Managing a Conflict
Alternative Dispute Resolution in Contests*

Benjamin Balzer† Johannes Schneider†

November 14, 2015

–JOB MARKET PAPER–

Abstract

We study the optimal design of alternative dispute resolution (ADR) mechanisms by a third-party mediator. ADR takes place before two litigants face each other in court. Litigation is a legal contest with players who are privately informed about the cost of collecting admissible evidence. Players update their beliefs *after* the mediation process, but *before* they decide on evidence collection. Different from standard mechanism design problems, the belief-system post-ADR is important for the outcome of the continuation game: within litigation, choice variables are similar to strategic complements and the evidence supplied is driven by the belief system. There is an incentive for parties to misreport in ADR to profit from this deviation in litigation should ADR fail to resolve the conflict. We show that optimal ADR has to break down on-path in some cases to screen the players with respect to their costs. Furthermore, ADR induces truthful reporting by creating post-breakdown beliefs which are independent of type-reports during ADR. To reduce inefficiency vis-à-vis symmetric litigation, optimal ADR induces asymmetric breakdown beliefs even for ex-ante symmetric types to increase the settlement rate compared to symmetric mechanisms. Independent of the set of parameters, ADR achieves settlement for the majority of cases.

Keywords: Conflict Resolution, Information Design, Mediation, All-Pay Contest, Mechanism Design, Endogenous Outside Option, Litigation
JEL codes: C72, D74, D82, D83, K41

---

*We are grateful to Volker Nocke and Thomas Tröger for support and encouragement. We would like to thank Andreas Engert, Christian Ewerhart, Yassine Lefouili, Simon Loertscher, Frank Rosar, André Stenzel, Emanuele Tarantino, Christoph Wolf, and the audiences at Mannheim, UCL, ESWC'15, EEA Annual Meeting '15, EARIE Annual Meeting '15, TSE-MaCCI-ENTER workshop Toulouse, SFB YRW Bonn and SFB YRW Munich for valuable comments and discussions. This paper has also benefited from discussions on an earlier version with Susanne Goldlücke, Bruno Jullien, Vasiliki Skreta, and Peter Vida.

†Center for Doctoral Studies in Economics, University of Mannheim. email: `josndr@outlook.de`. For the latest version of the paper please check `http://josndr.github.io`.

1 Introduction

Alternative Dispute Resolution (ADR) is a tool introduced into the legal system of many countries to increase the system's efficiency by settling as many cases as possible outside court. ADR itself can take many forms and describes a third-party mechanism other than formal litigation to solve the conflict. However, ADR typically cannot overturn the rule of law, such that parties return to the litigation track once ADR fails. Given that ADR and litigation remain thus connected, several questions arise. How does the information exchanged during ADR influence the behavior in litigation post ADR-breakdown? How does the threat of ADR-breakdown influence the litigants' willingness to release information during ADR? How should we design ADR "in the shadow of the court"?

The aim of this paper is to study the optimal third-party ADR-mechanism that uses litigation as the fall-back option in case no agreement is reached. We provide a model identifying the two-way channel that links an optimal mechanism (ADR) and an underlying contest (litigation). We show that optimal ADR and litigation cannot be considered as independent problems: the information revealed in the ADR-stage influences the choice of action in both ADR and litigation. Litigants' investment into evidence provision after breakdown depends on the beliefs about their opponent's action. The ADR-designer needs to be concerned about managing the players' beliefs in case ADR breaks down. Moreover, ADR cannot fully eliminate litigation as parties differ in their marginal cost of evidence provision. ADR breaks down sometimes to screen parties and to ensure truth-telling during ADR.

Most modern societies accept the concept of the "rule of law" despite an overburdened legal system: in 2014 each judge in the U.S. district courts received 658 new cases. At the same time the number of pending cases is even larger with 694 per judge. The large caseload leads to a median time from filing to trial of around 2 years. As litigation requires a lot of time and resources from courts, each case that forgoes litigation also has a positive externality on the functioning of the legal system as a whole.

Thus, most jurisdictions encourage parties to engage in some form of ADR before starting the formal litigation process. The U.S. Alternative Dispute Resolution Act of 1998 states that courts should provide litigants with ADR-options in all civil cases. ADR is defined as "any process or procedure, other than an adjudication by a presiding judge, in which a neutral third party participates to assist in the resolution of issues in controversy" (Alternative Dispute Resolution Act, 1998). However, ADR supplements the "rule of law" rather than replacing it. Ultimately, each party has the right to return to formal litigation.[1] Hence,

[1]For a detailed discussion on this, see Brown, Cervenak, and Fairman (1998).

ADR indeed happens “in the shadow of the court:” whenever no settlement is achieved via ADR, litigants return to the traditional litigation path.

Nonetheless, ADR is a very effective tool to settle conflicts and has success rates substantially above 50% across time, jurisdictions, and case characteristics. Furthermore, litigants report that ADR has an impact on the continuation of the trial even if unsuccessful (Genn, 1998; Anderson and Pi, 2004). The informational spillovers to post-breakdown litigation influences the design of optimal ADR: if the information a player receives during ADR depends on the information she provides, parties have an incentive to strategically extract information *within ADR* which they can use *in litigation* once ADR breaks down.

We follow a large literature dating back to Posner (1973) and consider litigation as a legal contest (for an overview on the litigation literature see Spier (2007)). The party providing the most convincing evidence wins the case. In such a contest, the optimal amount of evidence the plaintiff provides is a function not only of her own cost of evidence provision, but also of her beliefs about the defendant’s evidence choice and vice versa. Hence, litigation strategies after ADR-breakdown are a function of the players’ *belief system*.

Optimal ADR-design should take the belief-channel into account to ensure incentive compatibility: suppose a plaintiff who only has access to circumstantial evidence reports to the mediator instead that she has direct evidence. She then might gain from misreporting in two dimensions. First, through a direct effect: reporting better evidence can lead to a more favorable settlement. Second, there is an indirect effect: if the plaintiff misreports, she may also benefit if ADR fails to resolve the conflict. By misreporting in the ADR stage, the plaintiff may influence her post-breakdown expectation about the defendant’s type since breakdown is a function of both players’ reports. Changing the beliefs post-breakdown affects expected litigation outcomes and provides an additional incentive to misreport. While the direct effect is present in standard mechanism design models, we seem to be the first to consider the indirect effect as the outside-option of our mechanism depends on the belief system.

Our analysis highlights several important features of ADR in the shadow of a legal contest: we show that if ADR cannot promise full-settlement for *all* type-profiles, then ADR cannot promise full-settlement for *any* type-profile. The reason is that if the mediator promises settlement for a specific type-profile, it imposes an externality on the other types by influencing their breakdown beliefs.

We further show that the optimal mechanism is always asymmetric. It favors one player when ADR breaks down and the other when ADR is successful, even when players are fully symmetric ex-ante. At the time of participating, players only care about their expected valuation being the sum of the valuations in case of both settlement and breakdown. To keep the expected valuations constant, the valuation promised to players in settlement must increase the more competitive and therefore wasteful litigation post ADR-breakdown is. Consequently, optimal mediation makes the litigation process post-ADR less competitive by inducing asymmetric beliefs to save on resources needed for settlement.

While the optimal mechanism results in asymmetric beliefs, it ensures that beliefs are independent of the player's type-report. If a player could obtain different information from different reports, she could induce a situation without common knowledge of beliefs post-breakdown: the deviating player knows that she misreported, but her opponent does not. Each player's optimal action depends on both her own belief about the opponent and what the opponent thinks this belief is. Learning from reports can thus provide an incentive to misreport in hope of breakdown. If beliefs are independent of the report, however, such a problem does not arise because deviations do not create an information advantage.

We significantly differ from standard models of conflict resolution in that we consider a model in which investment into the conflict is made *after* the resolution mechanism broke down. Nonetheless, a key result derived by Hörner, Morelli, and Squintani (2015) carries over to our setting: If the mediator can talk to parties in private, the players' level of commitment is not important. Compared to a situation in which parties commit to the mechanism at an interim stage, the mediator can achieve (almost) the same result if parties are allowed to unilaterally opt-out of mediation after the settlement proposal. The reason is that private communication allows the mediator to conceals some information even at an ex-post stage.

Our findings contribute to the ongoing discussion of optimal ADR-design by pointing out several important aspects: (1) optimal ADR can settle most of the cases outside court independent of the cases' characteristics; (2) the level of commitment needed by the parties is not important if the mediator can communicate to parties in private; (3) regulators should be careful when preventing mediators from using asymmetric protocols as they increase the probability of ADR breaking down; and (4) to incentivize settlement, optimal ADR should predominantly manage beliefs in case a breakdown occurs.

We also contribute to the literature on mechanism design. If screening can happen only through an underlying game, on-path breakdown is informative for players and necessary for optimality. Our model emphasizes the relevance of belief management by the mechanism if the underlying game, and thus the outside option, is belief dependent. Our findings directly apply to other situations in which a wasteful contest is the last resort such as strikes, political lobbying, patent races, and standard setting organizations.

Outline. After discussing the literature in Section 2, we set up the model in Section 3 and derive the optimal mechanism in Section 4. Subsequently, we discuss the findings in Section 5 and several extensions in Section 6. Section 7 concludes.

2 Related Literature

We contribute to three strands of literature: (1) to the best of our knowledge we provide the first formal model in the law and economics literature that explicitly addresses the complementarity of litigation and ADR; (2) we add a new channel to the literature on mechanism design with endogenous outside option by showing that a mechanism which cannot fully avoid a post-mechanism game should be concerned about the information release during the process; and (3) we add to the existing literature of mechanism design and conflict resolution as we consider a setup in which parties make their decision on investment into the default game *after* the conflict arises.

We connect to the law and economics literature on settlement under asymmetric information dating back to the seminal paper by Bebchuk (1984). Spier (1994) is the first in this line to consider a mechanism design approach. She uses a model that applies to situations in which investment in evidence provision was made *prior* to negotiations and is interested in optimal fee-shifting between parties. We differ in two aspects: we hold the rules of litigation fixed and study a model in which the choice on how much evidence to present is made *after* settlement negotiations. This results in an optimal mechanism that conditions on informational spill-overs of ADR onto litigation.[2]

Brown and Ayres (1994) highlight that managing the information flow between litigants can be a rationale for ADR that goes beyond reducing psychological barriers to negotiation. There is, however, to the best of our knowledge no paper yet, that links information exchange in pre-litigation ADR with litigation as a strategic game. We model litigation in the tradition of Posner (1973) as a legal contest.[3] Our findings show that such a link is important as ADR and litigation should not be treated as two independent problems, but two stages of the same game.

The second strand of literature we relate to is that of mechanism design with endogenous outside options, i.e. mechanisms which cannot fully replace an underlying strategic game. Similar to Cramton and Palfrey (1995) and Celik and Peters (2011, 2013), we consider a mechanism that needs to be ratified by both parties. Without mutual consent, parties play the litigation game. In our model moreover, mediation sometimes breaks down and parties are referred to the underlying game. Breakdown is informative as in Cramton and Palfrey

[2]Another recent paper discussing third-party mediation is Doornik (2014) who studies the optimal use of a fixed mediation mechanism. Different from us, she is interested in *when to use* a certain ADR mechanism,while we focus on *the optimal design* of ADR.

[3]Examples include Katz (1988), Baye, Kovenock, and Vries (2005), Spier and Rosenberg (2011), and Prescott, Spier, and Yoon (2014). In addition, see Spier (2007) for a general discussion on litigation in the law and economics literature.

(1995) and Celik and Peters (2011). While Cramton and Palfrey (1995) are interested in finding worst off-path beliefs, Celik and Peters (2011) show that for some games it is optimal to design a mechanism without full participation. In our model, both channels are not present and full participation is optimal. Instead, we explore an additional channel: we ask how on-path references to the default game *by the mechanism* interact with the belief structure of the players after breakdown.

We also connect to the literature on conflict resolution as the two closest papers to ours are Bester and Wärneryd (2006) and Hörner, Morelli, and Squintani (2015). Bester and Wärneryd (2006) were the first to study conflict resolution in a mechanism design environment. Similar to us, they look for the conflict minimizing mechanism and find that it is typically stochastic. Hörner, Morelli, and Squintani (2015), building on Bester and Wärneryd (2006), study optimal mediation in the context of international relations. They show that limited commitment of the disputants does not change the outcome of the optimal mechanism as long as the mediator can talk to parties in private.

The main difference between our model and those of Hörner, Morelli, and Squintani (2015) and Bester and Wärneryd (2006) is the timing of events: through their fixed, type-dependent outside option, Hörner, Morelli, and Squintani (2015) implicitly assume that investment decisions take place *before* the conflict arises. While this assumption may apply to mediation attempts in international relations, it applies less to ADR negotiations as the collection of evidence typically happens *after* the conflict arises. Our results are thus a complement to Meirowitz et al. (2015) who study the relationship between dispute resolution and *pre-conflict investment*. Contrary to that, we study the relationship between dispute resolution and *post-mediation investment*. An important result of Hörner, Morelli, and Squintani (2015), however, carries over to our setting: limited commitment changes the result of the optimal mechanism arbitrarily little.

Although the result on limited commitment is similar, the optimal mechanism itself is qualitatively different: in Hörner, Morelli, and Squintani (2015), the result is always symmetric and involves full-settlement between weak types. In our setup neither occurs: the optimal mechanism is never symmetric and mediation has a positive breakdown probability for all type profiles as weak types are needed in the post-mediation contests to ensure full participation which is always optimal.

Our concept of mediation is based on Bester and Wärneryd (2006) and lies between pure communication devices as in Mitusch and Strausz (2005) and a mediator with independent sources of information (Fey and Ramsay, 2010). Pavlov (2013) shows that the former has no effect on the outcome in contests but, different to Fey and Ramsay (2010), the mediator can resolve the majority of conflicts without the need of an exogenous information source.

3 Model

Litigation Game. The underlying litigation game Γ of our model is an all-pay contest with asymmetric information as in Szech (2011) and Siegel (2014).[4] There are two risk-neutral players $i = 1, 2$ who compete for a good of a commonly known value of 1. Both players simultaneously decide on a score s_i and the player with the highest score wins the good. Ties are broken in favor of player 1.[5] Obtaining a score is costly. Players are ex-ante symmetric and have low marginal cost, c_l, with probability p, or high marginal cost, $c_h \equiv \kappa\ c_l$; $\kappa > 1$, with probability $(1-p)$. All but the realization of the cost, which is privately learned by each player, is common knowledge. To simplify notation, we denote the low-cost type "l" and the high-cost type "h". In line with this simplification, we are going to use the expressions "player i, type k" and "player ik" interchangeably.

Mediator. We model the mediator as a neutral third-party possessing no private information who announces a protocol $\mathcal{X}$ and has the ability to commit to it. The protocol is a mapping from a message profile, M, to triple (G, X_1, X_2) where G denotes the matrix of breakdown probabilities and X_i the matrix of settlement shares. It is without loss of generality to restrict the message space to the number of type-pairs once the mechanism has been ratified (Cramton and Palfrey, 1995; Celik and Peters, 2011). Thus, let

$$G = \begin{pmatrix} \gamma(l,l) & \gamma(l,h) \\ \gamma(h,l) & \gamma(h,h) \end{pmatrix},$$

and

$$X_i = \begin{pmatrix} x_i(l,l) & x_i(l,h) \\ x_i(h,l) & x_i(h,h) \end{pmatrix},$$

where $\gamma(M)$ denotes the probability of mediation breakdown after message profile $M = (m_1, m_2)$, that is the probability that players are sent back to the litigation game Γ after message M. Further, $x_i(M)$ denotes the share of the good assigned to player i after M.[6]

We assume budget balance and non-negative shares: the designer can only divide the good in question and allocate shares to players. These shares sum up to not more than one, that is $x_1(k_1, k_2) + x_2(k_1, k_2) \leq 1$.[7]

Formally, the mediator is a mechanism designer who has no ability to enforce actions

[4] We follow the terminology of Siegel (2009), indicating that players have heterogeneous cost of effort but a common perception of the prize.

[5] This technical assumption allows us to circumvent openness problems off-path. However, any other tie-breaking rule would work at cost of additional notation.

[6] For the ease of notation, we assume without loss of generality that the message k is assigned to the meaning "I am type k".

[7] If the good itself was indivisible, a lottery could implement the same result.

within the contest. In principle, we could also allow the mediator to send players non-binding recommendations within the contest. It is however without loss of generality to abstract from these recommendations as they would induce a communication equilibrium in the litigation game. All communication equilibria in all-pay contests are, however, payoff equivalent to the unique Nash equilibrium.[8]

We are looking for a mechanism that minimizes the ex-ante probability of mediation breakdown, $Pr(\Gamma)$. The solution concept is perfect Bayesian equilibrium.

Timing. For most of the analysis, we consider an interim individually rational mechanism.[9] Hence, the timing is as follows: first, the mediator commits to the mediation protocol $\mathcal{X}$ and players learn their type privately. Second, players simultaneously decide whether to participate in the mediation mechanism. If any player rejects, players update beliefs and play the litigation game. If both accept, players privately send a message m_i to the mediator.

Following her protocol $\mathcal{X}$, the mediator either implements an allocation (x_1, x_2) or initiates breakdown. In the latter case players update beliefs and go to litigation.

Discussion of the Assumptions. We follow a large strand of the literature in assuming that litigation is a legal contest. The all-pay contest, a limiting case of a general Tullock (1980) contest, is only assumed to ensure closed form solutions.

As expected, contest utilities are continuous for every action pair, hence adding noise would not change our results qualitatively.[10] The same is true for the constant marginal cost of evidence production. Results maintain if we assume a more sophisticated (monotonic) evidence provision function as used e.g. in Baye, Kovenock, and Vries (2005). Ex-ante symmetry is chosen for simplicity, too, and can be relaxed without changing the results.

The assumption that mediation is designed by a neutral third-party follows the U.S. Alternative Dispute Resolution Act of 1998. In practice, ADR is typically conducted by (retired) judges, law professors or private mediation companies all repeating the mediation services on a regular basis. Clearly, trust is a relevant issue for those mediators and provides a rationale for commitment.

Interim individual rationality of the players is assumed for the ease of notation, only. In Section 6 we show in line with the argument by Hörner, Morelli, and Squintani (2015) that assuming ex-post individual rationality changes results only arbitrarily little.

Finally, the assumption that the mediator aims to minimize breakdown is in line with

[8] See Pavlov (2013), especially Proposition 6 for details on this. Uniqueness of the Nash equilibrium is discussed in section 4.1.

[9] In Section 6 we show in an extension that assuming ex-post individual rationally can changes results arbitrarily little.

[10] See e.g. Baye, Kovenock, and Vries (1996), Che and Gale (2000), and Ewerhart (2015) for a detailed discussion.

the theoretical literature on conflict resolution. Courts have an enormous backlog in pending cases. Mainly because of the backlog, the time from filing to trial takes typically more than two years. Decreasing the number of court cases therefore has a positive effect on caseloads as well as on possible future conflicting parties and their ability to use the legal system effectively. Related to that, reducing the backlog is the main goal of ADR in practice: the success of dispute resolution programs is typically measured in the share of cases settled (see, e.g., Genn (1998) and Anderson and Pi (2004)). Moreover, the assumption that ADR minimizes the number of court cases adds to the tractability of the model: contest utilities are not well behaved in the mediator's choices. A different objective complicates the analysis substantially by adding non-convexities to the objective function.

