Mediation in the Shadow of an Audience: How Third Parties Use Secrecy and Agenda-setting To Broker Settlements

Shawn L. Ramirez*

Abstract

How does mediation work? When a leader faces domestic pressure in negotiating with an enemy, mediation with *secrecy and agenda-setting* helps by *reducing uncertainty* about enemy resolve and *locking in concessions.* As a result, mediation improves the prospects for peace, and should talks fail, the leader and her audience are more likely to win in any ensuing war. However, mediation involves costlier settlements. The theory holds implications for mediation, audience costs, and democracies in showing that when a leader faces domestic pressure, her enemy *with no audience costs* can demonstrate resolve credibly through a mediator. For the delegation literature, this research shows that when a principal faces external pressures, she can reduce her risk through an uninformed agent who, through secrecy and discretion over outcomes, can obtain credible information from an adversary.

Words: 9,017

*The author thanks Jeffrey Arnold, Brianna Bardos, Kyle Beardsley, Joseph Bongiovi, Margit Bussmann, Cliff Carrubba, Mary Beth Chopas, Kevin Clarke, Michael Colaresi, Dana D'Amelio, David Davis, Jacob Eisenstein, Mark Fey, Richard Frank, Hein Goemans, Lindsay Heger, Gretchen Helmke, Patrick Kuhn, Danielle Jung, Bethany Lacina, Vanessa Lefler, Eleonora Mattiacci, Richard Nielsen, John Patty, Eric Reinhardt, Dan Reiter, Kenneth Schultz, Curt Signorino, Beth Simmons, Jessica Weeks, the editors and anonymous reviewers comments and questions. Also, Rachel Harmon and Susanna Chi for their invaluable research assistance. The author is responsible for any errors.

Introduction

In December 1991, North and South Korea appeared to reach a conclusive end to four decades of hostility: they joined the United Nations, renounced the use of armed force, and signed a mutual pledge to never develop nuclear weapons. But the situation took a sudden turn for the worse. Between May '93 and June '94, North Korea successfully tested a midrange missile, cruise missiles capable of sinking ships within a 100-mile range, and expelled IAEA inspectors, signaling their intentions to divert fuel from their power program to create nuclear weapons. With the US's nuclear umbrella and deterrent capabilities in jeopardy, Clinton increased troops in South Korea and threatened economic sanctions. In response, Kim Il Sung threatened to "turn Seoul into a sea of flames."[1] Seeing war as increasingly likely, the US began to consider air strikes to destroy North Korea's nuclear reactor, even knowing this might provoke a North Korean invasion of the South.

Whilst deliberating in the White House Cabinet room, Clinton received a phone call. Former President Jimmy Carter had independently begun a mediation, and was notifying them that Kim agreed to freeze the nuclear program – "there'd be no reprocessing, no separation of plutonium, and we could go back to the negotiating table."[2] Within months, mediation produced a breakthrough pact: beyond halting the nuclear program, North Korea would dismantle its nuclear complex and allow international inspections of two secret military sites. In exchange, they would receive oil and light-water reactors that were less threatening and facilitated increased international monitoring. How did mediation halt this collision course to war?

This exemplifies a situation often encountered by leaders faced with a foreign enemy. First, *war would be costly* for both sides: even if the US destroyed the nuclear complex, the

[1]US or UN sanctions could be considered an act of war in and as a treaty violation since the Korean War was fought by American-led UN forces.

[2]Pbs.org. 2016c.

subsequent loss of civilian and military lives on the Korean peninsula would be tragic.[3] For North Korea, the conditions were poor even absent war with an economy on the verge of collapse, food and fuel shortages, and declining support from Soviet and Chinese allies. Surely, some concession could act as a countermand to quash North Korean nuclear aims, but *uncertainty* about North Korean resolve and Clinton's *domestic pressure* makes this complicated. Was North Korea strongly resolved and willing to risk war to become a nuclear power, or less resolved and willing to relinquish its aims for a token concession? In an ideal setting, Clinton might answer this by making some token offer, and using Kim's response to gauge whether raising that offer would be necessary. But raising that offer in the face of enemy resistance would invite sanctioning for weak, incompetent leadership. These three factors – a costly war, uncertainty, and domestic politics – put leaders in the proverbial 'tight spot' in crises: uncertainty makes it necessary to probe for acceptable bargains, but domestic politics turns any guessing game into a political endgame.

This article shows that mediation with secrecy and agenda-setting can help a leader in this situation through two mechanisms. First, a mediator can *lock in concessions* that all parties accept where a leader negotiating independently would risk war. This is because while a leader's domestic pressure forces a trade off that makes her accept some risk of war, a mediator can issue proposals solely based on what improves the prospects for peace. Second, a mediator can reduce uncertainty by making a small *screening offer* that only an enemy with low resolve will accept. This allows the leader to infer that

---

[3] "We were also confident in 1994 – and I'm sure we're very confident today – that we would, within just a few weeks, destroy North Korea's armed forces if they started that war, and we would destroy then their regime. We reckoned there would be many, many tens of thousands of deaths: American, South Korean, North Korean, combatant, non-combatant. So the outcome wasn't in doubt. But the loss of life in that war – God forbid that kind of war ever starts on the Korean Peninsula. The loss of life is horrific." Pbs.org. 2016a.

any enemy remaining in mediation must have higher resolve, thereby warranting greater concessions that the leader's audience permits.

As a result, leaders obtain peace with a greater probability when wars are more costly, the enemy is likely to have high resolve, and the problem of uncertainty is not as large. Further, when the problems caused by uncertainty are large, such that knowledge of the enemy's resolve would make a significant difference, mediation can reduce that uncertainty by winnowing away low resolve types. Thus, mediation brings two main benefits at the cost of a higher settlement. Mediation reduces the *ex ante* probability of war, and if talks fail, then mediation raises the probability that the leader and her audience will win in any ensuing war.

This research advances three streams of literature. First, this provides a novel explanation for how mediation succeeds without carrots, sticks, or independent information through its *process* (Beardsley 2013; Favretto 2009; Fey and Ramsay 2010; Kleiboer 2002; Kydd 2003; Rauchhaus 2006; Savun 2008; Smith and Stam 2003; Zartman 2008). Since any mediator with secrecy and agenda-setting can make peace more likely, this provides the first rational explanation for why increased communication, track-two diplomacy, and weak mediators succeed (Bercovitch and Gartner 2006; Böhmelt 2010; Fey and Ramsay 2010; Wallensteen and Svensson 2014; Beardsley 2009). The theory helps to explain what it means for a conflict to become ripe, since the results imply that mediators are more likely to lock in concessions if the costs of war are higher or sufficient uncertainty is lower over time. The results clarify how mediation helps when leaders need "face-saving" – by locking in concessions and reducing uncertainty (Allee and Huth 2006; Beardsley 2010; Beardsley 2011; Gent and Shannon 2010; Huth, Croco, and Appel 2011; Simmons 2002). Further, this research gives novel implications for multi-mediator episodes, and the complementary effects for powerful and informed mediators.

Second, the results speak to the literatures on audience costs and democracies, since the game involves one-sided audience costs, and thus mediation allows an enemy *with no audience costs* to demonstrate resolve credibly. This contributes to audience cost theories that typically enhance a leader's *own* credibility (Fearon 1994; Schultz 2012; Slantchev

2006; Weeks 2012). For democracies, this implies a double-edged sword. On the one hand, democratically-elected leaders might use mediation to learn about their enemies more efficiently, by resolving their conflicts through diplomacy that averts the need for a costly war. On the other hand, autocratic leaders might be more likely to escalate crises to prompt a mediation that can shift the status quo. This double-edged sword provides a consistent explanation for why we observe a democratic peace, institutions that follow democratic norms of conflict resolution, and dictatorships that escalate crises only to return to mediation.

Third, this research connects the political economy literature on delegation to international relations. The model bears resemblance to traditional models of delegation, wherein an agent acting on behalf of a principal can reduce the risk of worse outcomes by providing information (Epstein and O'Halloran 1994; Bendor, Glazer, and Hammond 2001; Bendor and Meirowitz 2004; Gailmard and Patty 2012; Gailmard and Patty 2013; Huber and Shipan 2006).[4] However, to our knowledge, this is the first model to show that how information is obtained by an uninformed agent endogenously when a principal faces external pressure. Specifically, the model shows that when the principal faces external pressure, an uninformed agent operating with considerable independence becomes empowered to extract information from an adversary; and further, the amount of information obtained is increasing in the amount of external pressure. Since principals in other contexts are likely to face adversarial relationships and external pressures – such as between an executive and legislature with voters, or a union and firm with stakeholders – the model here suggests novel considerations for delegation to uninformed agents with discretion over outcomes.

[4] We leave the central questions asked by the delegation literature – of when to delegate, and which mediators are the best delegates – to future research, since these choices can relay information and change bargaining in ways beyond the scope of this paper.

1 Mediation

Why is mediation important, and how does it succeed? Since 1945, mediators have intervened in over 70% of conflicts, achieving ceasefires and settlements with an estimated 35% success rate (Bercovitch 1996; Greig and Diehl 2012; Melin 2013).[5] Globally, mediation is the most prevalent form of conflict management; more frequent than adjudication or arbitration in both interstate and civil conflicts; and, as scholars argue, may be selected into to manage conflicts that are the most violent (Bercovitch and Gartner 2006; Gartner 2013; Greig 2005; Ruhe 2015).

Research establishes three primary mechanisms to explain how mediation succeeds. First, *powerful* mediators with bombs, sanctions, aid, and credible threats can push intransigent parties toward agreement, threaten to intervene, and enforce post-settlement outcomes (Beardsley 2013; Favretto 2009; Kleiboer 2002; Smith and Stam 2003; Svensson 2007; and Zartman and Touval 1985). Second, *informed* mediators with access to advanced intelligence who credibly convey that information can steer parties toward settlement by providing a 'reality check' to correct misperceptions, reducing incentives to bluff, and helping disputants overcome psychological biases (Fey and Ramsay 2010; Kydd 2003; Rauchhaus 2006; Savun 2008). Third, conflicts may become *ripe* for a mediator's timely intervention: when parties find themselves in protracted and deadly conflicts – a "mutually hurting stalemate" - unfavorable expectations make parties more susceptible to mediator leverage (See Greig and Diehl 2006; Greig 2013; and Regan and Stam 2000 on timing. See Kleiboer 1994; Zartman 2000; and Zartman and Touval 1985 on ripeness.). In short, war widens the bargaining surplus to make room for settlement, or powerful and informed mediators alleviate commitment problems and uncertainty.

Research also finds evidence that domestic politics are important. Scholars argue that

[5] Bercovitch and Langley (1993) find a 29% success rate. Dixon (1996) finds a 40% success rate. Regan, Frank, and Aydin (2009) find that 57% of all civil war mediations result in a ceasefire.

mediation provides political cover that enables leaders under pressure to enter mediation and accept mediated agreements (Beardsley 2010; Beardsley 2011; Beardsley and Lo 2013; Brown and Marcum 2011). Beardsley (2010) finds that mediation is more likely for leaders with *domestic audience costs*, i.e., when a leader will be punished for backing down to an enemy. Beardsley and Lo (2013) find that audience costs make asymmetric concessions more likely in mediation. Melin (2013) notes that mediation's face-saving helped leaders accept "unacceptable terms in Sinai (1974), El Salvador (1988), and Mozambique (1992)" (87). Zartman (2002) discusses how mediation via assembly meetings reduce conflict by furnishing opportunities for leaders "to meet without loss of face" (84). Novak (2009) argues that face-saving ended the 1979 conflict in Zimbabwe because leaders could use the mediator as a scapegoat: engaging in sharp public confrontation and conciliating in private 'shadow' negotiations (149). These arguments suggest that mediation helps by protecting leaders from domestic audiences through *procedural* aspects: *secrecy* to facilitate shadow negotiations, and *agenda-setting* so that the audience can blame the mediator for the settlement.

Relatedly, scholars find evidence that mediation succeeds in ways that cannot be explained by existing theories. Mediators who facilitate communication works, despite the lack of a rational explanation for why simply increasing communication should fare better than doing nothing at all (Bercovitch and Gartner 2006; Fey and Ramsay 2010; Wallensteen and Svensson 2014). Böhmelt (2010) finds that track-two diplomacy works, in which citizens and community members voice their opinions to mediators. Beardsley (2009) finds that weak mediators are used often; hypothesizing that this might be because weak mediators are in large supply. These mediations might also entail these helpful *procedural* aspects. Communication-facilitation can involve secrecy and agenda-setting if mediators propose offers and shuttle between parties. Track-two diplomacy can increase the opacity of mediation making it easier to maintain secrecy. Weak mediators might rely on secrecy and agenda-setting given their inabilities to wield power-based conflict management options.

It behooves one to note that focusing on secrecy and agenda-setting is new, despite a long tradition of specifying and categorizing mediation procedures, since a unique feature of mediation is its permissive environment. Wall (1981), for example, identified over 100 mediation procedures that Wall, Stark, and Standifer (2001) groups into three categories, those that affect: individual disputants; relations between disputants; and relations with the mediator. Kressel (1972) was the first to categorize mediation procedures as reflective (to identify issues), non-directive (to shape the climate of negotiations), and directive (to manipulate outcomes). Several of the most widely used categorizations followed suit: the International Conflict Management dataset classifies mediation as communication-facilitation, procedural, or directive; the International Crisis Behavior Project uses facilitative, formulative, or manipulative; and Zartman and Touval (1985) classify mediators as communicator, formulator, or manipulator.

In contrast, secrecy and agenda-setting are concepts used by the political economy literature to show how *procedural rules* governing political institutions (frequently, legislatures) influence *political behavior* that in turn make certain outcomes permissible or impermissible.[6] For example, Shepsle (1989) shows how the agenda-setting power of Congressional committees shapes legislative voting behavior in structure-induced equilibrium (Cox and McCubbins 2005; Romer and Rosenthal 1978; Shepsle and Weingast 1994). Stasavage (2004) shows that secrecy helps European parliamentary ministers make deals across the aisle by altering their strategic behaviors.[7] Given these results that agenda-setting affects whether policies are accepted, and secrecy can help political representatives

[6] This is distinct from the above classifications: for example, an agenda-setter can be considered loosely as both a formulator and manipulator, but a manipulator may also wield sanctions or offer aid. Similarly, secrecy can be involved in communication or formulation.

[7] He contrasts two models of open and closed door bargaining to show that representatives biased in favor of an opposing constituency use secrecy to obtain their own, and not their public's, most preferred outcomes. That outcome does not apply here since a

broker deals with adversaries, there is reason to believe that an agenda-setting mediator bargaining in secret can help a leader faced with domestic pressure in bargaining with a foreign enemy.

Why would a leader delegate bargaining to a mediator? A political leader faces considerable domestic pressure in international crises. Domestic audiences prefer to maintain a strong national reputation to enhance deterrence (Smith 1998). In audience cost theories, a leader who makes an initial offer and then backs down in the face of enemy resistance will be sanctioned domestically (Fearon 1994). In principal-agent theories, a domestic audience will sanction a political representative to obtain its best outcome given its constraints (Gailmard 2012). In general, these theories assume that the audience understands that the leader is an agent bargaining on their behalf, and knows that they have the power to sanction the leader (Fearon 1999; Ferejohn 1986; Miller 2005).[8] Here, given the audience's uncertainty about the policy-making environment (common priors about the enemy's resolve) and the bargaining that occurs within a mediation (beliefs about whether the leader raised or did not), the audience uses the policy outcome (the

---

leader could not be biased in favor of an enemy's constituency.

[8]As in models of American politics, we do not model *incumbent types* (see Ferejohn 1986). This means that the settlement (policy outcome) must surpass a certain threshold for the leader to avoid sanction. This threshold arises endogenously since the audience is modeled as a rational actor. The audience's threshold for sanction depends on its common prior about the enemy's resolve, such that when resolve is likely to be high, the audience knows that high concessions are likely merited; when resolve is likely to be low, the threshold for acceptable concessions shifts. The result of this modeling decision means that the audience is focused on a moral hazard problem (not an adverse selection problem): all politicians have the same preferences and abilities, and it is up to the audience to police its leader by constraining her to do what is in their best interest.

settlement) to infer whether to sanction its leader.

To see how this affects mediation, this next section takes a standard crisis bargaining model, commonly used to examine escalation between a leader and a foreign enemy, and adds a mediator and a domestic audience (Banks 1990; Levenotoglu and Tarar 2005; Morrow 1989; Powell 1987). As in the US-North Korean crisis, all actors are uncertain about the enemy's resolve (except the enemy). *Resolve* determines the minimum concessions necessary for peace such that a high resolve enemy requires greater concessions. A mediator *sets the agenda* by issuing an initial proposal, and then shuttles from the enemy to the leader.[9] To incorporate *secrecy*, all mediated bargaining is private in the sense that only the leader, mediator, and enemy are direct participants. The leader's domestic audience is unaware of mediation until a settlement is announced: as in the US-North Korean crisis, the US public became aware of mediation only once Carter announced his breakthrough on CNN. In the model, if a settlement is reached, the leader's domestic audience reacts by deciding whether to sanction its leader or not. If no settlement is reached, then the two countries collide to war. This provides the simplest model to see how a costly war, uncertainty, and domestic politics interacts with secrecy and agenda-setting in mediation.

[9]Similar to Romer and Rosenthal (1978) where the agenda-setter needs majority approval from voters, and depending on the status quo, agenda-control allows for a variety of outcomes. Here the mediator needs the approval of the bargainers, and depending on the uncertainty, agenda-control allows for outcomes. As in Romer and Rosenthal (1978), to focus on how agenda-setting affects outcomes, there are no dynamic or sequential aspects, no log-rolling or issue linkages, and no uncertainty about whether voters will vote (no incomplete turnout). This is to explore the implications of mediation process as it might interact with a domestic audience, leaving further exploration to future research.

Model

Consider a model of one-sided incomplete information in which two countries, 1 and 2, face a crisis and a mediator is involved. Country 1, the *home country*, consists of a leader and her domestic audience. Country 2, the *enemy*, is a unitary actor with private information about his resolve. The mediator prefers peace and receives a payoff of one for a settlement, and zero otherwise. We will use female pronouns for the leader and mediator, and male pronouns for the enemy.

To focus on mediation, we model war as a costly lottery with a prize normalized to one, and costs of war $c > 0$ for each country. Country 1 owns the prize at stake. If a war occurs and 2 wins, then country 1 pays the prize to country 2. If 1 wins, then no transfer is made, and each side receives a payoff of zero minus the costs of war. To prevent a war, countries 1 and 2 must reach a settlement, σ, which is an amount that 1 will pay 2 to avoid war.

