Managing A Conflict: Optimal Alternative Dispute Resolution*

Benjamin Balzer[†] Johannes Schneider[‡]

June 10, 2019

Abstract

We study optimal Alternative Dispute Resolution (ADR). ADR aims to achieve settlement avoiding the cost of evidence provision. Participation is voluntary. Disputants are privately informed about their marginal cost to increase their evidence quality. If ADR fails to settle, disputants use the information obtained during ADR and decide on the quality of evidence they provide in a legal contest. Disputants may exploit ADR as a source of information by participating in ADR only to extract information. Optimal ADR does not allow for information trading, fails to settle with positive probability independent of the type profile, and induces asymmetries between disputants.

1 Introduction

Alternative Dispute Resolution (ADR) is a tool to increase the legal system's efficiency. It aims at settling cases without the need of formal litigation. There is wide consensus that the legal system is overburdened. In 2018 a judge in the U.S. district courts received 575 new cases on average. In addition, there were 685 pending cases per judge. The large caseload leads to a median time from filing to trial of more than 27 months. Litigation requires time and resources from courts. Each case that forgoes litigation therefore has a positive externality on the efficiency of the legal system.

Most jurisdictions encourage parties to engage in ADR before the formal litigation process. The U.S. Alternative Dispute Resolution Act of 1998 states that courts should provide disputants with ADR options in all civil cases. ADR is defined as a "process or procedure, other than an adjudication by a presiding judge, in which a neutral

*Large parts of this paper originate from an older, discontinued paper circulated under the title "Managing a Conflict" by the same authors. We thank João Correia-da-Silva, Andreas Engert, Willie Fuchs, Daniel Garrett, Susanne Goldlücke, Johannes Hörner, Alon Klement, Dan Klerman, Nenad Kos, Aditya Kuvalekar, Yassine Lefouili, Antoine Loeper, Volker Nocke, Martin Peitz, Kathryn Spier, Thomas Tröger, Christoph Wolf, and audiences at seminars and conferences in Augsburg, Bergen, Faro, Geneva, Ithaca, Lisbon, Louvain-la-Neuve, London (UK), London (Ontario), Maastricht, Mannheim, Madrid, Melbourne, Miami, Milan, Montreal, Munich, New York, Oslo, Sydney, Toulouse, Warwick, and Palma for comments and suggestions. Johannes Schneider gratefully acknowledges support from the Ministerio Economia y Competitividad, Maria de Maeztu grant (MDM 2014-0431), and Comunidad de Madrid, MadEco-CM (S2015/HUM-3444).

[†]University of Technology Sydney, benjamin.balzer@uts.edu.au

[‡]Carlos III de Madrid, Department of Economics, jschneid@eco.uc3m.es

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. Each party has the right to enforce litigation.[1] Hence, ADR happens "in the shadow of the court:" if ADR does not achieve settlement, disputants engage in a legal contest.

In this paper we study the design of alternative dispute resolution (ADR). We provide a model in which the information revealed during ADR affects strategic choices in litigation. Any common ADR practice implies some information exchange between the disputants (see Genn, 1998; Anderson and Pi, 2004, for case studies). In addition, there is widespread consensus that litigation strategies are a function of disputants' information (see Spier, 2007, for an overview). Yet, classic approaches to ADR design ignore the strategic relevance of information revelation.

We address the interaction between strategies in ADR and strategies in litigation explicitly. We assume that the fall-back option if settlement fails is a legal contest. In that contest disputants invest in the quality of their evidence. The two-way channel linking ADR and litigation is the information that ADR reveals upon failed settlement. That channel influences disputants' behavior in both ADR and litigation. Disputants' investment into litigation depends on what they believe their opponent invests. ADR design influences how precise these beliefs are through the implied information revelation. If ADR cannot eliminate evidence provision, its effect on the information structure is of first-order importance. We characterize the optimal ADR mechanism and the corresponding optimal information structure.

We derive four main features of optimal ADR in the shadow of a legal contest. First, 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 ADR promises settlement for a specific type profile, it imposes an externality on other types by influencing their beliefs after failed settlement.

Second, we show that the optimal protocol is asymmetric. It implies that upon failed settlement, it is common knowledge which disputant has (on average) a better case. Asymmetry reduces inefficiencies in litigation. Higher welfare in litigation reduces the cost of failed settlement. That reduction facilitates initial agreement to join ADR. Under optimal ADR, litigation after failed settlement implies less litigation expenditure than litigation absent an ADR option.

Third, disputants cannot influence what they learn within ADR through their own behavior. Suppose a disputant could obtain different information from making different reports. Then she can induce a situation without common knowledge of beliefs. Suppose that a weak disputant mimics a strong disputant during ADR. She knows about her misreport but her opponent does not. Now suppose settlement fails. Continuation strategies depend on disputants' beliefs. Not aware of the deviation, the non-deviating

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

opponent best-responds to an incorrect belief. The deviator arbitrages on her information advantage about her own deviation. If the information structure is independent of a player's own report, such an information advantage to deviators is absent.

Finally, despite the induced asymmetries, optimal ADR is ex-ante fair. The disputant who expects being worse off conditional on escalation expects being better off conditional on settlement. Our findings provide implications in line with stylized facts about ADR.

ADR is effective. Optimal ADR settles more than 50% of the disputes independent of the parameter choices. That finding is in line with the empirical evidence we observe across jurisdictions and case characteristics.[2]

ADR has a positive effect on litigation expenditure even if unsuccessful. If optimal ADR fails to settle, the expected litigation expenditure is smaller than absent the ADR process. The result is in line with survey evidence stating that disputants appreciate ADR even if unsuccessful (Genn, 1998). It also matches the findings in the report to the president by the interagency ADR working group, or the corresponding report to the European commission.

We show that a way to implement optimal ADR is (voluntary) mediation. This result is in line with evidence that mediation has become the most common and successful form of ADR (e.g Stipanowich, 2004).

Related Literature. The law and economics literature studying out-of-court settlement under asymmetric information is large. Most of it emerges from the seminal work by Bebchuk (1984). A large part of the literature focuses on bilateral bargaining.[3] We allow for flexible bargaining mechanisms following Spier (1994) in approaching the problem from a mechanism-design perspective.

Spier (1994) uses a model that applies to situations in which investment in litigation is fixed (i.e. is made *prior* to settlement negotiations) and is interested in optimal fee shifting between parties.

Similarly, Klement and Neeman (2005) look for the optimal fee-shifting rule *and* the optimal settlement procedure jointly. They highlight the role of renegotiation proofness. Similar to us, information revelation plays an important role in their analysis. Different to us, information revelation does not affect parties behavior within litigation, but determines incentives to renegotiate.

One conceptional difference to both Spier (1994) and Klement and Neeman (2005) is that we keep the rules of the legal system fixed in our setting. We assume the American rule throughout. Our focus is on the design of the settlement procedures *within* a given legal system. This is motivated by the observation that the law often gives the suppliers of ADR full freedom in the design of their ADR protocol, but not in the choice of legal

[2] See e.g., for the US, `https://www.justice.gov/olp/alternative-dispute-resolution-department-justice`

[3] Examples include Schweizer (1989), Shavell (1995), Posner (1996), Spier (1997), and Hylton (2002) or recently, Vasserman and Yildiz (2018).

consequences should ADR fail to settle. A second difference to the above mentioned papers is that both implicitly assume that litigation strategies do not depend on the information structure. That is, the choices of a disputant in litigation do not depend on the information she obtained about her opponent. Our focus is, instead, precisely on the interaction between information and litigation choices.

Related to Klement and Neeman (2005), we assume a form of renegotiation proofness—credibility—on the ADR protocol. No disputant can be forced to participate in ADR. Instead, parties have to mutually agree on the ADR protocol to start the settlement procedure. We show that a second notion of renegotiation proofness from Klement and Neeman (2005)—durability—does not always apply if ADR is designed to minimize litigation. However, once ADR aims at maximizing welfare, durability can be guaranteed.[4]

In line with Spier (1994) and Klement and Neeman (2005), but different to Klement and Neeman (2013), we do not model the choice of ADR mechanisms by the parties explicitly. Instead, we assume that a single take-it-or-leave-it ADR protocol is provided. However, and in particular if failure of settlement is costly to the supplier of ADR mechanisms (e.g. by damaging their reputation), our mechanism naturally emerges when the ADR market is competitive.

Doornik (2014) takes the ADR protocol as given and studies in *which cases* ADR should be applied. In our model, we allow for any ADR mechanism and show that it is optimal to ensure full participation irrespective of the case characteristics.

In line with Brown and Ayres (1994) managing the information flow between litigants is the main rationale for ADR in our model. We view litigation as a legal contest in the tradition of Posner (1973), Katz (1988), Baye, Kovenock, and Vries (2005), and Spier and Rosenberg (2011). Assuming legal contests for litigation implies that managing the information flow becomes a first-order concern and strategies in ADR and litigation cannot be separated.

We connect to Prescott, Spier, and Yoon (2014) and Spier and Prescott (2019) who study contingent settlement contracts. Contingent contracts in these models provide an insurance to risk-averse disputants. In our model, instead, we assume that disputants are risk-neutral and have no insurance motive. Contingencies in the settlement contract can play a role if optimal ADR is implemented through mediation. Some settlement offers from optimal ADR may not be accepted because they contain sufficient information about the opponent's access to evidence. Having obtained that information some disputants prefer to litigate instead. Contingencies mitigate that problem and disputants are willing to accept the settlement offer coming from ADR.

On a technical level, we apply a general finding from our own methodological work (Balzer and Schneider, 2019). In Balzer and Schneider (2019) we document an envi-

[4] We show that optimal ADR can be implemented through mediation, which implicitly shows robustness to another special case of renegotiation proofness.

ronment in which the problem of finding the optimal mechanism can be transformed into a problem of finding the optimal information structure once that mechanism fails. The methodology applies in our setting. It simplifies both the analysis and the characterization. We provide both the optimal information structure and the corresponding properties of the optimal mechanism.

Outside the law and economics literature, our paper is related to the analysis of peace negotiations in international relations (Bester and Wärneryd, 2006; Fey and Ramsay, 2011; Hörner, Morelli, and Squintani, 2015; Meirowitz et al., 2017). These models use an environment related to Spier (1994). Compared to our setting results differ fundamentally. We discuss the relation in detail in Section 3.3 after presenting our result. The starting point of recent work by Zheng (2018) and Zheng and Kamranzadeh (2018) is related to ours, but both complement our findings. Zheng (2018) generalizes our first-best result to a larger environment. Zheng and Kamranzadeh (2018), discuss the welfare maximizing take-it-or-leave-it settlement offers in an environment identical to ours.

Roadmap. We describe our model in Section 2. We state our main findings and the intuition behind it in Section 3. In Section 4 we construct the optimal ADR protocol. Section 5 discusses robustness of our results to various extensions. Section 6 concludes. All proofs are relegated to the appendix.

2 Model

There are two risk-neutral disputants, A and B, having a conflict over a pie of size 1. Each disputant holds private information, θ_i. A disputant's private information captures her *marginal cost of increasing the quality* of her formal evidence. For simplicity we assume binary types. Without loss we normalize the low marginal cost to θ_i=1. High marginal cost are captured by the parameter θ_i=$K > 1$. Types are independent draws from the same distribution, where p is the probability that θ_i=1. We use the terms low-cost (high-cost) type and strong (weak) type interchangeably.

The default way to solve the conflict is through formal litigation. Alternatively, disputants can participate in a given ADR mechanism aiming to obtain a settlement solution instead. Disputants avoid litigation only if both participate in ADR *and* ADR settles the conflict. If at least one disputant vetoes ADR, her veto decision becomes public and disputants engage in litigation. Similarly, if ADR fails to settle the conflict, failure becomes public and the conflict escalates to litigation.

Litigation. Litigation is a legal contest. Disputants compete in providing evidence to a judge or jury.[5] Disputant i chooses the quality level of the evidence she provides,

[5]Another possibility is that evidence is provided in front of an arbitrator after "settlement through mutual agreement" fails.

$a_i \in [0, \infty)$. The highest quality of evidence wins the lawsuit. Increasing quality of the evidence is costly and type θ_i's ex-post utility from evidence profile (a_i, a_{-i}) is

$$u(a_i; a_{-i}, \theta_i) = \begin{cases} 1 - \theta_i a_i & \text{if } a_i > a_{-i} \\ -\theta_i a_i & \text{if } a_i < a_{-i} \\ 1/2 - \theta_i a_i & \text{if } a_i = a_{-i}. \end{cases}$$

ADR. ADR is a mechanism offered by an ex-ante uninformed, non-strategic third party—the designer—at the beginning of the game. We refer to it as the ADR protocol. The ADR protocol results in one of two outcomes: settlement or litigation. Settlement directly awards share $x_i \in [0, 1]$ to disputant i. ADR may destroy parts of the initial pie but cannot provide additional surplus. Thus, the designer is budget constrained and can only implement outcomes satisfying $x_A(\theta_A, \theta_B) + x_B(\theta_B, \theta_A) \leq 1$.

If ADR fails to to settle, or any party vetoes the proposed protocol, disputants move directly to litigation. Litigation is out of the designer's control and follows the rules described above.

The model we present here falls in the class of canonical arbitration problems we define in (Balzer and Schneider, 2019). There, we state and prove the revelation principle for these problems which applies in our setting. We restrict attention to direct revelation mechanisms, that is, ADR protocols in which

(i) disputants report their types by sending a private message $m_i \in \{1, K\}$ to the designer,
(ii) the ADR protocol is *incentive compatible* and disputants truthfully report their type in equilibrium, and
(iii) the ADR protocol has the structure $(m_A, m_B) \mapsto (x_A, x_B, \gamma)$,

where (x_A, x_B) is the *settlement outcome* described above, and $\gamma \in [0, 1]$ is the likelihood that *no settlement is found* and disputants move to litigation.[6]

Timing. First, each disputant privately observes her type and the ADR protocol is publicly announced. Then, disputants simultaneously decide whether to participate in ADR. If both participate, they report to the designer and ADR results in either settlement or in litigation. Settlement ends the game and disputants consume their settlement shares x_i. If ADR leads to litigation, disputants update their beliefs and decide on their strategy in litigation.

If some type decides not to participate in ADR, ADR does not take place and disputants move directly to litigation. They update their beliefs and decide on their strategy in litigation.

[6]For the moment, we abstract from additional information revelation by the designer. In principle such information revelation can be helpful. It turns out that in our model it is instructive to ignore it first as it is often not relevant. We address the (ir)relevance of additional information disclosure in Section 3.3 and Section 4.4.

Whenever litigation is played, it obeys the rules sketched above and disputants consume the resulting value u_i.

Solution Concept, Objective, and Commitment. Our solution concept is perfect Bayesian equilibrium (Fudenberg and Tirole, 1988). We look for the ADR protocol that can maximize the (ex-ante) likelihood of settlement. In our baseline model we assume commitment on all sides outside litigation.