4 Analysis

We proceed with the analysis in several steps. First, we characterize the equilibrium of the continuation game after on-path breakdown for a given information structure. Next, we characterize the properties of the continuation game following a misreport during the reporting stage. Breakdown after a false report essentially produces a situation without common knowledge of beliefs and provides the deviator with an informational advantage. We show that all players and types weakly prefer the on-path contest to the deviation contest only if beliefs are independent of their type reports. The third step is to rewrite the problem to overcome non-convexities and to make it tractable. Litigation is the only source of screening, and thus, the mediator is concerned about choosing the optimal information structure post-breakdown. This determines the solution of the problem up to a constant. We show that this constant is entirely determined by the fact that the optimal mechanism is budget balanced. Finally, we characterize the optimal mechanism. We show that it discriminates even between symmetric players, but involves a type-independent belief structure.

We organize the remainder of this section as follows: for each step we first state its result and provide an intuition thereafter. Formal proofs are provided in Appendix C.

4.1 Equilibrium Characterization of the Continuation Game

The continuation game after breakdown of mediation is an all-pay contest with type-dependent probabilities as defined in Section 3.

Let $p_i(k_i|m_{-i})$ denote the probability that *player i is of type k_i*, given that *player -i is of type m_{-i}*. For readability, we drop the player subscript in the arguments and write $p_i(k|m)$. In contests, the literature typically assumes some form of monotonicity condition which guarantees that having a low-cost type is desirable for all players. We follow Siegel (2014) and call the environment monotone if

$$\frac{p_i(k|l)}{p_i(k|h)} > \frac{c_l}{c_h} = \frac{1}{\kappa} \qquad \forall\ i, k. \tag{M}$$

In what follows, we are going to assume that (M) holds, i.e. we assume that it is optimal for the mediator to induce post-breakdown belief structures that satisfy (M). In the Appendix we show that this is indeed optimal even if the mediator could choose non-monotone environments.[11] Further, we assume throughout the paper that the probability that player 1 has low-cost, given player 2 reported low-cost, is weakly larger than the probability that player 2 has low-cost, given player 1 reported low-cost. Hence, player 1 is the stronger player in the contest or $p_1(l|l) \geq p_2(l|l)$. This assumption is without loss of generality.

Lemma 1. *Suppose (M) holds and $p_1(l|l) \geq p_2(l|l)$. Then, the all-pay contest has a unique equilibrium which has the following properties:*

- *the support of equilibrium strategies of each type is disjoint from but connected to the other type of the same player,*
- *the highest score played in equilibrium, $\Delta_{l,l}$, is in the strategy support of any l-type,*
- *the joint support of player 1's strategies is $(0, \Delta_{l,l}]$,*
- *the joint support of player 2's strategies is the same as that of player 1 plus an additional mass point at 0, in case $p_1(l|k) \neq p_2(l|k)$ for some k,*
- *both players play mixed strategies with piecewise constant densities on at most three subintervals of $(0, \Delta_{l,l}]$.*

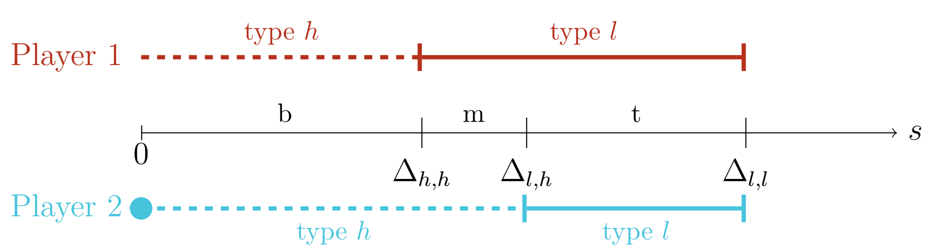

Figure 1: Strategy support of player 1 and 2 with type-dependent priors.

The Lemma is a direct application of Siegel (2014) to our setting. Figure 1 summarizes the equilibrium strategies. The horizontal axis depicts the score s. The dark-red and the light-blue line denote equilibrium strategy support for both players if player 1 is more likely

[11] Siegel (2014) shows that in principle little can be said if (M) is violated. In our setting, the mediator can only induces Bayes' plausible belief structures. Thus, it is actually possible to characterize the non-monotonic equilibria explicitly. We characterize them in the Appendix C.12.

to have low-cost. Player 1 (dark-red line at the top), type h (dashed part), is indifferent for all scores on the bottom interval b from 0 up to and including $\Delta_{h,h}$. This is the lower bound for the score of $1l$ (solid part) who is indifferent on all scores on intervals m and t up to and including $\Delta_{l,l}$ given the strategy of player 2. Player $2h$ (light-blue dashed line at the bottom) is indifferent between a score of 0 (indicated by the dot) and on intervals b and m up to and including $\Delta_{l,h}$. Player $2l$ is indifferent on interval t. If players become ex-ante symmetric, interval m vanishes and the mass point at 0 disappears and utilities and strategies become fully symmetric.

There are no pure-strategy equilibria: whenever one player scores on a singleton only, it is either optimal to marginal overscore this value or to score 0 instead. There are several relevant properties of this mixed-strategy equilibrium. First, the highest score obtained by both players is the same. If one player was to strictly overscore her opponent, she could always deviate by reducing her score to the highest possible score of her opponent. Such a deviation does not reduce the probability of winning, but reduces the cost of the score.

Second, choices in all-pay contests are similar to strategic complements: whenever the likelihood of player $1l$ increases, player $2l$ reacts by scoring more aggressively. As l-types share the upper bound in their strategies, $2l$ has a higher average score than player $1l$.

Third, for every information structure at least one h-type player receives 0-utility in expectations. This player is always the ex-ante weakest player-type combination, here player $2h$. If this is not the case, no player would score exactly 0 with positive probability. But then, whatever the lower bound of the joint support, scoring at this lower bound yields a negative utility, which can always be avoided by deviating to a score of 0.

If player $2h$ has a mass point at 0, player $1h$ receives strictly positive utility as every score arbitrarily close to 0 guarantees her to win if player $2h$ decides to score 0.

Overall, the equilibrium actions in the all-pay contest depend on the belief about both the opponent's type, and the opponent's action, where the latter is a function of the opponent's beliefs. Thus, expected utilities depend on the entire belief structure. The following corollary to Lemma 1 defines the expected contest utilities in closed form.

Corollary 1. *Under the assumptions of Lemma 1, and $p_i(l|k) > 0$, the expected contest utilities are*

$$\begin{aligned} U_1(l) = U_2(l) &= 1 - c_l\ \Delta_{l,l} > 0, \\ U_1(h) &= p_2(h|h) F_{2,h}(0), \\ U_2(h) &= 0. \end{aligned} \tag{U}$$

Moreover, utilities are linear in beliefs, if beliefs are type-independent. If beliefs are symmet-

ric, $F_{2,h}(0) = 0$.

The utility of the low-cost types is a direct consequence of the common highest score. Both players win with probability 1 if they score at $\Delta_{l,l}$ and have cost $c_l\Delta_{l,l}$. On all other scores in their support they must be indifferent. The utility of the high-cost type of player 1 is derived as she always wins against those high-cost types that score 0 even if she scores arbitrarily close to 0. High-cost types of player 2 score 0 with probability $F_{2,h}(0)$ which gives them utility 0. If beliefs become type-independent, that is $p_i(l|l) = p_i(l|h)$, the upper bound, $\Delta_{l,l}$, and the mass on 0, $F_{2,h}(0)$, is linear in beliefs. If beliefs are symmetric between players, that is $p_1(l|k) = p_2(l|k)$, the mass point on 0, $F_{2,h}(0) = 0$ and $U_1(h) = 0$.

4.2 Deviator Payoffs in the Continuation Game

As players in our model differ only with respect to their cost in the contest, it is important for incentive compatibility to characterize the post-deviation continuation game. It needs to be assessed how players' actions and utilities change in case of breakdown conditional on a false report during the reporting stage. A false report introduces non-common knowledge of beliefs between the players. The deviating player knows about her deviation and assigns correct beliefs to her opponent. The non-deviating player and the mediator, on the other hand, are unaware of the deviation and incorrectly predict the deviator's beliefs. The wrong prediction affects actions, expected contest utilities, and thus incentive compatibility.[12]

Lemma 2. *Assume (M) and* $p_1(l|l) \geq p_2(l|l) > 0$*. All player-type combinations but player 1h are weakly better off in their respective deviation contest. Player 1h is strictly worse off in the deviation contest if and only if the probability of facing a high-cost type in her deviation contest is strictly smaller than in her on-path contest.*

Lemma 3. *Assume (M) and* $p_1(l|l) \geq p_2(l|l) > 0$*. Then, exactly one type of each player is strictly better off in the deviation contest than in the on-path contest if and only if the beliefs the player holds are* not *type-independent. If beliefs are type-independent, no player is better off in the deviation contest.*

Lemmas 2 and 3 state that the only situation in which no player-type prefers the deviation contest to the on-path contest is when beliefs, $p_i(l|m)$, are independent of the reported type m. To understand the intuition, let us first define the two types of contest.

Definition 1. *On-path contest:* the contest is called on-path contest if the belief structure is such that any player i, type k, holds belief $p_{-i}(l|k)$ about player $-i$. Further, the belief that each player and type holds is common knowledge.

[12]The deviator of course correctly predicts the wrong prediction of the non-deviator, and so on.

Definition 2. *Deviation contest:* the contest is called deviation contest of player ik if player i, type k holds a belief $p_{-i}(l|\neg k)$ that is the same belief that player i, who is *not* k, holds on-path. This belief is called the deviator's belief. Player $-i$, however, holds her on-path belief $p_i(l|k)$ about player i. Thus, generically, there is no common knowledge of beliefs in this contest.

A direct consequence of non-common knowledge of beliefs is that the deviating player is no longer indifferent between several scores. The non-deviating player chooses her strategy to make an *on-path opponent* indifferent on some interval. The deviator, however, has a different belief about the non-deviator than the *on-path opponent* and is thus *not indifferent* as second-order beliefs differ. Decisions are similar to strategic complements, such that a too aggressive choice of the non-deviator leads the deviator to pick an aggressive response. If the choice is too soft, the deviator picks a soft response. The best response is generically a singleton.

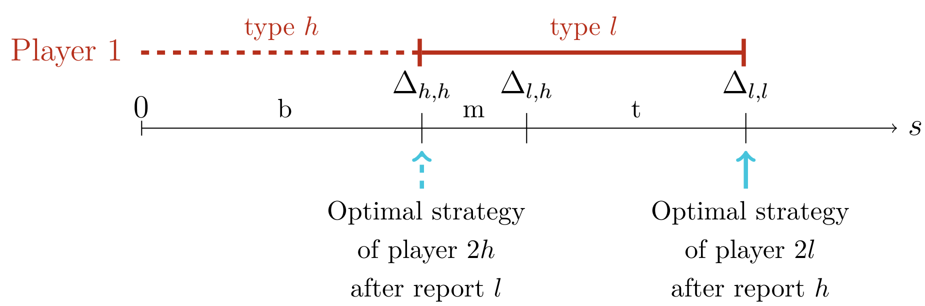

Figure 2: Optimal behavior in the deviation contest of player 2 if $p_1(l|h) > p_1(l|l)$. Notice that the deviation strategies are conditional on $2l$ reporting h and $2h$ reporting l without player 1 noticing.

Figure 2 illustrates the optimal strategies for player 2's deviation contest in case it is more likely that l-types appear after an h-report, i.e. $p_1(l|h) > p_1(l|l)$. The horizontal axis describes the scores, the dark-red line the strategy of player 1, which is the same as in equilibrium. The light-blue, dashed arrow points to the unique best response of player $2h$ who reported l, the solid arrow to that of player $2l$ who reported h.

If the probability that the opponent has low cost is larger in the deviation contest, the deviating l-type decides to score more aggressively. By the common upper bound in the strategy support, scoring above the highest score, $\Delta_{l,l}$, is never beneficial. Thus, her optimal strategy in the deviation contest is to score at $\Delta_{l,l}$ and to win with probability 1, if she is more likely to meet an l-type. Therefore, her utility is the same as on path, where she wins with probability 1 at a score $\Delta_{l,l}$ which is part of her equilibrium strategy.

Whenever reporting h increases the likelihood to meet an l-type opponent for player 2, reporting l must increase the likelihood to meet an h-type, i.e. $p_1(h|l) > p_1(h|h)$. Similar to the case of $2l$, a deviation by $2h$ makes her increase the score against an h-type (interval b in Figure 2), but decrease it against an l-type (interval m in Figure 2), since those occur less likely. Thus, her optimal response is $\Delta_{h,h}$ which leads to a win against all h-types. High-cost types occur with higher probability as $p_1(h|l) > p_1(h|h)$, and hence, $2h$ prefers the deviation contest to the on-path contest.

Low-cost players are never worse off in the deviation contest, as they can always score at the top. Moreover, player $2h$ is not worse off either as she can secure her on-path utility of 0. The only player that can be worse off in the deviation contest is player $1h$, if she expects to meet less $2h$. She then softens her bid to 0 and wins by the tiebreaker but suffers from the low probability of meeting $2h$.

Having discussed both on-path and post-deviation behavior in the continuation game, we shorten notation and use $U_i(k|m)$ to describe the expected utility that player i, type k enjoys in the contest stage if she reported to be type m and behaves optimally thereafter.

4.3 Rewriting the Problem

We now turn to the problem of the designer. Note that the problem is highly non-convex and standard techniques do not apply. To be able to characterize the solution we need to transform it to a tractable problem. We do so in several steps. As the transformation is a series of technical issues we proceed as follows. First, we state the proposition describing the reformulated problem. Second, we state the original problem. Third, we provide a brief, non-technical comment on each transformation step in the main text. We refer the interested reader to Appendix A for the corresponding detailed description of the transformation including the intermediate lemmas.

Proposition 1. *Any ex-post implementable, individually feasible and incentive compatible solution to*

$$\min_P Pr(\Gamma) = \min_P R(P)\gamma^*(P) \tag{P1'}$$

is also a solution to the mediator's problem if and only if $\gamma^*(P) \leq 1$*, where* $R(P) = Pr(\Gamma)/\gamma(l,l)$.

The proposition states that an equivalent formulation of the mediator's problem exists. In it, she optimizes over the set of breakdown beliefs, $P = \{p_1(l|l), p_2(l|l), p_1(l|h)\}$, instead of the set of shares and breakdown probabilities, $\mathcal{X} = (G, X_1, X_2)$. The remaining breakdown belief about player 2, $p_2(l|h)$, is implicitly defined by P and Bayes' rule. The rewritten problem comes at the cost of two additional, technical constraints, namely ex-post implementability and individual feasibility. We are going to discuss these constraints below.

The Original Problem of the Mediator. As the mechanism needs to pass a ratification stage it is not necessarily without loss of generality to assume full participation. Given the payoff structure of the litigation game, however, we can use a result of Celik and Peters (2011) to conclude that full participation is indeed optimal in our setting, the corresponding lemma stating this result is included in Appendix A. Given full participation, the mediator's problem is

$$\min_{\mathcal{X}} Pr(\Gamma) = \min_{\mathcal{X}} \Big(p, (1-p)\Big) \cdot G \cdot \begin{pmatrix} p \\ (1-p), \end{pmatrix} \tag{$P1$}$$

subject to the following sets of constraints for all $i \in \{1, 2\}$ and $k, m \in \{l, h\}$

$$\Pi_i(k|k) \geq V_i(k), \tag{PC_i^k}$$

$$\Pi_i(k|k) \geq \Pi_i(k|m), \tag{IC_i^k}$$

$$x_1(k_1, k_2) + x_2(k_1, k_2) \leq 1, \quad x_i(k_1, k_2) \geq 0,$$

$$0 \leq \gamma(k_1, k_2) \leq 1,$$

where $\Pi_i(k|m)$ describes the expected total payoff of a participating player i, type k given she reports m. $V_i(k)$ describes the value of vetoing the mechanism for player i, type k. The first set of constraints are participation constraints, (PC_i^k), indicating that each player and type should prefer to participate in ADR over vetoing. The second set, the incentive compatibility constraints (IC_i^k), state that it is optimal for each agent to announce her true type. The third set of constraints prohibits additional payments by the agents or the mechanism and ensures a balanced budget. Finally, the last set of constraints ensures that breakdown probabilities are between 0 and 1.

Value of vetoing. To determine the outside option we need to define the equilibrium of the litigation game after a veto by either of the parties in the ratification stage. High-cost types do not receive any payoff after a veto and are thus always at least indifferent to participate in ADR. Low-cost types' value of vetoing depends on the choice of beliefs after vetoing. In our case any choice of these off-path beliefs after vetoing which satisfy the intuitive criterion leads to the same value of vetoing: the expected litigation payoff under the prior p.[13]

Whenever the value of vetoing is smaller than 1/2 for low-cost types, however, the me-

[13]This is a direct consequence of the low-cost types' contest utilities being a function of the weaker players' probability to have low-cost in case of type-independent beliefs. Any deviation belief satisfying the intuitive criterion, makes the non-deviating player the weaker one. Thus, the relevant belief remains constant at p.

diator could offer parties a sharing rule of $(1/2, 1/2)$ for each type-realization and settle all cases. To make the problem interesting we make the following assumption.

Assumption 1. The low-cost types' value of vetoing is strictly above 1/2.

Assumption 1 translates into the following condition on parameters: $\kappa > (2-2p)/(1-2p)$.

Expected payoff. The expected payoff from participation, $\Pi_1(k|m)$, has two components: the expected value of successful settlement and the expected value of mediation breakdown and subsequent litigation. Thus,

$$\Pi_i(k|m) = z_i(m) + \gamma_i(m)U_i(k|m), \tag{1}$$

where message m leads to a value of settlement, $z_i(m)$, and a value of breakdown $\gamma_i(m)U_1(k|m)$. The expected contest probability, $\gamma_i(m)$, is a convex combination of the breakdown probabilities conditional on the opponents type

$$\gamma_1(m) = p\gamma(m,l) + (1-p)\gamma(m,h),$$

the value of settlement is a convex combination of realized shares and settlement probabilities

$$z_1(m) = p(1-\gamma(m,l))x_1(m,l) + (1-p)(1-\gamma(m,h))x_1(m,h),$$

and analogously for player 2. Equation (1) shows how optimal mediation relies on the litigation game. While the value of settlement, z_i, is similar to transfers in standard mechanism design, the utility of the contest continuation game is the screening device.

Step 1: Reduced-Form Problem à la Border (2007). In this step we make use of a procedure introduced by Border (2007) to reduce the problem from realized values to expected values. The reduced form problem has the advantage that the exact composition of the settlement shares, X_i, becomes irrelevant and we can use the settlement values, $z_i(\cdot)$, directly as choice variables. To ensure a feasible X_i, reducing the problem introduces two additional constraints: an individual feasibility constraint, (IF), and an ex-post implementability constraint, (EPI). The first constraint states that each player cannot get more than the whole good in case of settlement. The second constraint guarantees that the total amount of value distributed to a given type-profile does not exceed the total probability of any of the types within that profile occurring.