The settlement that 2 is willing to accept depends on his resolve, which is drawn by Nature at the start of the game. Country 2 can be a low or high resolve type, denoted by $\tau \in \{\tau_L, \tau_H\}$, where $\tau_L < \tau_H$, and τ gives the probability that 2 will prevail in war.[10] Only country 2 knows his type. The mediator, leader, and audience share common priors that 2 has high resolve with probability $1 - p$ and low resolve with probability p.

Given this setup, 2 prefers any settlement σ that is at least as high as his reservation value. A low resolve enemy accepts a smaller settlement, $\sigma_L = \tau_L - c$, than a high resolve enemy, $\sigma_H = \tau_H - c$. Let us call these the *low* and *high offers*.

Definition 1 (Low Offer, High Offer). *A low offer is the minimum settlement that a low type is willing to accept, denoted $\sigma_L = \tau_L - c$. A high offer is the minimum settlement that a high type is willing to accept, denoted $\sigma_H = \tau_H - c$.*

The sequence of the game captures the process of mediation.

[10] An alternative model of resolve assumes that high resolved types pay lower costs of war, and low resolved types pay high costs.

Sequence

After Nature draws 2's type, the mediator makes an initial proposal, m, that 1 will give to 2 if both parties agree. Next, 2 decides whether to *accept* or *reject* m.

If 2 accepts m, then the leader of country 1 updates her beliefs about 2's type using Bayes' Rule. Let q_L represent the probability that the low type accepts m, and q_H be the probability that the high type accepts m. Upon observing the enemy accept, the leader believes the enemy is a low type with probability λ_1,

$$\lambda_1 = P(\tau = \tau_L | accept\ m) = \frac{q_L * p}{q_L * p + q_H * (1-p)},$$

and a high type with probability $1 - \lambda_1$. Given these beliefs, the leader decides whether she too will *accept* or *reject* m. If the leader also accepts, then the settlement is the mediator's proposal, $\sigma = m$. If the leader rejects, then the two countries go to war.

On the other hand, if 2 rejects m, then the leader updates her beliefs about 2's type, and decides whether to *raise* or *exit* to war.[11] The leader believes that an enemy *who rejects* is a low type with probability λ_2,

$$\lambda_2 = P(\tau = \tau_L | reject\ m) = \frac{(1-q_L) * p}{(1-q_L) * p + (1-q_H) * (1-p)},$$

and a high type with probability $1 - \lambda_2$. If the leader *raises* the offer, then 2 decides whether to *accept* or *reject* this new offer, $m + \delta$. If 2 rejects, then the two countries go to war. If 2 accepts, then the raised offer becomes the settlement, $\sigma = m + \delta$.

If a settlement is reached, then the audience must decide whether to *sanction* the leader, or *not*. The audience does not know whether that settlement was proposed by the mediator alone, or if the leader raised. Further the audience does now know whether

[11]This exit option to war is necessary, otherwise the enemy can force the leader into accepting m: by initially rejecting m, and then accepting m only once the leader offers nothing extra, the leader is forced to accept m.

the enemy is a high or low type. The audience therefore updates its beliefs using Bayes' Rule about whether it was more likely that 1) the low type and leader accepted the mediator's offer, 2) the low type rejected, the leader raised, and the low type accepted the raised offer, 3) the high type and leader accepted the mediator's offer, or 4) the high type rejected, the leader raised, and the high type accepted that raised offer. Let each of these beliefs be represented by α_L^m, α_L^r, α_H^m, and α_H^r, respectively, where the subscript represents the enemy's type, and the superscript represents whether the mediator made the offer, or the leader raised.[12]

To incentivize the audience, we normalize the audience's payoff for sanctioning correctly to one: the audience receives a payoff of one for sanctioning when the leader raised, and for not sanctioning when the mediator proposed the settlement. Thus, in any equilibrium, the audience's best response is to *not sanction* if it is more likely that the settlement came from the mediator, $\alpha_L^m + \alpha_H^m \geq \alpha_L^r + \alpha_H^r$, and to *sanction* if it is more likely that the leader raised, $\alpha_L^m + \alpha_H^m < \alpha_L^r + \alpha_H^r$.

Lastly, if the audience sanctions, then the leader pays an audience cost. Let s represent the probability the audience sanctions, and let audience costs be given by $a(\sigma)$, where $a(\cdot)$ is positive and increasing in the settlement, σ.[13] This way if the settlement includes

[12]For example, the audience's belief that the low type and leader accepted the mediator's offer is:

$$\alpha_L^m = \frac{q_L * p}{q_L * p + (1 - q_L) * p * r + q_H * (1 - p) + (1 - q_H) * r * (1 - p)},$$

where r represents the probability the leader raised, and for brevity, this sample posterior assumes that the leader accepts the mediator's offer anytime the enemy accepts, and the high and low types accept a raised offer.

[13]Alternatively, the audience might sanction for raising and for concessions, modeled as two additive terms $A + a(\sigma)$. This alternative would carry through the model as two

high concessions, then the leader pays a large audience cost. If the settlement includes few concessions, the leader pays a smaller audience cost.

The game is presented in Figure 1.[14] The solution concept is a perfect Bayesian equilibrium.

Equilibrium

The model reveals that only two paths of play occur in equilibrium. In the first, the mediator proposes a *high* offer that is accepted by all, and peace is achieved. Otherwise, the mediator proposes a smaller offer – in between the low and high offers – that only a low resolve type accepts. Since only the low type accepts, the leader learns about the resolve of any enemy who remains, which commits the leader to playing a mixed strategy where she will sometimes *raise*. This second path of play breaks up the equilibrium space into two additional regions described in Propositions 2, where the audience does not sanction, and 3, where sanctioning can occur. Here we explain the intuitions underlying each of these regions. All proofs are found in the Appendix.

To understand the first path, consider what happens if the mediator makes the high offer, $m^* = \sigma_H$. Both types will accept it, since it meets their reservation values, as long as they expect the leader not to raise. Further, as long as the leader does not raise, a settlement is reached only through the mediator. The audience has rational reasons to believe that this costly settlement comes from the mediator, $\alpha_L^m + \alpha_H^m = 1$ and $\alpha_L^r + \alpha_H^r = 0$, and does not sanction the leader, $s^* = 0$. The mediator will obtain peace with probability one as long as the leader is willing to accept this high offer. Therefore, this is an equilibrium if two conditions are met: the leader must prefer to not raise given her off the equilibrium path beliefs; and the leader must be willing to accept the high

additive terms in place of $a(\sigma)$ here, which would shift the cut-points along p to make higher concessions ex ante less likely. The intuition remains the same.

[14]War payoffs are suppressed for space.

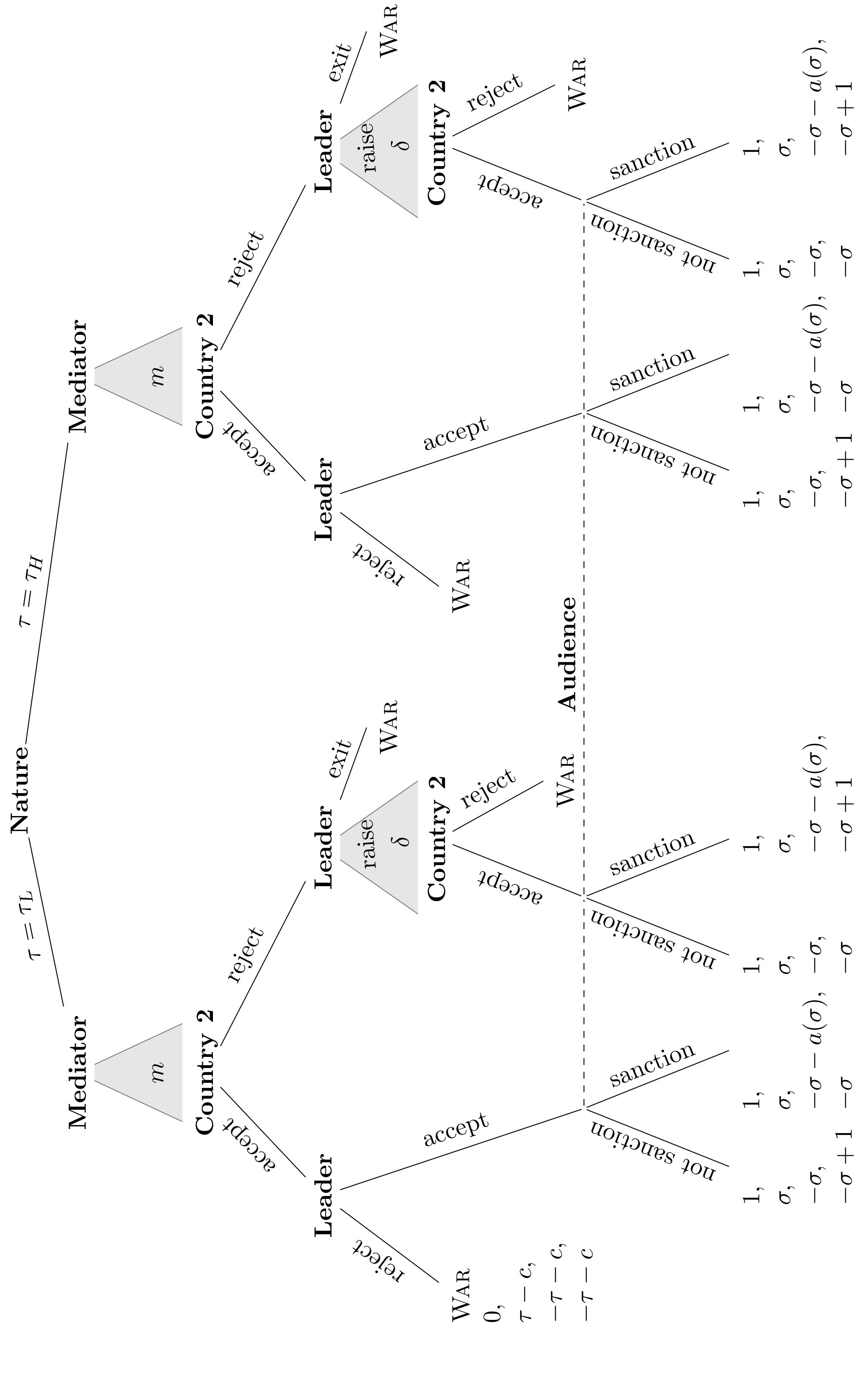


Figure 1: Model

offer.

The leader prefers not to raise as long as the mediator makes the high offer, and there is sufficient chance that she faces a low type off the equilibrium path; otherwise being certain that she faces a high type, the leader will prefer to raise by just a bit rather than face a strong chance of losing a costly war. For this to work, the leader must believe she faces a low type with probability λ_2:

$$\lambda_2 \geq \frac{2c - \tau_H}{\tau_H - \tau_L} \equiv \overline{\lambda_2}. \tag{1}$$

At the same time, the leader must prefer to pay the high offer rather than accept her expected payoff from war given her prior (since both types accept): $-(\tau_H - c) \geq -p\tau_L - (1-p)\tau_H - c$, which is true if:

$$p \leq \frac{2c}{\tau_H - \tau_L} \equiv p_H^* \tag{2}$$

Since $\overline{\lambda_2} < p_H^*$, the leader's off-path beliefs are supported: when the leader is willing to accept the high offer, she prefers to exit. These off-path beliefs satisfy condition D1, which requires that beliefs be supported on any type who stands to gain from deviation: the leader assigns positive weight to the high type, $\lambda_2 \neq 1$, and the mediator makes the high offer, $m^* = \sigma_H$ (Cho and Kreps 1987).[15]

Since both countries accept this high offer simply because the mediator proposes it, and it guarantees peace, we will refer to equilibrium as *locking in* peaceful concessions. The mediator, who receives her best outcome, locks in concessions wherever possible,

[15]Since the low type does not gain from deviation, $m > \sigma_L$, D1 can only apply to the high type, and m cannot be greater than σ_H, otherwise rejection would also be dominated for the high type, which would mean the leader could not assign positive weight to either type. Alternatively, universal divinity would result in the same high offer, $m^* = \sigma_H$, but would be more restrictive in needing more weight to be placed on the high type, $\lambda_2 < \frac{1}{2}$.

which is in Region I of Figure 2 where $p < p_H^*$.

Proposition 1 (Region I: High Offer)**.** *When $p \leq p_H^*$, the mediator proposes $m^* = \sigma_H$, and both types of enemy accept, where $p_H^* = \frac{2c}{\tau_H - \tau_L}$. If the enemy accepts, the leader accepts m^* with beliefs $\lambda_1 = p$. If the enemy rejects, the leader's beliefs are $\lambda_2 \geq \frac{2c - \tau_H}{\tau_H - \tau_L}$, and the leader exits to war. The audience does not sanction the leader, $s^* = 0$, with beliefs $\alpha_L^m + \alpha_H^m = 1$ and $\alpha_L^r + \alpha_H^r = 0$. The probability of war is zero.*

Outside this region, the mediator sets in pace a semi-separating equilibrium in which at least the low type accepts m sometimes, and the high type rejects.[16] To do so the mediator must make an offer in between the low and high offers that a low type is indifferent between accepting and rejecting.[17] This causes the leader, who thinks she likely faces a high type, to raise with a probability that keeps the low type indifferent.

For this to work, the leader must raise with a probability $r = \frac{m - \sigma_L}{m + \delta - \sigma_L}$ that keeps the

[16]Mediation has *no separating equilibrium* in which the low type accepts and the high type rejects the mediator's proposal. If this were to occur, then the leader would know that an enemy who rejects must be a high type, and can either raise or exit. Neither of these forms an equilibrium. If the leader raises, then the low type can profitably deviate to reject m. If the leader exits, then the audience believes the mediator is to blame, and the leader can profitably deviate to raise the offer.

[17]According to Harsanyi's purification theorem, the low type's mixed strategy is equivalently conceived of as pure strategies played by different types (a low and lowest type) in a nearby game with added incomplete information. In this equivalent model, a *lowest* type always accepts the mediator's offer, and the enemy who remains is one of the higher of *two or more types*. This allows us to interpret the mixed strategy equilibrium by perturbing each actor's payoffs, without the need for any actor to randomize their strategies.

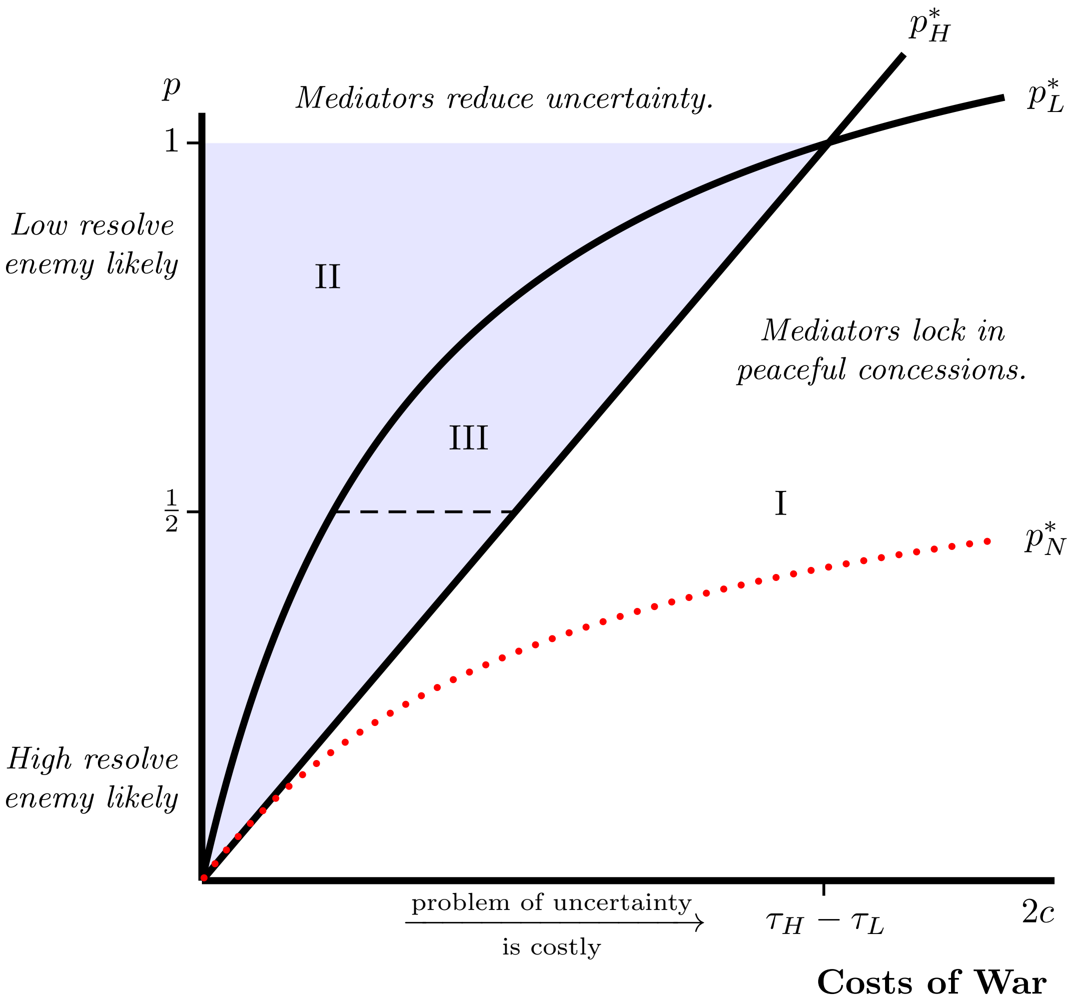


Figure 2: Mediation's Two Mechanisms

low type indifferent between accepting the mediator's offer, $m > \sigma_L$, and rejecting it for a gamble between his war payoff, σ_L, which is strictly worse, and a raised settlement, which is strictly better. The leader is willing to raise, in which case she will pay a higher price for peace, but only if she thinks it is unlikely that she will prevail in war. In other words, the leader must believe that it is sufficiently likely that she faces a high type, $\lambda_2 = p_H^* - \frac{sa(\sigma_H)}{\tau_H - \tau_L}$, which requires that the the low type accept the mediator's proposal only so often, with probability $q_L = \frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{p[\tau_H - \tau_L - 2c + sa(\sigma_H)]}$. Given the leader's beliefs that she likely faces a high type, she is willing to make the high offer, $m + \delta = \sigma_H$, which secures peace against both types. To support the leader's beliefs, as the probability of a low type increases, the low type must accept m more often. The mediator orchestrates all of this – within limits. The mediator's challenge is to make an offer that is high enough to pull the low type out from subsequent bargaining, but not so high that the leader will reject it.