Key Modeling Choices. We aim to highlight how information exchange during ADR interacts with the optimal choice of both evidence provision and the ADR protocol. Our model is simplified along many other dimensions. Here we discuss the underlying implicit assumptions behind the simplifications. After the analysis of our baseline model we provide an in-depth discussion on alternative scenarios.[7]

We assume the simplest possible objective—maximizing settlement. An alternative is to assume an objective that maximizes parties ex-ante expected welfare. Under both objectives, however, the low-cost types' participation constraint holds with equality.[8] Thus, the low-cost types are indifferent between the optimal mechanisms under either objective. Imagine now the limiting case of a (competitive) market for ADR protocols in which the designer has a fixed cost $\epsilon \to 0$ of sending parties to litigation. The limiting case is the protocol we present. Any non litigation-minimizing ADR protocol has to be offered at a higher price. Under common belief refinements, any disputant that reveals she is willing to pay a higher price implicitly reveals her cost, i.e. her private information. Such revelation is not beneficial and high-cost types do not deviate and accept the litigation minimizing mechanism. We discuss welfare maximizing ADR in Section 5.

We assume that parties have commitment power to obey any settlement decision once they decide to participate in the mechanism. One realistic scenario is that of two-step binding arbitration. Once disputants agree to arbitration they opt out of the litigation option. The designer then tries to settle the case without any evidence provision. If such settlement fails, however, the designer calls for disputants to provide evidence and "as-if litigation" starts *within* arbitration. It is not crucial to our results that litigation after rejecting ADR and litigation after failed settlement are identical games. We only need that the private information is ordered in the same way in both cases. We comment on that aspect in the discussion of the first-best solution. A crucial aspect of our model is that litigation strategies are a function of the belief a disputant holds about her opponent's relative strength. This assumption implies that

[7] Our simple model, however, captures the main features of ADR guidelines. See, e.g., `https://www.adr.gov/coreprin.htm`

[8] The reason under the alternative scenario is that welfare gains are higher for any marginal reduction of the high-cost type's quality of evidence provision. See also Balzer and Schneider (2019) for the respective argument.

disputants have a sufficient amount of time between failed ADR and the beginning of the litigation game to recalibrate their litigation strategy. Given the significant delays in the court system, which partly motivate this analysis, it seems reasonable to assume that disputants can indeed decide on litigation strategies after seeing the outcome of ADR.

We assume that the designer has full control over the ADR protocol, but cannot influence litigation directly. This assumption captures the prevelance of the rule of law.

Finally, the assumption that ADR is designed by a neutral third-party with commitment power 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 their services on a regular basis. Clearly, trust is a relevant issue and provides a rationale for the designer's commitment power.

3 Findings

In this part we present our main findings. We start at the end of the game and analyze the continuation game of litigation.

3.1 The Litigation Game

We analyze the continuation game of litigation after ADR fails to achieve settlement. Here we present on-the-equilibrium-path litigation to build intuition. We provide an analysis of potential litigation off-the-equilibrium path when we construct the optimal ADR protocol in Section 4.

Conditional on on-path litigation disputants' type distributions may no longer be symmetric. Depending on the ADR protocol it may be more likely that settlement fails if B is the low-cost type than if A is the low-cost type. The parameter ρ_i stands for the probability with which player i has a low-cost type, conditional on on-path litigation. Without loss we assume $\rho_B \geq \rho_A$.

Our construction essentially replicates Siegel (2014) who we refer to for a broader discussion. Since the equilibrium construction with interdependent types is somewhat involved we provide it here for completeness.

Once disputants interact through ADR type distributions may no longer be independent. Depending on the ADR protocol it may be more likely that settlement fails if B is the low-cost type and A is the low-cost type, than if B is the low-cost type and A is the high-cost type. *Beliefs*, $b_i(\theta_i)$, capture these potential interdependencies. A belief $b_i(\theta_i)$ is the likelihood that disputant i, type θ_i, attaches to disputant $-i$ being the low-cost type. Under the assumption that $\rho_B \geq \rho_A$ we can restrict attention to beliefs in the relation $b_A \geq b_B$, that is, all types of both players agree that player B is the "stronger" player.[9] Beliefs shape disputants' equilibrium behavior. We assume that

[9] We get back to the relation between distributions and beliefs in Section 4.

beliefs satisfy a monotonicity condition,

$$Kb_i(1) > b_i(K) > 1 - K(1 - b_i(1)). \tag{M}$$

The monotonicity condition implies that having low cost is good news to a disputant and low-cost types invest more than high-cost types. It ensures that payoffs are monotone in the sense that low-cost types expect a (weakly) larger payoff than high-cost types. We verify later (Lemma 7) that this assumption is without loss.

Below we construct the unique mixed strategy equilibrium. We sketch the equilibrium strategy support in Figure 1. At a high level, equilibrium behavior follows from an *encouragement/discouragement* effect resulting from beliefs $b_i(m_i)$. If the expected investment of an opponent with the same type increases, for example, because that type becomes more likely, this *encourages* a disputant and she increases her quality investment. In contrast, if a low-cost opponent's expected investment increases, this *discourages* a high-cost disputant and she reduces investment in quality.

Proposition 1. *Under* (M) *and* $\rho_B \geq \rho_A$*, disputants' expected payoffs in litigation,* $U_i(\theta_i)$*, are*

$$\begin{aligned}
U_i(1) &= 1 - b_B(1) - \frac{1}{K}\Big(1 - \frac{b_A(K)b_B(1)}{b_A(1)}\Big), \\
U_A(K) &= 0, \\
U_B(K) &= b_A(K) - b_B(K) - \frac{1}{K}\Big(1 - \frac{b_B(1)}{b_A(1)}\Big)\frac{b_A(K)(1 - b_B(K))}{1 - b_B(1)}.
\end{aligned} \tag{1}$$

We provide a constructive argument for the equilibrium after discussing the broad intuition. First $U_A(K) = 0$ because $\theta_A{=}K$ is discouraged by attaching a large probability mass on facing a low-cost opponent. Low-cost types' payoffs decrease in $b_A(1)$ and $b_B(1)$ as the encouragement effect leads to intense litigation. The increase in $b_B(1)$ implies more aggressive litigation behavior of $\theta_B = 1$. That behavior discourages $\theta_A{=}K$ and thereby encourages $\theta_B{=}K$. Vice versa for an increase in $b_A(1)$. Thus, $U_B(K)$ increases (decreases) in $b_B(1)$ ($b_A(1)$). Moreover, $U_1(1)$ and $U_B(K)$ increase in $b_A(K)$ through discouragement on $\theta_A{=}K$. Finally, $U_B(K)$ decreases in $b_B(K)$. Type $\theta_B{=}K$ faces a low-cost opponent with larger probability.

We now provide the main arguments behind the equilibrium construction. Let $F_i^{\theta_i} : [0, \infty) \to [0, 1]$ describe type θ_i's distribution of actions a_i. Fix F_{-i}^1 and F_{-i}^K. Then disputant i, type θ_i, holding belief $b_i(\theta_i)$ solves

$$\max_{a_i} \quad F_{-i}(a_i|b_i(\theta_i)) - \theta_i a_i, \tag{2}$$

where $F_{-i}(a_i|b_i(\theta_i))$ is the expected likelihood that $a_i > a_{-i}$ given belief $b_i(\theta_i)$. We can

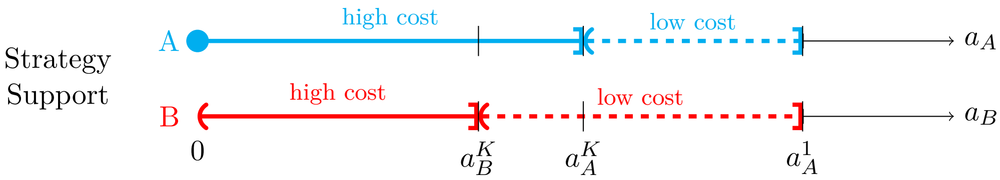


Figure 1: Equilibrium quality of evidence in the on-path continuation game. *All types (piece-wise uniformly) mix. Solid lines denote the intervals of θ_i=K, dashed lines those of θ_i=1. The equilibrium distribution is non-atomistic, apart from the mass point at* 0 *for disputant A, type K.*

decompose $F_{-i}(a_i|b_i(\theta_i))$ to

$$F_{-i}(a_i|\theta_i) = b_i(\theta_i)F^1_{-i}(a_i) + (1 - b_i(\theta_i))F^K_{-i}(a_i).$$

Fix a set of beliefs $b_i(m_i)$. An equilibrium is a fixed point solving each type's and player's maximization problem simultaneously. We provide a full characterization of monotone equilibria following Siegel (2014). Graphically, Figure 1 summarizes the equilibrium characterization. The upper bound of the joint support is the same for both disputants. Consistently outperforming the opponent by some margin cannot be optimal as investment in quality is costly. For the same reason the equilibrium is in mixed strategies and disputants make their opponent indifferent. Since marginal cost are constant, so are densities. If the distribution of cost is asymmetric, equilibrium strategies are asymmetric too. If $b_A(1) = b_B(1)$, then $a^K_B = a^K_A$ and the equilibrium is symmetric. Moreover, the mass point disputant A has at 0 vanishes in that case. Disputant $\theta_A = K$ is the weakest of all potential realizations. She has high cost and faces an opponent that is likely to have low cost. She expects zero payoff in equilibrium and is willing to abstain with positive probability. Analytically, the following lemma provides the characterization.

Lemma 1. *Assume* $1 > \rho_A(1) \geq \rho_B(1) > 0$ *and* $b_A(1) \in (0, 1)$*. Litigation has a unique equilibrium and is characterized by quality levels* $\bar{a}^1_A > \bar{a}^K_A \geq \bar{a}^K_B > 0$ *that partition the action space. The support of each disputant's equilibrium strategy is on the intervals*

- $(0, a^K_A]$ *for disputant A, type K, and* $(a^K_A, a^1_A]$ *for disputant A, type 1,*
- $(0, a^K_B]$ *for disputant B, type K, and* $(a^K_B, a^1_A]$ *for disputant B, type 1.*

In addition, disputant A, type K, has a mass point at 0 *if* $\rho_A(1) > \rho_B(1)$*. The density* $f^1_i(a) = \frac{\theta_{-i}}{b_i(\theta_{-i})}$ *for all quality levels a in the joint support of* $\theta_i = 1$ *and* θ_{-i}*. Similarly, type* $\theta_i = K$ *has density* $f^1_i(a) = \frac{\theta_{-i}}{(1-b_i(\theta_{-i}))}$ *for quality levels in the joint support with* θ_{-i}*.*

The equilibrium payoff then follows from evaluating (2) given the the above equilibrium behavior. Note that $U_A(K) = 0$ because that type concedes immediately (evidence quality= 0) with positive probability. At the infimum of her equilibrium strategy support type $\theta_B = K$ wins only against a type $\theta_A = K$ if that type concedes. Her expected payoff is thus $U_B(K) = (1 - b_B(K))F_A^K(0)$. Finally, low-cost types win with certainty if they choose to provide the upper bound of their equilibrium strategy support. Their expected payoff is $U_i(1) = 1 - a_A^1$. For the remainder of the analysis it suffices to keep the expected utilities $U_i(\theta_i)$ and Figure 1 in mind.

3.2 Optimal ADR

We now turn to the design of the optimal ADR protocol. We first determine under which conditions ADR achieves full settlement. We refer to these cases as the arbitrator's "first best." Then, we focus on the "second best" ADR protocol, that is, optimal ADR if full settlement is not implementable. We state the arbitrator's problem and present the economic features of optimal ADR. Thereafter, we derive the optimal ADR protocol at a more technical level.

3.2.1 First-Best ADR

As a benchmark we characterize when ADR achieves full settlement. We will see that full settlement is possible if participation in ADR is easy to guarantee, that is, if litigation without an ADR attempt is (expected to be) sufficiently costly for all types. Intuitively, litigation is costly for low-cost types if they expect a low-cost opponent with sufficiently high likelihood. In such a case disputants expect that high-quality evidence is necessary to win litigation, leading to intense litigation. In turn, no type has an incentive to veto a full-settlement proposal.

Under full settlement γ=0 and litigation never takes place. Disputants' private information, however, is payoff relevant only in litigation. Under full settlement a player's expected payoff from ADR collapses to the (expected) settlement share. That share is independent of the disputant's type θ_i. If several shares are offered each type naturally picks the highest among these. It is therefore without loss to assume that ADR offers a single share x_i to disputant i irrespective of her type. For an equilibrium the proposed share must be larger than each types' veto payoff, that is, the (off-path) payoff of not participating in arbitration.

Veto Payoff. Suppose a disputant vetoes x_i. Her veto decision becomes public before litigation takes place.[10] A vetoing disputant cannot learn from her own veto in a perfect Bayesian equilibrium. Thus, she continues to hold the prior belief about the non-vetoing disputant. The non-vetoing disputant, in turn, holds an off-path belief about the vetoing disputant. The belief about the vetoing disputant is in principle arbitrary but known. We assume without further loss that vetoing beliefs, p^V, are symmetric, that is, a non-vetoing disputant $-i$ believes that a vetoing i has low cost with probability p^V. The

[10]The fact that vetoes become public is not important. Since vetoes are off-path events a disputant i who vetoes always expects her opponent $-i$ not to have vetoed. In turn, because the optimum is with full participation, unilateral vetoes are always identified as such.

payoffs from off-path litigation follow from applying Proposition 1 to beliefs p for the deviator and p^V for the non-deviator regardless of their respective types.

Corollary 1. *Litigation following a rejection of i implies expected payoffs for i of $V^K = \max(p^V - p, 0)(K-1)/K$ for the high-cost type and $V^1 = (1 - \min(p, p^V))(K-1)/K$ for the low-cost type.*

A consequence of Corollary 1 is that the veto payoffs are type dependent but not information dependent. Thus, if the litigation procedure after rejected ADR is different (e.g. through contractual punishments agreed upon ex-ante) nothing substantial in our analysis changes. The only relevant aspect is that $V^1 \geq V^K$. To keep it simple we continue assuming that litigation following a veto and litigation following failed settlement are identical.

First-Best ADR. Disputant i accepts the proposal x_i if $x_i \geq V^{\theta_i}$. Corollary 1 implies $V^1 \geq V^K$. Full settlement is possible if there is a veto belief p^V such that $2V^1 \leq 1$. In that case a $(1/2, 1/2)$-split implements the arbitrator's first-best solution. It turns out that first-best ADR is implementable if $p \geq \overline{p} := \frac{K-2}{2(K-1)}$ as the following result shows.

Proposition 2 (First-Best ADR). *The first-best ADR protocol guarantees settlement, $\gamma = 0$, and awards $x_i = 1/2$ to each i independent of her type. Disputants participate in that mechanism if and only if $p \geq \overline{p}$.*

The result is in line with Compte and Jehiel (2009) and Zheng (2018). It is immediate that first-best ADR exists for any p if $K \leq 2$ and for any K if $p \geq 1/2$.

Proposition 2 identifies conditions for when litigation can be avoided completely. This happens if (i) the advantage of low-cost types to high-cost types is small, that is, K is sufficiently small; or (ii) if it is sufficiently likely that both disputants have a low-cost type, that is, p is sufficiently high. In (i) high-cost types are less discouraged and in (ii) low-cost types are encouraged to invest into quality. Either scenario implies intense litigation resulting in small payoffs for low-cost types. Full settlement is feasible. In any other scenario ADR implies on-path litigation.