Step 2: Backing out Expected Settlement Shares. In the second step, we make use of the fact that we can assume without loss of generality that both the high-cost types' incentive compatibility constraints and the low-cost types' participation constraints are binding. The latter follows naturally from the values of vetoing, that is the fact that low-cost types need to be compensated to take part in ADR. Binding incentive compatibility for high-cost types follows from their low expected payoff in litigation: it provides an incentive to mimic low-cost types to get their settlement value. The binding constraints allow us to eliminate all settlement values, as they can be expressed in terms of breakdown valuations.

Step 3: From Breakdown Probabilities to Breakdown Beliefs. This step uses that breakdown beliefs are homogeneous of degree 0 with respect to the set of breakdown probabilities, G by Bayes' rule. Thus, the set of breakdown *beliefs* defines the set of breakdown *probabilities* up to a constant. We choose this constant to be $\gamma(l,l)$ such that all other breakdown probabilities are defined relative to $\gamma(l,l)$. This allows us to eliminate all breakdown *probabilities* but $\gamma(l,l)$, and replace them by breakdown *beliefs*.

Step 4: Eliminate $\gamma(l,l)$ via expected feasibility. The final step is to eliminate $\gamma(l,l)$. We use the fact an ex-ante feasible settlement rule is a necessary condition for individual feasibility, (IF). All expected breakdown probabilities increase linearly in $\gamma(l,l)$ by Step 3. Therefore, the mediator wants to set $\gamma(l,l)$ as low as possible, as long as the problem remains feasible in expectation. This introduces an equality constraints $\gamma(l,l) = \gamma^*(P)$ by which we replace $\gamma(l,l)$. The additional constraint $\gamma^* \leq 1$ ensures that $\gamma(l,l)$ remains a probability. This concludes the rewriting of the problem.

4.4 Optimal ADR-Mechanism

Having established the reduced problem (P1'), which is a problem of three choice variables only, we can now state the main result:

Theorem 1. *Suppose Assumption 1 holds. Then, any optimal mediation protocol has the following properties:*

- *on-path breakdown beliefs are type-independent, that is for any i it holds that $p_i(l|l) = p_i(l|h) =: \rho_i$,*
- *on-path breakdown beliefs are asymmetric, that is $\rho_i \neq \rho_{-i}$,*
- *both player's on-path breakdown belief is weakly larger than the prior, that is $\rho_i \geq p \ \forall i$,*
- *all type profiles $\{k_1, k_2\}$ have a breakdown probability that is strictly positive.*

Theorem 1 states that, independent of the primitives, any optimal protocol induces an information structure that is report-independent. In addition, although parties start perfectly symmetric, the mediation protocol should always be set up asymmetrically. At the same time the ADR protocol ensures that both parties appear to be at least as strong after mediation breakdown as they appeared before mediation. Therefore, the fraction of low-cost types is at least as high in a post-mediation contest as before the start of the game. Finally, the mediator needs to ensure that in principle any type profile can lead to a breakdown of mediation to get the above mentioned features.

To build intuition we organize the remainder of the section as follows. We first discuss the optimal solution to (P1') ignoring (IC_i^l) and $\gamma^*(P)$. We then reintroduce (IC_i^l) and later $\gamma^*(P) \leq 1$. Finally, we verify that the solution is implementable in the sense of Border (2007).

Recall that the assumption of player 1 appearing weakly stronger in the contest implies the following expected litigation utilities of the high-types: $U_2(h|h) = 0$ and $U_1(h|h) \geq 0$ with strict inequality whenever player 1 appears strictly stronger. Further, litigation utilities, $U_i(k|m)$, depend on breakdown *beliefs* and all expected breakdown *probabilities*, $\gamma_i(m)$, are linear in $\gamma(l,l)$. In addition, the following technical lemma is useful to keep in mind. It states that whenever it is more likely for player 2 to meet $1l$ after a report of l, the same is true for player 1 and vice versa.

Lemma 4. *$p_1(l|l) > p_1(l|h) \Leftrightarrow p_2(l|l) > p_2(l|h)$ if $p_i(l|m) \in (0,1)$.*

Part 1: Neglecting (IC_i^l) and $\gamma^*(P) \leq 1$. First, we want to argue that beliefs are type-independent. The basic idea is straight-forward: if the mechanism does not allow parties to influence the opponent's type distribution in case of breakdown, then there is no incentive for a false report. Similar to a second price auction, where expected payments are independent of the type report, the mediator ensures that the type distribution the player faces, and by that her contest utility, is independent of her type report.

Proposition 1 states a problem with the three breakdown beliefs, $p_1(l|l), p_2(l|l), p_1(l|h)$ as choice variables. Given Lemma 4 we can fix $p_2(l|l)$ and $p_1(l|h)$ for the upcoming argument and concentrate on $p_1(l|l)$ without loss of generality.

As the mediator cannot achieve full settlement by the participation constraint of the low-cost types and the high-cost types' desire to mimic them, she needs to strategically fail mediation to screen types. High-cost types need to be present in the contest to guarantee some utility for the low-cost player and to match her participation constraint. However, the high-cost players should have an incentive to avoid the contest to report truthfully. Thus, the probability of a high-cost player meeting another high-cost player after mediation breakdown, $p_i(h|h)$, should be smaller than the ex-ante probability of a high-cost type, $1-p$. Without a belief dependent outside option this effect typically drives $p_i(l|h)$ to 1 as in e.g. Hörner, Morelli, and Squintani (2015).

There is, however, a second, non-standard effect, changing utilities after breakdown. If breakdown is informative, i.e. $p_i(l|l) \neq p_i(l|h)$, the expected utility in the deviation contest might differ from the expected utility in the on-path contest.

Recall from Lemmas 2 and 3 that $U_i(h|l) > U_i(h|h)$ whenever it is more likely to meet a low-cost type under truth-telling than under deviation, that is whenever

$$p_i(h|h) < p_i(h|l) \Leftrightarrow p_1(l|l) < p_1(l|h),$$

due to the information advantage effect in the contest. This advantage vanishes as $p_1(l|l) \to p_1(l|h)$. If $p_1(l|l)$ increases further, player 2 receives no utility in the contest and therefore also no marginal breakdown utility from lying. Player 1, on the other hand, actually starts gaining utility again, as an intimidation effect becomes dominant. Player 1 appears to be much stronger in expectation than player 2. Thus, player 2 invests less into the contest which increases player 1's utility. Therefore, both deviation utilities have a minimum at type-independent beliefs.

Deviation utilities have a kink at type-independent beliefs by the all-pay contest assumption. The kink is a direct consequence of Lemma 3 as deviating high-cost players are only indifferent for type-independent beliefs. High-cost types score at the upper end of their on-path equilibrium strategy set for lower values of $p_1(l|l)$ and at the lower end for higher values of $p_1(l|l)$. Hence, for type-independent beliefs their utilities are non-differentiable and obtain a minimum. The left panel of Figure 3 plots the deviation utilities as a function of $p_1(l|l)$.

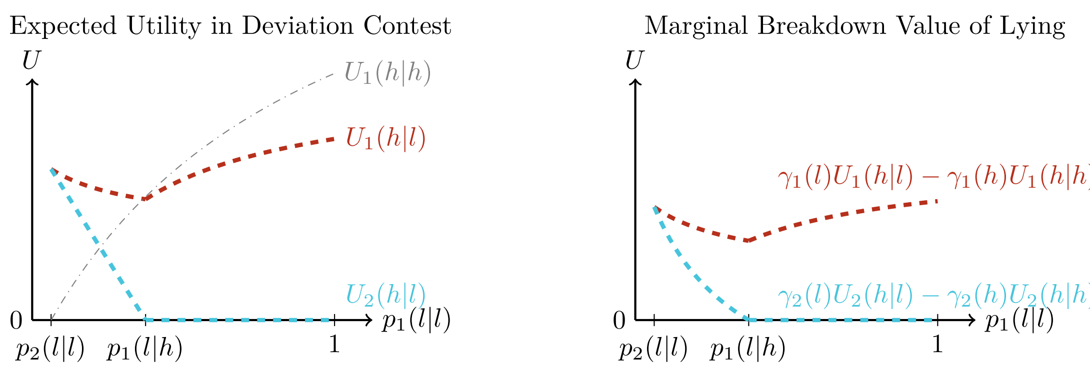

Figure 3: The left panel depicts the high-types deviation utilities as a function of $p_1(l|l)$. The right panel depicts the marginal breakdown-value of lying. Red is for player 1, blue player 3. The gray line in the right panel is the on-path utility of the high-cost type of player 1.

If we combine the effects on breakdown probabilities $\gamma_i(m)$ and contest utilities, we find that the minimum at type-independent beliefs prevails. The result can best be seen if we consider the marginal breakdown-value of lying. This breakdown value is the right-hand side of the following representation of the high-types incentive constraint,(IC_i^h),

$$z_i(h) - z_i(l) = \gamma_i(l)U_i(h|l) - \gamma_i(h)U_i(h|h).$$

The left hand side can be interpreted as the marginal settlement value of truth-telling which matches the right hand side being the marginal breakdown value of lying. The right panel of Figure 3 displays the marginal breakdown value of lying and illustrates how the minimum property prevails and type-independent beliefs are optimal. We can thus simplify notation and define ρ_i to be the probability that player i is the low-type post-mediation.

Having established that beliefs are type-independent we can simplify the analysis using a corollary to the derivation of the breakdown beliefs.[14]

Corollary 2. *If beliefs are type independent, breakdown probabilities can be simplified to*

$$\gamma_i(l) = \frac{p}{\rho_{-i}}\gamma(l,l), \qquad \gamma_i(h) = \frac{(1-\rho_i)}{(1-p)}\frac{p}{\rho_i}\gamma_i(l), \qquad Pr(\Gamma) = \frac{p^2}{\rho_1\rho_2}\gamma(l,l).$$

Moreover, Corollary 1 allows us to write contest utilities with type-independent beliefs as

$$U_i(l|m) = (1-\rho_2)\frac{\kappa-1}{\kappa}, \qquad U_1(h|m) = (\rho_1-\rho_2)\frac{\kappa-1}{\kappa}. \tag{2}$$

These expressions are useful in the argument for asymmetry of the optimal mechanism which we turn to next. We discuss the general argument non-formally to provide a good understanding of the qualitative results. A more detailed and formal analysis is in Appendix B.

The main argument for asymmetry lies in the structure of a contest. A symmetric contest is expected to be tight: parties expect to be matched with an opponent of similar strength and the marginal value of investment is high. By contrast, an asymmetric contest appears to be less tight, and the marginal value of investment is lower for both parties. This imposes an externality, especially for the high-cost type of the ex-ante stronger player. Her opponent's high-cost type is going to increase her investment but remains at a utility of 0 as she is the weakest of all player-types. Thus, the stronger player's h-type can reduce the investment and still has a reasonable chance to win the contest as the opponent believes she likely faces a low-cost type. This effect can be seen by inspecting equations (2). If we start in a symmetric setting and unilaterally increase the belief put on player 1, then l-types would not benefit in terms of expected utilities and neither would $2h$. However, player $1h$ actually achieves a positive utility in such a case which she would not under symmetry.

Although only concerned about the probability of contest, the optimal ADR-mechanism uses this property of the underlying game to increase the breakdown utility of one of the high-cost types. This allows the mediator to reduce the settlement value that needs to be paid to this player which in turn increases the available resource for settlement. There is,

[14]To be precise, Corollary 2 is a corollary to Lemma 11 which is stated in Appendix A.

however, a second effect that limits the extent to which the mediator can use this feature: as breakdown probabilities are derived in their *relative relation* to $\gamma(l,l)$ in problem $(P1')$, an increase in ρ_1 is effectively a decrease of the breakdown probability of high-cost types of player 1, $\gamma(h,l)$ and $\gamma(h,h)$. This implies, in turn, a decrease in the breakdown probability for *player 2l*, $\gamma_2(l)$, according to Corollary 2. While such a decrease has a positive effect on the objective, $Pr(\Gamma)$, it also leads to a decrease in player 2's breakdown utility. Thus, the mediator would need to increase player 2's settlement utility. Making the contest less resource intensive is therefore only optimal up to a certain point. This point balances the additional resources needed to finance the loss for player $2l$ and the gain from making the contest less resource-intensive. A similar argument is true for the other player-types.

To see the aggregate effect consider the expected settlement share paid to player i, z_i. The expected settlement share is a convex combination of the settlement share paid to the l-type to ensure participation and the settlement share paid to the h-type to ensure incentive compatibility. The shares are given by

$$z_2 = V(l) - \frac{1-\rho_2}{\rho_1}\frac{\kappa-1}{\kappa}p\gamma(l,l)$$

$$z_1 = \underbrace{V(l) - \frac{1-\rho_1}{\rho_2}\frac{\kappa-1}{\kappa}p\gamma(l,l)}_{\text{symmetric part}} + \underbrace{(\frac{p}{\rho_1} - \frac{p}{\rho_2})\frac{\kappa-1}{\kappa}p\gamma(l,l)\kappa.}_{\text{asymmetric part}}$$

The first part in z_1 is present in the symmetric case, too, while the second vanishes. Without the second part z_1 would be the anti-symmetric version of z_2 which would lead to endogenous symmetry. However, the second part provides a clear incentive for asymmetry driven by $U_1(h|h)$.[15] An increase in ρ_2 requires more resources to compensate the players than an increase in ρ_1. Thus, the optimal choice involves $\rho_1 > \rho_2$, that is player 1 appears relatively stronger in the contest. Finally, notice that the asymmetric part is always negative and thus, some asymmetry always saves resources. The next lemma states the findings up to this point.

Lemma 5. *Ignoring (IC_i^l), (IF), (EPI) and $\gamma(l,l) \leq 1$, and assuming that $\rho_1 \geq \rho_2$, the unconstrained optimum of (P1') is achieved at*

$$\rho_1^* = \frac{1+p}{2} \qquad \rho_2^* = \frac{1-p}{2}.$$

Moreover, the optimal breakdown belief ρ_i^* is independent of the opponents breakdown belief ρ_{-i}.

[15]Notice that this part can also be written as $-Pr(\Gamma)U(h|h)$.

Part 2: Reintroducing (IC_i^l)**.** Next, we reintroduce the low-cost type's incentive compatibility constraint, (IC_i^l). For type-independent beliefs and with (IC_i^h) satisfied this boils down to

$$\Big(\gamma_i(l) - \gamma_i(h)\Big)U_i(h|h) \leq \Big(\gamma_i(l) - \gamma_i(h)\Big)U(l|l). \tag{3}$$

A sufficient condition for this to hold is $\gamma_i(l) \geq \gamma_i(h)$, as $U(l|l) \geq U_i(h|h)$ by construction. For player 2 it is also necessary since $U_2(h|h) = 0$. Using Corollary 2, $\gamma_i(l) \geq \gamma_i(h)$ is equivalent to $\rho_2 \geq p$. Intuitively the reasoning is straightforward: suppose $\rho_2 \leq p$. The likelihood of breakdown must be larger when reporting to be an h type. By (IC_2^h), the value of settlement, $z_2(l) = z_2(h)$, is independent of the report and the low-cost type prefers to be sent to contest more often and would misreport. Thus, incentive compatibility requires $\rho_2 \geq p$.

Taking into account the results from Lemma 5, this means that (IC_i^l) is violated whenever $(1-p)/2 < p$ which holds if and only if $p > 1/3$. Note further that $\rho_1^* > p$ for all p and thus, (IC_1^l) never binds. As the optimal ρ_i does not depend on ρ_{-i}, we get the following lemma.

Lemma 6. *Ignoring* (EPI) *and* $\gamma^*(P) \leq 1$*, and assuming that* $\rho_1 \geq \rho_2$*,* (IC_i^l) *binds for player 2 if and only if* $p \geq 1/3$*. In this case the constrained optimum is achieved at*

- $\rho_1^* = \frac{1+p}{2}$
- $\rho_2^* = p$.

Lemma 6 states that the probability of breakdown for low-types is larger than the probability of breakdown for high-types, i.e. $\gamma_i(l) \geq \gamma_i(h)$. In such a case one individual feasibility, (EPI), which is one of the two constraints coming from the reduced form, is always satisfied. Appendix C.4.2 provides details on this.

Part 3: Full model. So far we have ignored that the scaling parameter γ^* is in fact always equal to the probability of breakdown for two low-cost types, $\gamma(l,l)$, in the original problem. Thus, we need to ensure that $\gamma^* \in [0,1]$ to guarantee that $\gamma(l,l)$ remains a probability.

Whenever the constraint $\gamma^*(P)$ binds, (IC_i^l) must hold, too. To see this recall

$$\gamma_i(l) = \frac{p}{\rho_{-i}}\gamma(l,l).$$

To ensure $\gamma_i(l) \in [0,1]$ even if $\gamma(l,l) = 1$ we need $p \leq \rho_{-i}$. Such a high post-breakdown belief ensures incentive compatibility by Lemma 6. If the ex-ante probability of low-cost types is high enough for (IC_i^l) to bind, the scaling parameter $\gamma^*(P) < 1$. Thus, $\gamma^* \leq 1$ does not change the results of Lemma 6. Next, recall that

$$\gamma^*(P) = \frac{\nu}{Q(P) - R(P)},$$

such that γ^* is increasing in ν for any P. The value of ν, in turn, is large for small p and large κ. Therefore, the solution computed in Lemma 5 violates $\gamma^* \leq 1$ if cost difference between low-cost and high-cost types are high, or the probability to have high-cost is small.

To compensate this, the mediator can decrease either ρ_i. As in the discussion of Lemma 5 such an operation increases the resources available for distribution in settlements and allows to reduce γ^*.

Given small values of the prior, p, the optimal breakdown belief ρ_i *without* considering the γ^*-constraint is strictly larger than p, and thus the mediator reduces both beliefs, ρ_1 and ρ_2, simultaneously up to the point at which one equals the prior, i.e. $\rho_2 = p$. If this does not suffice to make $\gamma(l,l)$ feasible, the mediator decreases the belief on player 1, ρ_1, further until $\gamma^*(P) = 1$. It turns out that the remaining Border-constraint, (EPI), holds at any such point and ex-post implementation is thus possible. Combining all results allows us to make a statement about any set of parameters, κ and p. The characterization is given in the next lemma which concludes the argument for Theorem 1.