In Region II, when a weak type is very likely, $p > p_L^*$, the mediator needs to pull the low type out as often as possible. Thus, the mediator offers as much as possible. Since the leader realizes that an enemy who accepts the mediator's offer must be a low type, $\lambda_1 = 1$, the mediator's offer is determined by the maximum settlement that the leader will tolerate against a low type, $\overline{\sigma_L} = \tau_L + c - sa(\sigma)$, minus her potential audience costs. In this region, a low type is sufficiently likely that the audience believes that any settlement is more likely because the likely low type accepted the mediator's offer, $\alpha_L^m > \alpha_L^r + \alpha_H^r$, and $\alpha_H^m = 0$. Therefore, the audience does not sanction the leader, $s^* = 0$. Substitution of s^* into each actor's best responses gives the equilibrium in Proposition 2, depicted in Region II, where $p > p_L^*$, of Figure 2.

Proposition 2 (Region II: No Sanction). *When $p \geq p_L^*$, the mediator offers $m^* = \tau_L + c$, where $p_L^* = \frac{4c}{\tau_H - \tau_L + 2c}$. The low type accepts with probability $q_L^* = \frac{p(\tau_H - \tau_L) - 2c}{p(\tau_H - \tau_L - 2c)}$, and the high type rejects. If the enemy accepts, the leader accepts m^* with beliefs $\lambda_1 = 1$. If the enemy rejects, the leader raises with probability $r^* = \frac{2c}{\tau_H - \tau_L}$ to offer $\delta^* = \tau_H - \tau_L - 2c$ with beliefs $\lambda_2 = \frac{2c}{\tau_H - \tau_L}$. Both types accept the raised offer. The audience does not sanction, $s^* = 0$, with beliefs $\alpha_H^m = 0$, and $\alpha_L^m > \alpha_L^r + \alpha_H^r$. The probability of war is $1 - p$.*

In Region III, the maximum that the leader will tolerate, $\overline{\sigma_L}$, can get the leader sanctioned – which can deter the leader from accepting the mediator's proposal. Thus, three things can happen.

When the audience is sufficiently weak, $a(\sigma_H) < 2c$, the mediator makes the maximum tolerable offer, $\overline{\sigma_L}$, and the leader raises even though she is sanctioned, $s^* = 1$. Since the probability of war is decreasing in the mediator's offer and the probability the audience sanctions, $P(war) = \frac{(1-p)(\tau_H - m - c)}{\tau_H - \tau_L - 2c + sa(\sigma_H)}$, this reduces the probability of war to less than $1 - p$.

When the audience is sufficiently strong, $a(\sigma_H) > 2c$, the leader cannot accept $\overline{\sigma_L}$ with audience costs. Therefore, the mediator is constrained to making a lower offer, $m^* = \widehat{\sigma_L} < \overline{\sigma_L}$, which is strictly better for country 1. Since the audience is strong, the leader is willing to raise (her indifference condition is met) only if low type accepts the mediator's proposal consistently, $q_L^* = 1$, and a high type is likelier, $p < \frac{1}{2}$. This occurs in Region III, below the dashed line at $p = \frac{1}{2}$ in Figure 2 for sufficiently strong audiences, where the probability of war is reduced to $1 - 2p$.

Otherwise, the mediator makes an offer that keeps the audience indifferent between sanctioning and not, and maintains the leader's mixed strategy to raise to result in a probability of war of $1 - p$. These three possibilities are not overlapping: there is a unique equilibrium for every pair $(p, a(\sigma_H))$. For the remaining discussion, we will focus on the potential for strong domestic audiences to reduce the probability of war in this region.

Proposition 3 (Region III: Sanctions)**.** *When $p \in (p_H^*, p_L^*)$, the perfect Bayesian equilibrium for any pair $(p, a(\sigma_H))$ is as follows:*

1. *If the audience is weak, $a(\sigma_H) < 2c$, the mediator offers $m^* = \overline{\sigma_L} = \tau_L + c - a(\sigma)$, and the leader is sanctioned, $s^* = 1$. The probability of war is $\frac{(1-p)[\tau_H - \tau_L - 2c + a(\sigma)]}{\tau_H - \tau_L - 2c + a(\sigma_H)} < 1 - p$.*

2. *If the audience is strong, $a(\sigma_H) > 2c$, a high type is likely, $p < \frac{1}{2}$, and the sanction for accepting the mediator's offer is not too high, $a(\sigma) \leq 2c - \frac{p(\tau_H - \tau_L)}{1-p}$, then the mediator offers $m^* = \widehat{\sigma_L} = \sigma_L + \frac{p(\tau_H - \tau_L)}{1-p}$. The low type accepts, $q_L^* = 1$, the high*

type rejects. If the enemy rejects, the leader raises with probability $r^* = \frac{p}{1-p}$ *and concessions* $\delta^* = \frac{(1-2p)(\tau_H - \tau_L)}{1-p}$ *with beliefs* $\lambda_2 = 0$*. The audience sanctions with probability* $s^* = \frac{2c}{a(\sigma_H)}$ *with beliefs* $\alpha_L^m = \alpha_H^r = \frac{1}{2}$*,* $\alpha_L^r = \alpha_H^m = 0$*. The probability of war is* $1 - 2p$*.*

3. *Otherwise, the mediator offers* $m^* = \overline{\sigma_L} = \tau_L + c - s^* a(\sigma)$*. The audience sanctions with probability* $s^* = \frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}$ *and beliefs* $\alpha_L^m = \alpha_L^r + \alpha_H^r = \frac{1}{2}$*, and* $\alpha_H^m = 0$*. The probability of war is* $1 - p$*.*

In each case, the leader accepts the mediator's offer with beliefs $\lambda_1 = 1$*, and both types accept a raised offer.*

In equilibria 1 and 3, the low type accepts with probability $q_L^* = \frac{p(\tau_H - \tau_L) - 2c + s^* a(\sigma_H)}{p[\tau_H - \tau_L - 2c + s^* a(\sigma_H)]}$*, the high type rejects. If the enemy rejects, the leader believes it is a low type with probability* $\lambda_2 = \frac{2c - s^* a(\sigma_H)}{\tau_H - \tau_L}$*, and raises with probability* $r^* = \frac{2c - s^* a(\sigma)}{\tau_H - \tau_L}$ *and concessions* $\delta^* = \tau_H - \tau_L - 2c + s^* a(\sigma)$*.*

Analysis

The model shows that domestic pressure gives rise to two mechanisms for peace. First, the mediator can *lock-in concessions* that all parties accept. Second, the mediator can make a *screening offer* that allows the leader to learn about the enemy's resolve, which warrants greater concessions. This section unpacks how each of these works, and establishes the benefits and costs of mediation.

Two Mechanisms for Peace

In the first mechanism, the mediator locks in concessions by proposing the high offer under specific circumstances. To understand why this is beneficial, consider a similar, as yet unrepresented, model of a *bilateral negotiation.*[18] This bargaining with audience

[18] The leader makes an offer, and the enemy chooses to accept or reject. If he rejects, then the leader can raise or exit. If she raises, she pays audience costs.

costs game results in a well known risk-return trade off: since the leader knows she will face audience costs for raising, she never raises, and instead makes the high offer when a high type is sufficiently likely, $p < p_N^* = \frac{2c}{\tau_H - \tau_L + 2c}$, and otherwise makes a low offer that risks war against the less likely high type (Powell 1999; Slantchev 2004; Tarar and Leventoğlu 2013).[19] Peace occurs from negotiation where $p < p_N^*$ indicated by the dotted line in Figure 2. Since the threshold for peace with negotiation is strictly lower than that of mediation, $p_N^* < p_H^*$, mediation obtains peace where a negotiation cannot for all $p \in (p_N^*, p_H^*)$. Given that the leader avoids sanction while reaching a settlement with costly concessions, the lock in mechanism can be interpreted as follows: mediation enables a leader to sign a settlement that she would agree to, but could not offer on her own.

This occurs because the mediator as agenda-setter does not face the same risk-return trade off as the leader. While a leader must accept some risk of war in being pressured to stand firm, the mediator does not face this pressure: she can offer concessions *solely based* on what improves the prospects for peace. This is similar to agenda-setting in Romer and Rosenthal (1978), where an agenda-setter can have considerable control over the outcome by presenting voters with a 'take-it-or-leave-it' choice against the status quo. Here, the mediator pressures the leader and enemy to choose between the status quo (war) and the mediator's proposal. Since a costly war is worse, and the mediator is not constrained by the domestic audience, the mediator achieves an outcome that would not arise if a leader bargained bilaterally.

Importantly, this provides one answer to the question of why mediate. The threshold for peace, p_H^*, is increasing in the total value destroyed by war, $2c$, and decreasing in this difference in types, $\tau_H - \tau_L$. One can think of this difference in types, $\tau_H - \tau_L$, as a measure of how much uncertainty matters. If types are very different, then knowing whether the enemy has high or low resolve significantly alters the expected outcome. If types are similar, then uncertainty does not matter as much. The model shows that as concerns about the destruction of war begin to outweigh the problems caused by uncertainty, as

[19] This is proven in the Appendix.

$\tau_H - \tau_L \to 0$ or c increases, the region under p^*_H expands – mediation is ex ante more likely to secure peace by allowing leaders to reach settlements they could not achieve otherwise.[20]

Result 1 (Preventing Costly Wars)**.** *As the destruction of war increases, c, or the problem caused by uncertainty decreases, $\tau_H - \tau_L \to 0$, then $p^*_H \to 1$. Mediation is ex ante more likely to result in settlement by allowing leaders to reach settlements they could not achieve otherwise.*

This provides intuition for how conflicts become ripe for mediation. If uncertainty about the enemy declines or the costs of war increase over time, then the costs of war will begin outweigh the problem of uncertainty. Mediators can step in at these *ripe* moments, and succeed by locking in peaceful concessions. Figure 3 shows how the lock-in region grows when the problem of uncertainty is small. At some point, when $2c \geq \tau_H - \tau_L$, then $p^*_H \geq 1$, mediators can step in and obtain peace for any $p \in [0, 1]$.

Result 2 (Conflict Ripeness)**.** *When the costs of war are sufficiently high, $2c \geq \tau_H - \tau_L$, or the difference between more and less resolved adversaries is sufficiently low, then $p^*_H \geq 1$, a mediator secures peace for any $p \in [0, 1]$ by proposing σ_H.*

On the other hand, when the problem of uncertainty is costly, $\tau_H - \tau_L \to 2c$, Regions II and III expand as seen in Figure 2, and mediation serves a different purpose in reducing that uncertainty – specifically when a low type is sufficiently likely (when $p > p^*_H$). There the mediator makes an initial offer that screens for low resolved adversaries. By making an offer that only a low-resolve type accepts, the mediator enables low types to select themselves out of subsequent bargaining. This allows the leader to update her beliefs that any adversary who remains in mediation is more likely to have high resolve, warranting greater concessions.

---

[20]The region under p^*_N also grows, but since $p^*_N < p^*_H$, for all $p < p^*_H$, mediation will always obtain peace more often.

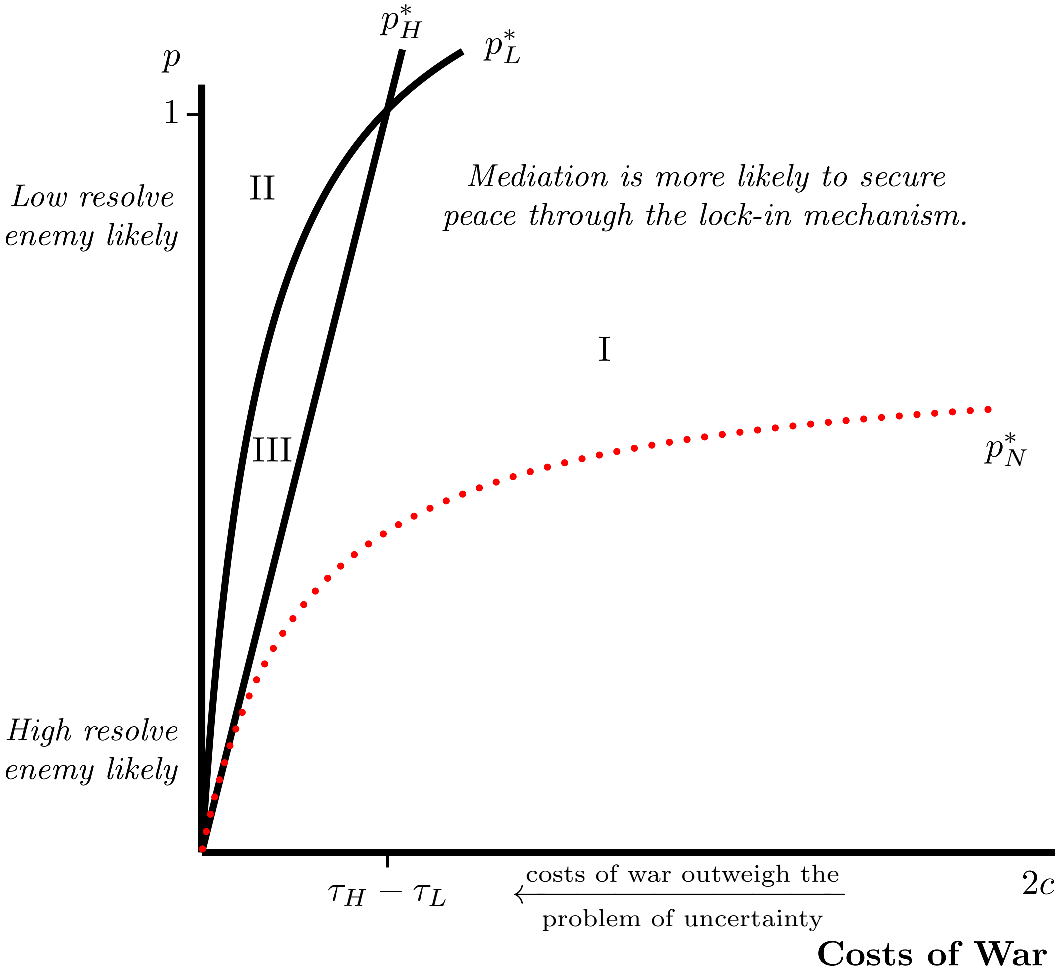


Figure 3: Mediation's Lock In

Audience costs, secrecy, and agenda-setting play key roles in making this screen work. Audience costs put direct pressure on the leader to stand firm against the enemy, which *puts indirect pressure on low resolve enemies to accept early.* This can be seen in two ways. The probability that the low type accepts, q_L, is increasing in the leader's audience costs, and in accord, the leader's posterior beliefs that an enemy who rejects is a low type, $\lambda_2 = p_H^* - \frac{sa(\sigma_H)}{\tau_H - \tau_L}$, is decreasing in audience costs.[21] This means that the mediator's screen is more effective at selecting low types out, and the leader learns more, when that leader has higher audience costs.

Result 3 (Effective Screening)**.** *Mediation is more effective at screening out low resolve enemies, and the leader learns more, when that leader has higher audience costs.*

Thus, if we compared two similar conflicts involving two separate leaders (and two adversaries), and one leader faced greater domestic pressure, then a mediator will be better able to reduce uncertainty for that leader with higher audience costs.

While audience costs pressure the low type to back out of mediation, secrecy and agenda-setting are necessary to allow the leader to raise. If an agenda-setting mediator does maintain secrecy, then the audience will observe whenever the leader offers new terms, and the leader will not raise the offer. If a mediator hosts secret talks, but does not control the agenda, then the audience can infer that any new concessions must be from the leader – again the leader cannot raise.[22] The leader will only raise, and the screening mechanism will succeed, if a mediator who sets the agenda keeps talks secret.

[21]The derivative of q_L with respect to audience costs, $a(\sigma_H)$ is positive: $\frac{dq_L}{da(\sigma)} = \frac{s(1-p)(\tau_H - \tau_L)}{p[\tau_H - \tau_L - 2c + sa(\sigma_H)]^2} > 0$.

[22]One might argue that the leader will raise if the audience is weak enough, as she does in Proposition 3. However, in that equilibrium, the leader must raise with probability less than one. If the audience is sufficiently weak such that the leader raises all the time, then we are back to a situation in which mediation makes no difference: both types will reject any mediator's offer knowing that the leader will raise. Secrecy, agenda-setting,

Costs and Benefits

Since each of these two mechanisms is more likely under specific circumstances – depending on whether the significant problem is the costs of war or the uncertainty problem – one benefit is that mediation is likely to serve the right purpose at the right time.[23] In general, when the costs of war outweigh the problems caused by uncertainty, mediation is ex ante more likely to lock in a settlement that prevents those costlier wars. When uncertainty becomes more problematic, mediators can reduce that uncertainty.

In addition, mediation reduces the probability of war as seen in Figure 4. In negotiations, a leader risks war whenever $p > p_N^*$, which occurs with probability $1-p$ should the enemy turn out to be a high type, as indicated by the dotted line. Mediation removes the risk of war in the shaded region where $p \in (p_N^*, p_H^*)$. Otherwise, mediation does no worse and can do better at reducing war. When $p < p_L^*$ the risk of war is the same. In between, if the leader faces a strong domestic audience, then mediation does strictly as seen by the dashed line. Otherwise, the probability of war from mediation in this region is no greater than the probability of war from negotiation, $1 - p$. Thus, in general, and especially when a high resolve enemy is likely, mediation reduces war.

Result 4 (Risk of War)**.** *Mediation reduces the ex ante probability of war.*

At the same time, if talks fail, then mediation increases the probability of prevailing in any ensuing war. This is because if a negotiation fails, then it fails because the enemy has higher resolve and was unwilling to accept the leader's offer. If mediation fails, then it fails because the leader refused to raise the offer and the enemy either has low or high resolve. Since there is some chance that the enemy has low resolve, the leader and her audience fight a less resolved enemy on average, which gives them a higher probability of winning in war.

