants’ ability to shop around for alternative mediators and conflict resolution avenues, leaving them with the option of following Annan’s lead or continuing a fight neither side wanted. This strategy fit the crisis well because it was clear that both parties wanted to end the conflict, but that the possibility of alternative means to attain their political goals tempted them away from conceding too much at the negotiating table.

Despite a mixed record of success (failure in Rwanda and Bosnia; success in Kenya), Annan was asked again to step into the mediator’s role as Special Envoy for the U.N and the Arab League in Syria’s escalating civil war. One of the prominent reasons cited for his appointment was his recent success in Kenya. But, unlike in the Kenya case, in Syria Annan faced a deeply divided international community. Though nominally unanimous in support of Annan’s role and his six-point peace plan, the competing world powers almost immediately took actions to undermine the mediation process. The U.S. and Western allies began arming rebel forces and continued to call for Assad to step down while Russia and China continued to prop up Assad’s regime. Without the international community united behind the process, the situation deteriorated. Annan quit in response to continued escalation of violence and both parties’ violation of the six-point plan. As such, Annan maintained his credibility to commit to withdraw from his role in the event disputants become uncooperative.

Mediation and Arbitration Without Commitment

We here explore the conflict resolution capabilities of mediators who pursue peace at all costs in all circumstances. To provide a framework to our analysis, we focus on truthful revelation mechanisms, as in the analysis of the mediators with commitment of section 3. Each disputant i privately sends a message $m_i \in \{h, l\}$ to the mediator, who then makes a (public) recommendation. The only difference with respect to section 3 is that, here, the mediator does not have any commitment capability: given her (equilibrium) beliefs with respect to the types of disputants as a function of the received message, she makes recommendations that induces beliefs which leads to the equilibrium that maximizes the chance of peace.

We argue that such mediators are ineffective. Specifically, we first show that there do not exist equilibria in which the players reports their types truthfully and agree to the mediator's recommendations. Hence, mediators who seek peace in all circumstances cannot improve upon unmediated communication by means of standard truthful revelation mechanisms. This result follows from Theorem 2 by Fey and Ramsay (2009), but because the setting is slightly different, we present a proof.

The first step is to note that if there were an equilibrium, it would be one in which peace is achieved with probability one. In fact, because the players would truthfully reveal their types, the mediator can achieve peace with probability one by making the proposals $(p\theta, 1 - \theta)$ after reports (h, l) , $(1 - \theta, p\theta)$ after reports (l, h) , and $(1/2, 1/2)$ otherwise.

Because the equilibrium achieves peace with probability one, the high-type's incentive compatibility constraint with double deviation reduces to:

$$\int_0^1 b dF(b|h) \geq \int_0^1 \max\{b, \Pr[l|b, l]p\theta + \Pr[h|b, l]\theta/2\} dF(b|l), \quad (3)$$

and the low-type's incentive compatibility constraint with double deviation reduces to:

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

These two constraints are stronger than the following high-type and low-type incentive compatibility constraints without double deviation:

$$\int_0^1 b dF(b|h) \geq \int_0^1 b dF(b|l) \text{ and } \int_0^1 b dF(b|l) \geq \int_0^1 b dF(b|h),$$

which evidently, can only be satisfied if

$$\int_0^1 b dF(b|h) = \int_0^1 b dF(b|l) \equiv B \leq 1/2.$$

Let us turn to the *ex post* participation constraint for the high type. For any $b \in (0, 1)$ such that $\Pr[b, h] > 0$, this constraint is:

$$b \geq \Pr[l|b, h]p\theta + \Pr[h|b, h]\theta/2, \quad (5)$$

Integrating this constraint, we obtain:

$$1/2 \geq B = \int_0^1 b dF(b|h) \geq \int_0^1 (\Pr[l|b, h]p\theta + \Pr[h|b, h]\theta/2) dF(b|h),$$

or, with a minor notational violation,

$$\begin{aligned}
1/2 &\geq \int_0^1 \left[\frac{f[l, b|h]}{f(b|h)} p\theta + \frac{f[h, b|h]}{f(b|h)} \theta/2 \right] f(b|h) db \\
&= \Pr[l|h] p\theta + \Pr[h|h] \theta/2 \\
&= q\theta/2 + (1 - q) p\theta
\end{aligned}$$

which contradicts

$$q\theta/2 + (1 - q) p\theta > 1/2, \text{ i.e., } \lambda < \gamma.$$

We have shown that there do not exist equilibria in which the players reports their types truthfully and agree to the mediator's recommendations. Hence, mediators who seek peace in all circumstances cannot improve upon unmediated communication by means of standard truthful revelation mechanisms.

We now turn to consider arbitration without commitment. For comparability with mediation without commitment, we set the analysis in a revelation game. The players choose whether to participate and then report their types to the arbitrator, who imposes peaceful splits. It is easy to show that there cannot exist any equilibrium in which the players agree to participate and play any strategy (including mixed strategies) at the report stage of the game.

In fact, arbitration without commitment establishes peaceful splits $(b, 1 - b)$ for any pair of messages (m_1, m_2) , according to the symmetric distribution $F(b, 1 - b|m_1, m_2)$. Let $F(b|m)$ be the distribution of either player i 's split b , given i 's message. In equilibrium, letting σ_t be the mixed strategy of a player of either type t , and any message $m \in \text{Supp}(\sigma_t)$ and it must be that:

$$\int_0^1 b dF(b|m) \geq \int_0^1 b dF(b|m'),$$

for any other message m' . As in the above argument, these inequalities can only be simultaneously satisfied if $\int_0^1 b dF(b|m) = \int_0^1 b dF(b|m') \leq 1/2$ for all m and m' in $\text{Supp}(\sigma_t)$. But then, the interim participation constraint of the hawk fails, because, again

$$q\theta/2 + (1 - q) p\theta > 1/2, \text{ when } \lambda < \gamma.$$

Hence, the finding of Proposition 1 is confirmed for the case of third party intervention without commitment. The arbitrator is no more effective than the mediator, even in absence of commitment.

Bibliography

1. Bebachik, B. (2002): “The Philosophy and Methodology of Ambassador Dennis Ross as an International Mediator,” *International Negotiation*, **7**, 115–131.
2. Brecher, M., and J. Wilkenfeld (2000): *A Study of Crisis* Ann Arbor, University of Michigan Press.
3. Curran, D. and J.K. Sebenius (2003): “The Mediator as Coalition Builder: George Mitchell in Northern Ireland,” *International Negotiation*, **8(1)**, 111–147.
4. 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.
5. Griffiths, M. (2008): “The Prisoner for Peace: An Interview with Kofi A. Annan,” *Centre for Humanitarian Dialogue*, Available from <http://www.reliefweb.int/rw/RWFiles2008.nsf/FilesByRW7QMD4C>
6. Hensel, P.R., S. McLaughlin Mitchell, T.E. Sowers II, and C.L. Thyne (2008): “Bones of Contention: Comparing Territorial, Maritime, and River Issues.” *Journal of Conflict Resolution* 117-143.
7. Lindenmayer, E. and J.L. Kaye (2009): “A Choice for Peace? The Story of Forty-One Days of Mediation in Kenya,” *International Peace Institute*, Available from <http://www.ipacademy.org/publication/policy-papers/detail/221-a-choice-for-peace-the-story-of-forty-one-days-of-mediation-in-kenya.html>
8. Watkins, M. (1998): “Building Momentum in Negotiations: Time-Related Costs and Action-Forcing Events,” *Negotiation Journal*, **14(3)**, 241–256.