Lemma 7. *Consider without loss of generality only $\rho_1 \geq \rho_2$. Fix some κ such that assumption 1 holds. Then there are three cutoff values p', p'' and p''' such that the optimum of the minimization problem is either 0 or satisfies*

- *(IC_2^l) and therefore $\rho_2 = p$ with equality only if $p \notin (p', p''')$,*
- *$\gamma(l,l) \leq 1$ with equality only if $p \leq p''$,*
- *$2p < \rho_1 \leq (1+p)/2$ where the last holds with equality only if $p \geq p''$.*

The cutoffs are given by:

$$
\begin{aligned}
p' &= \frac{1}{6(\kappa-1)}\left(\kappa - 8 + \sqrt{28 - 4\kappa + \kappa^2}\right),\\
p'' &= \frac{1}{2+3\kappa}\left(2(\kappa-1) - \sqrt{8 - 4\kappa + \kappa^2}\right),\\
p''' &= \frac{1}{3}.
\end{aligned}
$$

The cutoffs describe the main characteristic of the optimum. For low p the mediator offers low-cost types a litigation utility post breakdown which is smaller than their value of vetoing, i.e. $\rho_2 > p$. To do this l-types need a high enough settlement share which the mediator finances by reducing the overall breakdown probability by increasing γ^*. However, for very low p not even $\gamma^* = 1$ suffices as $V(l)$ is too high. To account for the constraint, the mediator decreases both breakdown probabilities, ρ_2 and ρ_1. However, ρ_2 cannot fall below p as this would violate both (IC_2^l) and $\gamma_i(k) \leq 1$. Thus, for very low p, the mediator chooses $\rho_2 = p$ and adjusts ρ_1 accordingly.

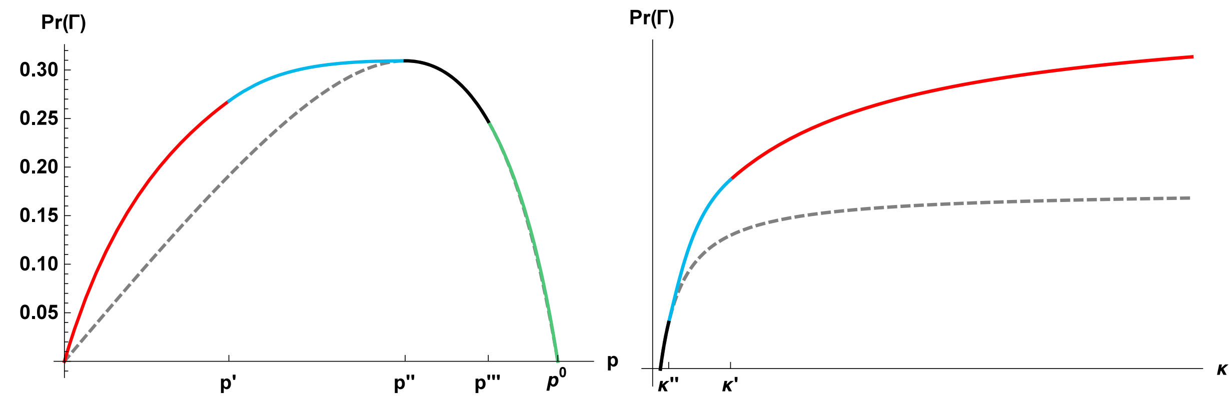

Figure 4: Ex-ante probability of the contest as a function of p (left panel) and κ (right panel). The dashed line describes the situation of the unconstraint problem ($P1'$) as in Lemma 5. The green solid line corresponds to Lemma 6. All solid lines together form display the result of Lemma 7.

As the prior p increases, the solution ρ_2 increases, too, and $\rho_2 \geq p$ does not bind anymore. The resource constraint, $\gamma^* \leq 1$, however, still does. If p is larger than p'', the solution of Lemma 5 can be implemented directly. For $p > 1/3$, on the other hand, low-cost types of player 2 have an incentive to misreport given the protocol from Lemma 5 which means that (IC_2^l) binds and the belief on player 2 is set to the prior, $\rho_2 = p$. The left panel of Figure 4 illustrates the findings. The dashed line plots the optimal protocol according to Lemma 5 whereas the solid line is the full model.

5 Discussion of the Results

Comparative Statics. Figure 4 depicts the probability of litigation under the optimal mechanism both as a function of the prior, p (left panel), and as a function of the distance between low and high cost, κ (right panel). The different colors indicate the different regimes as discussed in Lemma 7. Red and blue (for $p < p''$) denote the areas in which the resource constraint, $\gamma^* \leq 1$, binds; green (to the right of p''') is the area in which $2l$'s incentive constraint binds and black is the area in which ($P1'$) is solved "unconditionally" as in Lemma 5. p^0 indicates the point at which Assumption 1 starts to fail and the mediator achieves full settlement for $p > p^0$. For comparison, the dotted line depicts the solution ignoring (IC_i^l) and $\gamma^* \leq 1$.

As expected, the probability of litigation increases in the distance between high-costs and low-costs. As the low-cost type's cost advantage increases, it becomes more expensive to compensate her for participation and thus the mediator initiates breakdown more often. The relationship with respect to the prior is non-monotone. When chances to meet a low-cost type are small, litigation can effectively be avoided. Although low cost types require a large compensation for a settlement, the mediator can grant this as she needs to pay this compensation seldom. As the ex-ante probability of low-cost types increases the mediator must pay the compensation more often, but at the same time the amount decreases. The result is an inverse U-shaped relationship between the prior and the probability of litigation.

In addition, comparative statics show that ADR is a very effective tool. In our setup the mediator can settle the majority of the cases for any set of parameters, p and κ. The next proposition summarizes these findings.

Proposition 2. *Under the optimal mediation protocol, the ex-ante probability of breakdown is never greater than* $1/2$. *Moreover, the probability of breakdown is increasing and concave in* κ *while it takes the form of an inverse U-shape in* p.

Next, we want to discuss how the asymmetry translates to the different outcome variables. A first result is straightforward and a direct consequence of Theorem 1: low cost types experience breakdown more often than high-cost types. Moreover, player $1l$ is sent to court more often than player $2l$ as the belief on player 1 is larger than on player 2. Since the participation constraint binds, both low-cost type players experience the same utility in expectations. However, the contest utility is the same for both low-cost types and smaller than the value of vetoing, $V(l)$, as low-cost types are more likely after breakdown than in the initial population. Thus, player $2l$, who is sent to court less often, receives a smaller expected share than player $1l$. For high-cost types the intuition is the other way around. Player $2h$, who experiences no utility in contest post-mediation, is compensated with a larger amount than $1h$. The next proposition states that this is the case for all parameter values. Thus, player 1, who is stronger in the contest, expects a less favorable settlement contract than player 2 who, in turn, faces a more difficult task to win the litigation process after breakdown.

Proposition 3. *Both the pre-mediation probability of being sent to court during mediation and the expected share conditional on settlement are largest for player 1l and smallest for player 1h.*

Figure 5 illustrates the results of Proposition 3 as a function of the prior distribution. The left panel (a) describes breakdown utilities, the middle panel (b) expected shares conditional on settlement, $x_i(m) \equiv z_i(m)/(1-\gamma_i(m))$, and the right panel (c) the settlement valuation. Dark-red lines are for player 1 and light-blue lines for player 2. Dashed lines indicate high-cost types, solid lines indicate low-cost types. The linear, gray, dotted line in panel (b) denotes the value of vetoing for the l-type, $V(l)$.

If the probability of low-cost types is very small, the mediator sends one of the two low-cost types to litigation with certainty to ensure that the resource constraint holds. As the

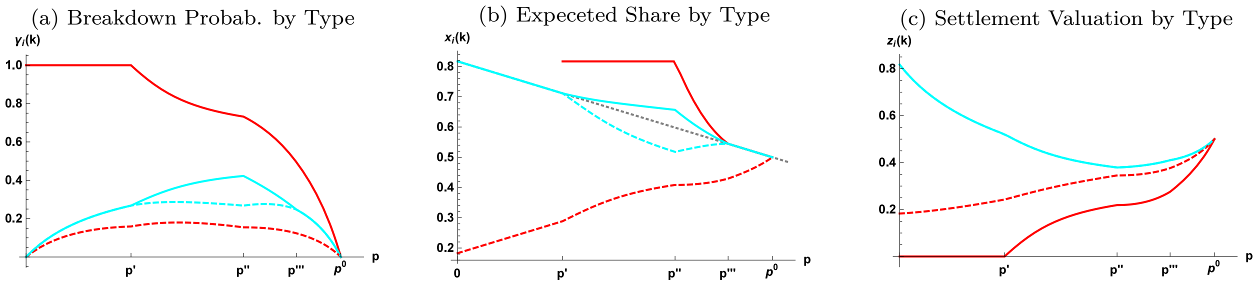

Figure 5: (a) *Expected Contest Probability,* (b) *Expected Share conditional on settlement taking place, and* (c) *Valuation of Settlement by player-type as a function of the ex-ante probability of being a low-cost type. Solid lines depict low-cost types, dashed lines depict high-cost types. Dark-red is player 1 and light-blue is player 2. The dotted gray line in* (b) *is the value of vetoing for low-cost type players. In* (c)*, player 2h has the same settlement value as 2l by incentive compatibility.*

probability of low-cost players increases, the pressure from the resource constraint relaxes as compensation for low-cost types declines. The mediator thus wishes to implement a less asymmetric solution. As p increases further, the mediator can in fact reduce the probability of litigation for all types up to the point where Assumption 1 seizes to hold and the problem therefore becomes trivial.

Another feature of our model is that we are able to evaluate the consequences of the mediation decision on the litigation process. As litigation in our model is a strategic game with actions that depend both on first and second order beliefs, we should not expect players to play the same strategies as in litigation without a preceding mediation stage. Indeed the mediation attempt changes the belief structure of the opposing parties in two ways: (1) it increases the likelihood for both players to meet a low-cost type in court and (2) it introduces an asymmetry that makes player 1 more likely to be the low-cost type than player 2.

The first effect clearly makes competition more intense as litigants are afraid that the opponent can and will produce good evidence. The second effect works in the other direction, since the high-cost type of player 2 has little chance of winning in court. She refuses to compete at all from time to time and gives away the good for free. The second effect exceeds the first, if player 2's likelihood of being the low type is the same as the prior, that is if $\rho_2 = p$ or $p \neq (p', p''')$. In such a case the incentive to downsize investment in evidence due to the asymmetry between players always supersedes the incentive to increase investment in evidence due to the higher probability of low-cost types and we would see lower legal expenditure post-mediation.

Proposition 4. *Assume parameters are such that $p < p'$ or $p > p'''$. Then, the sum of expected legal expenditures after breakdown never exceeds the sum of expected legal expenditures if mediation did not exist.*

Outside this range no clear statement can be made other than that for any κ there exists a possibly empty interval $(\hat{p}, \check{p})$, with $\hat{p} \geq p'$ and $\check{p} \leq p'''$. Only in this interval, the expected legal expenditure after breakdown is higher than without mediation.

6 Extensions

Pre-trial Bargaining. The traditional law and economics literature focuses mainly on bilateral settlement negotiations. Typically, these bargains are modeled as a simple take-it-or-leave it bargaining game (Schweizer, 1989; Shavell, 1995; Posner, 1996). For illustration assume the following bargaining procedure close to Schweizer (1989): one player (Sender) makes a take-it-or-leave-it offer to the other player (Receiver) who decides whether to accept or reject the offer. Upon rejection both players update their beliefs and proceed to litigation.

To compare our results, notice first that by the revelation principle and Lemma 8, the equilibrium rejection channel is absent. Pre-trial negotiations thus cannot out-perform the result of the mechanism.

As in the mediation mechanism, off-path beliefs play a crucial role in the bargaining game. The actions in the contest are based on the belief structure as discussed above.

The solution concept of perfect Bayesian Nash equilibrium allows to freely choose beliefs put on the deviator at the first node of deviation, but requires Bayes' rule thereafter. Any bargaining equilibrium that performs as well as the mediation mechanism replicates outcome utilities of the mechanism and is furthermore equipped with a set of off-path beliefs that deter any deviation by any player. It turns out that no off-path belief exists such that the bargaining can replicate the mediator's solution as long as Assumption 1 holds.

Proposition 5. *Independent of the off-path belief structure, take-it-or-leave-it bargaining leads to a strictly higher probability of litigation than the optimal mediation mechanism provided that Assumption 1 holds.*

The intuition behind the result is that a low-cost Sender could always profitably deviate by proposing an arbitrarily small share ϵ to Receiver. Then, given any belief Receiver holds after observing this deviation, she either accepts the share which gives Sender a higher utility than in the optimal mechanism, or rejects the share if she thinks Sender is weak. Assuming a weak Sender, however, induces her to score softer than in the litigation game under priors. By strategic complementarity, Sender scores softer as well. But then, Sender expects a higher utility as winning is less costly. Thus, it is not optimal for a low-cost Sender to reproduce the outcome of the optimal mechanism: the incentive to deviate from the mechanism leads to a higher breakdown probability in expectations.

This shows the importance of a third-party who manages the information flow. With direct bargaining, Receiver always interprets Sender's proposal as a signal and Sender cannot commit to abstain from signaling via her proposal. A neutral third-party can overcome this adverse selection problem and thus improves upon bilateral negotiations.

Asymmetric Players. Asymmetric players do not change any of the results obtained. The reason is that the mediator would always treat the ex-ante stronger player as "player 2", i.e. the player that gets the better settlement conditions. The ex-ante weaker player accepts a small settlement share, since she fears a strong opponent in litigation. The weaker player, however, is compensated for the small share with a favorable contest after breakdown. Thus, while the ex-ante weaker player is strong post-breakdown, she agrees to settlement-contracts that favor her opponent. With such a protocol the mediator is still able to solve the majority of the cases. A key result of our analysis is, however, that we get asymmetric results even with symmetric players.

Different forms of commitment. So far we have assumed that both players can fully commit to the proposed mediation protocol. In particular, once the mechanism is accepted, parties commit to only go back to litigation if the mediator tells them. In reality this is not always the case. Many jurisdictions demand that parties can unilaterally opt-out of ADR at any point to return to litigation. We discuss two stages at which parties can unilaterally decide to break down ADR. The first is a situation in which they can leave after the mechanism has told players' their *expected* share conditional on settlement. We call this commitment structure post-ADR individual rationality (PAIR). The second commitment structure is that parties can veto the mechanism after they have learned their *realized* share conditional on settlement. We call this ex-post individual rationality (EPIR).

The mediation protocol developed in Section 4 does not directly carry over to PAIR and EPIR. In fact, given these commitment schemes, the mediator profits from the ability to communicate to parties even after ADR breaks down. If this is the case, the mediator can give parties non-binding recommendations for the play of the contest and by that restore the outcome under full-commitment. The modified game thus follows a slightly enhanced timeline:

1. the mediator commits to $\mathcal{X}$ and recommendation structure Σ; players learn their types,
2. players send a message m_i to the mediator,
3. the mediator *privately* announces a share x_i according to $\mathcal{X}$ to each player i,
4. players accept/reject the share,
5. players receive a recommendation σ_i by the mediator,
6. if either of the players rejected her offer, the contest is played under updated beliefs.

Note that since the mediator first observes the behavior of the players with respect to the announced share she has the ability to detect a deviation in this stage (other than in the reporting stage). To restore the result of Section 4 the mediator uses the following slightly more sophisticated mechanism(s).

To find the optimal PAIR mechanism, we need to define a convex combination of the protocol derived in Section 4 and its mirror image switching roles of player 1 and 2. Define $\hat{\mathcal{X}}_\lambda$, a mediation protocol such that $\mathcal{X}_i$ applies with probability λ and $\mathcal{X}_{-i}$ with probability $(1-\lambda)$. $\mathcal{X}_i$ denotes a mediation protocol similar to the one discussed in Theorem 1. When mediation is successful, player i is treated as "player 1". To trigger litigation in this protocol, the mediator offers a share of 0 to at least one of the players. This share is going to be rejected such that parties move to the litigation game.

To ensure EPIR we need that, in addition, the mediator sends both parties to contest irrespective of their reports with probability $\epsilon > 0$. Thus, we define $\hat{\mathcal{X}}_\lambda^\epsilon$ to be a mediation protocol such that with probability ϵ players are send to court and with probability $(1-\epsilon)$ the mediator executes $\hat{\mathcal{X}}_\lambda$. This is sufficient to ensure the following two results.

Proposition 6. *There exists a signal* Σ*, such that an incentive compatible PAIR mechanism* $(\hat{\mathcal{X}}_{1/2}, \Sigma)$ *has the same breakdown probability* $Pr(\Gamma)$ *as the mechanism* $\mathcal{X}$ *under interim individual rationality.*

Proposition 7. *For any* $\delta > 0$*, there exists a signal* Σ *and an* $\epsilon > 0$*, such that an incentive compatible EPIR mechanism* $(\hat{\mathcal{X}}_{1/2}^\epsilon, \Sigma)$ *achieves a breakdown probability* $Pr(\Gamma)^\epsilon < Pr(\Gamma) - \delta$*, where* $Pr(\Gamma)$ *is the optimal breakdown probability of the mechanism* $\mathcal{X}$ *under interim individual rationality.*

To gain intuition observe the following. First, with both PAIR and EPIR the mediator can trigger the play of a contest by offering at least one party an unacceptable share as rejection leads to contest. Second, the mediator achieves the result by obfuscating two issues: the role of the player and the relevance of her decision.

The latter derives from the possibility that the mediator wants to trigger contest play and has offered the player's opponent an unacceptable share. As both do not know which litigant takes the role of player 1, and who is offered the trigger share 0, she cannot learn much from her own offer. As the conditional distribution post-breakdown is on-path revealed via the signal σ, obfuscation is only payoff relevant in deviation games. Deviation is, however, only detected by the mediator, *not* by the non-deviator. Thus, the mediator can react to deviation by sending the the deviator a signal of a strong non-deviator to punish her. This suffices to get the same result as under full-commitment.

In the case of EPIR the mediator is more constrained as revealing the ex-post share $x_i(k_1, k_2)$ to player i allows for more inference by the player. For some parameter values it might be the case that certain constellations do not settle on-path. Thus, the mediator might have a degenerate belief after some proposed realized shares which makes the procedure of PAIR impossible. The mediator can use another option instead, though. She can commit to initiate breakdown for any type-profile with a small probability ϵ and to send a fully informative signal thereafter. In such a case both parties can end up with 0 expected utility after breakdown. If the mediator commits to signal this event to the non-deviator after any deviation, the non-deviator will always invest an amount large enough to effectively punish the deviator. As $\epsilon \to 0$ the mechanism converges to $\hat{\mathcal{X}}_\lambda$ and the resulting probability of breakdown is arbitrarily close to that of the mechanism described in Theorem 1.

Nonetheless, allowing the types to go back to court after all uncertainty has unraveled would naturally lead to a different result. Typically however, once a detailed settlement agreement has been signed by both parties, it is hard to imagine a legal system that allows parties to overturn this contract simply because they have learned that they might have a good chance to beat the opponent.

7 Conclusion

In this paper we characterize optimal Alternative Dispute Resolution (ADR) in the shadow of the court. We show that optimal ADR is always asymmetric and offers one player an advantage after breakdown and the other one an advantage under settlement. We show that the optimal information structure post-ADR is completely independent of the players' report, but conditions only on their identity. Such a mechanism prevents players from misreporting to achieve an informational advantage.