---

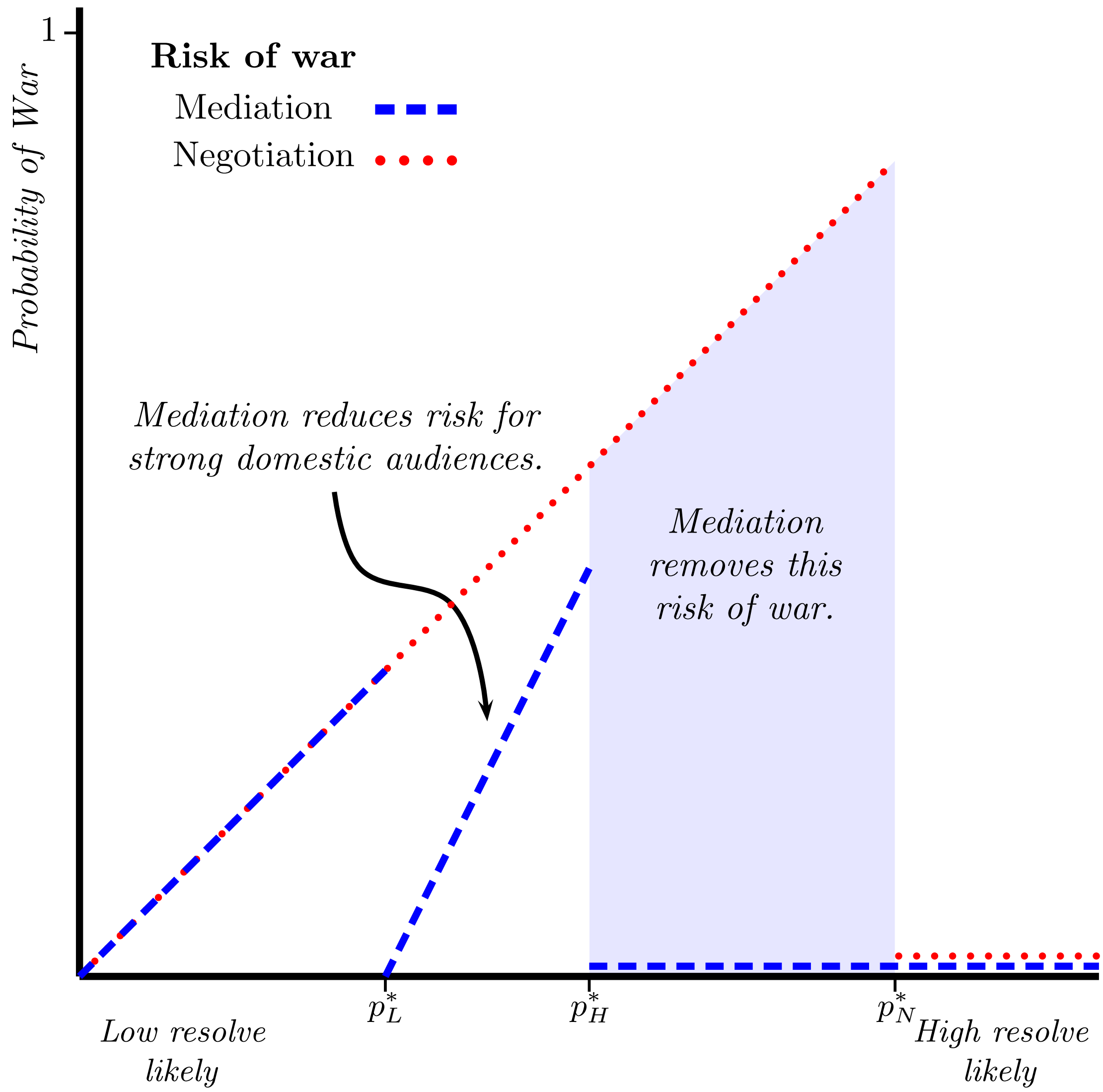
and audience costs are required for screening to work.

[23]This depends on the underlying distribution of p.

Result 5 (Victory in War). *Mediation results in a higher the probability of winning any ensuing war, than negotiation.*

Of course, these benefits come at a price as seen in Figure 5. While the leader makes only the low offer for any $p < p_N^*$, the mediator's offer and the leader's mediated raise are strictly larger as seen by the solid and dashed lines. Otherwise, when a high type is sufficiently likely, $p > p_N^*$, both the mediator and leader make the high offer. Any mediated settlement will be at least as costly as a negotiated one.

Result 6 (Settlement Cost). *A mediated settlement is always at least as costly as a bilaterally negotiated settlement:* $\sigma_M \geq \sigma_N$ *for all* p.


1
Probability of War
Risk of war
Mediation
Negotiation
Mediation reduces risk for strong domestic audiences.
Mediation removes this risk of war.
p_L^*
p_H^*
p_N^*
Low resolve likely
High resolve likely

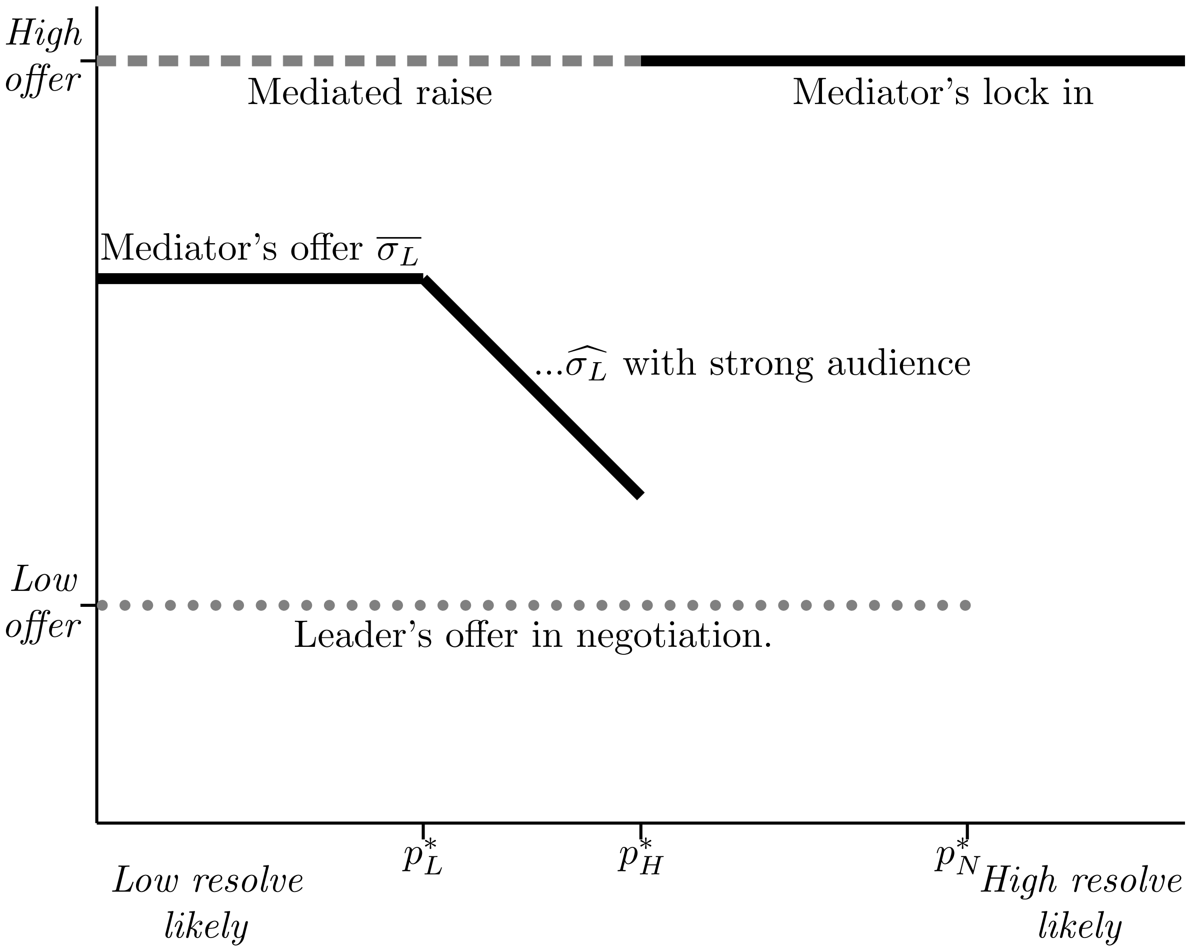
Discussion

When a leader faces domestic pressure, mediation with *secrecy and agenda-setting* achieves settlements by allowing for learning that reduces uncertainty about an adversary's resolve and by locking in concessions. With these, mediation can improve the prospects for peace, especially when an adversary is likely to be high-resolved; should talks fail, the leader and her audience are more likely to win in any ensuing war. However, mediation is no guarantee since settlements are more costly. This holds immediate implications for mediation, audience costs, and international organizations.

First, rather than focusing on powerful or informed mediators, the theory shows how *any* mediator can succeed with secrecy and agenda-setting. Since the leader faces domestic pressure, the enemy's response to the mediator is a credible signal of his resolve. In contrast to previous research, this information transmission does not rely on the mediator having an independent source of information or the appropriate bias to credibly transmit information.[24] At the same time, mediators can use power or information as complements to this mechanism. *Power-based* mediators can supplement the price of mediation for the leader and her audience – who pay a higher settlement. *Information-based* mediators who can convey an enemy has high resolve can make peace more likely by selecting into regions where peace is more likely.

This research also sheds light on the argument that mediators or other third parties broker settlements by helping leaders save face (Allee and Huth 2006; Beardsley 2010; Gent and Shannon 2010; Huth, Croco, and Appel 2011; Simmons 2002;). Here, leaders do not always save face – they are sometimes punished even for small settlements. Thus, a better interpretation is that when face-saving would be useful, third-parties make peace

[24]Existing research places several constraints on when mediators provide information: mediators must have an independent source of information; cannot merely prefer to avoid war; and must be biased in favor of the party for which information reduces that party's payoff (Fey and Ramsay 2010; Kydd 2003; Rauchhaus 2006).



High offer
Mediated raise
Mediator's lock in
Mediator's offer $\overline{\sigma_L}$
...$\widehat{\sigma_L}$ with strong audience
Low offer
Leader's offer in negotiation.
p_L^*
p_H^*
p_N^*
Low resolve likely
High resolve likely

more likely.

This provides one answer to the puzzle of how *weak* mediators succeed, and provides a role for strong mediators to employ secrecy and agenda-setting. Private citizens, small countries, and regional organizations might mediate successfully to the extent that they can limit media attention or fly under the radar until a settlement is reached. Further, agenda-setting and secrecy are not exclusive to mediation. The theory may apply to other third party interventions such as adjudication and arbitration, although one should note that these processes also have distinct features leaving room for future research.[25] Since other third-parties are similar, but distinct from mediation, this paper lays groundwork to explore how these institutional features compare.

This research also holds implications for democracies by showing how a leader's own audience costs enables an enemy *with no audience costs* demonstrate resolve. Since uncertainty can be reduced in mediation, then to the extent that democratically-elected leaders face greater audience costs, this provides one explanation for why democracies end their wars earlier: in short, less information needs to be learned on the battlefield if democratically-elected leaders mediate their disputes (Bennett and Stam 1998; Gartner 2008; Reiter and Stam 2002). Further, this helps to explain why democracies should be more likely to resolve conflicts using third-parties, establish international institutions that can serve as mediators, follow norms of compromise, and resolve their militarized conflicts with mutual concessions (Dixon 1994; Mitchell 2002; Mousseau 1998; Russett, Oneal, and Davis 1998). These institutions serve as mediators to reduce uncertainty and lock in settlements that prevent costlier wars.

[25]Keohane, Moravcsik, and Slaughter (2000) states that international and transnational courts and tribunals practice agenda-setting, relying on international laws and legal principals, however, these adjudicators are likely to differ in terms of secrecy. While transnational bodies are legally insulated from the state, international bodies are influenced by domestic constituencies.

This also provides one answer to the puzzle of why dictators escalate crises against powerful democracies only to return to the mediation table.[26] Autocracies should avoid war with democracies, since democracies are more likely to win and more effective at fighting. The theory here suggests that autocracies may be playing to their equilibrium advantage: escalating crises and pursuing mediation to demonstrate their resolve and shift the status quo.

Together, these imply a double-edged sword for democracies: able to more efficiently learn information and prevent costly wars through mediation, but at the cost of inviting dictatorships to escalate crises in effort to credibly demonstrate their resolve. This might explain North Korea's penchant for initiating threats only to return to the mediation table – not to suggest that threat of war is not real, but rather that it might behoove democracies to create ways for autocratic leaders to enter into third party processes, such that they may demonstrate resolve without resort to (in North Korea's case) nuclear brinksmanship.

Beyond this, the theory raises new questions about the potential for political actors faced with external pressures to strategically delegate bargaining to uninformed third parties. The model shows that a principal who faces external pressure can reduce her risk of a worse outcome by delegating bargaining to an uninformed agent, where that agent is given considerable discretion over the outcome. As in traditional models, delegation here serves an informational purpose (Bendor and Meirowitz 2004; Gailmard and Patty 2012). However, here the delegate does not have technical expertise to better inform the principal. Instead, information arises through the mechanism of external pressure and the process by which the agent extracts information from the adversary. Since principals in other contexts are likely to face external pressures and adversarial relationships, this suggests that uninformed agents acting in secret with considerable discretion over out-

[26]This puzzle was posed by Gelpi and Grieco (2001), who ask "Why do otherwise powerful formidable democracies disproportionately attract serious political-military challenges by authoritarian regimes?" (794).

comes may serve a special role in helping representative politicians reach agreements when battling over policies or budgetary controls, or in helping unions and firms reach compromises over labor laws and wages. In line with other delegation results, since the amount of information is increasing in external pressures, we expect delegation to be more likely in times of political, fiscal, or other instability (Huber and Lupia 2001; Volden 2002).

At the same time, the delegation literature can inform the mediation model. The questions of to whom the leader delegates and how much discretion is given to the mediator are central questions from the delegation literature not answered here (Bawn 1997; Gailmard 2002; Staton and Vanberg 2008). Here, we assume that mediation is accepted, but it is not clear that the informed party would agree ex ante to mediation. One might think that the acceptance of mediation reveals information, which requires future research to work through these consequences. One possibility is that a low-resolved type might seek *compensation* from a powerful mediator – perhaps pushing for the intervention of multiple mediators. If this is true, then powerful mediators may face a hold-out problem in which low-resolved adversaries refuse to enter mediation unless compensation is committed to in advance. It could also be that when a strong mediator faces a hold-out problem, a weak mediator must step in to facilitate secret agenda-setting talks – resulting in the coordination issues that complicate multi-mediator talks (Böhmelt 2012).

To illustrate the learning mechanism, this next section examines the US-North Korean crisis.

1994 US-North Korean Crisis

The North Korean crisis provides a *most likely* case since a mediated settlement resulted, and the actors and situation fit closely with the model: it provides confirmatory or disconfirmatory evidence.[27] First, North Korean resolve was unclear: the most significant unknown was whether North Korea was willing to go to war to become a nuclear state or willing to end this crisis diplomatically (Creekmore 2006). To the US, North Korea

[27] See Gerring 2006, Ch. 5.

was either resolved to obtain nuclear weapons, or less resolved and using this crisis as leverage to obtain some other security assurance or tangible benefit to promote regime survival (Wit, Poneman, and Gallucci 2004, 37).

Second, Clinton faced strong domestic pressure to avoid conciliation. Many in the US government wanted to pursue *confrontation*, a 'crime and punishment' approach to stopping proliferation, that made negotiation difficult (Sigal 1998a). Following North Korea's inflammatory "sea of flames" threat, the US became more serious about its military options.

Clinton used for public threats to pressure North Korea (Wit, Poneman, and Gallucci 2004, 28). The administration considered this to be steering *a middle course*:

> A strategy of gradual escalation that would seek to build a coalition, increase pressure on North Korea, and, hopefully, draw China into its ranks. The process would start with expressions of support for the IAEA and calls for North Korean compliance, and then shift to the enactment of sanctions by degrees. (Ibid., 32)

In terms of the model, this "middle course" was Clinton's *low offer* in a *public* negotiation. If Clinton was wrong, then the US and North Korea would go to war. Carter's concerns appear accurate: both sides had "maneuvered themselves into a diplomatic gridlock from which their respective policies offered no retreat" (Creekmore 2006, xxi). The US was engaging in serious military preparations while publicly threatening sanctions and pressing for renewed IAEA involvement (Sigal 1998b).

Evidence shows that mediation then followed the model's sequence. Carter independently pursued mediation with little prior approval from President Clinton, and met with Kim on June 16, 1994. Carter explained that he was there "as a private citizen but with the knowledge and support of the Clinton administration" (Wit, Poneman, and Gallucci 2004, 223). Carter went further than the administration's middle course by suggesting direct US-North Korean talks and offering US assistance to obtain new safer light-water reactors. In line with the model, Carter independently made a *small* offer that was in between the *low* and *high* offers.

Next, Kim accepts Carter's proposal, and Carter notifies Clinton. Wit, Poneman, and Gallucci (2004) write that Carter promptly addressed the nuclear crisis asserting that "the IAEA should be permitted to maintain constant and unbroken surveillance of the fuel rods" (Ibid., 223). In response, "Pyongyang was ready to dismantle its graphite-moderated reactors if the United States would help it get new light-water reactors" (Ibid, 224). Kim stated that his country required electricity for economic development, and pledged that North Korea would rejoin the Non-Proliferation Treaty, which would mandate inspections, if new reactors were received. Carter further requested that the North Korea allow the current IAEA inspectors to remain in the country, since they had not yet left, to which Kim agreed. Following Kim's agreement, Carter phoned the White House.

All historical accounts indicate that Carter set the agenda, and was surprised by the ease with which Kim agreed to concessions (Creekmore 2006; Sigal 1998b; Wit, Poneman, and Gallucci 2004). In an interview, Carter describes how he secured concessions beyond US initial interests: first, he sought to resolve all the issues presented to him in his briefing at Washington; then to incorporate additional issues that included a mutual reduction of military forces North and South of the demilitarized zone, direct peace talks at the summit level with South Korea, and a symbolic concession to help find the bodies of soldiers buried in North Korea from the Korean War (Pbs.org. 2016b). Carter stated, "All these were requests that I had made to him on my own initiative. He agreed to all of them." (Ibid).

Talks were secret. Carter then relayed the agreement to the White House in a phone call, inquiring about the possibility for resumed talks and no sanctions, *and* indicated that he was about to announce on CNN that peace was at hand in describing the terms to which North Korea agreed. All parties involved knew that this public statement would box-in in the sitting president to accept the deal and pressure the US to withdraw the sanctions resolution (Creekmore 2006; Pbs.org. 2016c; Sigal 1998b; Wit, Poneman, and Gallucci 2004).