3.2.2 Second-Best Arbitration: The Arbitrator's Problem

The payoff a disputant obtains from participating in ADR is a weighted sum of her expected settlement outcome and the expected payoff from the continuation game of litigation. Weights are determined by the expected probability that settlement fails,[11]

$$\gamma_i(m_i) := p\gamma_i(m_i, 1) + (1-p)\gamma_i(m_i, K).$$

The expected payoff coming from settlement is a function of the reports only. We

[11]We use the convention that $\gamma(\theta_A, \theta_B) \equiv \gamma_i(\theta_i, \theta_{-i})$ to shorten notation.

summarize it as the *settlement value*

$$z_i(m_i) := p(1-\gamma_i(m_i,1))x_i(m_i,1) + (1-p)(1-\gamma_i(m_i,K))x_i(m_i,K).$$

In equilibrium, the arbitrator's choice of γ determines the disputants' beliefs in litigation if ADR fails to achieve settlement. Thus, these beliefs are a function of the report only and read

$$b_i(m_i) := \frac{p\gamma_i(m_i,1)}{\gamma_i(m_i)}.$$

Analogous to the definition of $U_i(\theta_i)$ in Section 3.1 the expected continuation payoff conditional on failed settlement of type θ_i who reported to be type m_i is[12]

$$U_i(m_i;\theta_i) = \sup_{a_i} F(a_i|b_i(m_i)) - a_i\theta_i.$$

The expected payoff from litigation after failed settlement is $\gamma_i(m_i)U_i(m_i;\theta_i)$. Combining the two we obtain the payoff from participation in ADR which is

$$\Pi_i(m_i;\theta_i) := z_i(m_i) + \gamma_i(m_i)U_i(m_i;\theta_i).$$

An ADR protocol is *incentive compatible* if $\forall i, m_i : \Pi_i(\theta_i;\theta_i) \geq \Pi_i(m_i;\theta_i)$. An incentive compatible ADR protocol features *full-participation* only if $\Pi_i(\theta_i;\theta_i) \geq V^{\theta_i}$.

Lemma 2. *It is without loss to assume full participation at the optimum.*

The result follows as ADR can replicate a veto outcome by passing disputants on to litigation. It is not necessary for the designer to screen disputants by letting them veto ADR. Such screening can be replicated with an appropriate choice of γ. Full participation follows.

The arbitrator's optimization problem is as follows.

$$\begin{aligned} \min_{(\gamma_i,x_i)} \; & \underbrace{p^2\gamma(1,1) + p(1-p)\gamma(1,K) + (1-p)p\gamma(K,1) + (1-p)^2\gamma(K,K)}_{=:Pr(L)} \\ \text{s.t. } \forall\theta_i, i, m_i : \quad & \Pi_i(\theta_i;\theta_i) \geq \Pi_i(m_i;\theta_i) \text{ and} \\ & \Pi_i(\theta_i;\theta_i) \geq V^{\theta_i}. \end{aligned} \tag{3}$$

3.2.3 Second-Best ADR

Here we characterize optimal ADR for the case in which $p < \overline{p}$, that is, when full settlement is not possible. Characterizing optimal ADR is challenging because the design of ADR influences how litigation is played after failed settlement, which, in turn,

[12] We use the sup instead of the max to account for potential non-existence of a best-response. This, however, is relevant in off-path events only. Thus, non-existence of equilibrium is not an issue.

influences disputants' incentives within ADR. Indeed, there is a non-trivial feedback loop between the optimal choice of γ, the implied beliefs b_i in litigation, the equilibrium action choices a_i and disputants' behavior (i.e. incentive compatibility).

We provide a constructive approach to optimal ADR in Section 4. Here, we state our main findings and give a broad intuition of the economics behind it. We state necessary conditions for the optimal protocol and provide a full characterization for p sufficiently large. Define $\underline{p} := (2(K-1) - \sqrt{8-4K+K^2})/(2+3K) < 1/3$.

Proposition 3 (Second-Best ADR)**.** *Optimal ADR implies the following features.*

(Induced Asymmetry). *The distribution of types in litigation following ADR is asymmetric, $\rho_A \neq \rho_B$.*

(No Information Trading). *The information a disputant obtains within ADR is independent of her own behavior, that is, $b_i(\theta_i) = \rho_{-i}$.*

(No Guarantees). *Any pair of types, (θ_A, θ_B) escalates with positive probability, $\rho_i \in (0,1)$.*

Moreover, if $p \in [\underline{p}, \overline{p})$, optimal ADR is characterized in closed form as

$$(\rho_i, \rho_{-i}) = \left(\frac{1-p}{2}, \frac{1+p}{2}\right).$$

Proposition 3 provides a full characterization of optimal ADR if $p \in [\underline{p}, \overline{p})$. For the remaining cases a full characterization is also possible albeit not in closed-form. We provide a constructive argument leading to Proposition 3 in Section 4.[13]

Here we build an intuition behind Proposition 3 and discuss the economics. Observe first that information changes the equilibrium strategies in litigation and thus the expected payoffs from litigation. Next, recall the reasons why full settlement is not achievable: (i) low-cost types would prefer litigation right away to full settlement and (ii) full settlement implies that the designer cannot differentiate a low-cost type from a high-cost type. The features in Proposition 3 address these obstacles to full settlement. The first property, *induced asymmetry*, addresses (i). A failed settlement attempt provides disputants with information that reduces the intensity of litigation. The prospect of less intense litigation increases continuation payoffs and makes ADR a more attractive alternative. The other two properties, *no information trading* and *no guarantees*, address (ii). ADR should not provide disputants with too much information. Otherwise high-cost types have high-powered incentives to mimic low-cost types during ADR to extract information. Screening becomes harder. We elaborate on each property in detail.

Asymmetry. Asymmetry decreases the expected quality of evidence. Increasing the quality of evidence provision is costly and thus any such reduction is beneficial to aggregate expected welfare in litigation. If litigation happens with positive probability,

[13] We provide a computer program to calculate the optimal solution including $p < \underline{p}$ on our website.

the reduction also increases expected welfare from participating in ADR. Larger welfare from participation eases participation constraints, and is thus beneficial to the designer.

Asymmetry operates through the discouragement effect. If a low-cost type is sure to face a high-cost type and vice versa, both types are reasonably certain about the outcome. The high-cost type has little incentives to invest into evidence quality. She expects to lose with high probability. The low-cost type has little incentives to invest into evidence quality. She expects to win against the high-cost type also at a lower quality level.

The stronger the asymmetry, however, the larger the settlement share that ADR has to promise the disadvantaged party to compensate her for the worse prospects in the continuation game. That promise is costly to the designer. The trade-off implies an interior level of asymmetry.

No Information Trading. No information trading makes the amount of information conveyed to a particular type independent from the information that type provides.

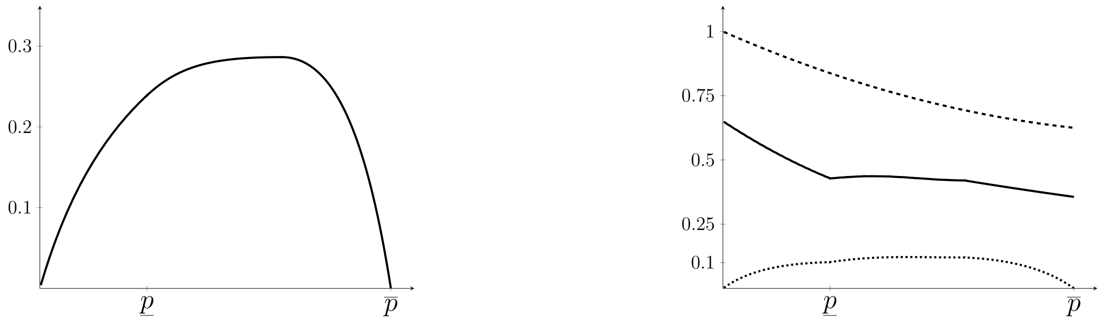
Absent no information trading a party that misreports receives an information advantage in litigation. Assume $\theta_A = K$ deviated and reported $m_A = 1$ and litigation followed. The deviation implies that she holds belief $b_A(1)$. The non-deviator B cannot detect that litigation follows a deviation by A. She thus believes that any type $\theta_A = K$ holds belief $b_A(K)$. That (incorrect) second-order belief implies that the deviator A being aware of her own deviation has an information advantage.

The deviator can leverage on that information advantage. The non-deviator follows her equilibrium strategy. Recall that the equilibrium is in mixed strategies. If $b_A(1) \neq b_A(K)$, the deviator is not indifferent between the actions in her *on-path* strategy support anymore. Instead she chooses a pure strategy. Moreover, she does not fear that her change in behavior influences B's strategy since B *expects* on-path beliefs and thereby on-path behavior.

No information trading implies that $b_A(1) = b_A(K)$. The information advantage disappears. Despite still being the only one aware of her deviation, A has now no incentive to adjust her strategy. Consequently, off-path behavior is not different from on-path behavior. Thus, expectations about behavior are correct on both sides and A cannot leverage on her superior knowledge.

The no-information trading condition resembles the intuition from a second-price auction. There, to ensure incentive compatibility, the payment conditional on winning is independent of a bidder's type report. Similarly here, to ensure incentive compatibility, the belief in litigation is independent of a disputant's type report.

No Guarantees. This property implies that no "easy settlements" exist. Suppose–to the contrary–that the designer guarantees settlement if both disputants have high cost. Further, assume that both A and B have high cost, but A mimics the low-cost type in ADR. If B observes the outcome litigation, she is sure to face a low-cost A. She is pessimistic about her chances of winning litigation. The pessimism discourages her to



(a) Probability of litigation

(b) Litigation Expenditure

Figure 2: Probability of litigation and litigation expenditure as a function of prior p. *The left panel shows the probability that settlement fails under optimal ADR. In the right panel the* ***solid line*** *shows the litigation expenditure* conditional on failed settlement, *the* ***dashed line*** *at the top shows the litigation expenditure without ADR, the* ***dotted line*** *at the bottom shows the* unconditional *litigation expenditure under optimal ADR. All results are computed using* $K = 5$.

invest into evidence provision. Disputant A can leverage on B's pessimism. A has to invest little into evidence to win against B simply because B *expects* A to have low cost.

The designer has two instruments to make these deviations less attractive. First, she can induce litigation also for a pair of high-cost types. High-cost types become less pessimistic and increase the quality of evidence. Second, the designer can increase the likelihood that two low-cost types face each other in litigation. After claiming low cost, a disputant faces a low-cost opponent more often if settlement fails. Moreover, that low-cost opponent expects intense litigation and increases investment. Both effects make deviations less profitable.

Both instruments relax incentive constraints, but have an adverse effect on the low-cost types' participation constraints. The first alternative has a smaller effect on participation constraints. As a result, ADR provides no guarantees to any type pair. It implies that ADR does not reveal disputants' types.

3.3 Properties of Optimal ADR

In this part we address implications of the optimal ADR protocol. We focus on the intuition and trade-offs behind the economic implications. Formally, all results follow from the construction of optimal ADR in Section 4.

Litigation Rates. According to the department of justice, ADR is successful with settlement rates substantially above 50% across time, jurisdiction and case characteristics. Our model is able to replicate such rates independent of the parameter setting.

Proposition 4. *The litigation rate is inversely u-shaped in p. It is smaller than* 50%.

The result is intuitive. For $p \to \overline{p}$ litigation without preceding ADR becomes sufficiently unattractive for low-cost types. Thus, their participation constraint becomes less and less binding and second-best ADR converges to first-best ADR. Thus, the probability of litigation goes to 0. Similar, for $p \to 0$ the environment converges to complete information. As the private information disappears the main obstacle to settlement disappears too and second-best ADR converges to first-best ADR. These countervailing effects of p on the probability of litigation results in an inverse u-shape. Overall ADR is effective with a settlement rate above 50%.

Litigation Expenditure. A second concern is litigation expenditures. Given our objective, the designer is agnostic about the litigation expenditure once settlement fails. However, the average litigation expenditure *conditional on failed settlement* is small.

Related to litigation expenditure is the question whether disputants feel regret after they have been through ADR and have observed its outcome. A disputant *regrets participation* after failed settlement if her continuation utility is smaller than that had she rejected ADR in the first place. It turns out that low-cost types can regret initial participation in ADR once settlement fails.

Proposition 5. *Expected litigation expenditure conditional on failed settlement in ADR is smaller than expected expenditure absent ADR if* $p \geq \underline{p}$*. Conditional on litigation as the outcome of ADR, low-cost types regret initial participation if* $p < 1/3$*, high-cost types do not regret participation.*

Combining Proposition 4 and 5 immediately implies that ADR leads to large welfare gains. Indeed, while ADR is able to settle more than 50% of the cases and thereby reduces litigation expenditure to 0, the remaining cases are also solved more efficiently. There are two factors driving the result. The first is that the composition of remaining litigants after failed settlement differs. In particular, there are more low-cost types. Thus, even for the same evidence quality cost are lower. Second, disputants adjust their strategies. In particular, high-cost types do not invest too much in evidence facing another high-cost type with larger probability.

Although ADR leads to welfare gains, parties might regret to having agreed to ADR in the first place once settlement fails. High-cost types never regret participation independent of the outcome. For low-cost types two effects interact. On one hand the asymmetry of litigation drives down expected quality of evidence. On the other hand the increase in the overall likelihood of low-cost types intensifies the competition between two low-cost types. It turns out that the second effect dominates if p is small. In contrast, for large p low-cost types do not regret participation conditional on the outcome.

A simple corollary to Proposition 5 and the binding participation constraint shows that the opposite is true for the settlement outcome.

Corollary 2. *Conditional on settlement as the outcome of ADR, and* $p > 1/3$*, only low-cost types regret participation. Otherwise no type regrets participation.*

(Ex-ante) Fairness and Additional Signals. In our model, no disputant is ex-ante *more correct.* That is, from an observers point of view no disputant "deserves" to win the case more than the other. Yet, the optimal mechanism induces asymmetries as demonstrated in Proposition 3. That is, conditional on litigation, it is common knowledge that disputant B has a "better case" than the other.

Such an asymmetric treatment may raise fairness concerns and lead to the presumption that the designer is *biased.*[14] However, asymmetric treatment is essential to implement the optimal mechanism because asymmetric litigation is less costly which increases incentives to participate in ADR.

We argue that asymmetric treatment does not jeopardize fairness for two reasons. First, if $p > \underline{p}$ the optimal mechanism is not asymmetric in terms of payoffs. That is, a disputant who has the "better case" after settlement fails can expect a less favorable settlement arrangement, and $\Pi_A(\theta;\theta) = \Pi_B(\theta;\theta)$. Second, the designer can augment her protocol by a simple coin flip that determines whether she carries out the optimal protocol resulting in $\rho_B > \rho_A$ or that resulting in $\rho_A > \rho_B$.

One way to implement such a coin flip is the following. Disputants report to the mechanism without knowing whether they are treated as A or B. After the reports the designer performs the coin flip, announces who takes which role and implements ADR accordingly. The mechanism is stochastic, but both disputants are treated (ex-ante) symmetric. The announcement of the coin flip's result is a public signal. We refer to it as the *symmetrizing signal.*

It turns out that the symmetrizing signal is the designer's optimal signal. Allowing for additional signals implies that the designer may wish to send the symmetrizing signal. Sending that signal never hurts the designer but strictly benefits her if $p > 1/3$.[15]

Proposition 6. *Suppose $p \geq \underline{p}$. Additional information revelation beyond the symmetrizing signal does not improve over the outcome without information revelation. If $p \leq 1/3$ no additional information revelation improves. Any ADR protocol that is optimal is also optimal when augmented by the symmetrizing signal.*

The symmetrizing signal is not only sufficient but also necessary if $p > 1/3$. A low cost-type has lower cost of increasing quality. Thus, high-cost types fear litigation more than low-cost types. In turn, if litigation was more likely for a high-cost type, the low-cost type would have an incentive to imitate the high-cost type. On average, the designer wants a larger probability measure of low-cost types in litigation than of high-cost types because failed settlement is the only screening device. After averaging over disputants the probability that settlement fails is larger for low-cost types than for

[14]Despite the common belief that symmetric procedural treatment is a fundamental part of fairness, most ADR guidelines state other aspects as their fairness and impartiality principles. Shapira (2012) provides a summary. The fairness principles mentioned there are satisfied for the optimal ADR protocol.