We find that a litigation-minimizing ADR-protocol is highly effective and solves the majority of cases. The effectiveness indicates that mandatory ADR should be considered by all courts to reduce the prevalent stress on judges and court's backlog of cases. In addition, the asymmetry of the optimal mechanism implies that regulators should act carefully when defining their notion of fairness for mediation protocols. The same holds true for discretionary policies: mediators should always have the possibility to talk to the disputants in private as this eliminates commitment problems on the disputants side. Finally, we show that mediators should not be forced to disclose all their information in the event of breakdown. Trust in the mediator's discretion is an important driving force of the success of a mechanism.

More broadly, we show that the most important aspect of the optimal ADR-protocol is the management of the information structure in litigation post-breakdown. The optimal protocol imposes type-independent beliefs to minimize the potential gain a deviator can earn in the litigation game following a misreport. In addition, the protocol is asymmetric to reduce resource intensity in case of breakdown.

We demonstrate that the standard assumption of fixed, type-dependent outside options in mechanism design is not innocuous when the following two conditions are satisfied: (1) the mechanism cannot replace the underlying default game completely and (2) the actions chosen in the underlying game depend on player's beliefs. We show that the behavior of the players in the mechanism and those in the underlying game are interconnected. For the case of contests, we show that players invest less resources post-breakdown for extreme type distributions compared to a situation in which no resolution mechanism is present. For intermediate type-distributions, however, the post-mediation contest can also be more resource intensive.

Not claiming that the actual ADR-mechanisms we observe in reality are optimal, we want to note that our findings are in line with some observations on ADR. Its success rates are beyond 50% across cases and jurisdictions and mediation is considered to be informative when breaking down. In addition, one reason why mediation is perceived to be successful is its ability to not rely on publicly observable actions of the mediator, but allowing for private settlement negotiations.

Our findings provide several interesting directions for future research. First of all, the assumption that the mechanism designer has full-commitment could be relaxed to allow for third-party renegotiation. Especially when mediators compete for clients this seems reasonable. Further, extending the analysis to a setup of more than two players and possibly correlated types might add several interesting channels to the model. In addition, many conflicts evolve around a variety of battlefields on different subjects or points in time. If types are correlated over time this adds an additional signaling dimension which is interesting to analyze further. Finally, although minimizing court appearances is optimal given the public good properties of the legal system, it is less clear in other contest situations whether this is the most suitable objective. Although a richer model is needed to address such issues properly, we are confident that the results of this papers provide a first step towards analyzing these problems.

References

Anderson, H. and R. Pi (2004). *Evaluation of the Early Mediation Pilot Programs*. Tech. rep. Judical Council of California.

Baye, M. R., D. Kovenock, and C. G. Vries (1996). “The all-pay auction with complete information”. *Economic Theory* 8, pp. 291–305.

Baye, M. R., D. Kovenock, and C. G. Vries (2005). "Comparative Analysis of Litigation Systems: An Auction-Theoretic Approach". *The Economic Journal* 115, pp. 583–601.

Bebchuk, L. A. (1984). "Litigation and settlement under imperfect information". *The RAND Journal of Economics*, pp. 404–415.

Bester, H. and K. Wärneryd (2006). "Conflict and the Social Contract". *The Scandinavian Journal of Economics* 108, pp. 231–249.

Border, K. C. (2007). "Reduced form auctions revisited". *Economic Theory* 31, pp. 167–181.

Brown, J. G. and I. Ayres (1994). "Economic rationales for mediation". *Virginia Law Review*, pp. 323–402.

Brown, S., C. Cervenak, and D. Fairman (1998). *Alternative Dispute Resolution Practitioners Guide*. Tech. rep. Conflict Management Group.

Celik, G. and M. Peters (2011). "Equilibrium rejection of a mechanism". *Games and Economic Behavior* 73, pp. 375–387.

— (2013). "Reciprocal relationships and mechanism design". *mimeo*.

Che, Y.-K. and I. Gale (2000). "Difference-form contests and the robustness of all-pay auctions". *Games and Economic Behavior* 30, pp. 22–43.

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

Doornik, K. (2014). "A rationale for mediation and its optimal use". *International Review of Law and Economics* 38, pp. 1–10.

Ewerhart, C. (2015). "Rent-seeking games and the all-pay auction". *mimeo*.

Fey, M. and K. W. Ramsay (2010). "When Is Shuttle Diplomacy Worth the Commute? Information Sharing through Mediation". *World Politics* 62 (04), pp. 529–560.

Fudenberg, D. and J. Tirole (1988). "Perfect Bayesian and Sequential Equilibria - A clarifying Note". *mimeo*.

Genn, H. (1998). *The Central London County Court Pilot Mediation Scheme*. Tech. rep. Department for Constitutional Affairs.

Hörner, J., M. Morelli, and F. Squintani (2015). "Mediation and Peace". *The Review of Economic Studies* 82, pp. 1483–1501.

Katz, A. (1988). "Judicial decisionmaking and litigation expenditure". *International Review of Law and Economics* 8, pp. 127–143.

Meirowitz, A., M. Morelli, K. W. Ramsay, and F. Squintani (2015). "Dispute Resolution Institutions and Strategic Militarization". *mimeo*.

Mitusch, K. and R. Strausz (2005). "Mediation in Situations of Conflict and Limited Commitment". *Journal of Law, Economics, and Organization* 21, pp. 467–500.

Pavlov, G. (2013). "Correlated equilibria and communication equilibria in all-pay auctions". *mimeo*.

Posner, R. A. (1973). "An economic approach to legal procedure and judicial administration". *The Journal of Legal Studies*, pp. 399–458.

— (1996). *The Federal Courts: Challenge and Reform*. Harvard University Press.

Prescott, J. J., K. E. Spier, and A. Yoon (2014). “Trial and Settlement: A Study of High-Low Agreements”. *Journal of Law and Economics* 57, pp. 699–821.

Schweizer, U. (1989). “Litigation and settlement under two-sided incomplete information”. *The Review of Economic Studies* 56, pp. 163–177.

Shavell, S. (1995). “Alternative dispute resolution: an economic analysis”. *The Journal of Legal Studies*, pp. 1–28.

Siegel, R. (2009). “All-Pay Contests”. *Econometrica* 77, pp. 71–92.

— (2014). “Asymmetric all-pay auctions with interdependent valuations”. *Journal of Economic Theory* 153, pp. 684–702.

Spier, K. E. (1994). “Pretrial bargaining and the design of fee-shifting rules”. *The RAND Journal of Economics*, pp. 197–214.

— (2007). “Litigation”. *Handbook of law and economics* 1, pp. 259–342.

Spier, K. E. and D. Rosenberg (2011). “A Contest Model of Separate Action versus Class Action Litigation”. *Journal of Legal Analysis*, pp. 11–28.

Szech, N. (2011). “Asymmetric all-pay auctions with two types”. *mimeo.*

Tullock, G. (1980). “Efficient rent seeking”. *Towards a Theory of the Rent-Seeking Society.* Ed. by J. Buchanan, G. Tullock, and R. Tollison. College Station, TX: Texas A and M University Press, pp. 97–112.

A Details on Rewriting the Problem

Full Participation. Full participatio is a consequence of the fact that litigation utility is convex in beliefs and Proposition 2 from Celik and Peters (2011).

Lemma 8. *It is without loss of generality to assume full participation in the optimal mechanism.*

Value of Vetoing. Any off-path belief structure that satisfies the intuitive criterion leads to a player independent value $V_i(k)$, which is

$$V(l) = (1-p)\frac{\kappa - 1}{\kappa}, \text{ and } V(h) = 0.$$

Given the constant outside option, the channels identified by Celik and Peters (2011) and Cramton and Palfrey (1995) are not present in our model as off path beliefs are less important.
Reduced Form Problem à la Border (2007). We reduce the problem by replacing the settlement shares, X_i, by the settlement values, z_i. For any given matrix of breakdown probabilities, G, this reduction is possible if and only if each settlement share is both individually feasible (condition (F), below) and ex-post implementable (condition (EPI), below). The following lemma states these conditions. With some abuse of notation, let $p(m)$ be the ex-ante probability that player i is of type m, that is $p(l) = p$ and $p(h) = 1 - p$.

Lemma 9. *For every message $m \in \{l, h\}$, let $m^c := \Big\{k \in \{l, h\} | k \neq m\Big\}$, and fix some feasible G and $z_i \geq 0$ for every i. Then there exists an ex-post feasible X_i that implements z_i if and only if the following constraints are satisfied:*

- $\forall \{m,n\} \in \{h,l\}^2$:

$$p(m)z_i(m) + p(n)z_{-i}(n) \leq \\ 1 - Pr(\Gamma) - \Big(1 - \gamma(m^c, n^c)\Big)p(m^c)p(n^c) \qquad (EPI)$$

- $\forall m \in \{h,l\}$ *and* $i = 1,2$*:*

$$z_i(m) \leq 1 - \gamma_i(m) \qquad (IF)$$

Moreover, if $\gamma_i(l) \geq \gamma_i(h)$ *then* $z_i(l) \leq 1 - \gamma_i(l)$ *and* (IC_i^l) *imply equation* (IF)*.*

Note that a necessary condition for individual feasibility (IF) is that it holds in expectations, that is the weighted sum of settlement values cannot exceed the probability of successful ADR,

$$\sum_{i\in\{1,2\}} \sum_{m\in\{l,h\}} p(m)z_i(m) \leq 1 - Pr(\Gamma). \qquad (AF)$$

The High-Cost's IC and the Low-Cost's PC bind. Next, we eliminate all settlement values with help of the following lemma stating that in the optimal mechanism the high-cost type's incentive constraint and the low-cost type's participation constraint bind for both players.

Lemma 10. *It is without loss of generality to assume that* (IC_i^h) *and* (PC_i^l) *hold with equality in the optimal mechanism.*

The result is a direct consequence of the different costs. High-cost types care more about settlement than about breakdown. Thus, incentive compatibility requires a large value of settlement, $z_i(h)$, for them. However, there is no reason for the mediator to set $z_i(h)$ too high, as the h-type would never veto ADR. We can express (IC_i^h) as

$$z_i(h) + \gamma_i(h)U_i(h|h) \geq z_i(l) + \gamma_i(l)U_i(h|l). \qquad (IC_i^h)$$

If this inequality is strict, the mediator can reduce the value of settlement, $z_i(h)$, without affecting the breakdown probability $Pr(\Gamma)$ or any of the other constraints.

Similarly the mediator can reduce the value of settlement, $z_i(l)$, if l-types' participation constraint is not binding, as any negative effect on l-types incentive constraint (IC_i^l) is of second order compared to the positive effect on h-types incentive constraint, (IC_i^h). By readjusting the settlement value for h-types, $z_i(h)$, incentive compatibility for both types can always be guaranteed. The l-types participation constraint is

$$z_i(l) + \gamma_i(l)U_i(l|l) \geq V(l). \qquad (PC_i^l)$$

Using (PC_i^l), (IC_i^h) and Lemma 10 we can eliminate all settlement values, z_i, and express the result only in terms of breakdown valuations, $\gamma_i(m)U_i(k|m)$.

Breakdown Probabilities and Beliefs. Breakdown beliefs $p_i(l|k)$ are a result of breakdown probabilities. The belief that player 1 is type l, given 2 reported m is

$$p_1(l|m) = \frac{p\gamma(l,m)}{p\gamma(l,m) + (1-p)\gamma(h,m)}.$$

Observation 1. Any $p_i(l|m)$ is homogeneous of degree 0 in G.

Thus, any set of beliefs $p_i(k|m)$ induced by some G is induced by $G' = \alpha G$, too.

Lemma 11. *Fix any feasible G with $1 \geq \gamma(l,h), \gamma(h,l), \gamma(l,l) \geq 0$ and define*

$$q_i(m) := \frac{p}{1-p} \frac{1-p_i(l|m)}{p_i(l|m)}.$$

Then the induced information structure $P > 0$ satisfies:

$$\begin{aligned} \gamma(h,l) &= q_1(l)\gamma(l,l) \leq 1 & \gamma(l,h) &= q_2(l)\gamma(l,l) \leq 1; \\ \gamma(h,h) &= q_2(h)q_1(l)\gamma(l,l) \leq 1 & q_2(h)q_1(l) &= q_1(h)q_2(l), \end{aligned} \tag{C}$$

where the last equation ensures consistency with the prior. Conversely, for any $\gamma(l,l) \in (0,1]$ and $P > 0$ satisfying (C) there exists a feasible G.

The Fully Reduced Problem. By Lemma 11 all breakdown probabilities are linear in $\gamma(l,l)$. If we plug all breakdown probabilities into the aggregate feasibility constraint, (AF), we get an expression of the form

$$\underbrace{2V(l) - \gamma(l,l)Q(P)}_{\text{LHS of } (AF)} \leq \underbrace{1 - \gamma(l,l)R(P)}_{1-Pr(\Gamma)}. \tag{4}$$

Assumption 1 implies $Q(P) \geq R(P)$ and we can reformulate

$$1 \geq \gamma(l,l) \geq \frac{\nu}{Q(P) - R(P)} =: \gamma^*(P)m, \tag{AF'}$$

with $\nu = 2V(l) - 1$ independent of P. Reducing $\gamma(l,l)$ reduces $Pr(\Gamma)$. Thus, constraint (AF') binds at the optimum, and $\gamma(l,l) = \gamma^*(P)$. Plugging into $Pr(\Gamma)$, we get

$$\min_P R(P)\gamma^*(P) \tag{P1'}$$

subject to the remaining constraints (IC_i^l), (IF), (EPI) and $\gamma^*(P) \leq 1$ and any solution to $(P1)$ is also a solution to $(P1')$.[16]

B Forces of Asymmetry

We first consider the optimal symmetric mechanism. Notice that the designer of a symmetric mechanism has only one choice variable $\tilde{\rho} := \rho_1 = \rho_2$. In a symmetric mechanism, Corollary 1 holds and any subscripts can be dropped. In combination with type-independent beliefs we get $U(h|h) = U(h|l) = 0$. By incentive compatibility, (IC^h), settlement values must thus be equal, i.e. $z(l)$=$z(h)$=z. Using the participation constraint, (PC^l), the settlement value z can be expressed as

$$z = V(l) - \gamma(l)U(l|l).$$

[16]Problem (P1') is in fact equivalent to problem $(P1)$ whenever $P > 0$. As every argument is continuous in P this limitation only becomes relevant once (P1') has no minimum.

Ignoring any effect on $U(l|l)$, an increase in $\tilde{\rho}$ increases the settlement-value the mediator needs to offer. This effect is strengthened as $\tilde{\rho}$ decreases $U(l|l)$. Next, consider the total resources distributed

$$2z = 1 - Pr(\Gamma). \tag{AF}$$

As $\tilde{\rho}$ increases, breakdown decreases and the mediator can distribute more resources in case of settlement.

Combining the two equations yields

$$\underbrace{2V(l) - 1}_{=\nu} = 2\gamma(l)U(l|l) - Pr(\Gamma). \tag{5}$$

Using Corollary 2 and (2) we can rewrite equation (5)

$$\begin{aligned} \nu =& \gamma(l)\left((1-\tilde{\rho})\frac{(\kappa-1)}{\kappa} - \frac{p}{\tilde{\rho}}\right) \\ \Leftrightarrow \nu =& 2\underbrace{\gamma(l)\frac{p}{\tilde{\rho}}}_{=Pr(\Gamma)}\left(\frac{(1-\tilde{\rho})\tilde{\rho}}{p}\frac{(\kappa-1)}{\kappa} - 1\right). \end{aligned}$$

Solving for $Pr(\Gamma)$ yields

$$Pr(\Gamma) = \frac{\nu}{2}\left(\frac{(1-\tilde{\rho})\tilde{\rho}}{p}\frac{(\kappa-1)}{\kappa} - 1\right)^{-1}$$

which is minimized for $\tilde{\rho} = 1/2$. Thus, the optimal symmetric solution to $(P1')$ is obtained for breakdown probability $\tilde{\rho} = 1/2$.

A symmetric mechanism is, however, never optimal. This follows from the differences in the resources needed to sustain a certain level of either ρ_i. First, observe that despite any asymmetry, (IC_2^h) still requires that the settlement value of the high type $z_2(h) = z_2(l)$. As $U_2(h|h) = 0$, the breakdown value is 0 and expected settlement valuation of player 2 is

$$z_2 := z_2(l) = V(l) - \gamma_2(l)U(l|l) = V(l) - \frac{(1-\rho_2)}{\rho_1}\frac{(\kappa-1)}{\kappa}p\gamma(l,l).$$

The first equality comes from (PC_2^l) and second from the results of Corollary 2 and the equations in (2).

For player 1, on the other hand the results change more substantially under asymmetry. Player $1h$'s incentive constraint is

$$z_1(h) = z_1(l) + \Big(\gamma_1(l) - \gamma_1(h)\Big)U_1(h|h). \tag{IC_1^h}$$

As $U_1(h|h) \neq 0$ the mediator pays an information rent to player 1 if $\gamma_i(l) \neq \gamma_i(h)$. Thus, the ex-ante expected valuation of player 1 under settlement is

$$\begin{aligned}pz_1(l)+(1-p)z_1(h) &= z_1(l)+(1-p)\Big(\gamma_1(l)-\gamma_1(h)\Big)U_1(h|h)\\ &= z_1(l)+\gamma_1(l)\left(1-\frac{p}{\rho_1}\right)U_1(h|h) \qquad (6)\end{aligned}$$

where the first uses (IC_1^h) and the second uses Corollary 2 to simplify. Simplifying this using (PC_1^l), (2), and $U_i(\cdot,\cdot)$ yields

$$z_1 := \underbrace{V(l)-\left(\frac{(1-\rho_1)}{\rho_2}\right)\frac{\kappa-1}{\kappa}p\gamma(l,l)}_{\text{symmetric part}}+\underbrace{\left(\frac{p}{\rho_1}-\frac{p}{\rho_2}\right)\frac{\kappa-1}{\kappa}p\gamma(l,l)}_{\text{asymmetric part}}.$$

While the symmetric part is always present, the asymmetric is only non-zero in asymmetric cases. As $\rho_1 > \rho_2$ in such cases the asymmetric part is genereically negative. Marginal effects on the second part cancel out with those on z_2. As the asymmetric part is additive separable in ρ_i, the optimum of ρ_i is independent of the choice of ρ_{-i}.

C Proofs

C.1 Proof of Lemma 1

Proof. The proof is along the lines of Siegel (2014). However, as the proof is instructive and our setup differs slightly, we spell it out here. We first show that at least one type of one player has 0 expected utility. Second, we show that at most one type has an atom at 0. Third, we constructively show that the equilibrium exists and then show that it is indeed unique given (M). Then we calculate Δ to state Corollary 1.

Step 1: One player has 0 expected utility and no atoms at positive scores. We prove this by contradiction. Suppose that both players and both types expect a utility larger 0. That means the smallest score $\underline{s} > 0$ in the union of the best-responses of all players wins the contest with positive probability as otherwise it is no best response. As a result, the smallest score is an atom in the strategy of at least one type of each player. But then, there exists an ϵ in the neighborhood of $\underline{s}$ such that the probability of winning increases with more than $\epsilon * \kappa c_l$. Deviating to $\underline{s}+\epsilon$ is profitable for that type of player, and thus $\underline{s}$ cannot be an atom in her strategy. Therefore, at least one player earns an expected utility of 0 for sure. Note that this player may very well have an atom at 0 as there is no need to win the good with positive probability for an atom at 0. However, if both players had a type with an atom at 0 at least one of them can profitably deviate to a positive neighborhood of 0 winning against the atom scoring opponent with a probability that exceeds the cost of scoring. Thus, at most one player has an atom at 0.