In line with the model, Carter was locking-in concessions from the US President. The Cabinet room adjourned to watch Carter's announcement that Kim Il Sung promised to not expel IAEA inspectors, to increase transparency, and to discard its old reactors in exchange for new reactors and high-level direct negotiations. Carter indicated that the next move, the acceptance or rejection of this agreement, was up to the Clinton.

The Clinton administration, however, responded in two ways. First, "President Clinton and his advisers, who had originally said Mr. Carter was on a private trip and then became televised participants in the delicate talks with the North Korean leader, Kim Il Sung, clearly distanced themselves from the former President's initiative."[28] While Carter reported that "We've reached complete agreement between us [the United States and North Korea] on the major issues," administration officials were far more cautious about the prospect of resuming talks. Second, the Clinton administration raised the bar to improve terms for the United States. In an official statement delivered by the President, Clinton publicly announced an additional requirement:

> Today there have been reports that the North Koreans, in discussion with President Carter, may have offered new steps to resolve the international community's concerns... If North Korea means by this, also, that it is willing to freeze its nuclear program while talks take place, this could be a promising development. (Wit, Poneman, and Gallucci 2004, 229)

Freezing the nuclear program meant that beyond ensuring that no plutonium would be separated, North Korea would not be able to produce any further plutonium. Clinton stated that any high-level talks and removal of sanctions were conditional on North Korean acceptance; otherwise, the US would continue to pursue sanctions, noting that Ambassador Albright continued to discuss these with the Security Council that day.

This move to increase demands on North Korea raised conspicuous risks. Gallucci had to "trod carefully on the question of whether the United States had 'raised the bar,' the traditional kiss-of-death to any new public proposal" in media questions immediately

[28] Sanger, David. "Carter Visit to North Korea: Whose Trip Was It Really?" The New York Times, June 18th, 1994.

following Clinton's address (Ibid., 230). "To rub North Korean noses in this new condition US risked jeopardizing this agreement and squandering this diplomatic opening" (Ibid.).

Why was the US able to credibly distance itself from Carter's efforts? As the model suggests, because Clinton would never make this offer on his own accord, he *could* credibly distance himself from the settlement terms, even if he agreed.

Why did the US raise the bar? One answer is that the US wanted to prevent North Korea from stalling, and this demand would force inspectors to remain and ensure the freeze took place (Beardsley 2011). Certainly, immediate IAEA monitoring would improve compliance. However, this new demand also raised the risk of dismantling the agreement – unless, of course, *the US had learned that North Korean resolve was low.* The model suggests that the US learned that North Korea was low-resolved: when Kim accepted the mediator's small offer, the enemy winnowed himself out of subsequent bargaining. The US could believe that North Korea was a low resolved type, and would likely accept reduced concessions.

What does this indicate about the model? In the model, when the enemy *accepts* the mediator's low offer, as North Korea did, then the leader learns that the enemy has *low resolve*, and the leader can choose whether to accept. The case shows that the world is more complex: in learning that an enemy has low resolve, Clinton used additional moves to renege on the mediator's offer and start a new negotiation that reduced Carter's offer. This makes sense. In learning that one's enemy is not as resolved as originally feared, the leader may choose to initiate a new iteration of this bargaining game on refined terms.

The screening mechanism that allowed for learning remains present. In moving *theoretically* down the model's path in which the mediator's proposal was accepted, Clinton could be confident that North Korea had low resolve. As Gallucci later indicated, the Cabinet quickly decided to raise the bar, giving no indication that the administration worried at all that North Korea might reject:

> What we decided was to raise the bar just a bit higher than President Carter had set it, and insist that, if we go back to the table, the North Koreans agree not to produce any more plutonium by not restarting the five-megawatt

> reactor. So we raised the bar a little above where President Carter had set it, but then said, yes... The North Koreans very quickly agreed to that one change in the arrangement. We got ourselves back to negotiating in Geneva in July. (Pbs.org. 2016c)

How did Carter's mediation halt this collision course to war? Carter's mediation revealed credible information about North Korean resolve.

Conclusion

How does mediation help? When a leader faces uncertainty, a costly war, and domestic pressure, mediation with secrecy and agenda-setting can promote peace in two ways. First, a mediator's offer provides a screen that winnows away less resolved adversaries to warrant greater concessions. Second, a mediator can lock in concessions that a leader is willing to accept, but would not offer independently. As a result, mediation makes peace more likely against high-resolved adversaries, and should talks fail, the leader and her audience are more likely to win in any ensuing war. However, mediation also entails more costly settlements.

The theory advances the study of mediation by showing how any mediator can succeed without information or material inducements. The theory helps to explain cases or factors in mediation long viewed as important such as conflict ripeness, privacy, back-channels, caucusing, and shuttle diplomacy. The theory provides the first explanation of why weak mediators succeed in their ability to set the agenda in secret. It explains how power and information can be used in conjunction with these mechanisms to make peace more likely. The theory also gives implications for multi-mediator episodes in that strong mediators are likely to face hold-out problems where weak mediators might best intervene.

This research also links the literature on audience costs to mediation, provides a new role for audience costs, and supplies numerous testable implications. By showing how mediation helps an enemy *with no audience costs* demonstrate resolve credibly, mediation can allow for learning and reduce the lengths of wars even if all talks fail. Democracies ought to learn credible information about their enemy's resolve through mediation, giving them advantages in reducing their costs of war off the battlefield. However, dictators may respond by escalating crises only to return to the mediation table. These theoretical implications can be tested to expand knowledge of when crises start and how they end.

The theory also improves understanding of face-saving: concessions are more likely, because when political cover is useful, a leader can use a third party *to obtain information* or *lock in concessions* that improve the prospects for peace. Since different institutions embody these rules to differing degrees, this paper lays groundwork for research to improve understanding of when and how arbitration and adjudication succeeds.

The research links the literature on delegation to international relations in showing how a principal faced with external pressure benefits from delegating to an uninformed agent who bargains in secret and sets the agenda. The theory gives reason for other principals faced with external pressures and adversarial relationships to delegate bargaining to third party mediators, especially in times of political, fiscal, or other instability. The model raises new questions about how bureaucrats or agencies mediate in other contexts, as well as how leaders at war choose their delegates and how much control are they given.

The case study illustrates the learning mechanism and that a mediator may attempt to lock in concessions. It augments the model in demonstrating how international diplomacy is complex with multiple iterated bargaining instances: since the enemy's acceptance of the mediator's proposal demonstrates low-resolve, the political leader can start a new round of bargaining by unilaterally demanding more against a low-resolved enemy. This appears to have been the case: in learning that North Korea had low resolve, Clinton quickly issued higher demands in a public announcement in a new round of bilateral negotiations.

Finally, for practitioners and policymakers, this research shows that the public announcement of potential agreements such as an Israeli-Palestinian peace plans, prior to the start of mediation, can be detrimental to peace. Public announcements undo the ability for secrecy and agenda-setting to help mediation succeed.

References

Allee, Todd L, and Paul K Huth. 2006. "Legitimizing dispute settlement: International legal rulings as domestic political cover." *American Political Science Review* 100 (02): 219–234.

Banks, Jeffrey S. 1990. "Equilibrium behavior in crisis bargaining games." *American Journal of Political Science* 34 (3): 599–614.

Bawn, Kathleen. 1997. "Choosing strategies to control the bureaucracy: Statutory constraints, oversight, and the committee system." *Journal of Law, Economics, and Organization* 13 (1): 101–126.

Beardsley, Kyle. 2009. "Intervention without leverage: Explaining the prevalence of weak mediators." *International Interactions* 35 (3): 272–297.

———. 2010. "Pain, pressure and political cover: Explaining mediation incidence." *Journal of Peace Research* 47 (4): 395–406.

———. 2011. *The mediation dilemma.* Ithaca, N.Y.: Cornell University Press.

———. 2013. "Using the right tool for the job: Mediator leverage and conflict resolution." *Penn St. JL & Int'l Aff.* 2:57–65.

Beardsley, Kyle, and Nigel Lo. 2013. "Third-party conflict management and the willingness to make concessions." *Journal of Conflict Resolution* 58 (2): 363–392.

Bendor, Jonathan, Amihai Glazer, and Thomas Hammond. 2001. "Theories of delegation." *Annual review of political science* 4 (1): 235–269.

Bendor, Jonathan, and Adam Meirowitz. 2004. "Spatial models of delegation." *American Political Science Review* 98 (2): 293–310.

Bennett, Scott D, and Allan C Stam. 1998. "The Declining Advantages of Democracy A Combined Model of War Outcomes and Duration." *Journal of Conflict Resolution* 42 (3): 344–366.

Bercovitch, Jacob. 1996. "Understanding mediation's role in preventive diplomacy." *Negotiation Journal* 12 (3): 241–258.

Bercovitch, Jacob, and Scott Sigmund Gartner. 2006. "Is there method in the madness of mediation? Some lessons for mediators from quantitative studies of mediation." *International Interactions* 32 (4): 329–354.

Bercovitch, Jacob, and Jeffrey Langley. 1993. "The nature of the dispute and the effectiveness of international mediation." *Journal of Conflict Resolution* 37 (4): 670–691.

Böhmelt, Tobias. 2010. "The effectiveness of tracks of diplomacy strategies in third-party interventions." *Journal of Peace Research* 47 (2): 167–178.

———. 2012. "Why many cooks if they can spoil the broth? The determinants of multiparty mediation." *Journal of Peace Research* 49 (5): 701–715.

Brown, Jonathan N, and Anthony S Marcum. 2011. "Avoiding audience costs: Domestic political accountability and concessions in crisis diplomacy." *Security Studies* 20 (2): 141–170.

Cho, In-Koo, and David M Kreps. 1987. "Signaling games and stable equilibria." *The Quarterly Journal of Economics* 102 (2): 179–221.

Cox, Gary W, and Mathew D McCubbins. 2005. *Setting the agenda: Responsible party government in the US House of Representatives.* Cambridge, NY: Cambridge University Press.

Creekmore, Marion V. 2006. *A moment of crisis: Jimmy Carter, the power of a peacemaker, and North Korea's nuclear ambitions.* New York, NY: PublicAffairs.

Dixon, William. 1994. "Democracy and the peaceful settlement of international conflict." *American Political Science Review* 88 (01): 14–32.

———. 1996. "Third-party techniques for preventing conflict escalation and promoting peaceful settlement." *International Organization* 50 (04): 653–681.

Epstein, David, and Sharyn O'Halloran. 1994. "Administrative procedures, information, and agency discretion." *American Journal of Political Science:* 697–722.

Favretto, Katja. 2009. "Should peacemakers take sides? Major power mediation, coercion, and bias." *American Political Science Review* 103 (02): 248–263.

Fearon, James D. 1994. "Domestic political audiences and the escalation of international disputes." *American Political Science Review* 88 (03): 577–592.

———. 1999. "Electoral accountability and the control of politicians: selecting good types versus sanctioning poor performance." *Democracy, accountability, and representation* 55:61.

Ferejohn, John. 1986. "Incumbent performance and electoral control." *Public choice* 50 (1): 5–25.

Fey, Mark, and Kristopher W Ramsay. 2010. "When is shuttle diplomacy worth the commute? Information sharing through mediation." *World Politics* 62 (04): 529–560.

Gailmard, Sean. 2002. "Expertise, subversion, and bureaucratic discretion." *Journal of Law, Economics, and Organization* 18 (2): 536–555.

———. 2012. "Accountability and principal-agent models." *Chapter prepared for the Oxford Handbook of Public Accountability.*

Gailmard, Sean, and John W Patty. 2012. "Formal Models of Bureaucracy." *Annual review of political science* 15:353–377.

———. 2013. "Giving Advice vs. Making Decisions: Transparency, Information, and Delegation."

Gartner, Scott Sigmund. 2008. "The multiple effects of casualties on public support for war: An experimental approach." *American Political Science Review* 102 (01): 95–106.

Gartner, Scott Sigmund. 2013. "Deceptive results: Why mediation appears to fail but actually succeeds." *Penn St. JL & Int'l Aff.* 2 (1): 27–37.

Gelpi, Christopher, and Joseph M Grieco. 2001. "Attracting Trouble Democracy, Leadership Tenure, and the Targeting of Militarized Challenges, 1918-1992." *Journal of Conflict Resolution* 45 (6): 794–817.

Gent, Stephen E, and Megan Shannon. 2010. "The effectiveness of international arbitration and adjudication: Getting into a bind." *The Journal of Politics* 72 (02): 366–380.

Gerring, John. 2006. *Case study research: principles and practices.* New York, NY: Cambridge University Press.

Greig, Michael. 2005. "Stepping into the fray: when do mediators mediate?" *American Journal of Political Science* 49 (2): 249–266.

———. 2013. "Nipping Them in the Bud: The Onset of Mediation in Low-intensity Civil Conflicts." *Journal of Conflict Resolution* 59 (2): 336–361.

Greig, Michael, and Paul Diehl. 2006. "Softening up: Making conflicts more amenable to diplomacy." *International Interactions* 32 (4): 355–384.

———. 2012. *International mediation.* Malden, MA: Polity.

Huber, John D, and Arthur Lupia. 2001. "Cabinet instability and delegation in parliamentary democracies." *American Journal of Political Science:* 18–32.

Huber, John D, and Charles R Shipan. 2006. "Politics, Delegation, and Bureaucracy." *The Oxford Handbook of Political Economy* 7:256–272.

Huth, Paul K, Sarah E Croco, and Benjamin J Appel. 2011. "Does international law promote the peaceful settlement of international disputes? Evidence from the study of territorial conflicts since 1945." *American Political Science Review* 105 (02): 415–436.

Keohane, Robert O, Andrew Moravcsik, and Anne-Marie Slaughter. 2000. "Legalized dispute resolution: Interstate and transnational." *International Organization* 54 (03): 457–488.

Kleiboer, Marieke. 1994. "Ripeness of Conflict: A Fruitful Notion?" *Journal of Peace Research* 31 (1): 109–116.

———. 2002. "Great power mediation: Using leverage to make peace." In *Studies in international mediation: Essays in honor of Jeffrey Z. Rubin,* edited by Jacob Bercovitch, 127–140. New York, NY: Palgrave Macmillan.

Kressel, Kenneth. 1972. *Labor mediation: An exploratory survey.* New York, NY: Association of Labor Mediation Agencies.

Kydd, Andrew. 2003. "Which side are you on? Bias, credibility, and mediation." *American Journal of Political Science* 47 (4): 597–611.

Levenotoglu, Bahar, and Ahmer Tarar. 2005. "Prenegotiation public commitment in domestic and international bargaining." *American Political Science Review* 99 (03): 419–433.

Melin, Molly M. 2013. "When states mediate." *Penn St. JL & Int'l Aff.* 2 (1): 78–90.

Miller, Gary. 2005. "The political evolution of principal-agent models." *Annu. Rev. Polit. Sci.* 8:203–225.

Mitchell, Sara McLaughlin. 2002. "A Kantian system? Democracy and third-party conflict resolution." *American Journal of Political Science* 46 (4): 749–759.

Morrow, James D. 1989. "Capabilities, uncertainty, and resolve: A limited information model of crisis bargaining." *American Journal of Political Science* 33 (4): 941–972.

Mousseau, Michael. 1998. "Democracy and compromise in militarized interstate conflicts, 1816-1992." *Journal of Conflict Resolution* 42 (2): 210–230.

Novak, Andrew. 2009. "Face-Saving Maneuvers and Strong Third-Party Mediation: The Lancaster House Conference on Zimbabwe-Rhodesia." *International Negotiation* 14 (1): 149–174.

Pbs.org. 2016a. *Interviews - Ashton Carter — Kim's Nuclear Gamble — FRONTLINE — PBS.* Accessed January 20, 2016.

———. 2016b. *Interviews - Jimmy Carter — Kim's Nuclear Gamble — FRONTLINE — PBS.* Accessed January 20, 2016.

———. 2016c. *Interviews - Robert Gallucci — Kim's Nuclear Gamble — FRONTLINE — PBS.* Accessed January 20, 2016.

Powell, Robert. 1987. "Crisis bargaining, escalation, and MAD." *American Political Science Review* 81 (03): 717–735.

———. 1999. *In the shadow of power: States and strategies in international politics.* Princeton, NJ: Princeton University Press.

Rauchhaus, Robert W. 2006. "Asymmetric information, mediation, and conflict management." *World Politics* 58 (02): 207–241.

Regan, Patrick M, and Allan C Stam. 2000. "In the nick of time: Conflict management, mediation timing, and the duration of interstate disputes." *International Studies Quarterly* 44 (2): 239–260.

Regan, Patrick, Richard Frank, and Aysegul Aydin. 2009. "Diplomatic interventions and civil war: A new dataset." *Journal of Peace Research* 46 (1): 135–146.

Reiter, Dan, and Allan C Stam. 2002. *Democracies at war.* Princeton, NJ: Princeton University Press.

Romer, Thomas, and Howard Rosenthal. 1978. "Political resource allocation, controlled agendas, and the status quo." *Public choice* 33 (4): 27–43.

Ruhe, Constantin. 2015. "Anticipating mediated talks Predicting the timing of mediation with disaggregated conflict dynamics." *Journal of Peace Research* 52 (2): 243–257.

Russett, Bruce, John R Oneal, and David R Davis. 1998. "The third leg of the Kantian tripod for peace: International organizations and militarized disputes, 1950–85." *International Organization* 52 (03): 441–467.

Savun, Burcu. 2008. "Information, Bias, and Mediation Success*." *International Studies Quarterly* 52 (1): 25–47.

Schultz, Kenneth A. 2012. "Why we needed audience costs and what we need now." *Security Studies* 21 (3): 369–375.

Shepsle, Kenneth A. 1989. "Studying Institutions Some Lessons from the Rational Choice Approach." *Journal of Theoretical Politics* 1 (2): 131–147.

Shepsle, Kenneth A, and Barry R Weingast. 1994. "Positive theories of congressional institutions." *Legislative Studies Quarterly* 19 (2): 149–179.

Sigal, Leon. 1998a. "Averting a train wreck with North Korea." *Arms Control Today* 28 (8): 11–15.

———. 1998b. *Disarming Strangers: Nuclear Diplomacy with North Korea.* Princeton, NJ: Princeton University Press, Princeton.

Simmons, Beth A. 2002. "Capacity, Commitment, and Compliance International Institutions and Territorial Disputes." *Journal of Conflict Resolution* 46 (6): 829–856.

Slantchev, Branislav L. 2004. "How initiators end their wars: The duration of warfare and the terms of peace." *American Journal of Political Science* 48 (4): 813–829.

———. 2006. "Politicians, the media, and domestic audience costs." *International Studies Quarterly* 50 (2): 445–477.