[15]We implicitly restrict ourselves to settings in which the designer cannot disclose information about disputant A in private to disputant B without the consent of A (or vice versa). Indeed most guidelines on ADR protocols consider such a practice a breach of confidentiality and explicitly prohibit it.

high-cost types in optimal ADR. Under the symmetrizing signal, low-cost types decide about their reports to ADR based on that average. Their incentive constraint holds.[16]

We have seen in the discussion of Proposition 3 that information in case settlement fails is a first-order concern of the designer. In fact, in Balzer and Schneider (2019) we show in a general setting that the optimal information structure implies the optimal mechanism. Proposition 6 clarifies that the expected information structure is relevant at the reporting stage. To ensure compliant behavior, ADR has to promise disputants sufficient information in case settlement fails.

Mediation. In our baseline model we assume that parties can commit to accept any outcome the ADR protocol implements. Only when settlement fails they make choices again. In practice often weaker forms of commitment are present. For example, Stipanowich (2004) notes that almost all corporations (98 %) in the Fortune 1000 Corporate Counsel Survey have experienced mediation as a version of ADR.

Technically, mediation provides parties with an additional option to opt-out of ADR *after* the ADR process. We assume that in such cases parties move to litigation. More precisely, parties can only opt out of the ADR process after they have learned that the outcome is "settlement," but without having observed the precise amount of the share x_i they receive under settlement. Such a scenario may be credible if the settlement outcome is vague in that the proposed contract contains contingent payments. Such contracts are common in legal environments (Spier and Prescott, 2019).

Mediation announces the *expected shares* x_i privately to each disputant. Disputants then decide if they accept that share. If so, they are committed to that decision conditional on acceptance from the opponent. If either rejects her share, litigation follows.

Proposition 7. *Optimal mediation achieves the same settlement rate as optimal ADR.*

The intuition behind the result is the following. The induced asymmetry implies that one player always prefers to accept the settlement agreement. The designer makes her pivotal by proposing a share 0 if the designer wants to trigger litigation. The other disputant who may have an incentive to deviate once she learns the outcome of ADR receives the same offer regardless whether litigation or settlement should be implemented. Thus, no information is revealed to her and she has no incentive to deviate from acceptance. For this construction to work, separate sessions with the mediator are crucial to the success of mediation. Indeed, what is called *shuttle mediation* is an important tool to implement mediation. In addition, the asymmetry in optimal ADR helps to construct the (equally) asymmetric implementation through mediation.

Relation to Classical Dispute Resolution Mechanism. The mechanism-design approach has been applied to dispute resolution in various contexts. Spier (2007) pro-

[16] In Appendix A.3.8 we discuss the case $p < \underline{p}$. Unfortunately, the problem becomes discretely more complex if $p < \underline{p}$.

vides an overview of the (early) literature in law and economics. More recently Fey and Ramsay (2010), Hörner, Morelli, and Squintani (2015), and Meirowitz et al. (2017) use a similar approach as Spier (1994) and apply it to international conflicts. However, in those models action choices in litigation are invariant to the information structure.

Here we revisit the model of Hörner, Morelli, and Squintani (2015) to contrast results. Their model is close to ours, but they assume a type-dependent lottery if settlement fails. Failure reduces the size of the pie by a constant amount c and awards the remaining part $1 - c$ to disputant i with probability $F(\theta_i, \theta_{-i})$ and to $-i$ with the remaining probability $1 - F(\theta_i, \theta_{-i})$. That is, information revelation has no effect on behavior or the expected loss in surplus due to litigation. Results differ drastically.

Proposition 8. *Optimal ADR in the model of Hörner, Morelli, and Squintani (2015) has the following features if high-cost types cannot be guaranteed settlement.*

(Symmetry) *The information disputants obtain is independent of their identity, $b_A{=}b_B$.*

(Information trading). *The information a disputant obtains depends on her behavior in the mechanism, $b_i(1) \neq b_i(K)$.*

(Weak types settle). *Whenever two high-cost types meet, they settle and $b_i(K) = 1$.*

Proposition 8 is the arbitration result in Hörner, Morelli, and Squintani (2015), adapted to our solution approach. The results in Proposition 8 oppose those from Proposition 3. This highlights the important role the escalation game has on optimal arbitration.

Hörner, Morelli, and Squintani (2015) obtain a *sorting mechanism*. "Weak dyads" enjoy guaranteed settlement. Intermediate dyads settle sometimes. "Strong dyads" are guaranteed litigation. Proposition 3 demonstrates that an effect of information on behavior in the escalation game overturns that results. Once ADR fails, disputants reason about the cause and adapt their continuation strategies accordingly. Noiseless contests such as our model of litigation are particularly sensitive to changes in the information structure. The change in behavior becomes the primary concern of the arbitrator. It leads to the results from Proposition 3.

When can we expect information-sensitive escalation games? The effect on continuation strategies is important if disputants have sufficient time to react to the information they obtain within arbitration. Adjusting strategies may be difficult in international conflicts. Failure of settlement negotiation may immediately lead to war, leaving disputants no time to re-optimize military strategies. Strategies are only functions of the information prior to arbitration and the continuation game has an ex-post equilibrium. Legal disputes are different. Disputants face a sufficient time lag between failed settlement and the beginning of formal litigation.[17] That time lag allows for adjustments on litigation strategies.

[17] Litigation follows a strict procedure overseen by the court. Courts typically do not have excess capacities which leads to long waiting times between failed ADR and litigation. Consequently, parties can adjust strategies before entering formal litigation.

4 Construction of Optimal ADR

In this section we construct the optimal ADR protocol. We restrict attention without loss to protocols that imply $\rho_B \geq \rho_A$. Throughout, we focus on the case $p \geq \underline{p}$ such that closed-form solutions exists. Graphs are, however, produced for all values. For the remaining cases $p < \underline{p}$ we provide the solution in the appendix.

The problem is challenging because the continuation payoffs U_i, and thus incentive constraints, can be non-convex in γ. Non-convexities make brute force solutions hard to calculate. To characterize optimal ADR we invoke a change of variables argument similar to the one proposed in Balzer and Schneider (2019). That change of variables states that there is a one-to-one relationship between the optimal information structure and the parameters of the optimal ADR protocol.[18]

We proceed as follows. First, we define what we refer to as an information structure and show how it translates to the parameters needed to obtain both on-path and off-path litigation payoffs, $U_i(m_i; \theta_i)$. Second, we address the role of the information advantage after a deviation. It constitutes a key step to obtain the result of Proposition 3. We use the results from Section 3.1 to construct a measure of the deviator's information advantage. Finally, we map the outcome back into the properties of the optimal mechanism.

4.1 Information Structure

We represent an information structure $\mathcal{B}$ by the three variables $(\rho_A, \rho_B, b_A(1))$. Each ρ_i describes the likelihood that $\theta_i = 1$. Statistically, the belief $b_A(1)$ captures the entire correlation between types.

Figure 3 illustrates the relationship between distributions and information. In the left panel we plot a distribution of *type pairs*. In total there are four different type pairs, $(1, 1), (1, K), (K, 1), (K, K)$. The likelihood of each event is described via $\mathcal{B}$. The right panel shows how these distributions add up to individual type distributions of disputant A and B.

The domain of $\mathcal{B}$ is determined by *internal consistency*. This means that given ρ_A and ρ_B, $b_A(1)$ can be rationalized by some correlation.

Definition 1 (Internal Consistency)**.** An information structure $\mathcal{B}$ with $\rho_A > 0$ is internally consistent if $b_A(1) \in [\max\left(0, 1 - \frac{1-\rho_B}{\rho_A}\right), 1]$.

For the case of $\rho_A = 0$, the value of $b_A(1)$ can be chosen arbitrarily because the left panel of Figure 3 is independent of $b_A(1)$. The information structure determines all beliefs that we used in Section 3.1. They are determined as follows.

[18] To be precise, in Balzer and Schneider (2019) we show that this relationship holds assuming a *finite* action space. However, in our model existence is guaranteed and all arguments apply also under the assumption of a continuous action space

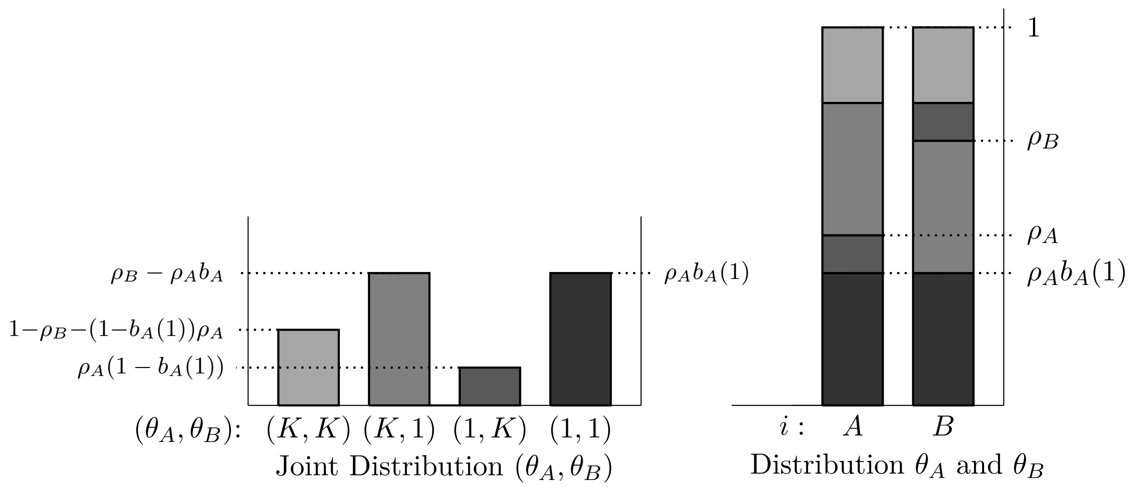


Figure 3: Distributions and Information Structure. *The left panel shows a distribution of type pairs* (θ_A, θ_B). *Each element is a function of the information structure* $\mathcal{B}$. *The right panel shows the distribution of types by disputants. It stacks the elements of the left panel in different order. The likelihood that* $\theta_A = 1$ *is the joint likelihood of* $(1,1)$ *and* $(1,K)$. *The correlation* $b_A(1)$ *determines the fraction of* ρ_A *attributed to* $(1,1)$. *The likelihood that* $\theta_B = 1$ *is the sum of the likelihoods of* $(1,1)$ *and* $(K,1)$.

Lemma 3. *Fix an information structure* $\mathcal{B}$ *with* $\rho_B \geq \rho_A$. *If that information structure arises on-path after some ADR protocol it is internally consistent. The associated on-path beliefs imply* $b_A(\theta_i) \geq b_B(\theta_i)$ *and are given by*

$$b_A(K) = \frac{\rho_B - \rho_A b_A(1)}{1-\rho_A}, \quad b_B(K) = \frac{\rho_A}{1-\rho_B}(1-b_A(1)), \quad \text{and } b_B(1) = \frac{\rho_A}{\rho_B} b_A(1).$$

Using Lemma 3 we can perform comparative statics on the objects in Figure 3.

First, consider an increase in $b_A(1)$. It increases correlation. The likelihood of symmetric type pairs $(1,1)$ and (K,K) (first and last column on the left; top and bottom interval on the right) becomes larger at the expense of that of asymmetric type pairs $(1,K)$ and $(K,1)$ (middle columns and intervals). If $b_A(1)$ increases, litigation in which disputants have similar cost of increasing the quality of their evidence is more likely to occur.

Second, consider an increase in ρ_A. Again, all columns on the left react to such a change. The last two columns increase at the expense of the first two columns. We see on the right panel that such an increase makes it more likely that any type of disputant B faces a low-cost opponent. For the low-cost type of disputant A no changes occur by construction, while for the high-cost type of disputant A we need to distinguish the following two cases. Suppose $\rho_B > b_A(1)$. That is, it is *more* likely for the $\theta_A = K$ to face $\theta_B = 1$ than it is for $\theta_A = 1$. Then an increase in ρ_A leads to a further *in*crease in that likelihood. In contrast, suppose $\rho_B < b_A(1)$. That is, it is *less* likely for the $\theta_A = K$ to face $\theta_B = 1$ than it is for $\theta_A = 1$. Then an increase in ρ_A leads to a further *de*crease in that likelihood.

Finally, consider an increase in ρ_B. Referring to the left panel, it increases the size of the second column at the expense of the first column, while the last two columns are not affected. Referring to the right panel, that makes it more likely for the high-cost disputant A to face a low-cost opponent. It becomes more likely for both types of disputant B to face a high-cost opponent.

In light of the discouragement effect, information structures that imply type-dependent beliefs seem attractive. Indeed, fix the probability mass on $\theta_B{=}1$, ρ_B, and suppose that $b_A(K) > \rho_B > b_A(1)$, that is, $\theta_A{=}K$ expects to face $\theta_B{=}1$ more often than 1_A does. The average (expected) payoff in litigation is larger than the one which results if $b_A(K) = \rho_B = b_A(1)$ because of the discouragement effect (see the discussion below Proposition 1). To implement such an information structure, however, asymmetric type pairs have to occur disproportionately often in litigation. This impacts players' incentives within ADR as we will see next.

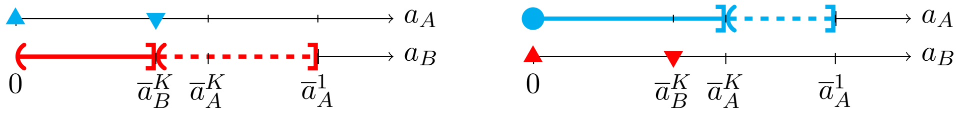
4.2 Off-Path Litigation

On-path litigation in second-best ADR is identical to that characterized in Section 3.1. In contrast, off-path litigation after a non-truthful report can be fundamentally different. After misreporting her own type, disputant i either receives the settlement share of her reported type or litigation is announced following her misreport. In neither case does the opponent (nor the arbitrator) suspect that a deviation had occurred in the reporting stage. In particular, the opponent believes that litigation follows as an on-path event. As a consequence she follows her equilibrium strategy. The deviator on the other hand is aware of her own deviation. As a result she optimizes taking into account that (i) she deviated previously and (ii) the opponent is unaware of that deviation.