Step 2a: Construct the equilibrium. First, consider the following strategy of player $2l$: she uniformly mixes on $(\Delta_{l,h}, \Delta_{l,l}]$ with density $f_{2,l}(t) = c_l/p_2(l,l)$. Then, player $1l$ is indifferent between playing any point on $s \in (\Delta_{l,h}, \Delta_{l,l}]$ as

$$\begin{aligned} U_1(l,s) &= F_2(\Delta_{l,h}) + p_2(l|l)(s-\Delta_{l,h})\frac{c_l}{p_2(l,l)} - c_l s = \\ &= F_2(\Delta_{l,h}) - \Delta_{l,h} c_l. \end{aligned}$$

We want to construct strategies with constant density and non-overlapping strategies, thus the length of the top interval $L(t)$ is the solution to

$$L(t) f_{2,l}(t) = 1.$$

To make player $2l$ indifferent as well, player $1l$ plays a similar strategy only flipping the probabilities from p_1 to p_2. As we assumed $p_1(l|l) \geq p_2(l|l)$, the mass of player $1l$ is only fully exhausted on the top interval iff $p_1(l|l) = p_2(l|l)$. If this is not the case, player 1 has some mass left to place. She does so on the middle interval $(\Delta_{h,h}, \Delta_{l,h}]$. For the same reasons as above, she assigns density $f_{1,l}(t) = c_l/p_1(l|h)$ to this interval to make player $1h$ indifferent.

The length of the medium interval can be calculated by acknowledging that player $1l$ needs to place all mass available to her and not placed on the top interval on this interval.

By a similar exercise we can find the length of the interval $(0, \Delta_{h,h})$ and by this the absolute values of all Δ.

Step 2b: Show that no (global) deviation is possible. What remains to be shown is that any player that scoring on more than one interval is in fact indifferent between those and that no global deviation is possible.

Note that the indifference across intervals follows from the intervals being connected. Consider for example player $1l$. From the above we know that

$$U_1(l, s = \Delta_{l,h}) = U_1(l, s = \Delta_{l,l})$$

but also that

$$U_1(l, s = \Delta_{h,h}) = U_1(l, s = \Delta_{l,h}).$$

Thus, it must be the case that

$$U_1(l, s = \Delta_{h,h}) = U_1(l, s = \Delta_{l,l}).$$

The same holds true for player $2h$. The two other player-type tuples place there scores on a single interval only. Note that, since player $1h$ has positive mass only on $(0, \Delta_{h,h}]$ it can in fact earn an expected utility greater 0 if and only if player $2h$ does not enter the auction with positive probability.

To exclude global deviation observe that player $2h$ would only deviate to anything on the interval $(\Delta_{l,h}, \Delta_{l,l}]$ if the probability of winning increases faster in the top interval than in the middle interval, that is the density is smaller in the top interval,

$$f_{1,l}(m) = \frac{c_l}{p_1(l|l)} \geq \frac{\kappa c_l}{p_1(l|h)} = f_{1,l}(t),$$

which is ruled out by (M).

For $1h$, the deviation could be made into the middle or the top interval if

$$\frac{\kappa c_l}{p_2(h|h)} \geq \frac{c_l}{p_2(h|l)},$$

which again is ruled out by monotonicity. As player $1h$ prefers the bottom interval to anything in the m she must prefer scoring at $\Delta_{l,l}$ to $\Delta_{h,l}$. However as player $2h$ does not prefer to score at $\Delta_{l,l}$ it follows that $\Delta_{l,l} > 1/\kappa c_l$. Thus player $1h$ does not want to deviate. Similar arguments hold for the second player, such that we can conclude that global deviations are not beneficial.

Step 3: Uniqueness. For uniqueness observe first that there is only one monotonic equilibrium, that is an equilibrium such that the lowest score of player i, type l, is weakly above the highest score of player i, type h. This follows directly from the equilibrium construction.

Second, we need to show that no non-monotonic equilibrium exists. We do so by contradiction, that is suppose there exists a score $s_i^h > s_i^l$ such that s_i^k is in the set of best responses for player i type k, $BR(k)$. Then, it must hold that

$$\begin{aligned} & U_i(h, s = s_i^h) \geq U_i(h, s = s_i^l) \\ \Leftrightarrow \quad & \sum_k p_i(k|h) F_{-i,k}(s_i^h) - \kappa c_l s_i^h \geq \sum_k p_i(k|h) F_{-i,k}(s_i^l) - \kappa c_l s_i^l \\ \Leftrightarrow \quad & \sum_k p_i(k|h)(F_{-i,k}(s_i^h) - F_{-i,k}(s_i^l)) \geq \kappa c_l (s_i^h - s_i^l). \end{aligned} \tag{7}$$

Similarly, as s_i^l is a best response for l it must hold that

$$\sum_k p_i(k|l)(F_{-i,k}(s_i^h) - F_{-i,k}(s_i^l)) \leq c_l (s_i^h - s_i^l). \tag{8}$$

But, as $F_{-i,k}(\cdot)$ is always positive and $p_i(h|\cdot) = 1 - p_i(l|\cdot)$, inequalities (7) and (8) only hold if

$$\frac{p_i(l|h)}{\kappa c_l} \sum_k (F_{-i,k}(s_i^h) - F_{-i,h}(s_i^l)) \geq \frac{p_i(l|l)}{c_L} \sum_k (F_{-i,k}(s_i^h) - F_{-i,h}(s_i^l)).$$

As the sum is identical on both sides, this boils down to the inverse of (M), a contradiction. □

(Addendum) Step 4: Equilibrium expected utilities. The length of the top interval, $(\Delta_{l,h}, \Delta_{l,l}]$, is $p_2(l|l)/c_l$ that of the bottom interval, $(0, \Delta_{h,h})$, is $p_1(h|h)/\kappa c_l$ and that of the middle interval $(\Delta_{h,h}, \Delta_{l,h}]$ is

$$\frac{p_1(l|h)}{\kappa c_l}(1 - \frac{f_{1,l}(t)}{f_{2,l}(t)}) = \frac{p_1(l|h)}{\kappa c_l}(1 - \frac{p_2(l|l)}{p_1(l|l)}).$$

Putting the respective probability masses on the different intervals leaves player 2 with some mass $\mu \geq 0$. This is placed on scoring 0 and constitutes $F_{2,h}(0)$.

Notice that scoring $\Delta_{l,l}$ wins the auction for sure at cost of $\Delta_{l,l} c_l$ for both players, type l, and player 1 scoring (arbitrarily close) to 0 wins the auction with probability $F_{2,h}(0)$ for no cost.

C.2 Proof of Lemma 2

Proof. First, consider player $2h$. She earns an expected utility of 0 on-path. Post-deviation she can always choose a score of 0 to secure this utility.

Second, consider player il. Independently of her report she can always choose a score $\Delta_{l,l}$ and win with probability 1. As this is also part of the best response on-path and the probability is 1 in that case as well, she can only be better of by choosing a score different than $\Delta_{l,l}$.

Finally consider player $1h$ after reporting to be type l. She holds belief $p_2(h|l)$ while her opponent plays the equilibrium strategies. If she were to score 0, then by our tie-braking assumption she would enjoy a utility at least as good as the equilibrium utility if $p_2(h|l) \geq p_2(h|h)$. Thus, in those cases she is weakly better of.

If, however, $p_2(h|l) < p_2(h|h)$ then player 1 suffers whenever scoring against an h-type compared to the on-path game as the probability of winning decreases while costs stay the same. However, scoring against the low-cost type and at the same time earning a higher expected utility than in the default game can, by the constant density of player 2's low-cost type on the support of her equilibrium strategy, only mean scoring to the very top, that is $\Delta_{l,l}$ which yields negative utility to a high type by the construction of the equilibrium. □

C.3 Proof of Lemma 3

Proof. First, notice that player il benefits if and only if $p_{-i}(l|l) > p_{-i}(l|h)$. The if part follows directly from the density of the opposing player on the top interval which is $f_{-i,l}(t) = c_l/p_{-i}(l|l)$. As $p_{-i}(l|h)$ is smaller than this, scoring at $\Delta_{l,h}$ is strictly preferred to $\Delta_{l,l}$, but $\Delta_{l,l}$ yields the same result as the on-path game.

The only-if party follows as for $p_{-i}(l|l) = p_{-i}(l|h)$ would induce type independent beliefs and therefore the same result as the on-path game. For $p_{-i}(l|l) < p_{-i}(l|h)$, however, scoring at the top, i.e. $\Delta_{l,l}$ is preferred leading to no changes in expected utilities at all.

As $p_{-i}(l|l) < p_{-i}(l|h)$ implies $p_{-i}(h|h) < p_{-i}(h|l)$ we know that player 1, type h is better off, as scoring 0 yields him already a higher payoff by $p_2(h|l)F_{2,h}(0) > p_2(h|h)F_{2,h}(0)$. Player 2 strictly prefers to score at $\Delta_{h,h}$ compared to 0 as the density of her opponent is given by $f_{1,l}(b) = c_l/p_1(h|h)$ which leads to a (strictly) increasing utility on the bottom interval. Thus scoring at $\Delta_{h,h}$ must yield strictly positive utility.

The only setup in which neither party has a type that strictly profits from deviating is that of type-independent beliefs. □

C.4 Proof of Proposition 1 (together with lemmas 8 to 11)

The proof of the proposition is along the lines described in appendix A.

C.4.1 Proof of Lemma 8

Proof. We show that the condition stated on the optimality of full participation stated in Proposition 2 of Celik and Peters (2011) is satisfied. That is, given the independent prior p, there is no Bayes' plausible belief structure $\tilde{p} = (\underline{p}, \overline{p})$ such that the expected utility $U_i(k, \tilde{p}, p) < U_i(k, p, p)$ for any type k. The condition is a direct consequence of expected contest utilities under a type-independent prior as defined in Corollary 1. For type independent priors utilities are in fact linear in beliefs except for a kink at the point where utilities become flat. However, around that point utilities are convex and Jensen's inequality yields the desired result. □

C.4.2 Proof of Lemma 9

Proof. We apply theorem 3 of Border (2007) which says the following:

> **Border (2007), Theorem 3:** The list $\mathbf{P}= (P_1, ..., P_N)$ of functions is the reduced form of a general auction $\mathbf{p}= (p_1, ...p_n)$ if and only if for every subset $A \subset \mathcal{T}$ of individual-type pairs (i, τ) we have
>
> $$\sum_{(i,\tau)\in A} P_i(\tau)\mu^{\bullet}(\tau) \leq (\{t \in T : \exists (i, \tau) \in A, t_i = \tau\}).$$

An individual type pair in our setting is given by (m, i), in what follows we are going to abuse notation slightly by treating $p(m)$ such that $p(l) = p$ and $p(h) = 1 - p$. The general auction $\mathbf{p}$ in our setup is defined by a list

$$q_i(m, n) := x_i(m, n).$$

We want to implement $\mathbf{p}$ by the list $\mathbf{P}$ containing

$$Q_i(m) := q_i(m, l)\mu_i(l|m) + q_i(m, h)\mu_i(h|m)$$

where

$$\begin{aligned}
\mu_i(n|m) &:= \frac{\mu(m, n)}{\mu_i^{\bullet}(m)}, \\
\mu(m, n) &:= p(l)p(m)\frac{1 - \gamma(m, n)}{1 - Pr(\Gamma)}, \\
\mu_i^{\bullet}(m) &:= p(m)\frac{1 - \gamma_i(m)}{1 - Pr(\Gamma)}.
\end{aligned}$$

Plugging in yields,

$$Q_i(m) = \frac{p(l)(1 - \gamma(m, l))x_i(m, l) + p(h)(1 - \gamma(m, h))x_i(m, h)}{1 - \gamma_i(m)} = x_i(m).$$

To state the conditions let in addition

$$m^c := \Big\{y \in \{l, h\}|y \neq m\Big\}.$$

Applying the above quoted theorem of Border (2007) to this and reformulating everything in terms of z_i allows us to conclude that $\mathcal{X}$ can be implemented via $z_i \geq 0$ if and only if the following conditions are satisfied:

- $\forall \{m, n\} \in \{h, l\}$:
$$p(m)z_i(m) + p(n)z_{-i}(n) \leq \qquad (EPI)$$
$$1 - Pr(\Gamma) - (1 - \gamma(m^c, n^c))p(m^c)p(n^c)$$
- $\forall m \in \{h, l\}$ and $i = 1, 2$:
$$z_i(m) \leq 1 - \gamma_i(n) \qquad (IF)$$

- $\forall i = 1, 2$

$$z_i(l)p(l) + z_i(h)p(h) \leq 1 - Pr(\Gamma) \tag{BC_2}$$

$$\sum_{i\in\{1,2\}} \sum_{k\in\{l,k\}} p(k)z_i(k) \leq 1 - Pr(\Gamma) \tag{AF}$$

- $\forall \{m, n\} \in \{h, l\}^2$ and $i = 1, 2$:

$$\sum_{k\in\{l,h\}} p_i(k)z_i(k) + pz_{-i}(n) \leq 1 - Pr(\Gamma). \tag{BC_4}$$

Note that in our setup equation (IF) implies (BC_2) and equation (AF) which implies (BC_4). For the second claim, recall (IC_i^l), that is

$$\gamma_i(h)U_i(h|l) + z_i(h) \leq \gamma_i(l)U_i(l|l) + z_i(l).$$

Hence,

$$z_i(h) \leq \gamma_i(l)U_i(l|l) - \gamma_i(h)U_i(l|h) + z_i(l) \leq (\gamma_i(l) - \gamma_i(h))U_i(l|l) + z_i(l), \tag{9}$$

where the last equality follows from Lemma 2.

If $\gamma_i(l) \geq \gamma_i(h)$ and $z_i(l) \leq 1 - \gamma_i(l)$ we can rewrite (9) to

$$z_i(h) \leq (\gamma_i(l) - \gamma_i(h))U_i(l|l) + z_i(l) \leq 1 - \gamma_i(h),$$

which indeed is equation (IF). □

C.4.3 Proof of Lemma 10

Proof. We proof this by contradiction. Suppose there exists a feasible $\mathcal{X}$ that forms an optimal mediation protocol without (IC_i^h) binding for some i. That is, without loss of generality assume that for player 1 it holds that

$$z_1(h) - z_1(l) > \gamma_1(h)U_1(h|h) - \gamma_1(l)U_1(h|l).$$

Recall that

$$z_1(h) = p(1 - \gamma(h, l))x_1(h, l) + (1 - p)\gamma(h, h)x_1(h, h),$$

but then if $\mathcal{X}$ was feasible before, it remains feasible if we reduce $x_1(h, l)$ such that (IC_1^h) holds with equality. Changing this has no effect on the right hand side of the inequality and (IC_i^l) gets relaxed as it is

$$z_1(h) - z_1(l) \leq \gamma_1(h)U_1(l|h) - \gamma_1(l)U_1(l|l)$$

Similarly, suppose (PC_i^l) is not binding, then

$$z_i(l) > V_i(l) - \gamma_i(l)U_i(l|l).$$

Provided that $z_i(l) > 0$ the mediator could react, by changing $z_i(l)$ such that the participation constraint is binding. Then, she can reduce $z_1(h)$ such that the high-cost types incentive constraint is binding which leads to another $\mathcal{X}$ with both (PC_i^l) and (IC_i^h) binding that is feasible and delivers the same value to the objective.

If $z_i(l) = 0$, this procedure is not possible, but then the mediator could use the homogeneity of degree 1 of $\gamma_i(k)$ and the homogeneity of degree 0 w.r.t. G to satisfy (PC_i^l) by multiplying all elements of G by $\alpha < 1$. Again, if $z_i(h) > 0$ the increase in in $z_i(h)$ can always be off-set by reducing X_i appropriately which is always possible. If $z_i(h)$ is indeed 0, then multiplying G by α has if at all only a positive effect on incentive compatibility. Thus, it is without loss of generality to assume that (PC_i^l) holds indeed. $\square$

C.4.4 Proof of Lemma 11

Proof. Recall that the elements of P can be rewritten such that e.g. the probability of meeting player $1l$, given a report $m_2 = l$ is

$$p_1(l|l) = \frac{p\gamma(l,l)}{p\gamma(l,l) + (1-p)\gamma(h,l)}. \tag{10}$$

As $p_1(l|l) > 0$ which is guaranteed by $\gamma(l,l) > 0$ the probability representation for $\gamma(h,l)$ follows immediately, that is

$$\gamma(h,l) = \frac{1 - p_1(l|l)}{p_1(l|l)} \frac{p}{1-p} \gamma(l,l).$$

Repeating the same exercise for any $\gamma(k,m)$ yields the desired representation.

The last equation of (C) can be obtained noticing that given we have established all other results from (C) and using the homogeneity of degree 0 of P w.r.t G we can rewrite G as

$$G = \gamma(l,l)G' = \gamma(l,l) \begin{pmatrix} 1 & q_2(l) \\ q_1(l) & q_2(h)q_1(l) \end{pmatrix}.$$

We know that G' induces the same P as G in particular we know that

$$p_1(l|h) = \frac{p\gamma(h,l)}{p\gamma(h,l) + (1-p)\gamma(h,h)} = \frac{pq_2(l)}{pq_2(l) + (1-p)q_2(h)q_1(l)}$$

which after rearranging yields the desired

$$q_1(l)q_2(h) = q_1(h)q_2(l). \tag{C}$$

As all we have done have been rearrangements, the converse holds as well, that is, for a given P and $\gamma(l,l) > 0$ that satisfy equation (C) we can establish a feasible G such that P and $\gamma(l,l)$ is induced by G. $\square$

C.5 Proof of Theorem 1 (together with Lemmas 4 to 7)

We proof the proposition in several steps. In line with the text, we first solve the "unconstrained problem" $(P1')$ which is also the proof of Lemma 5.[17] After that we introduce

[17] Recall that "unconstrained" refers to $(P1')$ which includes all constraints that bind at all points already in the problem definition.

(IC_i^l) and proof Lemma 6 before finally introducing the remaining constraints with the proof of Lemma 7. Throughout this proof we make use of the following lemma

Lemma 12. *At any optimum of $(P1')$, the monotonicity condition (M) is always satisfied.*

The proof of this lemma can be found at the end of the appendix as it is neither constructive nor relevant to understand the main argument. However, with help of this lemma, we can restrict the choice set of the mediator to the set of induced beliefs that result in monotonic equilibria as discussed in Lemma 1.