Smith, Alastair. 1998. "International crises and domestic politics." *American Political Science Review* 92 (03): 623–638.

Smith, Alastair, and Allan Stam. 2003. "Mediation and Peacekeeping in a Random Walk Model of Civil and Interstate War." *International Studies Review* 5 (4): 115–135.

Stasavage, David. 2004. "Open-door or closed-door? Transparency in domestic and international bargaining." *International Organization* 58 (04): 667–703.

Staton, Jeffrey K, and Georg Vanberg. 2008. "The value of vagueness: delegation, defiance, and judicial opinions." *American Journal of Political Science* 52 (3): 504–519.

Svensson, Isak. 2007. "Mediation with muscles or minds? Exploring power mediators and pure mediators in civil wars." *International Negotiation* 12 (2): 229–248.

Tarar, Ahmer, and Bahar Leventoğlu. 2013. "Limited Audience Costs in International Crises." *Journal of Conflict Resolution* 57 (6): 1065–1089.

Volden, Craig. 2002. "Delegating power to bureaucracies: Evidence from the states." *Journal of Law, Economics, and Organization* 18 (1): 187–220.

Wall, James A. 1981. "Review Section: Mediation An Analysis, Review, and Proposed Research." *Journal of Conflict Resolution* 25 (1): 157–180.

Wall, James A, John B Stark, and Rhetta L Standifer. 2001. "Mediation a current review and theory development." *Journal of Conflict Resolution* 45 (3): 370–391.

Wallensteen, Peter, and Isak Svensson. 2014. "Talking peace: International mediation in armed conflicts." *Journal of Peace Research* 51 (2): 315–327.

Weeks, Jessica L. 2012. "Strongmen and straw men: Authoritarian regimes and the initiation of international conflict." *American Political Science Review* 106 (02): 326–347.

Wit, Joel S, Daniel Poneman, and Robert L Gallucci. 2004. *Going critical: The first North Korean nuclear crisis.* Washington, D.C.: Brookings Institution Press.

Zartman, William. 2000. "Ripeness: The hurting stalemate and beyond." In *International conflict resolution after the Cold War,* edited by Paul C Stern and Daniel Druckman, 225–250. Washington, D.C.: Washington, DC: National Academy Press.

Zartman, William. 2002. "Mediation by regional organizations: the OAU in Chad and Congo." In *Studies in international mediation: Essays in honor of Jeffrey Z. Rubin,* edited by Jacob Bercovitch. New York, NY: Palgrave Macmillan.

———. 2008. "Introduction Bias, Prenegotiation and Leverage in Mediation." *International Negotiation* 13 (3): 305–310.

Zartman, William, and Saadia Touval. 1985. "International mediation: Conflict resolution and power politics." *Journal of Social Issues* 41 (2): 27–45.

1 Appendix

To solve the game, we first ask whether there are pooling or separating equilibria, where both types or one type, accepts the mediator's proposal. This establishes constraints on what the mediator can propose, and allows us to deduce the leader and enemy's sequentially rational best responses. From this we can determine the audience's beliefs and best response, since by Bayes' Rule, its beliefs must be consistent with all other players' strategies. Finally, we check that there are no profitable deviations on or off the equilibrium path to establish the equilibrium.

To assist the analysis, we assume that the mediator does not make an offer that obtains peace with probability zero. We prove that this assumption holds in equilibrium by showing that the mediator obtains peace with positive probability everywhere in the parameter space.

Assumption 1. *For any initial offer, m, if the probability of war is one, $P(war|m) = 1$, then the mediator does not propose m.*

1.1 Pooling on the high offer

Proposition 1 (Region I: High Offer)**.** *When $p \leq p_H^*$, the mediator proposes $m^* = \sigma_H$, and both types of enemy accept, where $p_H^* = \frac{2c}{\tau_H - \tau_L}$. If the enemy accepts, the leader accepts m^* with beliefs $\lambda_1 = p$. If the enemy rejects, the leader's beliefs are $\lambda_2 \geq \frac{2c - \tau_H}{\tau_H - \tau_L}$, and the leader exits to war. The audience does not sanction the leader, $s^* = 0$, with beliefs $\alpha_L^m + \alpha_H^m = 1$ and $\alpha_L^r + \alpha_H^r = 0$. The probability of war is zero.*

Proof of Proposition 1. To see that there a pooling equilibrium in which both types accept, we can deduce a few things. First, for both types to accept, the mediator must propose at least the high offer, $m \geq \sigma_H$. Further, the leader must not raise, otherwise both types will reject m in favor of $m + \delta$, which means the leader must exit. Since the leader exits, a settlement is reached only through the mediator, and therefore consistency requires that the audience's beliefs are the mediator proposed the settlement, $\alpha_L^m + \alpha_H^m = 1$, and the leader did not raise, $\alpha_L^r + \alpha_H^r = 0$. Therefore, the audience does not sanction, $s^* = 0$.

What conditions are required to maintain these strategies? If the enemy rejects m, then rejection is off the equilibrium path. The leader knows that both types will accept any raised offer, since the mediator's offer is already high. Therefore, this can form an equilibrium only if there exists off-path beliefs, λ_2, such that the leader prefers to exit rather than secure a raised settlement. To refine the leader's off-path beliefs, we require that the equilibrium satisfy condition D1, which requires that the leader assign positive weight to the chance that the enemy is a high type, $\lambda_2 \neq 1$, and the mediator make the high offer, $m^* = \sigma_H$.[1] The leader will exit, rather than raise, if war provides a better payoff than the mediated settlement:

$$\begin{aligned}
\lambda_2(-\tau_L - c) + (1-\lambda_2)(-\tau_H - c) &\geq -(\tau_H - c) \\
\lambda_2(\tau_H - \tau_L) - c &\geq c - \tau_H \\
\lambda_2 &\geq \frac{2c - \tau_H}{\tau_H - \tau_L} \equiv \overline{\lambda_2}. \qquad (1)
\end{aligned}$$

The leader has a credible threat to exit as long as there is sufficient probability she faces a low type.

If the enemy accepts m^*, then since both types accept, the leader's beliefs are given

[1] We opt for the fewest restrictions on off-path beliefs. D1 requires that beliefs be supported on any type who stands to gain from deviation (Cho and Kreps, 1987). The low type never stands to gain from deviation, since knowing that the leader plans to exit, accepting the mediator's offer strictly dominates the low type's war payoff from rejecting it. Therefore, $\lambda_2 \neq 1$. Further, $m^* = \sigma_H$ because otherwise rejecting $m' > \sigma_H$ would be strictly dominated for the high type as well, and the leader could not assign positive weight to either type. Alternatively, universal divinity would result in the same high offer, $m^* = \sigma_H$, but would be more restrictive in needing more weight to be placed on the high type, $\lambda_2 < \frac{1}{2}$. The intuitive criterion would be even more restrictive in requiring that zero weight be put on the low type.

by her prior, $\lambda_1 = p$. Given this, the leader will accept m^* if:

$$-(\tau_H - c) \geq -p\tau_L - (1-p)\tau_H - c$$
$$p \leq \frac{2c}{\tau_H - \tau_L} \equiv p_H^* \tag{2}$$

The leader accepts the high offer, m^*, as long as there is sufficient probability she faces a high type. Since $\overline{\lambda_2} < p_H^*$, the leader's off-path beliefs are reasonable given her priors, and the above strategies can be supported. Since m^* guarantees peace, the mediator makes this proposal whenever $p \leq p_H^*$.

□

Lemma 1 (No Separating Equilibrium)**.** *There exists no separating equilibrium in which the low type accepts and the high type rejects the mediator's proposal.*

Proof of Lemma 1. To see that there is no separating equilibrium, suppose that the low type accepts an offer m and the high type rejects it. Then the leader believes that an enemy who accepts must be a low type, $\lambda_1 = 1$, and that an enemy who rejects must be a high type, $\lambda_2 = 0$. There are two possibilities: either the leader raises the offer, or exits to war. If the leader raises, then the low type will have a profitable deviation to reject m; thus, the leader must exit. However, if the leader exits, then settlement occurs only through the mediator, and by consistency, the audience does not sanction the leader, $s = 0$. To see that this is not an equilibrium, observe that the leader will raise as long as there exists some δ such that raising is preferred to exiting:

$$U_L(Exit|\lambda_2) \leq U_L(Raise|\lambda_2)$$
$$-\tau_H - c \leq -m - \delta$$
$$\delta \leq \tau_H + c - m.$$

For all $m < \tau_H + c$, since the leader believes the enemy is a high type, and knows the audience will not sanction, there exists some $\delta > 0$ such that the leader deviates to raise. The only way that the leader exits is if the mediator offers $m \geq \tau_H + c$, but then the high type profitably deviates to accept m, since $\tau_H + c > \tau_H - c$. $\square$

1.2 Semi-separating equilibrium

The following lemmas specifies the best responses for each actor, before characterizing the semi-separating equilibria.

Lemma 2 (Enemy Response to m)**.** *In any semi-separating equilibrium, the low type must mix between accepting and rejecting m, while the high type rejects m.*

Proof of Lemma 2. To form a semi-separating equilibrium, it must be that the low type mixes between accepting and rejecting the mediator's offer, m, while the high type always rejects m. The reverse – for the high type to mix, and the low type to reject m – would not make sense.

To see this, let r represent the probability that the leader raises, and $1 - r$ the probability the leader exits. For the high type to mix, he must be indifferent between accepting and rejecting m, $U_{\tau_H}(\textit{accept } m) = m = r(m+\delta)+(1-r)(\sigma_H) = U_{\tau_H}(\textit{reject } m)$. But if that is true, then the low type will deviate to accept m, since for any r, m, and δ, the low type's payoff for rejecting m is strictly lower than the high type's, $U_{\tau_L}(\textit{reject } m) = r(m + \delta) + (1 - r)(\sigma_L) < U_{\tau_H}(\textit{reject } m)$, and therefore $m > U_{\tau_L}(\textit{reject } m)$.

Further, it would not make sense for the low and high type to semi-separate in response to the leader's raise. That would require that the raised offer be equivalent to the low type's reservation value for war, $m + \delta = \sigma_L$, to make the low type indifferent. The probability of peace would be less than p, since the low type is mixing. But then, the mediator could make an offer in between the low type and leader's reservation values, $m \in (\sigma_L, \tau_L + c]$, that the low type would strictly prefer and the leader would be willing to accept. The mediator would strictly prefer this outcome in securing peace with probability p. Thus, semi-separation must occur about m.

Therefore, let q_L represent the probability that the low type accepts m, and $1 - q_L$ represent the probability the low type rejects m. $\square$

Lemma 3 (Low Type)**.** *For the low type to mix between accepting and rejecting m, the leader must raise with probability $\overline{r} = \frac{m-\sigma_L}{m+\delta-\sigma_L}$.*

Proof of Lemma 3. For the low type to mix, the low type must be indifferent between accepting and rejecting m, $U_{\tau_L}(accept\ m|\cdot) = U_{\tau_L}(reject\ m|r,\delta)$. This section proves that for the low type to mix: 1. the leader must accept m following the enemy's acceptance of m, 2. the leader must raise with probability $\overline{r} = \frac{m-(\tau_L-c)}{m+\delta-(\tau_L-c)}$.

1. To see that the leader must accept m, consider the following proof by contradiction. Suppose that the leader rejects m. Then the low type knows that by accepting m, he receives his war payoff, σ_L. To keep the low type indifferent, the leader must not raise: if the leader raises with any positive probability r, then the low type would not be indifferent since a settlement $m+\delta$ with any positive probability is strictly preferred to war with certainty, $r(m+\delta)+(1-r)\sigma_L > \sigma_L$. But then war occurs with probability one, since when the enemy rejects m the leader also rejects, and when the enemy accepts m, the leader exits to war. By Assumption 1, this is not an equilibrium. Therefore, the leader must accept m.

2. To see that the leader must raise, consider the following. For the low type to mix, the low type must be indifferent between accepting and rejecting m. By the argument above, the low type will receive a utility of m if he accepts. Given this, the low type's indifference condition is:

$$\begin{aligned} U_{\tau_L}(accept\ m) &= U_{\tau_L}(reject\ m|r,\delta) \\ m &= r(m+\delta)+(1-r)(\tau_L-c) \end{aligned} \tag{1}$$

There are two ways to satisfy this indifference condition: either a) the leader never raises, $r=0$, and $m=\tau_L-c$; or b) the leader raises with positive probability that keeps the low type indifferent,

$$r = \frac{m-(\tau_L-c)}{m+\delta-(\tau_L-c)} \equiv \overline{r}. \tag{2}$$

To see that a) is not an equilibrium note that If the leader does not raise, $r = 0$, settlement is reached only through the mediator, and the audience will not sanction the leader. But then the leader can profitably deviate to raise, because:

$$
\begin{aligned}
U_L(exit|\lambda_2) &\leq U_L(raise|\lambda_2) \\
\lambda_2(-\tau_L - c) + (1 - \lambda_2)(-\tau_H - c) &\leq -m - \delta \\
\lambda_2(\tau_H - \tau_L) - \tau_H - c &\leq -\tau_L + c - \delta \\
\lambda_2(\tau_H - \tau_L) &\leq \tau_H - \tau_L + 2c - \delta \\
\delta &\leq (\tau_H - \tau_L)(1 - \lambda_2) + 2c.
\end{aligned}
$$

there exists $\delta > 0$ such that the leader prefers to deviate. Since a) is not an equilibrium, it must be that b) the leader raises with probability $\overline{r}$.

□

Lemma 4 (Leader's response to acceptance)**.** *When the enemy accepts m, the leader will accept if $m \leq \overline{\sigma_L}$ with beliefs $\lambda_1 = 1$, where $\overline{\sigma_L} = \tau_L + c - sa(\sigma)$.*

Proof of Lemma 4. Upon observing the enemy accept m, the leader's beliefs are that the enemy must be a low type, $\lambda_1 = 1$. The leader's best response is to accept m if the mediator's proposal no greater than the leader's maximum settlement against the low type:

$$
m \leq \tau_L + c - sa(\sigma) \equiv \overline{\sigma_L}. \tag{3}
$$

□

Lemma 5 (Leader's response to rejection)**.** *When the enemy rejects, the leader's beliefs that the enemy is a low type are $\lambda_2 = \frac{p(1-q_L)}{1-pq_L}$. The leader will mix between raising with $\delta^* = \sigma_H - m$ and exiting, if the low type accepts m with probability $\overline{q_L} = \frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{p[\tau_H - \tau_L - 2c + sa(\sigma_H)]}$. The leader's beliefs when the low type plays this strategy is $\lambda_2 = \frac{2c - sa(\sigma_H)}{\tau_H - \tau_L}$.*

Proof of Lemma 5. When the enemy rejects m, then the leader updates her beliefs that 2 is a low type, λ_2. Since the low type rejects m with probability $1 - q_L$, and the high type always rejects m, the leader's posterior belief that the enemy is a low type is:

$$\lambda_2 = \frac{p(1 - q_L)}{p(1 - q_L) + 1 - p} = \frac{p(1 - q_L)}{1 - pq_L}. \tag{4}$$

If the leader raises, she must raise with $\delta^* = \sigma_H - m$. Why? The leader will not raise with anything higher, $\delta' > \delta^*$, because then the leader overpays for peace against both types. The leader will not raise with anything lower, $\delta' < \delta^*$, because then the leader overpays for peace against the low type: since m is acceptable to the low type, the leader can offer any $\delta > 0$ and secure peace against the low type for a lower price, thus any $\delta' < \delta^*$ cannot form an equilibrium (the leader can always deviate to ϵ lower).[2] Therefore, the only reason for the leader to raise is to change the outcome by securing peace against the high type with $\delta^* = \sigma_H - m$. By sequential rationality, both types will accept this raised offer.

We can now plug these components into the leader's indifference condition. If the leader exits to war, she fights either the low or high type, and if the leader raises, then she offers a total settlement σ_H and pays audience costs with probability s:

$$\begin{aligned} U_L(exit|\lambda_2) &= U_L(raise|\lambda_2, m + \delta^* = \sigma_H) \\ \lambda_2(-\tau_L - c) + (1 - \lambda_2)(-\tau_H - c) &= -\sigma_H - sa(\sigma_H) \\ \lambda_2(\tau_H - \tau_L) - \tau_H - c &= -\tau_H + c - sa(\sigma_H) \\ \lambda_2 &= \frac{2c - sa(\sigma_H)}{\tau_H - \tau_L} \equiv p_H^* - \frac{sa(\sigma_H)}{\tau_H - \tau_L}. \end{aligned} \tag{5}$$

Given the leader's beliefs, λ_2, from (4), we can rearrange the leader's indifference condition

[2]By definition, it would not make sense for the leader to "back down" by offering zero concessions, thus, we do not allow $\delta = 0$.

as follows:

$$
\begin{aligned}
\frac{p(1-q_L)}{1-pq_L} &= \frac{2c - sa(\sigma_H)}{\tau_H - \tau_L} \\
p(1-q_L)(\tau_H - \tau_L) &= (2c - sa(\sigma_H))(1 - pq_L) \\
p(\tau_H - \tau_L) - pq_L(\tau_H - \tau_L) &= 2c - sa(\sigma_H) - pq_L(2c - sa(\sigma_H)) \\
p(\tau_H - \tau_L) - 2c + sa(\sigma_H) &= pq_L[\tau_H - \tau_L - 2c + sa(\sigma_H)] \\
\overline{q_L} &= \frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{p[\tau_H - \tau_L - 2c + sa(\sigma_H)]}. \qquad (6)
\end{aligned}
$$

This indicates that for the leader to be indifferent, the low type must accept m with probability $\overline{q_L} = \frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{p[\tau_H - \tau_L - 2c + sa(\sigma_H)]}$.