Recall from above that continuation utility of type θ_i reporting type m_i is

$$U_i(m_i;\theta) = \sup_{a_i} \quad \underbrace{b_i(m_i)F^1_{-i}(a_i) + (1 - b_i(m_i))F^K_{-i}(a_i)}_{=F_{-i}(a_i|b_i(m_i))} - \theta_i a_i.$$

From Lemma 1 we know that low-cost types invest in higher levels than high-cost types, thus $F^1_{-i} \neq F^K_{-i}$. Moreover, if $b_i(1) \neq b_i(K)$, then the optimal action a^*_i is different off the equilibrium path than on the equilibrium path. For any quality level a_i the marginal cost θ_i of increasing quality is the same as on the equilibrium path. The marginal benefit, that is, the change in the likelihood of winning, however, is different because the belief differs. In addition, the deviation does not trigger any response of the opponent. Indeed, her opponent does not detect the deviation and therefore does not change her behavior. If $b_i(K) > b_i(1)$, this is to the benefit of deviator $\theta_i{=}K$: If $\theta_{-i}{=}K$ knew that $\theta_i{=}K$ holds belief $b_i(1)$ rather than $b_i(K)$, $\theta_{-i}{=}K$ would increase her quality level by the encouragement effect (see the discussion below Proposition 1). In



(a) Quality of evidence in the continuation game after disputant A, type K deviates. *The deviator chooses action ▲ if $b_A(1) > b_A(K)$ and ▼ if $b_A(1) < b_A(K)$. The non-deviator follows her equilibrium strategy.*

(b) Quality of evidence in the continuation game after disputant B, type K deviates. *The deviator chooses action ▲ if $b_B(1) > b_B(K)$ and ▼ if $b_B(1) < b_B(K)$. The non-deviator follows her equilibrium strategy.*

Figure 4: Continuation strategies for different histories if $b_A(1) \geq b_B(1) \neq b_B(K)$.

turn, the gains from the change in beliefs would decrease.

The equilibrium constructed in Lemma 1 is in mixed strategies. Disputants are indifferent between any quality level in their equilibrium strategy support. In the interior of θ_i's equilibrium support U_i is differentiable and $b_i(\theta_i)f^1_{-i}(a_i) + (1 - b_i(\theta_i))f^K_{-i}(a_i) = \theta_i$ on the equilibrium path. If $b_i(1) \neq b_i(K)$ that indifference does not hold off the equilibrium path. Instead the deviator puts full mass on a single quality level. Figure 4 displays the optimal deviation strategies for type K deviators.

Suppose that $b_i(1) < b_i(K)$. High-cost types achieve a higher expected payoff from litigation after a deviation than from on-path litigation. Off path they face high-cost types more often than on path. As depicted in Figure 4 their optimal post-deviation quality level is positive. Thus, they obtain a higher utility than on path by the argument from above. The next lemma provides the statement behind that observation.

Lemma 4. *Suppose that $b_i(1) \neq b_i(K)$. A deviator's optimal action in the continuation game is a singleton. Moreover, if $b_i(1) < b_i(K)$, then $U_i(1; K) > U_i(K; K)$.*

4.3 The Designer's Trade-Off

Binding Constraints. We now formalize the designer's trade-off. We begin by stating a set of binding constraints.

Lemma 5. *It is without loss to assume that the high-cost type's incentive constraint and the low-cost type's participation constraint binds at the optimum.*

The intuition behind the result is as follows. Start with first-best ADR. To make the low-cost type participate the designer has to offer her a higher share $x_i > 1/2$. Increasing the share is costly because the designer has to generate the necessary funds. At the optimum the low-cost type's participation constraint binds. However, as types do not matter under settlement, the high-cost type is eager to mimic a low-cost type who expects a high share. The designer has to promise the high-cost type sufficient compensation to deter such mimicking behavior. At the optimum the high-cost type's incentive constraint binds. The binding constraints imply

$$z_i(1) = V^1 - \gamma_i(1)U_i(1; 1) \tag{IR}$$

and

$$z_i(K) = \gamma_i(1)U_i(1;K) - \gamma_i(K)U_i(K;K) + z_i(1). \tag{IC^K}$$

Finally, the designer's resource constraint implies

$$\underbrace{1 - Pr(L)}_{Prob.\ of\ settlement} \geq \underbrace{\sum_i (pz_i(1) + (1-p)z_i(K))}_{expected\ settlement\ shares\ ,\ \mathbb{E}z}. \tag{B}$$

The last inequality is a necessary condition for the designer's resource constraint, $x_A + x_B \leq 1$ to hold. Indeed, she can only have sufficient funds to distribute the *expected shares* $\mathbb{E}z$, if the *expected rate of settlement* is at least as high. Then the designer has the necessary funds available at least on average.

Lemma 6. *It is without loss to assume that* (B) *holds with equality at the optimum*

We hypothesize that the inequality is also sufficient for that task. In that case it is without loss to assume that (B) holds with equality at the optimum.[19] Using Bayes' rule we can represent the expected probability that settlement fails as

$$\gamma_i(1) = \frac{Pr(L)\rho_i}{p} \quad \textit{and} \quad \gamma_i(K) = \frac{Pr(L)(1-\rho_i)}{(1-p)}.$$

Substituting (IR), (IC^K), and $\gamma_i(m_i)$ into (B) under equality and rearranging yields

$$p\left(1 + \frac{(2V^1 - 1)}{Pr(L)}\right) = \sum_i \rho_i U_i(1;1) + \sum_i p(1-\rho_i)U_i(K;K) - \sum_i (1-p)\rho_i U_i(1;K), \tag{4}$$

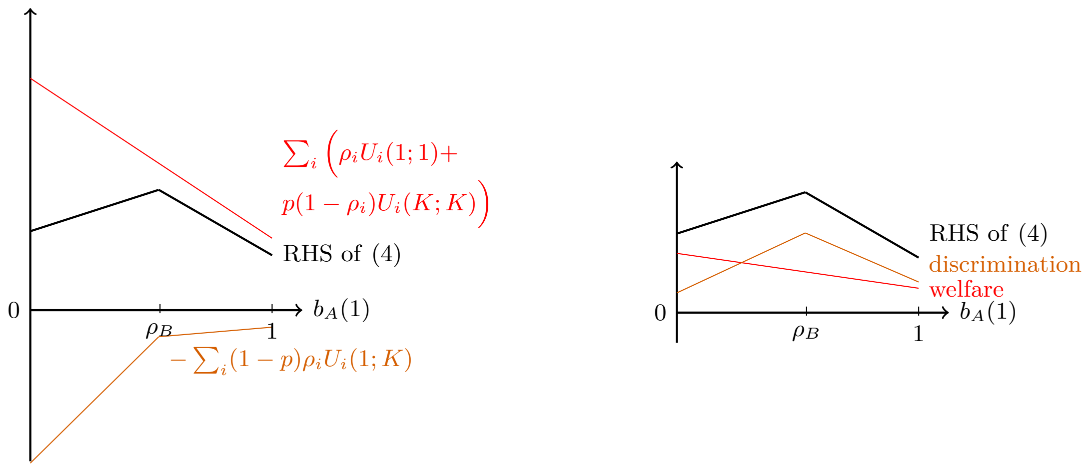
with $2V^1 > 1$ if $p < \overline{p}$ by Proposition 2.

Our final hypothesis is that the low-cost type's incentive constraint and the high-cost type's participation constraint are redundant. Then, maximizing the right hand side (RHS) of (4) minimizes $Pr(L)$.

No Information Trading. We start by addressing the optimal correlation between types in litigation. Figure 5 decomposes the RHS of equation (4) in two different ways. The left panel depicts the decomposition according to the formula stated in the RHS of equation (4). The parts containing on-path continuation payoffs enter positively, those with off-path continuation payoffs enter negatively.

The right panel provides an alternative decomposition. It is along the trade-off the designer solves in $b_A(1)$. She wants to increase the on-path aggregate payoff associated with litigation, $p\sum_i \rho_i U_i(1;1) + (1-\rho_i)U_i(K,K)$. Increasing that payoff reduces the resources needed to convince disputants to participate. The aggregate payoff decreases in the correlation between types, $b_A(1)$. Indeed, low-cost type A expects to face low-cost

[19] If the designer makes a positive payoff she can redistribute that money to settle more cases by relaxing either of the constraints.



(a) Decomposition as in equation (4)

(b) Decomposition in economic channels

Figure 5: Designer's objective as function of $b_A(1)$ (with $K = 3$). *The left panel decomposes the RHS of equation* (4). *The right panel decomposes the RHS alongside two economic channels,* discrimination, $(1-p)\sum_i \rho_i(U_i(1;1)-U_i(1;K))$, *and* welfare, $p\sum_i \rho_i U_i(1;1)+(1-\rho_i)U_i(K,K)$. *Discrimination measures how much better a low-cost type performs compared to a high-cost deviator. The deviator suffers from her higher cost, but benefits from the information advantage. If* $b_A(1) = \rho_B$ *the information advantage is* 0 *and discrimination is the highest. Welfare decreases in* $b_A(1)$ *as increased correlation in types implies more intense litigation.*

type B more often and intensifies her investment into quality. This triggers the same response from low-cost type B and leads to intense litigation.

At the same time, however, the designer wants to discourage high-cost types to mimic low-cost types within ADR by making the litigation play discriminatory. The more discriminatory the litigation play, the better her ability to screen types within ADR. The aggregate discriminatory power of litigation is $(1-p)\sum_i \rho_i(U_i(1;1)-U_i(1;K))$. It is the (weighted) difference in continuation payoffs between an actual low-cost type and a high-cost type who mimics the low-cost type within ADR. It measures how much of the payoff difference between the two types is associated to the fundamental difference in their underlying types. This difference depends on the types' optimal actions, which are functions of the distribution a disputant faces. In addition, the deviator makes use of her information advantage. As described above that information advantage disappears if $b_A(1) = \rho_B$. The information advantage is large and dominates in the discrimination term. Thus, litigation is most discriminatory if the information advantage disappears.

Asymmetry. Above we held fix the distribution of types within litigation and looked at the optimal correlation. We now address the optimal distribution of types in litigation. We fix ρ_A and $b_A(1)$ and vary ρ_B. Changing ρ_B increases the level of asymmetry. Figure 6 depicts that change in a similar way as Figure 5.

Consider the left panel of Figure 6 first. Increasing ρ_B implies that B profits on the equilibrium path as it becomes more likely for her to meet a high-cost type. This

(a) Economic channels

(b) Economic channels for $b_A(1) \equiv \rho_B$

Figure 6: Designer's objective as function of ρ_B (with $K = 3$). *The left panel decomposes the RHS of* (4) *along side the economic channels. For a description of the terms see Figure 5. The right panel uses the same decomposition but varies ρ_B keeping $\rho_B \equiv b_A(1)$.*

encouragement of B in turn leads to a discouragement for A. The discouragement reduces aggregate litigation expenditure and thus increases expected on-path welfare in litigation. Yet again, the information advantage effect dominates in the discrimination channel.

The right panel imposes $\rho_B \equiv b_A(1)$, and then varies ρ_B. It shows that assuming optimal correlation, the distribution channel balances out a negative effect in discrimination with a positive effect in welfare when deciding on ρ_B. It turns out that (independent of the choice of ρ_A), that trade-off is optimally balanced at $\rho_B = (1+p)/2$. Performing a similar task on ρ_A reveals that the optimal choice is $\rho_A = (1-p)/2$. Thus, the optimal degree of asymmetry is a spread of size p around $1/2$ for the two disputants.

Before we turn to the benefits of additional information revelation through the designer, we recall the hypotheses we made so far: In our construction we assume that the type distribution satisfies the monotonicity constraint, (M), that the high-cost types' participation constraints are redundant, and that the expected resource constraint, (B), is sufficient to satisfy the ex-post resource constraint. The next lemma states that our hypothesizes were indeed correct.

Lemma 7. *At the optimum on-path litigation satisfies the monotonicity constraint, high-cost types' (IR) constraints are redundant, and* (B) *guarantees that a sharing rule x_i exists that implements z_i and γ.*

To verify the last claim we invoke the general implementation condition for reduced-form mechanisms from Border (2007). Note that we have not verified the redundancy of the low-cost type's incentive constraint. The reason is that this constraint can become relevant. We address it in the next part together with the designer's desire to reveal information.

4.4 Additional Signals

We now construct the optimal signal from Proposition 6. We provide reasons why the symmetrizing signal is optimal, but often not even needed.

Consider a mechanism with signals. When reporting, disputants expect a continuation payoff conditional on escalation of $\sum_s Pr(s)U_i(m_i;\theta_i|\mathcal{B}(s))$, where $\mathcal{B}(s)$ is the information structure following announcement s and $U_i(m_i;\theta_i|\mathcal{B}(s))$, by a small abuse of notation, is type θ_i's payoff in litigation if she reported type m_i and the information structure is $\mathcal{B}(s)$.

From the designer's perspective we employ once again the RHS of (4). When using signals, that objective turns into

$$\sum_s Pr(s)\left(\sum_i \rho_i U_i(1;1|\mathcal{B}(s)) + \sum_i p(1-\rho_i)U_i(K;K|\mathcal{B}(s)) - \sum_i (1-p)\rho_i U_i(1;K|\mathcal{B}(s))\right).$$

Notice that any element of the first sum is, however, directly implementable through the designer by the arguments from the previous part. The designer can simply implement the most preferred $\mathcal{B}(s)$ with probability one instead of combining it with (less preferred) information structures.

Thus, if no further constraints restrict the designer's choice set in $\mathcal{B}$ she does not want to release additional information. The following lemma summarizes that finding.

Lemma 8. *Suppose $p \geq \underline{p}$. If the low-cost types' incentive constraint is satisfied at the maximum of the RHS of equation* (4)*, then the designer does not benefit from disclosing additional information.*

We now turn to the case in which the designer wishes to disclose additional information. By Lemma 8 a necessary condition for such a case is that a low-cost type's incentive constraint is violated when maximizing the RHS of equation (4). The low-cost type's incentive constraint is

$$z_i(1) + \gamma_i(1)U_i(1;1) \geq z_i(K) + \gamma_i(K)U_1(K;1).$$

Rearranging and substituting using (IC^K) we obtain

$$\gamma_i(1)\Big(U_i(1;1) - U_i(1;K)\Big) = z_i(K) - z_i(1) \geq \gamma_i(K)\Big(U_i(K;1) - U_i(K,K)\Big) \tag{IC}$$

which simplifies to

$$\gamma_i(1) \geq \gamma_i(K) \Leftrightarrow \frac{\rho_i}{p} \geq \frac{1-\rho_i}{1-p} \Leftrightarrow \rho_i \geq p,$$

via the no-information-trading result of Proposition 3 and its implication that $U_i(1;\theta_i) = U_i(K;\theta_i)$. The result is intuitive. Because a low-cost type is better off under litigation than a high-cost types, the former has to expect more often litigation than the latter. An implication is that we should expect more low-cost types in litigation after ADR than under the prior. This requirement, however, is challenged by the designer's desire to implement asymmetric information structures. The optimal distribution $(\frac{1-p}{2}, \frac{1+p}{2})$, from Proposition 3 violates the low-cost incentive constraints for disputant A if $p > 1/3$. However, the low-cost incentive constraint of B is satisfied with slack for any $p < \overline{p}$.

The designer can augment the optimal ADR protocol by use of a the simple coin flip that implements the symmetrizing signal. If the coin shows heads she implements the optimum for $\rho_B \geq \rho_A$. If the coin shows tails she implements it assuming $\rho_A \geq \rho_B$.

Parties are symmetric ex-ante and the litigation rate is the same independent of the outcome of the coin flip. However, low-cost types incentive constraints are always satisfied under the symmetrizing signal.

4.5 Payoffs

The expected payoff in ADR can be decomposed into three parts.