C.5.1 Proof of Lemma 4

Proof. Rewrite $p_2(l|h)$ with help of Lemma 11

$$p_2(l|h) = \frac{\Big(1 - p_1(l|l)\Big)\, p_2(l|l)p_1(l|h)}{p_2(l|l)p_1(l|h) - p_1(l|l)\Big(1 - p_2(l|l) - p_1(l|h)\Big)}. \tag{11}$$

$p_2(l|h) > p_2(l|l)$ if equation (11) divided by $p_2(l|l)$ is larger 1 that is

$$\frac{p_2(l|h)}{p_2(l|l)} = (1 - p_1(l|l))\frac{p_1(l|h)}{p_2(l|l)p_1(l|h) - p_1(l|l)\Big(1 - p_2(l|l) - p_1(l|h)\Big)} > 1.$$

Rewriting yields,

$$\begin{aligned} & (1 - p_1(l|l))p_1(l|h) > p_2(l|l)p_1(l|h) - p_1(l|l)(1 - p_2(l|l) - p_1(l|h)) \\ \Leftrightarrow \quad & p_1(l|h) - p_1(l|l) > (p_1(l|h) - p_1(l|l))p_2(l|l), \end{aligned}$$

which holds if and only if $p_1(l|h) > p_1(l|l)$. □

C.5.2 Proof of Lemma 5

Proof. Notice that the unconstrained problem is $(P1')$ is a problem of three elements $P = (p_1(l|l), p_2(l|l), p_1(l|h)$ only, as the fourth is directly defined via consistency equation (C). We calculate the unconstrained optimum in several steps. First, we show that at the optimum the objective is not differentiable with respect to at least one of the three choices variables. Second, we show that if $p_1(l|l)$ is either $p_1(l|h)$ or $p_2(l|l)$m then it is $p_1(l|l) = p_1(l|h)$ and calculate this optimum. Finally, we show that a deviation to $p_1(l|l) = 1$ is not optimal.

Step 1: No optimum in the differentiable interior exists. To proof this claim we are going to proof that the objective $Obj(P) := R(P)v/(Q(P) - R(P))$ is locally concave at any critical point in $p_1(l|l)$ in what we call the "differentiable interior", meaning that such a critical point is in fact a local maximum in $p_1(l|l)$, which is sufficient to proof the claim. Let us begin with defining the differentiable interior.

Definition 3 (Differentiable Interior)**.** The differentiable interior of problem (P1') is the set of all P such that for each $1 > p(k|m) > 0$ the left-derivative and the right-derivative of $Obj(P)$ with respect to all variables coincides.

Next, for the ease of notation define $\boldsymbol{\rho} = (\rho_1(l), \rho_2(l), \rho_1(h)) := \Big(p_1^*(l|l), p_2^*(l|l), p_1^*(l|h)\Big)$

Step 1a: Transform $R(P)$ and $Q(P)$.

Observe that

$$R(P) = \frac{Pr(\Gamma)}{\gamma(l,l)} = \frac{p^2}{\rho_1(l)\rho_2(l)\rho_1(h)} \left(\rho_1(l)(1-\rho_2(l)) + \rho_2(l)\rho_1(h)\right).$$

Defining the function

$$\tilde{Y} := Y * \frac{\rho_1(l)\rho_2(l)\rho_1(h)}{p^2}$$

allows us to rewrite (dropping the argument to simplify notation)

$$Obj = \frac{\tilde{R}v}{\tilde{Q} - \tilde{R}}.$$

Notice that $\tilde{R}$ is linear in any variable of $\boldsymbol{\rho}$. **Step 1b: Define necessary conditions for an optimal interior point.** Suppose $(P1')$ has indeed an optimal point in the differentiable interior. Then a necessary condition on this point is that it is indeed a critical point in all three variables, that is

$$Obj'(\rho) := \frac{\partial Obj(\rho)}{\partial \rho} = \underbrace{\frac{\nu}{(\tilde{Q}(\rho) - \tilde{R}(\rho))^2}}_{=:f(\rho)} \quad \underbrace{\left(\tilde{R}'(\rho)\tilde{Q}(\rho) - \tilde{Q}'(\rho)\tilde{R}(\rho)\right)}_{=:g(\rho)} = 0 \tag{FOC}$$

for every $\rho \in \boldsymbol{\rho}$. Noticing that $f(\rho) \neq 0$ for any ρ by definition, the necessary first order condition boils down to $g(\rho) = 0$. Another necessary condition for a local minimum is that any critical point in any ρ is not locally concave in this variable. If it was locally concave in any ρ this means that we are at a local maximum in this variable ρ and that the second order conditions for a minimum are never fulfilled. Formally, this means that at any critical point ρ^{cp} it needs to hold that

$$Obj''(\rho^{cp}) = \underbrace{f'(\rho^{cp})g(\rho^{cp})}_{=0 \text{ by equation (FOC)}} + f(\rho^{cp})g'(\rho^{cp}) \geq 0 \tag{12}$$

for every $\rho^{cp} \in \boldsymbol{\rho^{cp}}$. The first term is 0 by the standard envelope argument, such that (12) boils down to

$$Obj''(\rho^{cp}) = f(\rho^{cp})g'(\rho^{cp}) = f(\rho^{cp})\left(\tilde{R}''(\rho^{cp})\tilde{Q}(\rho^{cp}) - \tilde{R}(\rho^{cp})\tilde{Q}''(\rho^{cp})\right) \geq 0.$$

By the linearity of $\tilde{R}$ and the obeservation that $\tilde{R} \geq 0$ by construction, a necessary and sufficient condition for (12) to hold is simply

$$\tilde{Q}''(\rho^{cp}) \leq 0 \tag{13}$$

for every $\rho^{cp} \in \boldsymbol{\rho^{cp}}$. **Step 1c: Show that the necessary conditions never hold for $\rho_1(l)$.** To complete the claim of step 1 we are now going to show, that $\tilde{Q}(\rho_1(l))$ is indeed a convex function.

To see this observe first by plugging in we can reduce $\gamma_i(l) = \gamma(l,l)p/\rho_{-i}(l)$ which in turn means that while $\tilde{\gamma}_2(l)$ is constant in $\rho_1(l)$, $\tilde{\gamma}_2(l)$ is linearly increasing in $\rho_1(l)$. In addition, we do not need to worry about $\gamma_2(h)$ as player $2h$ has no expected utility by Corollary 1. Further we can rewrite using Corollary 1 and Lemma 11

$$\tilde{\gamma}_1(h)U_1(h|h) = \frac{\gamma(l,l)}{1-p}(1-\rho_1(h))\rho_1(h)\ (\rho_1(l) - \rho_2(l))\frac{(\kappa-1)}{\kappa}$$

which is linearly increasing in $\rho_1(l)$ and positive. Rewriting yields

$$\gamma(l,l)\tilde{Q} = \sum_i \tilde{\gamma}_i(l)\left(U_i(l|l) - (1-p)U_i(h|l)\right) + \tilde{\gamma}_1(h)(1-p)U_1(h|h)$$

it suffices to show that

$$h_i(\rho_1(l)) = \tilde{\gamma}_i(l)\Big(U_i(l|l) - (1-p)U_i(h|l)\Big)$$

is convex for every i.

For h_2, observe that by Lemmas 2 and 3, player $2h$ only gains from deviating if $p_1(h|l) > p_1(h|h)$. In such a case player $2h$, best post-deviation strategy is to play $\Delta_{h,h}$ with probability 1, which yields utility

$$U_2(h|l) = p_1(h|l) - \Delta_{h,h}\kappa c_l. \tag{14}$$

Bidding the same on-path is in the best response set of player 2 yielding

$$U_2(h|h) = p_1(h|h) - \Delta_{h,h}\kappa c_l = 0. \tag{15}$$

Subtracting equation (15) from equation (14) yields

$$p_1(h|l) - p_1(h|h) = \rho_1(h) - \rho_1(l) = U_2(h|l) \tag{16}$$

and thus $U_2(h|l)$ is linear in $\rho_1(l)$. As $\tilde{\gamma}_2(l)$ is constant in $\rho_1(l)$, $h_2(\rho_1(l))$ is convex if and only if $U_2(l|l)$ is convex in $\rho_1(l)$ which can easily be verified by the utilities derived in (U). The last step is now to show that $h_1(\rho_1(l))$ is convex as well.

To see this, observe first that whenever deviation is profitable for player 1, type h, she would deviate by playing $\Delta_{h,h}$. But, $\Delta_{h,h}$ is in fact the lower bound of player 1, type l and thus in such a case we can rewrite

$$U_1\Big(h|l, p_2(h|l) > p_2(h|h)\Big) = U_1(l|l) + (1-\kappa)c_l\Delta_{h,h}.$$

As $\tilde{\gamma}_1(l) = \rho_1(l)\rho_1(h)p$ we can use the expression derived in Corollary 1 to establish that $\tilde{\gamma}_1(l)U_1(l|l)$ is linear in $\rho_1(l)$ and thus convex.

What remains is to show that $-\rho_1(l)\Delta_{h,h}$ is weakly convex. This can be established using that $\Delta_{h,h} = p_1(h|h)/\kappa c_l$ which is independent of $\rho_1(l)$ which proofs the claim.

Step 2: $\rho_1(l) \in \{\rho_2(l), \rho_1(h)\}$.

By assumption $\rho_1(l) \leq \rho_2(l)$ is ruled out. Second, fix some $\rho_2(l)$ and $\rho_1(h)$. If $\rho_1(l) \in [\rho_2(l), \rho_1(h)]$ then $Obj(\rho_1(l) = 1) > Obj(\rho_1(l) = \rho_1(h))$. Further we know that Obj is continuously differentiable on $\rho_1(l) \in (\max\{\rho_2(l), \rho_1(l)\}, 1)$. By Step 1 we know that every interior point is a maximum in $\rho_1(l)$.

Next, notice by Lemma 4 that for $\rho_1(l) > \rho_1(h) \Rightarrow \rho_2(l) > \rho_2(h) \Rightarrow p_i(h|h) > p_i(h|l) \Rightarrow U_2(h|l) = 0$ and $U_1(h|l) < U_1(h|h)$.

Now, notice that $\rho_1(l) = 1$ can only be optimal if Obj is (LHS-)decreasing at $\rho_1(l) = 1$ as there cannot be a local minimum in $\rho_1(l)$ by Step 1. To check this it suffices to look at the sign determining function of the derivative which is, by Step 1, $R'Q - Q'R$. Solving this for $\rho_1(l) > \rho_1(h)$ yields a quadratic function in $\rho_1(l)$.

The sign-determining function at $\rho_1(l) = 1$ is quadratic in $\rho_1(h)$, i.e. a condition

$$a\rho_1(h)^2 + b\rho_1(h) + c < 0 \tag{17}$$

where

$$a = (\kappa - 1 + \rho_2(l)^2) \tag{18}$$

$$b = 1 + 2\rho_2(l) - 2\Big(\rho_2(l)\Big)^2 + p(1-\kappa) \tag{19}$$

$$c = \Big(\rho_2(l)\Big)^2\kappa - \rho_2(l)\Big((\kappa-1)(1-p) + \kappa\Big) + (\kappa-1)(1-p). \tag{20}$$

Note first, that (17) is decreasing in $\rho_2(l)$, second note that for $\rho_2(l) = \rho_1(h)$ condition (17) becomes

$$(\kappa-1)\Big(1 - p - 2\rho_1(h)\Big) + \Big(\rho_1(h)\Big)^4 - 2\Big(\rho_1(h)\Big)^3 + \Big(\rho_1(h)\Big)^2(1+2\kappa) < 0. \tag{21}$$

Note that this is minimal if κ is minimal and p is maximal. Therefore, it must hold that

$$1/2 + \underbrace{\rho_1(h))^2 \underbrace{\Big((\rho_1(h))^2 - 2\rho_1(h) + 5\Big)}_{>4} - 2\rho_1(h)}_{>-1/4} < 0,$$

a contradiction. Thus, whenever $\rho_1(h) \geq \rho_2(l)$, choosing $\rho_1(l) = 1$ is not preferred to $\rho_1(l) = \rho_1(h)$. Solving the first order conditions given $\rho_1(l) = 1$ for $0 < \rho_1(h) < \rho_2(l)$ yields that no critical point in both variables exists and therefore no interior solution. As Obj is decreasing at $\rho_2(h) = 0$, there cannot be any solution with $\rho_1(l) = 1$. Thus, $\rho_1(l)$ must either be equal to $\rho_1(h)$ or to $\rho_2(l)$.

Step 3: Calculate the optimum if $\rho_1(l) \in \{\rho_2(l), \rho_1(h)\}$. By Step 1, we know that if $\rho_1(l) \in [\rho_2(l), \rho_1(h)]$ the optimum involves ρ_1 being equal to either of the bounds.

Therefore, we only need to consider the two cases for any $\rho_2(l)$ and $\rho_1(h)$.

Step 3a: The equilibrium for $\rho_1(l) = \rho_2(l)$. First, consider $\rho(l) = \rho_1(l) = \rho_2(l)$. By Lemma 11, $\rho_1(h) = \rho_2(h) = \rho(h)$.

All payoffs are symmetric and, by Corollary 1, $U_i(h|h) = 0$ and, by Lemma 2, $U_i(h|l) = \max\{0, \rho(h) - \rho(l))$. Finally, $U_i(l|l) = (\kappa-1)/\kappa + (\rho(h) - \rho(l)\kappa)/\kappa$.

In addition, $\gamma_1(l) = \gamma_(h) = \gamma_(l) = p/\rho(l)$ and therefore

$$\tilde{Q} = \frac{2\rho(l)\rho(h)}{p}(U_i(l|l) - (1-p)U_i(h|l)).$$

Finally, as $\tilde{R} = \rho(l)(1 - \rho(l) + \rho(h))$ we can simplify Obj to

$$Obj(\rho(l), \rho(h)) = \frac{p(1 - \rho(l) + \rho(h))}{\underbrace{2\rho(h)(U_i(l|l) - (1-p)U_i(h|l))}_{=:\hat{Q}} - \underbrace{p(1 - \rho(l) + \rho(h))}_{=:\hat{R}}}.$$

Employing the same technique as in Step 1, we know, as $\hat{R}$ is linear in both $\rho(k)$ any interior solution needs to have that $\hat{Q}$ is concave in $\rho(k)$.

Notice that the second derivative of $\hat{Q}$ when $U_i(h|l) = 0$ boils down to $4/\kappa$ as $U_i(l|l)$ is linearly increasing with factor $1/\kappa$ in $\rho(h)$. Thus, any solution with $\rho(l) \geq \rho(h)$ can be ruled out.

Second whenever $\rho(l) < \rho(h)$ observe that $\hat{Q}$ is linearly decreasing in $\rho(l)$ with factor $2\rho(h)p$. Hence, the sign determining function of the first derivative $\hat{R}'\hat{Q} - \hat{Q}'\hat{R}$ becomes

$$\hat{R}'(\rho(l))\hat{Q} - \hat{Q}'(\hat{(}l))\hat{R}|_{\rho(l)<\rho(h)} = -2\rho(h)p\Big(\big(U_i(l|l) - U_i(h|l)\big) - \hat{R}\Big). \tag{22}$$

Note that by construction Obj defines a probability and is thus in $[0, 1]$. Whenever equation (22)=0, then $\hat{Q} - \hat{R} = (2\rho(h)p - 1)\hat{R}$ which can only be positive if $2\rho(h)p = 1$. As $p < 1/2$ this condition never holds. Therefore, we do not find an interior solution when $\rho(h) > \rho(l)$.

What remains are then boundary solutions with either of the $\rho(k) \in \{0, 1\}$.

If $\rho(h) = 1$ we need to go back to the original Q and R as our modifications are not valid if $\rho_i(k) \neq (0, 1)$.

This is for $\rho(h) = 1$

$$R = p^2 \frac{2 - \rho(l)}{\rho(l)}$$

$$Q = p^2 \frac{2(1 - \rho(l))}{\rho(l)},$$

which obviously violates $Q > R$ and is thus not feasible. $\rho(l) = 0$ would violate monotonicity and is ruled out by Lemma 12.

Step 3b: The equilibrium for $\rho_1(l) = \rho_1(h)$. It remains to show that an equilibrium exists in which $\rho_1(l) = \rho_1(h) = \rho_1$. Note that again by consistency in Lemma 11 we get $\rho_2 = \rho_2(l) = \rho_2(h)$.

With this, we know that $U_i(k|m) = U_i(k|k)$ for every i and k and $U_2(h|l) = 0$, $U_i(l|l) = (1 - \rho_2)\frac{\kappa-1}{\kappa}$, and $U_1(h|h) = U_1(h|l) = (\rho_1 - \rho_2)\frac{\kappa-1}{\kappa}$. As $\tilde{\gamma}_i(l) = p/\rho_{-i}$ we get

$$\tilde{Q} = \frac{1}{\kappa p}\rho_1(\kappa - 1)\Big((\rho_1)^2 - \rho_1(1 + p) - \rho_2(1 - \rho_2 - p)\Big)$$

and

$$\tilde{R} = \rho_1.$$

Note that this means that for an optimum in ρ_1 we need $\tilde{Q} = \rho_1 \tilde{Q}'(\rho_1)$ and for an optimum in ρ_2 we would need $\tilde{Q}'(\rho_2) = 0$.

Notice that

$$\tilde{Q}'(\rho_2) = \frac{\rho_1(\kappa - 1)}{\kappa p}(1 - p - 2\rho_2)$$
$$\Rightarrow \quad \rho_1 \tilde{Q}'(\rho_1) - \tilde{Q} = \frac{(\rho_1)^2(\kappa - 1)}{\kappa p}, (1 + p - 2\rho_1)$$

and thus we arrive at the desired results. Checking second order conditions in each variable yield that the function is convex in both arguments. As cross derivatives are 0 at the optimum, the critical point is a minimum by the second order derivative test. □

C.5.3 Proof of Lemma 6

Proof. **Step 1: The unconstrained optimum satisfies (IC_i^l) for $p \leq 1/3$.** As $U_i(l|h) = U_i(l|l)$ by Lemma 3 and with the help of Lemma 10 stating that (IC_i^h) binds, we can rewrite (IC_i^l)

$$(\gamma_i(l) - \gamma_i(h))U_i(l|l) \geq (\gamma_i(l) - \gamma_i(h))U_i(h|h). \tag{23}$$

As $U_i(l|l) \geq U_i(l|h)$ by construction this holds if and only if $(\gamma_i(l) - \gamma_i(h)) > 0$.

Calculating the difference yields

$$\gamma_i(l) - \gamma_i(h) = \frac{p}{\rho_{-i}} \frac{\rho_i - p}{1 - p} \tag{24}$$

which is positive if and only if $\rho_i \geq p$.

Recall from Lemma 5 that the optimal unconstrained $\rho_2 = \frac{1-p}{2}$ which is larger p if and only if $p < 1/3$.

Step 2: Describe the equilibrium including (IC_i^l) for $p > 1/3$.

Step 2a: No solution with $\rho_1(l) > \rho_1(h)$. First, we show that we do not want to deviate to any $\rho_1(l) > \rho_1(h)$ for $p > 1/3$. To do so, consider (IC_2^l). By Lemma 3 the RHS remains at 0, and $U_2(l|h) > U_2(l|l)$. Thus, for (IC_2^l) to hold we would still need that $\gamma_2(l) \geq \gamma_2(h)$. However, then also $\tilde{\gamma}_2(l) - \tilde{\gamma}_2(h)$ needs to be positive. Plugging in and simplifying, we find that

$$\tilde{\gamma}_2(l) - \tilde{\gamma}_2(h) = \rho_1(h)\rho_2(l)(1 - p) - \rho_1(l)p^2(1 - \rho_2(l)) \tag{25}$$

which is decreasing in $\rho_1(l)$. Hence, no deviation to $\rho_1(l) > \rho_1(h)$ is profitable since whenever IC holds for this deviation, it also holds for $\rho_i(l) = \rho_i(h)$ which is preferred by Lemma 5.