□

Lemma 6 (Audience)**.** *The audience's best response is sanction if $m > \widehat{\sigma_L}$, to not sanction if $m < \widehat{\sigma_L}$, with beliefs $\alpha_H^m = 0$, and α_L^m, α_L^r, and α_H^r given by (7), (8), and (9). The audience is indifferent when $m = \widehat{\sigma_L}$, where $\widehat{\sigma_L} = \frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{1-p} + \tau_L - c$, with beliefs $\alpha_H^m = 0$, and $\alpha_L^m = \alpha_L^r + \alpha_H^r = \frac{1}{2}$.*

Proof of Lemma 6. Given Lemmas 3, 4, and 5 settlement occurs on the equilibrium path. The audience updates its beliefs that upon observing a settlement, and believes that under no condition has the high type accepted the mediator's offer, $\alpha_H^m = 0$. The audience believes that the low type accepted the mediator's offer with probability:

$$
\alpha_L^m = \frac{pq_L}{pq_L + p(1-q_L)r + (1-p)r}, \qquad (7)
$$

the low type accepted the leader's raised offer with probability

$$
\alpha_L^r = \frac{p(1-q_L)r}{pq_L + p(1-q_L)r + (1-p)r}, \qquad (8)
$$

and the high type accepted the leader's raised offer with probability

$$
\alpha_H^r = \frac{(1-p)r}{pq_L + p(1-q_L)r + (1-p)r}. \qquad (9)
$$

Given these beliefs, the audience's best response is to sanction if the following holds:

$$\begin{aligned}U_A(sanction|\cdot) &\geq U_A(not\ sanction|\cdot)\\ \alpha_L^r + \alpha_H^r &\geq \alpha_L^m + \alpha_H^m\\ p(1-q_L)r + (1-p)r &\geq pq_L\\ pr - q_L pr + r - pr &\geq pq_L\\ r - q_L pr &\geq pq_L\\ r &\geq \frac{pq_L}{1-pq_L}. \end{aligned} \tag{10}$$

Since we know the probability that the low type accepts m, $\overline{q_L}$, and the probability the leader raises, $\overline{r}$, we can plug these values into (10) to determine the audience's best response.

Substitution of $\overline{q_L} = \frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{p[\tau_H - \tau_L - 2c + sa(\sigma_H)]}$ gives

$$\begin{aligned}r &\geq \frac{\frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{\tau_H-\tau_L-2c+sa(\sigma_H)}}{1-\frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{\tau_H-\tau_L-2c+sa(\sigma_H)}}\\ &\geq \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{\tau_H-\tau_L-2c+sa(\sigma_H)-p(\tau_H-\tau_L)+2c-sa(\sigma_H)}\\ &\geq \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{(\tau_H-\tau_L)(1-p)}.\end{aligned}$$

Then, substitution of $\overline{r} = \frac{m-\tau_L+c}{\tau_H-\tau_L}$ gives

$$\begin{aligned}\frac{m-\tau_L+c}{\tau_H-\tau_L} &\geq \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{(\tau_H-\tau_L)(1-p)}\\ m-\tau_L+c &\geq \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{1-p}\\ m &\geq \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{1-p} + \tau_L - c \equiv \widehat{\sigma_L}.\end{aligned} \tag{11}$$

The audience's best response is sanction if $m > \widehat{\sigma_L}$, not to sanction if $m \leq \widehat{\sigma_L}$, and to be indifferent if $m = \widehat{\sigma_L}$, where $\widehat{\sigma_L} = \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{1-p} + \tau_L - c$. □

Lemma 7 (Mediator)**.** *The mediator's best response is to offer* $m^* = \min\{\overline{\sigma_L}, \widehat{\sigma_L}\}$*, which means the mediator offers* $m^* = \overline{\sigma_L}$ *when* $p > p_L^*$*, and offers* $m^* = \widehat{\sigma_L}$ *when* $p < p_L^*$*, where*

$p_L^* = \frac{4c}{\tau_H - \tau_L + 2c}$. *The probability of war is* $P(war) = \frac{(1-p)(\tau_H - m - c)}{\tau_H - \tau_L - 2c + sa(\sigma_H)}$.

Proof of Lemma 7. Given these best responses, the mediator makes a proposal that minimizes the probability of war. War occurs in two ways. Either the enemy is a low type who rejected m, and the leader did not raise, or the enemy is a high type who rejected m, and the leader did not raise. Therefore, the probability of war is:

$$P(war) = p(1 - q_L)(1 - r) + (1 - p)(1 - r), \tag{12}$$

which reduces to $P(war) = (1 - r)(1 - pq_L)$. Substitution of $\overline{r}$ and $\overline{q_L}$ gives:

$$\begin{aligned} P(war) &= \left(1 - \frac{m - \tau_L + c}{\tau_H - \tau_L}\right)\left(1 - \frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{\tau_H - \tau_L - 2c + sa(\sigma_H)}\right) \\ &= \left(\frac{\tau_H - m - c}{\tau_H - \tau_L}\right)\left(\frac{(\tau_H - \tau_L)(1 - p)}{\tau_H - \tau_L - 2c + sa(\sigma_H)}\right) \\ &= \frac{(1 - p)(\tau_H - m - c)}{\tau_H - \tau_L - 2c + sa(\sigma_H)}. \end{aligned} \tag{13}$$

The probability of war is decreasing in m and s. If $s = 0$, then the probability of war is only decreasing in m, and the mediator proposes the most that the leader will tolerate, $m^* = \overline{\sigma_L}$. In order for $s = 0$, by Lemma 6, the audience will not sanction if $m^* \leq \widehat{\sigma_L}$, which is true if:

$$\begin{aligned} \overline{\sigma_L} &\leq \widehat{\sigma_L} \\ \tau_L + c &\leq \frac{p(\tau_H - \tau_L) - 2c}{1 - p} + \tau_L - c \\ (2c)(1 - p) &\leq p(\tau_H - \tau_L) - 2c \\ 4c &\leq p(\tau_H - \tau_L + 2c) \\ p &\geq \frac{4c}{\tau_H - \tau_L + 2c} \equiv p_L^*, \end{aligned} \tag{14}$$

where $p_L^* > p_H^*$ since $2c < \tau_H - \tau_L$.[3] Therefore, when $p \geq p_L^*$, the mediator offers $m^* = \overline{\sigma_L}$

[3]

$$p_L^* > p_H^*$$

and the audience does not sanction, $s^* = 0$. Let us call this Region II. Proposition 2 specifies this equilibrium.

When $p \in (p_H^*, p_L^*)$, the relationship between $\overline{\sigma_L}$ and $\widehat{\sigma_L}$ depends on s. In other words, there exists an s such that these values are equal. This means that mediator can either offer: 1) $m = \overline{\sigma_L} > \widehat{\sigma_L}$, which gets the leader sanctioned, $s = 1$, 2) $m = \widehat{\sigma_L} < \overline{\sigma_L}$, which keeps the audience indifferent, or 3) $m = \widehat{\sigma_L} = \overline{\sigma_L}$, which also keeps the audience indifferent but is the highest offer the leader will tolerate. We examine each in Lemmas 8, 9, and 10. Let us call this Region III. Proposition 3 specifies this equilibrium.

□

Proposition 2 (Region II: No Sanction)**.** *When $p \geq p_L^*$, the mediator offers $m^* = \tau_L + c$, where $p_L^* = \frac{4c}{\tau_H - \tau_L + 2c}$. The low type accepts with probability $q_L^* = \frac{p(\tau_H - \tau_L) - 2c}{p(\tau_H - \tau_L - 2c)}$, and the high type rejects. If the enemy accepts, the leader accepts m^* with beliefs $\lambda_1 = 1$. If the enemy rejects, the leader raises with probability $r^* = \frac{2c}{\tau_H - \tau_L}$ to offer $\delta^* = \tau_H - \tau_L - 2c$ with beliefs $\lambda_2 = \frac{2c}{\tau_H - \tau_L}$. Both types accept the raised offer. The audience does not sanction, $s^* = 0$, with beliefs $\alpha_H^m = 0$, and $\alpha_L^m > \alpha_L^r + \alpha_H^r$. The probability of war is $1 - p$.*

Proof of Proposition 2. When $p > p_L^*$, $\overline{\sigma_L} < \widehat{\sigma_L}$, which implies $\alpha_L^m > \alpha_L^r + \alpha_H^r$. Therefore, by Lemma 6, the audience does not sanction, $s^* = 0$, and the mediator offers $m^* = \overline{\sigma_L} = \tau_L + c$. By Lemma 5, the low type accepts m^* with probability $q_L^* = \frac{p(\tau_H - \tau_L) - 2c}{p(\tau_H - \tau_L - 2c)}$, which maintains the leader's indifference, while the high type rejects m. By Lemma 4, the leader updates her beliefs, $\lambda_1 = 1$, and accepts m, since this meets her reservation value against the low type, $m = \overline{\sigma_L}$. By Lemmas 3 and 5, when the enemy rejects m, the leader updates her beliefs, $\lambda_2 = \frac{2c}{\tau_H - \tau_L}$, and raises with probability $r^* = \frac{2c}{\tau_H - \tau_L}$, and offers

$$\frac{4c}{\tau_H - \tau_L + 2c} > \frac{2c}{\tau_H - \tau_L}$$
$$\frac{2}{\tau_H - \tau_L + 2c} > \frac{1}{\tau_H - \tau_L}$$
$$2(\tau_H - \tau_L) > \tau_H - \tau_L + 2c$$
$$\tau_H - \tau_L > 2c$$

$\delta^* = \tau_H - \tau_L - 2c$, which maintains the low type's indifference. Both types accept the raised offer, since $m^* + \delta^* = \sigma_H$. By Lemma 7, the probability of war is

$$P(war) = \frac{(1-p)(\tau_H - m^* - c)}{\tau_H - \tau_L - 2c + s^* a(\sigma_H)} \tag{15}$$

$$= \frac{(1-p)(\tau_H - \tau_L - 2c)}{\tau_H - \tau_L - 2c} \tag{16}$$

$$= 1 - p. \tag{17}$$

□

Lemma 8 (Region II: Audience Sanctions, $m = \overline{\sigma_L} > \widehat{\sigma_L}$). *When $p < p_L^*$, and the audience is weak, $a(\sigma_H) < 2c$, then the mediator can offer $m = \tau_L + c - a(\sigma)$. The low type accepts with probability $\overline{q_L} = \frac{p(\tau_H - \tau_L) - 2c + a(\sigma_H)}{p[\tau_H - \tau_L - 2c + a(\sigma_H)]}$, the high type rejects. If the enemy accepts, the leader accepts with beliefs $\lambda_1 = 1$. If the enemy rejects, the leader believes it is a low type with probability $\lambda_2 = \frac{2c - a(\sigma_H)}{\tau_H - \tau_L}$, and raises with probability $\overline{r} = \frac{2c - a(\sigma)}{\tau_H - \tau_L}$ and concessions $\delta = \tau_H - \tau_L - 2c + a(\sigma)$. Both types accept the raised offer. The audience sanctions the leader, $s = 1$. The probability of war is $\frac{(1-p)[\tau_H - \tau_L - 2c + a(\sigma)]}{\tau_H - \tau_L - 2c + a(\sigma_H)}$ which is less than $1 - p$. This forms an equilibrium as long as the mediator does not prefer another offer.*

Proof of Lemma 8. When $p < p_L^*$, the mediator can offer $m = \overline{\sigma_L} = \tau_L + c - a(\sigma)$, and by Lemma 6, the audience will sanction the leader, $s = 1$. If $s = 1$, then by Lemma 5, the low type must accept with probability $\overline{q_L} = \frac{p(\tau_H - \tau_L) - 2c + a(\sigma_H)}{p[\tau_H - \tau_L - 2c + a(\sigma_H)]}$ for the leader to mix. The low type is willing to accept m as long as $\tau_L + c - a(\sigma) > \tau_L - c$, which is true if $2c > a(\sigma)$. The high type rejects m.

If the enemy accepts m, then by Lemma 4, the leader believes the enemy is a low type with probability $\lambda_1 = 1$, and accepts since m is the most she will tolerate against the low type. If the enemy rejects, then by Lemma 5, the leader's beliefs are $\lambda_2 = \frac{2c - a(\sigma_H)}{\tau_H - \tau_L}$, where $\lambda_2 \in [0, 1]$ if $2c > a(\sigma_H)$. We will refer to this as the "weak audience" requirement.[4]

[4]Note that if $2c > a(\sigma_H)$, then $2c > a(\sigma)$, since $a(\sigma_H) > a(\sigma)$. Therefore, both the low type and leader's strategies are maintained if the weak audience requirement is satisfied.

By Lemma 3, the leader must raise with probability $\overline{r} = \frac{2c-a(\sigma)}{\tau_H-\tau_L}$ and concessions $\delta = \tau_H - \tau_L - 2c + a(\sigma)$ to keep the low type indifferent, where $\delta > 0$ since $2c < \tau_H - \tau_L$. Both types accept the raised offer.

These strategies accord with the best responses required for a semi-separating equilibrium. By Lemma 7, the probability of war is:

$$P(war) = \frac{(1-p)[\tau_H - \tau_L - 2c + a(\sigma)]}{\tau_H - \tau_L - 2c + a(\sigma_H)},$$

which is less than $1-p$ since $a(\sigma) < a(\sigma_H)$. This forms an equilibrium as long as the mediator does not prefer another offer. $\square$

Lemma 9 (Region III: Audience Indifferent, $m = \widehat{\sigma_L} < \overline{\sigma_L}$)**.** *When $p < \min\{p_L^*, \frac{1}{2}\}$, if the audience is sufficiently strong, $a(\sigma_H) > 2c$, and sanctioning for accepting the mediator's offer is not too high, $a(\sigma) \leq 2c - \frac{p(\tau_H-\tau_L)}{1-p}$, then the mediator can offer $m = \sigma_L + \frac{p(\tau_H-\tau_L)}{1-p}$. The low type accepts, $\overline{q_L} = 1$, the high type rejects. If the enemy accepts, then the leader accepts the mediator's offer with beliefs $\lambda_1 = 1$. If the enemy rejects, the leader raises with probability $\overline{r} = \frac{p}{1-p}$ and concessions $\delta = \frac{(1-2p)(\tau_H-\tau_L)}{1-p}$ with beliefs $\lambda_2 = 0$. The audience sanctions with probability $s_H^* = \frac{2c}{a(\sigma_H)}$ and beliefs $\alpha_L^m = \alpha_H^r = \frac{1}{2}$, $\alpha_L^r = 0$, and $\alpha_H^m = 0$. The probability of war is $1-2p$. This forms an equilibrium as long as the mediator does not prefer another offer.*

Proof of Lemma 9. Alternatively, when $p < p_L^*$, the mediator can offer $m = \widehat{\sigma_L} = \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{1-p} + \tau_L - c$, which makes the audience indifferent. However, notice that m is a function of s.

Therefore, plugging m into the $P(war)$ will yield an expression for the $P(war)$ that only depends on s:

$$\begin{aligned}
P(war|\widehat{\sigma_L}) &= \frac{(1-p)\left(\tau_H - c - \left(\frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{1-p} + \tau_L - c\right)\right)}{\tau_H - \tau_L - 2c + sa(\sigma_H)} \\
&= \frac{(1-p)(\tau_H - \tau_L) - p(\tau_H - \tau_L) + 2c - sa(\sigma_H)}{\tau_H - \tau_L - 2c + sa(\sigma_H)} \\
&= \frac{(1-2p)(\tau_H - \tau_L) + 2c - sa(\sigma_H)}{\tau_H - \tau_L - 2c + sa(\sigma_H)}. \qquad (18)
\end{aligned}$$

This is decreasing in s, which can be seen by sketching a similar function, $y = \frac{1-x}{2+x}$, or by taking the derivative of the $P(war|\widehat{\sigma_L})$ with respect to s.[5] Therefore, the mediator's best option is to offer the value of m that corresponds to the maximum value of s.

The maximum value of s is given by constraints in the best responses of the other actors, which also depend on s. The low type accepts m with probability $\overline{q_L} = \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{p[\tau_H-\tau_L-2c+sa(\sigma_H)]}$, which requires that $p(\tau_H - \tau_L) - 2c + sa(\sigma_H) > 0$. The leader accepts if $m^* < \overline{\sigma_L} = \tau_L + c - sa(\sigma)$. If the enemy rejects, the leader raises with probability $\overline{r} = \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{(1-p)(\tau_H-\tau_L)}$,[6] with additional concessions $\delta^* = \frac{(1-2p)(\tau_H-\tau_L)+2c-sa(\sigma_H)}{1-p}$,[7] and beliefs $\lambda_2 = \frac{2c-sa(\sigma_H)}{\tau_H-\tau_L}$.

Given these strategies, the constraints on the audience's probability of sanctioning

---

[5]Using the quotient rule, where $\frac{\delta[(1-2p)(\tau_H-\tau_L)+2c-sa(\sigma_H)]}{\delta s} = -a(\sigma_H)$ and $\frac{\delta[\tau_H-\tau_L-2c+sa(\sigma_H)]}{\delta s} = a(\sigma_H)$, we obtain:

$$\begin{aligned}\frac{\delta P(war|\widehat{\sigma_L})}{\delta s} &= \frac{-a(\sigma_H)[\tau_H - \tau_L - 2c + sa(\sigma_H)] - a(\sigma_H)[(1-2p)(\tau_H - \tau_L) + 2c - sa(\sigma_H)]}{[\tau_H - \tau_L - 2c + sa(\sigma_H)]^2} \\ &= \frac{-a(\sigma_H)[2(1-p)(\tau_H - \tau_L)]}{[\tau_H - \tau_L - 2c + sa(\sigma_H)]^2} < 0.\end{aligned}$$

Since the denominator is positive, $\tau_H - \tau_L - 2c + sa(\sigma_H) > 0$, we know that the derivative of $P(war|\widehat{\sigma_L})$ with respect to s is negative. Therefore, $P(war|\widehat{\sigma_L})$ is minimized by the maximum value of s.

[6]By Lemma 3, the leader must raise with the following probability to maintain the low type's indifference:

$$\begin{aligned}\overline{r} &= \frac{m^*_{\widehat{\sigma_L}} - \tau_L + c}{\tau_H - \tau_L} \\ &= \frac{\frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{1-p} + \tau_L - c - \tau_L + c}{\tau_H - \tau_L} \\ &= \frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{(1-p)(\tau_H - \tau_L)}.\end{aligned}$$

[7]By Lemma 5, the leader raises with additional concessions given by:

$$\delta^*_{\widehat{\sigma_L}} = \sigma_H - m^*_{\widehat{\sigma_L}}$$

can be summarized as follows:

$$\lambda_2 \in [0,1] \longleftrightarrow s \in \left[\frac{2c-(\tau_H-\tau_L)}{a(\sigma_H)}, \frac{2c}{a(\sigma_H)}\right], \tag{19}$$

and

$$\overline{r} \in [0,1] \longleftrightarrow s \in \left[\frac{2c-p(\tau_H-\tau_L)}{a(\sigma_H)}, \frac{2c+(1-2p)(\tau_H-\tau_L)}{a(\sigma_H)}\right], \tag{20}$$

where one can verify that these constraints also satisfy $m^* > \sigma_L$, $\overline{q_L} \in [0,1]$, and $\delta > 0$.