One part is the continuation payoffs in case litigation occurs. The no-information-trading property implies that $U_i(1, \theta_i) = U_i(K, \theta_i)$ and continuation payoffs are

$$
\begin{aligned}
U_i(1;1) &= (1-\rho_A)\frac{K-1}{K} = \frac{(1+p)(K-1)}{2K}, \\
U_A(K;K) &= 0, \\
U_B(K;K) &= (\rho_B - \rho_A)\frac{K-1}{K} = p\frac{(K-1)}{K}.
\end{aligned}
$$

Another part are the probabilities that settlement fails, $\gamma(\theta_A, \theta_B)$. They are

$$
G = \begin{pmatrix} \gamma(1,1) & \gamma(1,K) \\ \gamma(K,1) & \gamma(K,K) \end{pmatrix} = \alpha \begin{pmatrix} 1 & \frac{p}{1+p} \\ \frac{p(1+p)}{(1-p)^2} & \left(\frac{p}{1-p}\right)^2 \end{pmatrix},
$$

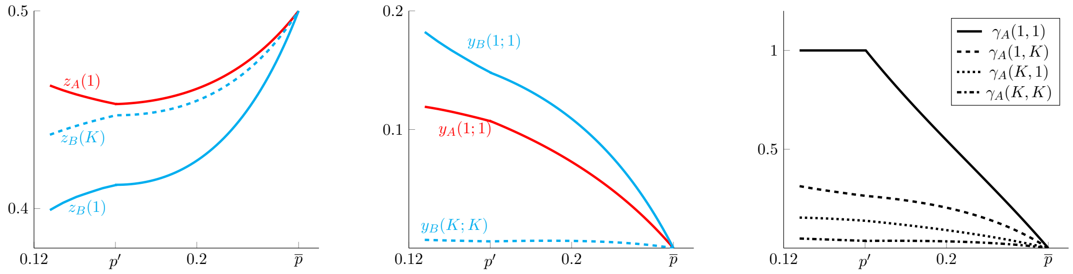
where α is a scalar in $[0, 1]$. We discuss the derivation of α later. First we derive the expected probability that settlement fails as a function of a disputant's report,

$$
\gamma_A(1) = \frac{2p}{1+p}\alpha, \quad \gamma_B(1) = \frac{2p}{1-p}\alpha, \quad \gamma_A(K) = 2\left(\frac{p}{1-p}\right)^2\alpha, \quad \gamma_B(K) = \frac{2p^2}{1-p^2}\alpha,
$$

and the litigation rate is

$$
Pr(L) = p\gamma_i(1) + (1-p)\gamma_i(K) = \frac{p^2}{\rho_A\rho_B}\alpha = \frac{4p^2}{1-p^2}\alpha.
$$

The third part, the expected settlement value $z_i(m_i)$, is determined through the



(a) Settlement Values $z_i(m_i)$

(b) Litigation Values $y_i(m_i;\theta_i) = \gamma_i(m_i)U_i(m_i;\theta_i)$

(c) Failure Rates $\gamma(m_A, m_B)$

Figure 7: Properties of the optimal mechanism as a function of p (with $K = 3$). To the left of p' no interior solution exists and instead a boundary solution with $\gamma(1,1) = 1$ is optimal. To the right of $\overline{p}$ full settlement is possible. In any case, $z_A(K) = z_A(1)$ and $U_A(K;K) = 0 \Rightarrow y_A(K,K) = 0$.

low-cost type's (IR) constraint and the high-cost type's (IC^K) constraint.[20] They are

$$\begin{aligned}
z_A(1) = z_A(K) = V^1 - \gamma_A(1)U_A(1;1) &= (1 - p(1+\alpha))\frac{K-1}{K}, \\
z_B(1) = V^1 - \gamma_B(1)U_B(1;1) &= \left(1 - p\left(1 + \alpha\frac{1+p}{1-p}\right)\right)\frac{K-1}{K}, \\
z_B(K) &= z_B(1) + (\gamma_B(1) - \gamma_B(K))U_B(K;K) \\
&= z_B(1) + 2\alpha\frac{p^2}{(1-p^2)}\frac{(K-1)}{K}.
\end{aligned} \tag{5}$$

The (ex-post) settlement shares $x_i(\theta_i, \theta_{-i})$ are not uniquely defined. However, the expected share conditional on a disputant's own report is

$$x_i(m_i) = z_i(m_i)/(1 - \gamma_i(m_i)).$$

Finally, α follows from substituting into equation (4) and rearranging,

$$Pr(L) = \frac{4p^2}{1-p^2}\alpha = \frac{2p(1-p)\frac{K-1}{K} - p}{\frac{1}{2}(1+p^2)\frac{K-1}{K} - p}. \tag{6}$$

For any $p \in [\underline{p}, \overline{p})$, some $\alpha \leq 1$ solves the above equation. If $p < \underline{p}$, that $\alpha > 1$ and G is not a matrix of probabilities. In that case we implement a corner of the set of feasible $\mathcal{B}$ such that $\alpha = 1$. Other than the existence of a closed-form solution, the properties of Proposition 3 remain, see Appendix A.3.8 for the implicit solution. The value of α *decreases* in p and *increases* in K.

Welfare. Using the results from above, we present the expected payoff from participation in ADR.

[20] It is straightforward to show that V^1 is minimized for $p^v = p$.

$$\begin{aligned}\Pi_i(1;1) = V^1 = (1-p)\frac{K-1}{K},\\ \Pi_A(K;K) = z_A(K) = (1-p(1+\alpha))\frac{K-1}{K}\\ = z_B(K) + \gamma_B(1)U_B(K;K) = \Pi_B(K;K).\end{aligned}$$

Despite the asymmetric treatment within the mechanism, the designer promises each disputant the same expected payoff conditional on their type. While one disputant is better off conditional on settlement, the other one is better off conditional on escalation. The cost and benefits cancel each other out, and payoffs are symmetric from an ex-ante perspective.

Albeit when looking at litigation only it appears that the ADR protocol is biased, when looking at welfare the protocol itself is unbiased. The reasons is that the disadvantaged disputant in litigation obtains a higher settlement value. Both biases cancel each other out in magnitudes too.

5 Robustness

In this part we address how our results are robust to changes in the enviornment.

5.1 Settlement Bargaining

The traditional law and economics literature following Bebchuk (1984) focuses mainly on bilateral settlement negotiations. A version with two-sided private information is Schweizer (1989). Different to our model, in that literature the court outcome is independent of the disputants' choices. Instead, once parties move to litigation they pay a fixed cost c_i and expect to win the pie with probability $y_i(\theta_i, \theta_{-i})$. Thus, the payoff for a given type pair (θ_i, θ_{-i}) is $y_i(\theta_i, \theta_{-i}) - c_i$, which is independent of the information structure. The bargaining works as follows. One of the disputants (say A) makes a take-it-or-leave-it offer to the other (B). If B accepts they settle, otherwise they proceed to litigation.

Two questions arise naturally in light of these models: (i) is it possible to *implement* the optimal mechanism through settlement bargaining, and (ii) do parties have an incentive to bargain over settlement *before ADR*. We address both questions separately.

Settlement bargaining cannot replace the optimal mechanism. For any $p < \overline{p}$, settlement bargaining performs strictly worse in any of its equilibria. The main reason is that B interprets A's offer as a signal. This makes truth-telling in unmediated communication harder than communication with help of a mediatior.[21]

[21] See also Goltsman et al. (2009) and Hörner, Morelli, and Squintani (2015) fo similar results in other models.

Proposition 9. *Private settlement bargaining leads to a strictly higher settlement rate than optimal ADR in any equilibrium of the bargaining game.*

The second, perhaps more relevant, question asks if A has an incentive to bypass ADR by making an offer to B ex ante.

We show that this is sometimes the case by means of an example. Suppose both types of A propose a share $\hat{x}$ to B. Further, assume that $\theta_B = 1$ rejects that offer. Then, B accepts it if she is the high-cost type and $\hat{x} > (1-p)^2 \frac{K-1}{K}$. Disputant A prefers to make the proposal to optimal ADR if

$$1 - \hat{x} = 1 - (1-p)^2 \frac{K-1}{K} > (1-p)\frac{K-1}{K} \Leftrightarrow 1 > (1-p)(2-p)\frac{K-1}{K}.$$

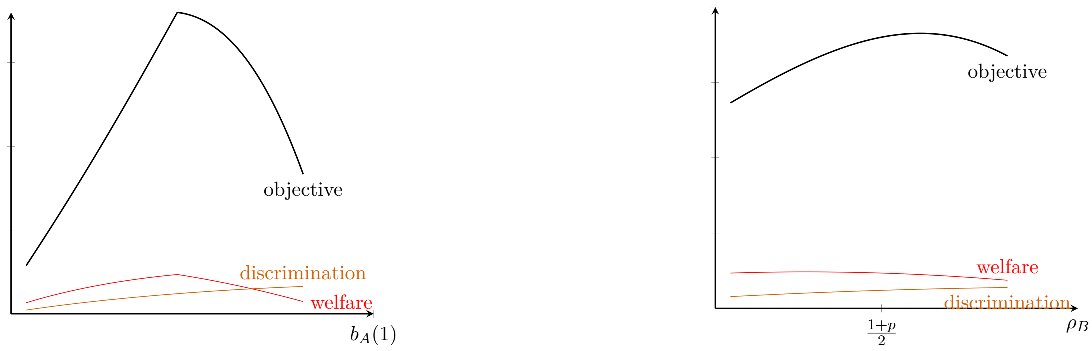
The right-hand side of the inequality increases in K and decreases in p. For example, if $p = 1/3$ and $K > 10$ the proposed bargaining would not be an equilibrium. More generally, for any K there is a $\hat{p}$ such that A prefers not to make the offer proposed. Notice that $\theta_B = K$ only accepts the settlement offer if she is better off from acceptance than from mimicking $\theta_B = 1$ in the negotiations which gives her an information advantage in litigation after the rejection. It is therefore necessary that A proposes a sufficiently large share to B. It turns out that the above proposal structure is the most threatening to ADR.[22]

Why is it possible for A to undermine ADR? The reason lies in the different objectives of ADR. While the focus of our designer is to minimize the likelihood of litigation, A aims at selfishly maximizing her own welfare. At the same time the bilateral bargaining structure suffers from the same caveat that drives Proposition 9. Both A and B have to take into account that their opponent interprets their actions as signals of their type. If litigation eventually occurs, this gives rise to double deviations much in the sense we described in the section on off-path litigation.

5.2 Surplus Maximizing ADR

To avoid going through the cumbersome computation of all possible bargaining structures, we employ a mechanism-design approach once more. We keep all elements of the mechanism as before, but change the objective to maximizing welfare instead of maximizing settlement. In Balzer and Schneider (2019) we show that generally we can represent the welfare-maximization objective by sticking to our familiar measures of welfare and discrimination. However, instead of adding the two we now have to optimize a fraction. Formally, the problem is described using

[22] We omit formally showing this result. It is available upon request.



(a) no information trading

(b) Asymmetry

Figure 8: (Ex-ante) welfare maximization (with $K = 3$). *The left panel shows discrimination, welfare, and the welfare maximzing objective. Welfare and discrimination are exactly as in Figure 5 (left) and Figure 6 (right) respectively. The objective is as described in the text. On the left ρ_A, and ρ_B are held constant. On the right $b_A(1) \equiv \rho_B$ and ρ_A is held constant.*

$$\max_{\mathcal{B}} \frac{\overbrace{(1-p)\sum_i \rho_i(U_i(1;1) - U_i(1;K))}^{\text{discrimination}}}{\underbrace{p - p\sum_i \rho_i U_i(1;1) + (1-\rho_i)U_i(K,K)}_{\text{welfare}}}$$

s.t. (IC[1]). When we use the welfare maximizing approach results do not change qualitatively. However, due to the increased complexity of the objective closed-form solutions may not exist. Figure 8 plots the objective for the same values as Figure 5 and 6. We see that both the no-information trading and the asymmetry property prevail. Yet, asymmetry becomes stronger. The result is intuitive. Under welfare maximization the designer assigns a higher weight on disputants continuation payoffs. By the revelation principle, no initial offer can improve total welfare over the welfare maximizing ADR protocol.

Low-types' participation constraints hold with equality at the optimum and the optimum is again unique (up to symmetrization). Moreover, Proposition 9 did not rely on the objective at all and therefore applies directly. No settlement-bargaining equilibrium exists that implements the outcome of the optimal mechanism.

Moreover, because no initial offer can improve total welfare the following game would result in selection of ADR: If some player rejects the proposed ADR mechanism, a settlement bargaining occurs in which one player is selected at random. That player makes a take-it-or-leave-it offer to her opponent. If the opponent rejects litigation occurs. Proposition 9 implies that total welfare is (strictly) lower for any settlement bargaining. The player making the offer is selected at random, so no player benefits on average in the continuation game. Thus, ADR is preferred to the bargaining procedure.

5.3 Verifiable Types, Correlation, and Assymetric Disputant

If parties could (costlessly) disclose their marginal cost of evidence quality, a first-best solution always exist. The well-known unraveling arguments of Grossman (1981) and Milgrom (1981) apply. Whenever a disputant decides not to disclose her type she is assumed to be a high-cost type. The mechanism offers $(1/2, 1/2)$ if types are the same, otherwise the low-cost type gets the entire pie. Then, low-cost types always disclose their type and obtain a payoff of $1 - p/2$, high-cost types obtain $1/2 - p/2$ and all types settle.

However, as discussed e.g. in Spier (2007), a reason for court procedures is that private types are not necessarily costlessly verifiable. In fact, under costless verification no third-party is needed and private bargaining can solve the dispute efficiently. We believe ADR is more likely required if there is some residual, non-verifiable uncertainty about how costly it is for parties to present evidence in the litigation process.

Moreover, in such cases it is also likely that the correlation between the residual uncertainty is small which is why we focus on independent types. Allowing for correlation, however, does not overturn the results and improves the prospects of arbitration. If types are correlated, the type report of A not only contains information about A's private information but also about the private information of B. That effect relaxes incentive constraints. However, since our model involves no direct utility transfers, incentive constraints do not become redundant as in the model of Crémer and McLean (1988). For more discussion on the effect of correlation and the relation of our model class to that of Crémer and McLean (1988) see Balzer and Schneider (2019).

Assuming that disputants are ex-ante asymmetric does not influence the results qualitatively. Interestingly, the designer still would keep asymmetries but would assign the ex-ante *stronger* disputant the role of the *weaker* player A, i.e. the disputant with the better settlement conditions. The low-cost type of the weaker disputant, in contrast, experiences litigation and its profits more often. That way, the designer can offer the ex-ante weaker disputant the lower settlement share while still providing her with incentives to participate in ADR. Assigning the lower share to the low-cost type of the ex-ante weaker opponent is important because it allows the designer to offer the high-cost type of that player the lower share without affecting incentives. Since the ex-ante weaker opponent is more likely to have high cost such a protocol saves on settlement resources. With such a protocol ADR 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 disputants.

6 Conclusion

We characterize optimal Alternative Dispute Resolution (ADR) in the shadow of the court. We show that optimal ADR is asymmetric and offers one disputant an advantage if settlement fails. The other disputant obtains an advantage under settlement. The optimal information structure post-ADR is completely independent of a disputant's own report, but conditions on her identity. That independence prevents disputants from misreporting to achieve an informational advantage.