Step 2b: The proposed solution is indeed an optimum. Next, we need to show that also no deviation to $\rho_1(l) < \rho_1(h)$ is optimal. For this we use a guess and verify approach to show that the proposed equilibrium with $\rho_2 = p$ is indeed an optimum.

To do this, this solution needs to satisfy the first order conditions of the Lagrangian at the proposed point. As we know from Step 2a we do not need to consider $\rho_1(l) > \rho_1(h)$. Define $g(\boldsymbol{\rho}) \leq 0$ to be the incentive constraint, reformulated such that if $g \leq 0$, (IC_2^l) holds.[18]

[18] As $\rho_1 \geq p$ at the imposed constrained optimum, we do not worry about (IC_1^l) which always has slack.

The Lagrangian is given by

$$\mathcal{L}(\lambda, \mu, \boldsymbol{\rho}) = Obj(\boldsymbol{\rho}) + \lambda g \boldsymbol{\rho}) + \mu(\rho_1(l) - \rho_1(h)). \tag{26}$$

Any solution to the constrained minimization problem $\boldsymbol{\rho}^*$ must be such that it solves the following problem

$$\min \mathcal{L}(\cdot) \tag{27}$$

and

$$\lambda, \mu \geq 0. \tag{28}$$

It turns out that the proposed solution is such a point and further $\mathcal{L}$ is strictly concave at this point, thus the problem is indeed locally minimized at $\boldsymbol{\rho}^*$.

Step 2c: Show that no other solution exists. It is not clear whether the problem is also globally minimized at this point, as both the objective as well as the constraint do not satisfy the usual assumption needed for global optimality, in particular they are not globally convex. However, fixing k we know the following two aspects:

(a) at $p = 1/3$ the solution is the same as the "unconstrained" optimum considered in Lemma 10. For $p > 1/3$ the solution is worse than the unconstrained optimum,

(b) as all functions are continuous in p the functional value and thus the equilibrium value must be continuous in p.

This means that if another solution (strictly better than the candidate) exists for some $\hat{p} > 1/3$ then there also must exist some $\check{p} \in [1/3, \hat{p}]$ such that the equilibrium values $\hat{\boldsymbol{\rho}}$ of $\hat{p}$ as a function of p yield the same outcome as the proposed equilibrium.

Further, as $\mathcal{L}$ is strictly convex at the proposed optimum, this alternative value $\hat{\boldsymbol{\rho}}$ must be bounded away in at least one of its variables.

Suppose the other optimal point is at some $\rho_i(k)$ not in the neighborhood of ρ_i^*. Then by continuity, the mean value theorem, and the strict convexity of $\mathcal{L}$ at the proposed point this point can only be optimal if the derivative of Obj w.r.t. $\rho_i(k)$ is 0 at some point on $(\rho_i(k), \rho_i^*)$.

As $\rho_1(l)$ has no extreme value on the interval $(\rho_2(l), \rho_1(h))$ by Step 1 in appendix C.4.3, $\rho_1(l)$ must be the same in both optima.

But then, if $\rho_1(l)$ is constant, $\rho_1(h)$ is increasing on $(a, 1)$. Then again $\rho_1(h) = 1$ cannot be optimal. Thus, no other minimum exists and our proposed minimum is the only and therefore global minimum. $\square$

C.5.4 Proof of Lemma 7

Proof. Finally, introducing $\gamma(l,l) \leq 1$ to the problem it is straightforward to compute that the constraint has slack for any $p \geq 1/3$.

Also, by computing $\nu/(Q(P) - R(P))$ one can verify that it holds at ρ_1^*, ρ_2^* whenever

$$k \leq \frac{2 - 4p - 2p^2}{1 - 4p + 3p^2}.$$

Further, if the constraint $\gamma(l,l) \leq 1$ binds, we can use Lemma 11 to see that $\gamma(l,h), \gamma(h,l) \leq 1$ if and only if $\rho_i(l) \geq p$.

We know that at the unconstrained optimum with $\rho_1(l) = \rho_1(h)$ and thus, we have a boundary solution in those variables for a given $\rho_2(l)$. However, the solution with respect to $\rho_2(l)$ is such that $Obj'(\rho_2(l)) = 0$.

In addition we know by strict concavity that in fact the regime change happening at $\rho_1(l) = \rho_1(h)$ (from high-cost types having a beneficial deviation payoff to low cost types having one), must be such that around the unconstrained optimum we would not change the equation $\rho_1(l) = \rho_1(h)$ as this would either provide us with a free lunch lowering $\rho_1(h)$ to put slack on $\gamma(l,l) \leq 1$. Then, as we change the regime to $\rho_1(l) > \rho_1(h)$ it must be that $Obj'(\rho_1(l)) > 0$ as we started at the optimum. Thus, we could lower $\rho_1(l)$ at no cost on the constraint to $\rho_1(l) = \rho_1(h)$ as the constraint can be rewritten as

$$\nu/(Q(P) - R(P)) - 1 = Obj - R \leq 0,$$

and $R|_{\rho_1(l)=\rho_1(h)} = p^2/\rho_1(l)\rho_2(l)$.

As $\rho_1(l) = \rho_1(h)$ remains to hold the problem

$$\min_{\rho_1,\rho_2} Obj$$

s.t. $\rho_2 \geq p$ and $\gamma(l,l) \leq 1$ is well-behaved such that we get the desired solution of the lemma.

Finally, plugging the solution for every regime into the Border constraints (EPI) and (IF) shows that they hold at the optimum. □

C.6 Proof of Proposition 2

Proof. For the first result, observe that for $p < 1/3$, the solution at which $\rho_2 = p$ and $\rho_1 = 2p+1/(\kappa-1)$ is always feasible and in line with $\gamma(l,l) \leq 1$ and (IC_i^l). The corresponding probability of contest is given as

$$Pr(\Gamma, \boldsymbol{\rho}^*) = \frac{(\kappa+1)p}{1+2(\kappa-1)p}, \tag{29}$$

which is increasing in p and κ and becomes $1/2$ for $p = 1/3$ and $\kappa \to \infty$.

Second, the optimal probability of a contest for $p > 1/3$ is

$$4p\frac{\kappa p - (1-p)(\kappa-2)}{(\kappa-1)(7p^2-2p-1)+4p}, \tag{30}$$

which is falling in p for $p > 1/3$. Thus, it suffices to look at the probability at $p = 1/3$. But at this point it becomes

$$\frac{\kappa-4}{2\kappa-5}, \tag{31}$$

which again is bounded by $1/2$.

The inverse u-shape follows from $Pr(\Gamma, \boldsymbol{\rho}^*)$ being concave on all intervals and that the derivative and smooth pasting at p', p'', p'''.

Finally, monotonicity (and concavity) in κ follows from monotonicity and concavity in κ for all regions as well as smooth pasting at the transition of the regions. □

C.7 Proof of Proposition 3

Proof. The results for the probability of being send to contest follow immediately from the ex-ante symmetry and the equilibrium beliefs specified in Theorem 1 and 7.

The result on the expected share follows from Theorem 1 and Lemma 1. The low-cost types expected utility from contest is weakly below her outside option V. In order to fulfill the participation constraint in expectations, the player needs to be compensated by a higher share if mediation fails. As player $1l$ has a higher probability to enter the contest, she also needs to receive a higher share than player $2l$. A weakly higher share for any l-type compared to the same player's corresponding h-type follows from h-types binding incentive compatibility. Finally, as player $1h$ gains a positive expected utility in case of the contest her expected share can be pushed down the most completing the proof. $\square$

C.8 Proof of Proposition 4

Proof. The expected legal expenditure of player ik is by the uniform equilibrium scoring functions given by

$$E[LE_i^k] = \sum_{r \in \{b,m,t\}} Prob(s_i^k \in r)\frac{\underline{r} + \overline{r}}{2}$$

where b, m and t are the scoring ranges used in Figure 1 and the proof of Lemma 1. Further, $\underline{r}$ denotes the upper bound of range r and $\underline{r}$ denotes the lower bound of range r.

The expected scoring function of player 1 entirely depends on ρ_2, that is

$$\begin{aligned}
\rho_1 E[LE_1^l] + (1-\rho_1)E[LE_1^h] &= \rho_1 \frac{\rho_1(2-\rho_1) + (\rho_2)^2(\kappa-1)}{2\rho_1 c_l \kappa} + (1-\rho_1)\frac{(1-\rho_1)}{2c_l\kappa} \\
&= \frac{1+\rho_2(\kappa-1)}{c_l\kappa}.
\end{aligned}$$

Thus, the equilibrium expected contest score of player 1 is the same as in a contest without mediation whenever $\rho_2 = p$.

The expected score of player 2 is computed in a similar manner but depends on both ρ_1 and ρ_2. It is given by:

$$\frac{1}{2c_l\kappa}\left(\frac{(\kappa-1)}{\kappa}\left(\rho_1(\rho_1-2) + (\rho_2)^2(\kappa-1) + 2\rho_2\right) + 1\right).$$

The derivative of this function w.r.t. to ρ_1

$$\frac{\kappa-1}{\kappa^2}(\rho_1 - 1) < 0.$$

As $\rho_1 > p$ by Lemma 7 and $\rho_2 = p$ for $p \notin (p', 1/3)$, it follows that total legal expenditures post-mediation are indeed smaller than under the prior belief p. $\square$

C.9 Proof of Proposition 5

Proof. As participation is optimal by lemma 8 and the optimal mechanism is unique, no bargaining protocol can achieve a better result than Theorem 1. By convexity of contest utilities in beliefs, no Bayes plausible signal structure over the prior can make the receiver worse-off than the prior. Thus, the participation constraint of the mechanism holds in the bargaining game as well.

To show that take-it-or-leave-it bargaining performs worse in environments that satisfy 1 we show that the low-cost type of Sender has always an incentive to deviate to some offer $0 < \epsilon < 1 - V(l)$ that yields a utility higher than $V(l)$ which is her on-path utility. We do so by considering the possible response of Receiver to such an offer given any off-path β_S describing the probability assessment of Receiver on Sender in the contest game.

Any Receiver type accepts. As $\epsilon < 1 - V(l)$, Sender earns a utility larger $V(l)$.

Any Receiver type rejects. The high-type only rejects an offer of ϵ if she expects a utility $U_R(h|\beta_S) > \epsilon$, given her off-path belief β_S. By Lemma 1 $U_R(h|\beta_S, \beta_R) > 0$ only if $\beta_S < \beta_R$. Since any Receiver type rejects the offer, the belief on the receiver is the same as the prior, that is $\beta_R = p$. But $\beta_S < p$ implies via lemma 1 that $U_S(l|\beta_S, \beta_R) > V(l)$.

h-type Receiver rejects and l-type Receiver accepts. This case doesn't exist, as any offer that the h-type rejects is also rejected by the l-type as PBE requires type-independent beliefs after the deviation (Fudenberg and Tirole, 1988) and l-types have lower cost of evidence provision.

l-type Receiver rejects and h-type Receiver accepts. h-types only accept if $\epsilon \geq U_R(h|\beta_S, \beta_R)$ that is

$$\epsilon \geq (\beta_R - \beta_S)\frac{\kappa - 1}{\kappa}.$$

If Receiver h, type l rejects, then Sender, type l gains $(1-p)(1-\epsilon)$ which is larger $V(l) = (1-p)(\kappa-1)/\kappa$ as ϵ goes to 0. Thus Receiver, type h must be indifferent. Rewriting the above equation yields

$$\beta_R = \frac{\epsilon\kappa}{\kappa - 1} + \beta_S.$$

In order to induce a belief of β_R, Receiver, type h must choose to reject the offer with probability

$$\gamma_{R,h} = \frac{p}{1-p}\frac{1-\beta_R}{\beta_R},$$

which follows analogously to Lemma 11.

Plugging this into Sender l-types yields:

$$\begin{aligned}(1-p)(1-\gamma_{R,h})(1-\epsilon) + \Big(p + (1-p)\gamma_{R,h}\Big)(1-\beta_S)\frac{\kappa-1}{\kappa} &= \\ (1-p)(1-\epsilon) + \frac{p}{\beta_R}\left((1-\beta_S)\frac{\kappa-1}{\kappa} - (1-\beta_R)(1-\epsilon)\right).\end{aligned}$$

Taking into account that β_R is a function of β_S this expression is continuous and monotone in β_S. β_S is naturally bounded by 1 and β_R. As we are looking for the lowest utility, we can assign for any $\epsilon > 0$ it suffices to consider an upper and a lower bound. For ϵ close to 0 however, both $\beta_S = \beta_R$ as well as $\beta_S = 1$ yield a utility larger $(1-p)(\kappa-1)/\kappa$. Thus, Sender, type l always has an incentive to deviate to some ϵ irrespective of the out-of-equilibrium beliefs of Receiver resulting in an inferior solution which is actually strict as long as the case is not trivial by the uniqueness of the proposed mechanism as shown in the proof of Theorem 1. □

C.10 Proof of Proposition 6

The proof relies on three features of the model which can be exploited to guarantee a weaker participation constraint:

- the mediator can ex-ante commit to probabilistic private messages she sends to parties following any given message profile (but before the acceptance decision),
- the mediator can ex-ante commit to an additional probabilistic private message she sends to parties following any message and acceptance profile (that is after the acceptance decision),
- yll type profiles lead to on-path to litigation with positive probability.

Proof. For PAIR we need that the expected share given one's own type, that is $x_i(l)$ is larger than the expected utility of a contest that occurs upon rejection of this share. Suppose without loss of generality that an offer of 0 is rejected by all parties and is used by the mediator to trigger litigation.

Two aspects facilitate the analysis: First, the mediator can choose a signal $\sigma(\mathbf{m}, \mathbf{d})$ that depends on the received messages $\mathbf{m}$ as well as on the acceptance decision $\mathbf{d}$ of both players. That is, the mediator has the possibility to define a post-mediation protocol, too.

Recall from Theorem 1 that any type profile leads to litigation with positive probability. At the same time rejection by one party is enough to trigger litigation. Thus, as we allow for private communication, the mediator is free to choose one of the two messages sent to one party if she triggers rejection by the other party. The mediator can therefore randomize not only between who takes the role of player 1, that is which $\mathcal{X}_i$ to use, but also between whom of the two player's receives the "trigger message" 0. For the non-triggering player the mediator can in fact randomizes between all messages the player could receive on-path when the conflict is settled. This way the player does not know whether she is treated as player 1 or player 2 in the mediation protocol at the time of making her decision as to whether to accept or reject the offer. She does in fact not even know whether rejecting the offer makes any difference at all (as the opponent might have received an offer of 0 anyways). By Proposition 3, the mediator can choose $\mathcal{X}_i$ such that for any offer $x_i(k)$ there exists an on-path continuation game in which the player is worse off than $x_i(k)$. Hence, it is possible for the mediator to choose a signal σ_i conditional on deviation that signalling the deviator is in this on-path subgame deterring deviation altogether. □

C.11 Proof of Proposition 7

Proof. Whenever $\gamma_i(k) \neq 1$ the proof is the same as that of Proposition 6.

The situation is however different if either of the players is sent to court with probability $\gamma_i(k) = 1$. According to Theorem 1 and Lemma 7, $\gamma_i(h) < 1$. In addition at most one of the l-types has $\gamma_i(l) = 1$ on path.

This way the player knows that in one of the two mediation protocols she is always going to litigate anyways. Thus if $x_1(l,h) \neq x_1(l,l)$, player 1 might have a strong incentive to deviate as she knows whom she is facing in case her decision is relevant at all. In all other cases she is going to litigate anyways and receives $V(l)$ as litigation payoff by Theorem 1 together with Lemma 1. Thus, it might be optimal for her to reject anything but $x_1(l,l)$.

Suppose instead the mediator announces a mediation protocol $\mathcal{X}_\lambda^\epsilon$ in which reporting two l-types follows mediation breakdown with full information disclosure with probability ϵ and a protocol as that derived in Section 4 otherwise. As $\epsilon \to 0$, the result gets arbitrarily close to that of Theorem 1. However, the mediator can signal any l-type deviator that in fact the low-cost vs. low-cost litigation game is played, causing the l-type to also except ex-post shares. □

C.12 Proof of Lemma 12

Proof. If condition (M) is violated, the equilibrium is no-longer monotonic but instead overlapping strategies might be possible. The reason for this is that if, e.g. $p_1(l|l)\kappa < p_1(l|h)$ the likelihood of meeting a low-cost type when being a high-cost type is too high compared to being a low-cost type, such that the high-cost type has a strong incentive to *overscore* the low-cost type. Further, by the consistency condition equation (C) whenever the high-cost type faces a low-cost type, she faces indeed a low-cost type that *thinks* she herself is facing a high-cost type with very high probability. This provides an incentive for the h-type to compete more aggressive and for the l-type to compete softer than under condition (M). The equilibrium scores in the non-monotonic equilibrium are as depicted in figure 6. Player $1l$ and player $1h$ overlap on the middle interval but are otherwise “close to monotonic”. While the high-cost type of player 2 has a support covering the whole scoring interval, player $2l$ only competes in the middle interval. In addition player $2h$ also has a mass point at 0.

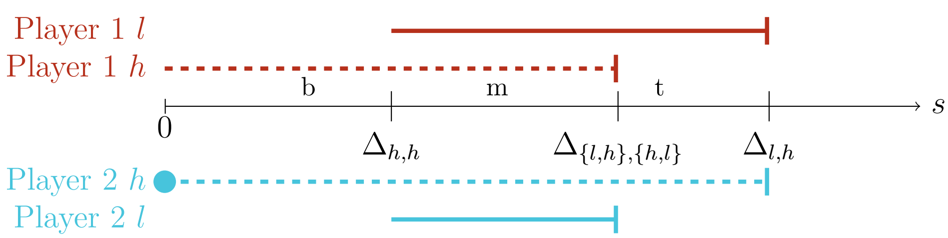

Figure 6: Strategy support of player 1 and 2 if monotonicity fails.

Solving for the optimal mechanism, it turns out that there is still no interior solution in $p_1(l|l)$. The mediator would set $p_1(l|l)$ equal to any discontinuity point or at the respective borders. That is either $p_1(l|l) = 0$ or $p_1(l|l) = \max\{p_2(l|l), p_1(l|h)/\kappa\}$. If $p_1(l|l) = p_2(l|l) = \rho(l)$ under non-monotonicity, the first order condition of the mediator's problem is monotone in $\rho(l)$ and thus, we would need $\rho(l) = 0$ which is never optimal. If $p_1(l|l) = p_1(l|h)/\kappa$ utilities converge to their monotone counterparts and thus, the solution is no different than that for monotonicity. Finally, $p_1(l|l) = 0$ is never optimal as the objective is always decreasing at this point. □