These constraints are ordered thus creating two possibilities.[8]

1. When a high type is more likely, $p < \frac{1}{2}$, the audience may sanction with any probability in the following range to maintain this equilibrium. Let us denote the set of probabilities s_H.

$$s_H = \left\{s : \frac{2c-p(\tau_H-\tau_L)}{a(\sigma_H)} < s < \frac{2c}{a(\sigma_H)}\right\}.$$

2. When a low type is more likely, $p > \frac{1}{2}$, the audience may sanction with any prob-

$$\begin{aligned}
&= \tau_H - c - \left(\frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{1-p} + \tau_L - c\right) \\
&= \tau_H - \tau_L - \frac{p(\tau_H-\tau_L)-2c+sa(\sigma_H)}{1-p} \\
&= \frac{(1-2p)(\tau_H-\tau_L)+2c-sa(\sigma_H)}{1-p}.
\end{aligned}$$

[8]We know that $\frac{2c-(\tau_H-\tau_L)}{a(\sigma_H)} < \frac{2c-p(\tau_H-\tau_L)}{a(\sigma_H)} < \frac{2c}{a(\sigma_H)}$, since $p \in [0,1]$. Further, when $p < \frac{1}{2}$, we know that $1-2p > 0$, and therefore $\frac{2c}{a(\sigma_H)} < \frac{2c+(1-2p)(\tau_H-\tau_L)}{a(\sigma_H)}$. This gives the ordering for possibility 1: $\frac{2c-(\tau_H-\tau_L)}{a(\sigma_H)} < \frac{2c-p(\tau_H-\tau_L)}{a(\sigma_H)} < \frac{2c}{a(\sigma_H)} < \frac{2c+(1-2p)(\tau_H-\tau_L)}{a(\sigma_H)}$. When $p > \frac{1}{2}$, $1-2p < 0$, and therefore, $\frac{2c+(1-2p)(\tau_H-\tau_L)}{a(\sigma_H)} < \frac{2c}{a(\sigma_H)}$. Further, it can be shown that $\frac{2c-p(\tau_H-\tau_L)}{a(\sigma_H)} < \frac{2c+(1-2p)(\tau_H-\tau_L)}{a(\sigma_H)}$ for all $p < 1$, and when $p = 1$, $\frac{2c-p(\tau_H-\tau_L)}{a(\sigma_H)} = \frac{2c+(1-2p)(\tau_H-\tau_L)}{a(\sigma_H)}$. This gives the ordering for possibility 2: $\frac{2c-(\tau_H-\tau_L)}{a(\sigma_H)} < \frac{2c-p(\tau_H-\tau_L)}{a(\sigma_H)} < \frac{2c+(1-2p)(\tau_H-\tau_L)}{a(\sigma_H)} < \frac{2c}{a(\sigma_H)}$.

ability in the following range to maintain this equilibrium. Let us denote this set s_L.

$$s_L = \Big\{ s : \frac{2c - p(\tau_H - \tau_L)}{a(\sigma_H)} < s < \frac{2c + (1-2p)(\tau_H - \tau_L)}{a(\sigma_H)} \Big\}.$$

Since the mediator chooses the offer that induces the highest probability of sanctioning, $s_H^* = \frac{2c}{a(\sigma_H)}$, and $s_L^* = \frac{2c+(1-2p)(\tau_H-\tau_L)}{a(\sigma_H)}$.

One can quickly show that s_L^* does not form an equilibrium. Substitution indicates that the mediator's offer is $m^* = \frac{p(\tau_H-\tau_L)+(1-2p)(\tau_H-\tau_L)}{1-p} + \tau_L - c = \tau_H - c$. The leader accepts m^* if

$$\tau_H - c \leq \tau_L + c - sa(\sigma)$$
$$\tau_H - \tau_L - 2c \leq -sa(\sigma).$$

Since $2c < \tau_H - \tau_L$, the left side of this equation is positive, while the right side is negative. Therefore, the leader rejects this offer, and s_L^* is not in equilibrium.

s_H^* forms an equilibrium: When $p < \frac{1}{2}$, the audience sanctions with probability $s_H^* = \frac{2c}{a(\sigma_H)}$, and the mediator offers $m = \frac{p(\tau_H-\tau_L)}{1-p} + \tau_L - c = \sigma_L + \frac{p(\tau_H-\tau_L)}{1-p}$. For $s_H^* \in (0,1)$, the audience must be sufficiently strong, $a(\sigma_H) > 2c$. We will refer to this as the "strong audience" requirement.

By Lemma 5, the low type must accept m with probability $\overline{q_L} = 1$ to maintain the leader's indifference. The leader's beliefs are $\lambda_1 = 1$ and $\lambda_2 = 0$. The leader accepts the mediator's offer as long as:

$$m \leq \tau_L + c - a(\sigma)s_H^*$$
$$\sigma_L + \frac{p(\tau_H - \tau_L)}{1-p} \leq \tau_L + c - a(\sigma)\left[\frac{2c}{a(\sigma_H)}\right].$$

Since this must hold for $a(\sigma_H) > 2c$, the right side of the inequality is strictly greater than $\tau_L + c - a(\sigma)$. Therefore, a sufficient condition for the above to hold is:

$$\sigma_L + \frac{p(\tau_H - \tau_L)}{1-p} \leq \tau_L + c - a(\sigma)$$

$$a(\sigma) \leq 2c - \frac{p(\tau_H - \tau_L)}{1-p}.$$

The sanction for accepting the mediator's punishment must be sufficiently low.

By Lemma 5, the leader must raise with probability $\overline{r} = \frac{p(\tau_H - \tau_L)}{(1-p)(\tau_H - \tau_L)} = \frac{p}{1-p}$, with additional concessions $\delta = \frac{(1-2p)(\tau_H - \tau_L)}{1-p}$. Given these strategies, the audience's beliefs are $\alpha_L^m = \frac{p}{p+(1-p)r} = \frac{p}{p+(1-p)\left[\frac{p(\tau_H-\tau_L)}{(1-p)(\tau_H-\tau_L)}\right]} = \frac{1}{2}$, $\alpha_L^r = 0$, $\alpha_H^m = 0$, $\alpha_H^r = \frac{1}{2}$, which maintains its indifference. By (18), the probability of war is $P(war) = \frac{(1-2p)(\tau_H-\tau_L)}{\tau_H-\tau_L} = 1-2p$. This is an equilibrium possibility for p where $p < \min\{p_L^*, \frac{1}{2}\}$ as long as the mediator does not prefer another offer. □

Lemma 10 (Region III: Audience Indifferent, $m = \widehat{\sigma_L} = \overline{\sigma_L}$)**.** *When $p < p_L^*$, the mediator can offer $m = \tau_L + c - a(\sigma)s_G^*$. The low type accepts with probability $\overline{q_L} = \frac{p(\tau_H-\tau_L)-2c+a(\sigma_H)s_G^*}{p\left[\tau_H-\tau_L-2c+a(\sigma_H)s_G^*\right]}$, and the high type rejects. If the enemy accepts, then the leader accepts the mediator's offer with beliefs $\lambda_1 = 1$. If the enemy rejects, the leader raises with probability $\overline{r} = \frac{2c-a(\sigma)s_G^*}{\tau_H-\tau_L}$ and concessions $\delta = \tau_H - \tau_L - 2c + a(\sigma)s_G^*$ with beliefs $\lambda_2 = \frac{2c-a(\sigma_H)s_G^*}{\tau_H-\tau_L}$. The audience sanctions with probability $s_G^* = \frac{4c-p[\tau_H-\tau_L+2c]}{a(\sigma_H)+a(\sigma)(1-p)}$ and beliefs $\alpha_L^m = \frac{1}{2}$, $\alpha_L^r + \alpha_H^r = \frac{1}{2}$, and $\alpha_H^m = 0$. The probability of war is $1-p$. This forms an equilibrium as long as the mediator does not prefer another offer.*

Proof of Lemma 10. If $p < p_L^*$ and $p > \frac{1}{2}$, or $p < p_L^*$ and the weak or strong audience requirements are not met, then the best that the mediator can do is to minimize the probability of war given the highest value of $m = \widehat{\sigma_L}$ that the leader is willing to accept, $m = \widehat{\sigma_L} \leq \overline{\sigma_L}$. Note that this is different from Lemma 9: here the mediator minimizes the probability of war subject to the maximum the leader will accept, which permits some value $s \in (0,1)$ but not necessarily the highest permissible value of s. Therefore, since the probability of war is decreasing in s, this will yield a larger probability of war than the equilibrium in Lemma 9. While this is worse for the mediator than the options in Lemmas 8 and 9, this is the best that the mediator can do, since any offer of m for which $s = 0$ in this region would make war more likely.

To solve for this value of s, the leader accepts $m = \widehat{\sigma_L}$ if:

$$
\begin{aligned}
\widehat{\sigma_L} &\leq \overline{\sigma_L} \\
\frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{1-p} + \tau_L - c &\leq \tau_L + c - sa(\sigma) \\
\frac{p(\tau_H - \tau_L) - 2c}{1-p} + \frac{sa(\sigma_H)}{1-p} + sa(\sigma) &\leq 2c \\
s\left[\frac{a(\sigma_H)}{1-p} + a(\sigma)\right] &\leq 2c - \frac{p(\tau_H - \tau_L) - 2c}{1-p} \\
&\leq \frac{2c(2-p) - p(\tau_H - \tau_L)}{1-p} \\
s\left[\frac{a(\sigma_H) + a(\sigma)(1-p)}{1-p}\right] &\leq \frac{4c - p[\tau_H - \tau_L + 2c]}{1-p} \\
s &\leq \frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)} \equiv s_G^*,
\end{aligned}
$$

which makes sense, since as we move toward Region II, $p \to p_L^*$, the audience's strategy converges to its equilibrium strategy in Region II, $s_G^* \to 0$.

Note that substitution of s_G^* into $m = \widehat{\sigma_L}$, where by construction $\widehat{\sigma_L} = \overline{\sigma_L}$, gives:

$$
m = \tau_L + c - a(\sigma)\left[\frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}\right].
$$

Additional substitution of s_G^* into the semi-separating equilibrium's best responses in Lemmas 3 to 7 establishes this lemma.

By Lemma 5, the leader mixes her strategies if the low type accepts m with probability:

$$
\begin{aligned}
\overline{q_L} &= \frac{p(\tau_H - \tau_L) - 2c + sa(\sigma_H)}{p[\tau_H - \tau_L - 2c + sa(\sigma_H)]} \\
&= \frac{p(\tau_H - \tau_L) - 2c + a(\sigma_H)\left[\frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}\right]}{p\left[\tau_H - \tau_L - 2c + a(\sigma_H)\left[\frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}\right]\right]}.
\end{aligned}
$$

The low type mixes, and the high type rejects m. The leader accepts m if the enemy accepts, with beliefs $\lambda_1 = 1$.

The leader's beliefs upon observing the enemy reject are

$$\begin{aligned}\lambda_2 &= \frac{2c - s_G^* a(\sigma_H)}{\tau_H - \tau_L} \\ &= \frac{2c - a(\sigma_H)\left[\frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}\right]}{\tau_H - \tau_L}.\end{aligned}$$

By Lemma 3, the leader must mix with the following probability r for the low type will to be indifferent:

$$r = \frac{2c - a(\sigma)\left[\frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}\right]}{\tau_H - \tau_L}.$$

The leader raises with concessions:

$$\begin{aligned}\delta &= \sigma_H - m \\ &= \tau_H - \tau_L - 2c + a(\sigma)\left[\frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}\right].\end{aligned}$$

The audience sanctions with probability s_G^* and beliefs $\alpha_H^m = 0$, $\alpha_L^m = \frac{1}{2}$, $\alpha_L^r + \alpha_H^r = \frac{1}{2}$. The probability of war is

$$\begin{aligned}P(war) &= \frac{(1-p)(\tau_H - m - c)}{\tau_H - \tau_L - 2c + s_G^* a(\sigma_H)} \\ &= \frac{(1-p)\left(\tau_H - \tau_L - 2c + a(\sigma)\left[\frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}\right]\right)}{\tau_H - \tau_L - 2c + a(\sigma_H)\left[\frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}\right]} \\ &= 1 - p.\end{aligned}$$

□

Proposition 3 (Region III: Sanctions)**.** *When $p \in (p_H^*, p_L^*)$, the perfect Bayesian equilibrium for any pair $(p, a(\sigma_H))$ is as follows:*

1. *If the audience is weak, $a(\sigma_H) < 2c$, the mediator offers $m^* = \overline{\sigma_L} = \tau_L + c - a(\sigma)$, and the leader is sanctioned, $s^* = 1$. The probability of war is $\frac{(1-p)[\tau_H - \tau_L - 2c + a(\sigma)]}{\tau_H - \tau_L - 2c + a(\sigma_H)} < 1 - p$.*

2. *If the audience is strong, $a(\sigma_H) > 2c$, a high type is likely, $p < \frac{1}{2}$, and the sanction for accepting the mediator's offer is not too high, $a(\sigma) \leq 2c - \frac{p(\tau_H - \tau_L)}{1-p}$, then the mediator offers $m^* = \widehat{\sigma_L} = \sigma_L + \frac{p(\tau_H - \tau_L)}{1-p}$. The low type accepts, $q_L^* = 1$, the high type rejects. If the enemy rejects, the leader raises with probability $r^* = \frac{p}{1-p}$ and concessions $\delta^* = \frac{(1-2p)(\tau_H - \tau_L)}{1-p}$ with beliefs $\lambda_2 = 0$. The audience sanctions with probability $s^* = \frac{2c}{a(\sigma_H)}$ with beliefs $\alpha_L^m = \alpha_H^r = \frac{1}{2}$, $\alpha_L^r = \alpha_H^m = 0$. The probability of war is $1 - 2p$.*

3. *Otherwise, the mediator offers $m^* = \overline{\sigma_L} = \tau_L + c - s^* a(\sigma)$. The audience sanctions with probability $s^* = \frac{4c - p[\tau_H - \tau_L + 2c]}{a(\sigma_H) + a(\sigma)(1-p)}$ and beliefs $\alpha_L^m = \alpha_L^r + \alpha_H^r = \frac{1}{2}$, and $\alpha_H^m = 0$. The probability of war is $1 - p$.*

In each case, the leader accepts the mediator's offer with beliefs $\lambda_1 = 1$, and both types accept a raised offer.

In equilibria 1 and 3, the low type accepts with probability $q_L^* = \frac{p(\tau_H - \tau_L) - 2c + s^* a(\sigma_H)}{p[\tau_H - \tau_L - 2c + s^* a(\sigma_H)]}$, the high type rejects. If the enemy rejects, the leader believes it is a low type with probability $\lambda_2 = \frac{2c - s^* a(\sigma_H)}{\tau_H - \tau_L}$, and raises with probability $r^* = \frac{2c - s^* a(\sigma)}{\tau_H - \tau_L}$ and concessions $\delta^* = \tau_H - \tau_L - 2c + s^* a(\sigma)$.

Proof of Proposition 3. To summarize Lemmas 8, 9, and 10, there are three options in the Region III.

- Lemma 8: The mediator offers $\overline{\sigma_L}$, which is the most the leader will accept and the leader is sanctioned, $s = 1$. This is possible only if the audience is weak, $a(\sigma_H) < 2c$. The probability of war is less than $1 - p$.

- Lemma 9: The mediator offers $\widehat{\sigma_L}$ that corresponds to the maximum probability the indifferent audience will sanction, $s_H^* = \frac{2c}{a(\sigma_H)}$. Here the price of the mediator's offer is strictly less than the maximum settlement the leader will accept against the low type, $\widehat{\sigma_L} < \overline{\sigma_L}$. This is possible only if a high type is likely, $p < \min\{p_L^*, \frac{1}{2}\}$, the audience is sufficiently strong, $a(\sigma_H) > 2c$, and sanctioning for accepting the mediator's offer is not too high, $a(\sigma) \leq 2c - \frac{p(\tau_H - \tau_L)}{1-p}$. The probability of war is $1 - 2p$.

- Lemma 10: The mediator offers $\widehat{\sigma_L} = \overline{\sigma_L}$, which is the most the leader will accept that keeps the audience indifferent between sanctioning and not. The audience sanctions with s_G^*. The probability of war is $1 - p$.

Since the first and second options result in a strictly lower probability of war, the mediator prefers to make those offers when possible. These options are not simultaneously available, since the first option relies on a weak audience, $a(\sigma_H) < 2c$, and the second option relies on a strong audience, $a(\sigma_H) > 2c$. If these options are not available, then the mediator resorts to the third option.

Since none of these equilibria overlap, there is only one equilibrium for every pair $(p, a(\sigma_H))$: the equilibrium is unique.

What is necessary to maintain this equilibrium? To maintain this equilibrium, the mediator must not deviate to another offer m'. To see that the mediator will not deviate to a lower offer $m' < m^*$, recall that the probability of war is decreasing in m, and thus, the mediator will not deviate to a lower offer.

To prevent the mediator from deviating to a higher offer $m' > m^*$, note that if the mediator deviates to a higher offer, then in the first case the leader will reject this offer thereby increasing the probability of war. In the second case, the audience will no longer be indifferent, and thus since the leader is sanctioned, she will exit to war rather than raise. This also increases the probability of war. In the third case, the leader will reject the mediator's offer, again increasing the probability of war. Since in all three cases, a higher proposal increases the probability of war, the mediator will not deviate to a higher offer.

□

Proposition 4 (Negotiation)**.** *When $p < p_N^*$, the leader offers σ_H and both types accept, where $p_N^* = \frac{2c}{\tau_H - \tau_L + 2c}$. Otherwise, the leader offers σ_L, which risks war against the high type. The probability of war is $1 - p$.*

Proof. Since the leader knows she will face audience costs, the leader never makes an offer knowing that she will raise and pay audience costs. Thus, she makes the high offer if paying that high settlement is better than her odds of a low type accepting the low offer and war against a high type:

$$\begin{aligned}
U(\sigma_H) \geq& U(\sigma_L) \\
-\tau_H + c \geq& p(-\tau_L + c) + (1-p)(-\tau_H - c) \\
-\tau_H + c \geq& p(\tau_H - \tau_L + 2c) - \tau_H - c \\
p \leq& \frac{2c}{\tau_H - \tau_L + 2c} \equiv p_N^*.
\end{aligned}$$

When $p < p_N^*$, the leader makes a high offer that secures peace, and otherwise, she makes a low offer that risks war against the less likely high type. War occurs against the high type when the low offer is made, with probability $1-p$. $\square$