The settlement-maximizing ADR protocol is highly effective and settles the majority of cases. The effectiveness supports the use of ADR to reduce the caseload of courts. Moreover, ADR has a positive effect on those cases that fail to settle. The litigation expenditure of the remaining cases is lower than that without the ADR option. The asymmetry of the optimal mechanism implies that imposing notions of fairness on ADR protocols could come with real cost on society and thus should be carefully considered. The same holds for disclosure policies: the ADR-designer should have the possibility to talk to disputants in private to eliminate any privacy concerns. Trust in the ADR designer's discretion is an important driving force of the success of an ADR protocol.

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, asymmetries reduce resource intensity in case of breakdown.

We demonstrate that the standard assumption "lotteries over outcomes" as outside option to settlement is not innocuous when the following two conditions are met: (1) evidence provision by the parties is relevant on the equilibrium path and (2) the choice of evidence is a strategic variable and depends on the expected evidence provided by the opponent. Under these two conditions the behavior of the disputants in ADR and litigation is interconnected.

Not claiming optimality of real-world ADR-protocols, our findings are in line with some stylized facts on ADR. Optimal ADR has success rates beyond 50% independent of characteristics. ADR is informative and helpful to disputants even when not resulting in settlement. Yet, ADR allows disputants to retain some privacy and does not fully disclose disputants' types.

Our findings provide several interesting directions for future research. The assumption that the ADR designer has full-commitment power is important in our framework. Yet, in reality if interactions between the same disputants occur frequently, the designer has an incentive to re-optimize after settlement fails. We expect that this form of renegotiation-proofness (with the designer) has a severe impact on the potential of ADR and may reduces its effectiveness substantially.

Extending the analysis to a setup of more than two disputants for example by considering a class-action framework might add several interesting channels to the model. If a group of disputants has to coordinate to act jointly during ADR concerns such as blameshifting and the opponent's bargaining power may influence the outcome of ADR.

A classic question that arises naturally is how our findings interact with the effects of fee shifting and ex-ante contracting on the negotiation procedure. Both provide additional strategic incentives that are intertwined with the information revelation within ADR. Although we expect results to be similar, a careful description of the environment is essential. Making explicit statements is thus beyond the scope of this paper.

Finally, 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. Although a richer model is needed to address these issues properly, we are confident that the channel and results we present in this paper provide a helpful first step.

Appendix

A Proofs

A.1 Proof of Proposition 1

We organize the proof as follows. First we prove Lemma 1. We then apply these findings to prove Proposition 1. We identify disputants through marginal cost θ_i. We focus on interior beliefs $b_i \in (0,1)$. The extension to the boundary cases is straightforward.

A.1.1 Proof of Lemma 1

Proof. The proof follows Siegel (2014). We omit proving uniqueness and the following properties: (i) the equilibrium is in mixed strategies, (ii) the equilibrium support of both disputants shares a common upper bound, and (iii) the equilibrium support is convex and at most one disputant has a mass point which is at 0. All arguments apply exactly as in Siegel (2014).

Each disputant θ_i holds belief $b_i(\theta_i)$, and maximizes

$$(1 - b_i(\theta_i))\, F^K_{-i}(a) + b_i(\theta_i) F^1_{-i}(a) - a\theta_i,$$

over a. Define the partitions $I_1 = (0, \overline{a}^K_B]$, $I_2 = (\overline{a}^K_B, \overline{a}^K_A]$ and $I_3 = (\overline{a}^K_A, \overline{a}^1_A]$. We define indicator functions $\mathbb{1}_{\in I_l}$ with value 1 if $a \in I_l$ and 0 otherwise. Similar the indicator function $\mathbb{1}_{>I_l}$ takes value 1 if $a > \max I_l$ and 0 otherwise. Disputant θ_i mixes such that the opponent's first-order condition holds on the joint support. The densities are

$$\begin{aligned}
f^1_B(a) &= \mathbb{1}_{\in I_2} \frac{K}{b_A(K)} + \mathbb{1}_{\in I_3} \frac{1}{b_A(1)}, \qquad & f^K_B(a) &= \mathbb{1}_{\in I_1} \frac{K}{1 - b_A(K)}, \\
f^1_A(a) &= \mathbb{1}_{\in I_3} \frac{1}{b_B(1)}, & f^K_A(a) &= \mathbb{1}_{\in I_1} \frac{K}{1 - b_B(K)} + \mathbb{1}_{\in I_2} \frac{1}{1 - b_B(1)}.
\end{aligned}$$

This leads to the following cumulative distribution functions:

$$F^1_2(a) = \mathbb{1}_{\in I_2} a \frac{K}{b_A(K)} + \mathbb{1}_{\in I_3} \left(\frac{a}{b_A(1)} + F^1_B(\overline{a}^K_B) \right) + \mathbb{1}_{>I_3},$$

$$F_2^K(a) = \mathbb{1}_{\in I_1} a\frac{K}{1-b_A(K)} + \mathbb{1}_{>I_1},$$
$$F_1^1(a) = \mathbb{1}_{\in I_3} \frac{a}{b_B(1)} + \mathbb{1}_{>I_3},$$
$$F_1^K(a) = \mathbb{1}_{\in I_1} \left(a\frac{K}{1-b_B(K)} + F_A^K(0) \right) + \mathbb{1}_{\in I_2} \left(\frac{a}{1-b_B(1)} + F_B^K(\overline{a}_B^K) \right) + \mathbb{1}_{>I_2}.$$

Disputants' Strategies: Interval Boundaries. The densities define the strategies up to the intervals' boundaries. These boundaries are determined as follows

1. $\overline{a}_B^K$ is determined using $F_B^K(\overline{a}_B^K) = 1$, i.e. $\overline{a}_B^K f_B^K(a) = 1$ for $a \in I_1$. Substituting yields
$$\overline{a}_B^K = \frac{1-b_A(K)}{K}.$$
2. For any $\overline{a}_A^K$, $\overline{a}_A^1$ is determined using $F_A^1(\overline{a}_A^1) = 1$, i.e. $\left(\overline{a}_A^1 - \overline{a}_A^K\right) f_A^1(a) = 1$ with $a \in I_3$. Substituting yields
$$\overline{a}_A^1 = \overline{a}_A^K + b_B(1).$$
3. $\overline{a}_A^K$ is determined by $F_B^1(\overline{a}_A^K) = 1$. That is, $\left(\overline{a}_A^K - \overline{a}_B^K\right) f_B^1(a) + \left(\overline{a}_A^1 - \overline{a}_A^K\right) f_B^1(a') = 1$ with $a \in I_2, a' \in I_3$. Substituting yields
$$\overline{a}_A^K = \overline{a}_B^K + \left(1 - \frac{b_B(1)}{b_A(1)}\right) \frac{b_A(K)}{K}.$$
4. $F_A^K(0)$ is determined by the condition $F_A^K(\overline{a}_A^K) = 1$, i.e. $F_A^K(0) = 1 - \overline{a}_B^K f_A^K(a) - \left(\overline{a}_A^K - \overline{a}_B^K\right) f_A^K(a')$ with $a \in I_1, a' \in I_2$. Substituting yields
$$F_A^K(0) = 1 - \frac{1-b_A(K)}{1-b_B(K)} - \left(1 - \frac{b_B(1)}{b_A(1)}\right) \frac{b_A(K)}{1-b_B(1)} \frac{1}{K}. \quad \square$$

A.1.2 Proof of Proposition 1

Proof. First, $U_A(K) = 0$ since the high-cost type of A puts positive mass on investment 0. Second, $U_B(K) = (1 - b_B(K))F_A^K(0)$. Substituting for $F_A^K(0)$ from the proof of Lemma 1, yields the result. Finally, $U_i(1) = 1 - \bar{a}_A^1$. Again, substituting $\bar{a}_A^1$ from the proof of Lemma 1 yields the result. $\square$

A.2 Proof of Proposition 2

Proof. Take a full-settlement mechanism. Then, the sum of the low-cost types expected payoffs cannot exceed 1. A necessary condition is full participation which implies $1 \geq 2V^1$ and
$$1 \geq \frac{2(K-1)}{K}(1-p) \Leftrightarrow p \geq \frac{K-2}{2(K-1)}.$$
Conversely, suppose that $p \geq \frac{K-2}{2(K-1)}$. Then, the equal-split full-settlement mechanism satisfies the disputants participation constraints for any $p^V \geq p$. $\square$

A.3 Proof of Proposition 3

Proposition 3 follows from the construction in Section 4.

What remains is to prove Lemma 2 to 8 leading up to that construction and the steps omitted in the main text. We do so in order.

A.3.1 Proof of Lemma 2

Proof. The veto payoffs from Corollary 1 are convex in p^V. Applying Proposition 2 in Balzer and Schneider (2019) implies that it is without loss to assume full participation at the optimum under PBE. $\square$

A.3.2 Proof of Lemma 3

Proof. Take any rule γ and suppose settlement fails. Let $\rho_i := Pr(1|L) = \frac{p\gamma_i(1)}{Pr(L)}$ which is determined by γ. Bayes' rule implies that

$$b_A(1) = Pr(\theta_B = 1|\theta_A = 1) = \frac{p\gamma(1,1)}{p\gamma(1,1) + (1-p)\gamma(1,K)} = \frac{\rho_B}{\rho_A} b_B(1).$$

An equivalent relation for any $b_i(\theta_i)$ exists. By the law of probability, one of these equations is redundant and we are left with three independent equations and six unknowns. Solving for $b_B(\theta_B)$ and $b_A(K)$ provides the relations in the lemma. Because $b_i(\theta_i) \in [0,1]$, $b_A(1)$ is internally consistent. $\square$

A.3.3 Proof of Lemma 4

Proof. Each θ_i is indifferent over her strategy support on the equilibrium path. By monotonicity, off the equlibrium path she faces strict incentives when holding different beliefs.

If $b_i(1) < b_i(K)$ for some i, then $b_{-i}(1) < b_{-i}(K)$ by the relation in Lemma 3. If the deviating high-cost type chooses any action in $(0, \overline{a}_B^K)$, she has the same cost as on the equilibrium path, however she wins with larger probability, as $(1 - b_i(1)) > (1 - b_i(K))$. Thus, her payoff increases compared to on-path litigation. $\square$

A.3.4 Proof of Lemma 5

We prove the Lemma assuming $z_i > 0$ at the optimum. That guess is verified in (5) in the main text. Suppose that disputant i's participation constraint holds with strict inequality. Then, the designer can decrease both $z_i(1)$ and $z_i(K)$ by the same amount until the participation constraint binds without violating any other constraint.

Second, the high-cost types' incentive constraints hold with equality at the optimum. Otherwise, the designer could reduce $z_i(K)$ without violating any other constraint.

A.3.5 Proof of Lemma 6

Proof. Suppose condition (B) holds with strict inequality. Then, the designer could increase the share of each disputant and type. In turn, the low-cost type's expected payoff would increase. This allows her to decrease all γ's proportionally without changing (i) beliefs in litigation and (ii) incentives within ADR. $\square$

A.3.6 Proof of Lemma 7

Proof. $b_i(1) = b_i(K)$ implies that condition (M) holds. High-cost types' payoffs $\Pi_i(K) > 0$ because $z_i > 0$ by (5). Their (IR) constraint are redundant. For the last claim, we have to invoke Theorem 3 in Border (2007).

For every message $m \in \{1, K\}$, let $m^c := \{k \in \{1, K\} | k \neq m\}$ be its complement. Finally let $p(1) \equiv p$ and $p(K) \equiv (1-p)$. Fix some γ 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:

- $\forall m, n \in \{1, K\}$:

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

- $\forall m, i$:

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

Plugging in the values at the optimum defined in Section 4 verifies the inequalities. □

A.3.7 Proof of Lemma 8

Proof. A public signal implies a lottery over several (internally consistent) information structures.

First, take the set $\{\rho_A, \rho_B, b_A(1)\}$ that maximizes (4). Assume that it violates neither (IC^1) and is feasible. By the definition of an optimum this implies that no other information structure provides a higher value to (4). Thus, no lottery over information structures can improve upon that optimum either. Hence signals have no use.

For the case $p \leq \underline{p}$ see the solution at the end of Appendix A.3.8. □

A.3.8 Omitted Steps to prove Proposition 3

The regular case: $p \geq \underline{p}$**.** We first address the case $p \geq \underline{p}$.

Proof. **(Piecewise-)Linearity in** $b_A(1)$**.** Fix some (ρ_A, ρ_B). A disputant's winning probability, $F_i(\overline{a}_B^K|m)$, is linear in $b_A(1)$ since $(1 - b_i(m))$ is linear in $b_A(1)$. $\overline{a}_B^K$ is linear in $b_A(1)$ too and so are the payoffs. Finally, $\gamma_i(\theta_i)$ is linear in b_i and thus in $b_A(1)$. Thus, the RHS of (4) is linear in $b_A(1)$. Observe that due to the change of action the deviator's utility $U_i(1; K)$ has a kink at $b_A(1) = b_A(K)$. According to Lemma 3 $b_A(1) = b_A(K)$ implies $b_A(1) = \rho_B$.

No interior optimum. Linearity implies that it is sufficient to consider the boundary points of each interval for $b_A(1)$. That is, the optimal $b_A(1)$ is on one of these points:

$$\underline{b} = \frac{\rho_A}{K(1-\rho_B) + \rho_B}, \qquad \overline{b} = \frac{(K-1)(1-\rho_A) + \rho_B}{K(1-\rho_A) + \rho_A}, \qquad b^* = \rho_B.$$

We guess it is at $b^* = \rho_B$ and proceed.

Solving for ρ_i**.** Replacing $b_A(1)$ by ρ_B in (4) reveals a concave quadratic function for the RHS with independent first-order conditions. The unique solution is $(\rho_A, \rho_B) = ((1-p)/2, (1+p)/2)$. The derivative with respect to $b_A(1)$ is

$$\frac{\partial \text{RHS of } (4)}{\partial b_A(1)}|_{\rho^*} = \begin{cases} \frac{K(1-(p)^2)-(1-(p)^2)}{K(1+p)} & \text{if } b_A(1) < \rho_B \\ -\frac{K(1-(p)^2)-(1-(p)^2)}{K(1+p)} & \text{if } b_A(1) > \rho_B \\ \text{undefined} & \text{if } b_A(1) = \rho_B, \end{cases}$$

and $(\rho_A, \rho_B, b_A(1)) = ((1-p)/2, (1+p)/2, (1+p)/2)$ is a local optimum of the RHS of (4). Assuming $b_A(1) = \underline{b}$ and $b_A(1) = \bar{b}$, solving for the optimal ρ_i and comparing results implies that the solution is also a global maximizer for the RHS of (4). Under the symmetrizing signal (IC^1) redundant as shown in the main text and all other constraints are redundant by Lemma 6.

We have found an optimum whenever $p \geq \underline{p}$. Combined with Lemma 8 apart from the symmetrizing signal if $p > 1/3$, no other signals improve. Equation (6) provides the final argument verifying that $\alpha < 1$ if $p > \underline{p}$. Thus $\gamma \leq 1$. Section 4.5 provides the respective closed form terms.

The irregular case: $p < \underline{p}$. We now turn to the case $p < \underline{p}$ which implies $p < 1/3$. We first solve for the optimum ignoring additional information revelation by the designer.

If $p < \underline{p}$ we cannot find $\gamma \leq 1$ to implement $(\rho_A, \rho_B, b_A(1)) = ((1-p)/2, (1+p)/2, (1+p)/2)$. Therefore $\gamma(\theta_A, \theta_B) = 1$ for some (θ_A, θ_B). It is straightforward to show that $\gamma(1,1) \geq \gamma(\theta_A, \theta_B)$. Thus $p < \underline{p} \Rightarrow \gamma(1,1) = 1$ at the optimum. Moreover, $\gamma(1,1) = 1 \Rightarrow Pr(L) = \frac{p^2}{\rho_A \rho_B}$ (see e.g. (6) in combination with Section 4.5 to verify). Linearity in $b_A(1)$ remains to hold and so does the result $b_A(1) = \rho_B$. Finally, Lemma 6 implies that

$$Pr(L) = \frac{p^2}{\rho_A \rho_B} = \frac{p(2V-1)}{U_i(1;1) + (p-\rho_B)U_i(K;K) - p} \tag{7}$$

The designer's problem can be reduced to

$$\max_{\rho_A, \rho_B} \underbrace{U_i(1;1) + (p-\rho_B)U_i(K;K)}_{=\text{RHS of } (4)}$$
$$\text{s.t. } (7).$$

This is a well-defined biconvex problem that has a solution. The reason that the problem is biconvex (rather than convex) lies in the constraint. Biconvexity means that the problem is convex in each dimension, but not necessarily in any trajectory. Biconvexity is thus a generalization of convexity. However, biconvex problems preserve many properties from convex problems. For the specific case at hand, the only (potential) for non-convexities is in a trajectory in which we change ρ_A and ρ_B in opposite direction. That is, if we increase ρ_A and decrease ρ_B simultaneously.

The particular biconvex structure also implies that disclosing additional information structure is only beneficial to the designer if it is feasible (under equation (7)) to implement both an information structure with more asymmetry compared to the unconstrained optimum (that is, $\rho_A < (1-p)/2$ and $\rho_B > (1+p)/2$), and an information structure with less asymmetry compared to the unconstrained optimum (that is, $\rho_A > (1-p)/2$ and $\rho_B < (1+p)/2$). In such case the optimum including additional revelation can improve over the optimum without the revelation. The reason is that additional information revelation is able to convexify the constraint, since that constraint needs to hold only on average. However, the additional information also convexifies the objective which is concave and thus revelation hurts its value.

We conjecture that additional information revelation is not optimal based on simulations, but are unable to prove it analytically. The solution (including signals) follows from solving the problem

$$\min_{\rho_A, \rho_B, \lambda} vex \left[-\left(U_i(1;1) + (p - \rho_B)U_i(K;K)\right)(1+\lambda) + \lambda \left(\rho_A \rho_B \frac{(2V-1)}{p} - p \right) \right]$$

where $vex[f]$ denotes the lower convex envelope of f. □

A.4 Proof of Proposition 4

Take the following information structure (assuming $b_A = \rho_B$):

$$(\rho_A, \rho_B) = \left(p, 2p + \frac{1}{K-1} \right).$$

That information structure is feasible and consistent if $p < \overline{p}$. It implies $\alpha = 1$.

Moreover, under that information structure

$$Pr(L) = \frac{p}{2p + \frac{1}{K-1}} < 1/2.$$

A.5 Proof of Proposition 5

Proof. If $p > \underline{p}$, then $U_i(1;1) = (1 - \rho_A)\frac{K-1}{K}$ and $U_B(K;K) = (\rho_B - \rho_A)\frac{K-1}{K}$. Thus, the expected welfare conditional on litigation is

$$(\rho_a + \rho_B)U_i(1;1) + (1 - \rho_B)U_B(K;K) = 1/2(1 + 2p - p^2)\frac{K-1}{K}$$

which is larger than the ex-ante welfare

$$2p(1-p)\frac{K-1}{K},$$

because $p < 1/2$.

Moreover, observe that if disputants' expected payoffs conditional on litigation are larger after failed settlement, then litigation expenditures are smaller.

Finally, a low-cost type's conditional payoff is larger if $1/2 + p/2 \geq 1 - p$ which holds if $p \geq 1/3$. Otherwise low-cost types regret participation conditional on the litigation outcome. □

A.6 Proof of Proposition 6

Proof. The proposition follows from Lemma 8 and the fact that the symmetrizing signal eliminates the low-cost types incentive constraint. For $p \leq \underline{p}$ see the discussion at the end of Appendix A.3.8. □

A.7 Proof of Proposition 7

We augment the optimal ADR protocol from the main text by the following.

1. After reports (m_A, m_B) ADR announces $x_A(m_A)$ to disputant A.

2. ADR announces 0 to disputant B with probability $\gamma_B(m_B)$.
3. ADR announces $x_B(m_B)$ to disputant B with probability $1-\gamma_B(m_B)$.
4. ADR publicly announces if B rejects a non-zero proposal.

Disputant A learns nothing from the designer's proposal $x_A(m_A)$ and attaches the prior probability p to B's type distribution upon rejecting her proposal. She therefore has no incentive to reject the proposal.

Define $q_i(m_i) = Pr(\theta_i = 1|m_{-i}$ and settlement). If A learns about the deviation by B she holds the off-path belief q^D. We are looking for an off-path belief q^D such that B accepts her share and no double deviation (misreport and reject) occurs.

The law of iterated expectations implies that after the report $m_B = 1$ the following holds.

$$(1-p) = (1-\gamma_B(1))(1-q_A(1)) + \gamma_B(1)(1-\rho_A).$$

Multiplying both sides with $\frac{K-1}{K}$,

$$\underbrace{(1-p)\frac{K-1}{K}}_{=V^1} = (1-\gamma_B(1))(1-q_A(1))\frac{K-1}{K} + \gamma_B(1)\underbrace{(1-\rho_A)\frac{K-1}{K}}_{=U_i(1;1)}.$$

The low-cost type's participation constraint binds, and therefore $x_B(1) = (1-q_A(1))\frac{K-1}{K}$. If $q^D \geq q_A(1)$ the share a low-cost type receives from accepting is equal to her expected payoff from deviating and rejecting the share. Disputant $\theta_B = 1$ has no incentive to reject the proposal. Similar, type $\theta_B = K$ has no incentive to reject the proposal after pretending to be type 1. The high-cost types incentive constraint at the reporting stage are not affected.

After a report of $m_B = K$ the following holds by the law of iterated expectations.

$$(1-p) = (1-\gamma_B(K))(1-q_A(K)) + \gamma_B(K)(1-\rho_A).$$

We first show that the low-cost type does not gain by imitating the high-cost type at the reporting stage and then rejecting the proposed share. Multiplying both sides of the above equation with $\frac{K-1}{K}$ implies

$$V^1 - \gamma_B(K)U(1;1) = (1-\gamma_B(K))(1-q_A(K))\frac{K-1}{K} \geq z_B(K).$$

Where the last inequality follows from the low-cost type's incentive constraint at the reporting stage. If $\theta_B = 1$ reports K and rejects her continuation payoff is $(1-\min(q^D, q_A(K)))\frac{K-1}{K}$. Setting $q^D \geq q_A(K)$ implies an expected payoff at the reporting stage,

$$(1-\gamma_B(K))(1-q_A(K))\frac{K-1}{K} + \gamma_B(K)U(1;1) = V^1 = \Pi_B(1;1),$$

and provides her no incentives to deviate.

Moreover, if type $\theta_B = K$ deviates by rejecting the proposal after truthfully reporting, it is observed and ADR announces that deviation. Since $q^D \geq q_A(K)$, $\theta_B = K$ obtains utility $(q^D - q_A(K))\frac{K-1}{K}$ if she rejects the proposal. She prefers to accept her share if $(q^D - q_A(K))\frac{K-1}{K} \leq x_B(K)$. For an off-path belief $q^D = q_A(K)$ she is willing to accept any share.

The symmetrizing signal does not affect the outcome as players learn their assigned role upon observing x_i.

A.8 Proof of Proposition 8

Proof. We describe the result in Hörner, Morelli, and Squintani (2015) and how they map in the variables we are interested in. First, the result in Hörner, Morelli, and Squintani (2015, in particular their Lemma 1.) is symmetric throughout. Thus, $b_A(\theta) = b_B(\theta)$. Moreover, depending on the parameter values, Hörner, Morelli, and Squintani (2015) distinguish between two cases. In the first case, high-cost dyads settle for sure and litigation occurs only between low-cost types. In the second case, settlement fails for high-cost types with positive probability. Then, they face a low-cost type opponent with probability 1. Thus, $b_i(K) = 1$ if settlement fails for high-cost types. But since low-cost dyads never settle and sometimes face a high-cost type in litigation, it follows $b_i(1) < 1$. Thus, $b_i(1) \neq b_i(K)$. □

A.9 Proof of Proposition 9

Proof. By the revelation principle, no mechanism can outperform that of *Proposition* 3. The optimal mechanism is unique (up to the symmetrizing signal). A bargaining mechanism that implements the result of Proposition 3 implements the same payoffs as the optimal mechanism, because combining Lemma 5 and 6 implies that constraints have to hold with equality.

Take an equilibrium in which A obtains utility V^1. Suppose the low-cost type deviates by proposing an alternative $\hat{V}$. The following scenarios could result.

Any type of B accepts. In that case the sender obtains utility $1 - \hat{V} > V^1$. She is better off for sure and the initial offer cannot be an equilibrium.

Any type of B rejects. High-cost types only reject if their continuation utility is larger than $\hat{V}$. The continuation utility is $\max((p - p^{\hat{V}})\frac{K-1}{K}, 0)$ with $p^{\hat{V}}$ the belief after observing $\hat{V}$. Since any type rejects, A holds the prior belief p about B. Thus, rejection implies that $p^{\hat{V}} < p$. But then a low-type deviator A obtains continuation utility $(1 - p^{\hat{V}})\frac{K-1}{K}$ after that rejection. That implies she is better off for sure and the initial offer cannot be an equilibrium.

Only type 1 of B accepts. If 1 accepts $\hat{V}$ so does K because for any belief type 1 obtains (strictly) higher continuation payoff (since A keeps the prior belief p). She is better off for sure and the initial offer cannot be an equilibrium.

Only type K of B accepts. If type K always accepts the proposal, then A's payoff is larger than $(1-p)(1-\hat{V})$. For $\hat{V} < 1/K$ that payoff is larger than V^1. Hence always accept cannot be the high type's best response, otherwise that is a profitable deviation. Now assume that type K accepts only sometimes. This implies type K is indifferent between accepting and rejecting. In turn,

$$\hat{V} = (p^D - p^{\hat{V}})\frac{K-1}{K},$$

where p^K is the deviator A's belief about K's acceptance decision. Note that by the requirements of PBE, p^D has to be consistent with B's (the non-deviator's) behavior, which in turn has to be consistent with $p^{\hat{V}}$ the off-path belief about A

after observing her deviation. Solving the above equation implies that

$$p^D = \frac{K}{K-1}\epsilon + \beta_S.$$

Now suppose type K of B rejects with probability γ. To induce the belief p^D we need

$$\gamma = \frac{p}{1-p}\frac{1-p^D}{p^D}.$$

Calculating the low-cost deviator A's payoff from proposing $\hat{V} < 1/K$ we obtain

$$\begin{aligned}(1-p)(1-\gamma)(1-\hat{V}) + \Big(p + (1-p)\gamma\Big)(1-p^{\hat{V}})\frac{K-1}{K} &= \\ (1-\hat{V})\left(1-p-p\frac{(1-p^D)}{p^D}\right) + p\frac{(K-1)}{K}\frac{(1-p^{\hat{V}})}{p^D}&.\end{aligned}$$

The above described behavior can only be an equilibrium if any deviation is unprofitable for A. For that to hold we need

$$\begin{aligned}(1-\hat{V})\left(1-p-p\frac{(1-p^D)}{p^D}\right) + p\frac{(K-1)}{K}\frac{(1-p^{\hat{V}})}{p^D} \leq V^1 \\ \Leftrightarrow (1-\hat{V})\left(1-p-p\frac{1-p^D}{p^D}\right) \leq \frac{K-1}{K}\left(1-p-p\frac{1-p^{\hat{V}}}{p^D}\right)\end{aligned}$$

We know that $\hat{V} < 1/K \Leftrightarrow \frac{K-1}{K} \leq (1-\hat{V})$ which implies that a necessary condition for such an equilibrium to exist is

$$\begin{aligned}1-p-p\frac{1-p^D}{p^D} &\leq 1-p-p\frac{1-p^{\hat{V}}}{p^D} \\ &\Leftrightarrow p^{\hat{V}} \leq p^D,\end{aligned}$$

which contradicts indifference between accepting and rejection by type $\theta_B = K$. Thus, she is better off for sure and the initial offer cannot be an equilibrium.

Thus, we have shown that private bargaining performs strictly worse than optimal ADR under *any* off-path belief. □

References

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

Balzer, B. and J. Schneider (2019). "Belief Management and Optimal Arbitration". *mimeo*.

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.

Compte, O. and P. Jehiel (2009). “Veto constraint in mechanism design: inefficiency with correlated types”. *American Economic Journal: Microeconomics* 1, pp. 182–206.

Crémer, J. and R. P. McLean (1988). “Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions”. *Econometrica* 56, pp. 1247–1257.

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

Fey, M. and K. W. Ramsay (2011). “Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict”. *American Journal of Political Science* 55, pp. 149–169.

— (Oct. 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.

Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani (2009). “Mediation, arbitration and negotiation”. *Journal of Economic Theory* 144, pp. 1397–1420.

Grossman, S. J. (1981). “The informational role of warranties and private disclosure about product quality”. *The Journal of Law and Economics* 24, pp. 461–483.

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

Hylton, K. N. (2002). “An asymmetric-information model of litigation”. *International Review of Law and Economics* 22, pp. 153–175.

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

Klement, A. and Z. Neeman (2005). “Against compromise: A mechanism design approach”. *Journal of Law, Economics, and Organization* 21, pp. 285–314.

— (2013). “Does Information about Arbitrators’ Win/Loss Ratios Improve Their Accuracy?” *The Journal of Legal Studies* 42, pp. 369–397.

Meirowitz, A., M. Morelli, K. W. Ramsay, and F. Squintani (2017). "Dispute Resolution Institutions and Strategic Militarization". *Journal of Political Economy*, forthcoming.

Milgrom, P. R. (1981). "Good News and Bad News: Representation Theorems and Applications". *Bell Journal of Economics* 12, pp. 380–391.

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.

Shapira, O. (2012). "Conceptions and Perceptions of Fairness in Mediation". *S. Tex. L. Rev.* 54, p. 281.

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

Siegel, R. (2014). "Asymmetric all-pay auctions with interdependent valuations". *Journal of Economic Theory* 153, pp. 684–702.

Spier, K. E. (1997). "A note on the divergence between the private and the social motive to settle under a negligence rule". *The Journal of Legal Studies* 26, pp. 613–621.

— (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 J. Prescott (2019). "Contracting on litigation". *The RAND Journal of Economics* 50, pp. 391–417.

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

Stipanowich, T. J. (2004). "ADR and the "Vanishing Trial": the growth and impact of "Alternative Dispute Resolution"". *Journal of Empirical Legal Studies* 1, pp. 843–912.

Vasserman, S. and M. Yildiz (2018). "Pretrial Negotiations Under Optimism". *RAND Journal of Economics* forthcoming.

Zheng, C. (2018). "A necessary and sufficient condition for Peace". *mimeo.*

Zheng, C. and A. Kamranzadeh (2018). "The Optimal Peace Proposal When Peace Cannot Be Guaranteed". *mimeo.*