Essays in Information Economics

Inauguraldissertation zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften der Universität Mannheim

Johannes Schneider

Frühjahrssemester 2016

Abteilungssprecher: Prof. Dr. Carsten Trenkler
Referent: Prof. Volker Nocke, PhD
Korreferent: Prof. Dr. Martin Peitz

Tag der mündlichen Prüfung: 30.August 2016

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig angefertigt und die benutzten Hilfsmittel vollständig und deutlich angegeben habe.

Mannheim, 4. Juli 2016 Johannes Schneider

Acknowledgements

First and foremost, I am indebted to my advisor Volker Nocke. I thank him for his constant support and encouragement during the years in Mannheim. I am grateful for the many fruitful discussions we had and the detailed and constructive feedback he provided. His stimulus to engage in new ideas and to creatively revise existing ones was a tremendous help in developing my research. Without him this thesis would not have been possible.

Furthermore, I would like to thank Martin Peitz for agreeing to be my second supervisor and for constantly providing me with high quality feedback and encouragement on my ideas and ongoings. I thank Thomas Tröger for the many helpful comments and ideas he provided to guide the direction of my ideas. All of the three mentioned above exerted tremendous effort to help me getting started as a researcher and I am deeply thankful for all they have done.

I am deeply grateful to my co-author Benjamin Balzer. Working with him was always a pleasure and enlightenment. The uncountable discussions we had provided me with lots of new ideas and insights not only related to our joint work.

I would also like to thank the whole Micro Group at the University of Mannheim for an extraordinary research environment. My research benefited incredibly from seminars and informal discussions. I am truly grateful for the many comments received. The never ending patience to attend yet another presentation or to look at yet another draft is highly appreciated. A special thank you goes to the early micro lunch group consisting of Benjamin Balzer, Felix Jarman, Vincent Meisner, and Christoph Wolf for daily discussions as well as to Takakazu Honryo, Andras Niedermayer, Stefan Penczynski, Emanuele Tarantino, Philipp Schmidt-Dengler, and Andre Stenzel for constructive comments.

I would also like to thank my fellow doctoral students for setting up such a great environment and for the interesting discussions I had over the years both about life and economics. It has been a great experience to discuss economic ideas across fields and to receive feedback from an engaged audience. In particular, I would like to thank Florian Exler, Jasper Haller, Wieland Hoffmann, and Henning Roth for the extensive discussions we had about economics, life and everything.

I would like to thank my parents without whom I would not be at this stage in many more dimensions than the obvious existence problem. Most of all, I would like to thank Leonie for her love and support before, during, and beyond this project.

Contents

1 General Introduction

Information plays an important role in many economic problems. A backbone of economic theory is that agents optimize their behaviour subject to the information they possess at the time of decision making. Due to several frictions information is often incomplete in a sense that not everyone acting in an economic environment has access to the same information set. Such information asymmetries may in fact be responsible for the breakdown of markets in which trade would otherwise be beneficial. The most famous example of this is the "market for lemons" going back to Akerlof (1970). Broadly speaking, economic theory has dealt with problems of asymmetric information in two different ways: positively and normatively. Positive analysis mainly falls into the area of game theory making use of so called "Bayesian Games" describing games where players hold a "belief" about the state of the world (see Fudenberg and Tirole (1991) for a discussion on this). The scope in this area is to analyse how the rules of a game and asymmetric information of the participants interact. The second branch, normative analysis, is to a large extend inspired by the theory of mechanism design, asking which rules on a game are best to achieve a given goal. An important corner stone of the mechanism design literature is the so-called "revelation principle". The revelation principle states that any possible indirect mechanism implements an allocation that can also be achieved using a direct revelation mechanism. A direct revelation mechanism is a mechanism in which players simply report their entire private information directly and privately to the *mechanism*. The mechanism enforces an allocation based on the agents reports, according to a certain rule. The mechanism can be seen as a contract, the players agreed upon ex-ante such that both the mechanism and the players fully commit to obeying this contract.

This thesis consists of three self contained articles. Chapters 2 and 3 are joint work with Benjamin Balzer, while Chapter 4 is single authored. All three chapters address issues of information economics both from a normative (chapters 2 and 3) and a positive (chapter 4) perspective. A common theme over the three chapters is that they centre around the question how a particular bit of information is interpreted by economic agents in light of the surrounding environment. I show that the interpretation of information, can be very different and depends on the surrounding. The underlying reason for this is that preferences over outcomes are often misaligned to some extend in economic environments. Thus, any form of communication between the different players is likely to have a strategic component. The receipients of the information take this knowledge, and the knowledge of the (sometimes constraint) information transmission technologie, into account when making their decisions. In Chapter 2, I address this within the field of Law and Economics, while chapter 4 is within the field of Industrial Organization. Chapter 3 serves as a bridge between the two by considering issues present in both fields on a more abstract level. Both, in Law and Economics and Industrial Organization, one can find many situations in which competitors interpret any sort of communication as a signal about the competitor's type. At the same time, private information is typically valuable to competitors within the fixed environment and communication mechanisms are embedded in the context of a greater institution which is why the issues discussed above are relevant in these areas. Next, I briefly comment on the three different chapters.

Chapter 2 contributes to the outlined theme by showing how litigation shapes the design of an optimal alternative dispute resolution mechanism (ADR). ADR is present in many legal disputes and describes any attempt to settle the dispute outside court with support of a third party. ADR helps to overcome overburdened legal systems that can be observed in many countries. An overburdened legal system imposes large costs on society as it limits access to the legal system. Thus, many countries aim at minimizing the number of cases that enter formal litigation and try to settle as many as possible outside court. At the same time it is a pillar of most modern societies that anybody who wants to bring her case to litigation can do so. This is called the rule of law. There is a debate among practitioners and policy makers that centres around the following question: How can one best design an alternative dispute resolution mechanism that minimizes the number of cases that enter formal litigation while keeping the rule of law in place? We use the techniques of mechanism design to characterize the litigation minimizing ADR ensuring that the rule of law is kept in place. The rule of law provides ADR with a non-cooperative outside option, the formal litigation process. In this chapter we investigate how the design of ADR interacts with the players actions in the litigation process following a breakdown of ADR. The interaction is a consequence of the expensive and highly competitive nature of litigation: The player providing the better argument in her favour is going to win litigation. Thus, the actions players choose very much depend on their beliefs about the other player. These beliefs are not only updated in light of the existence of ADR, but also by the exact specification of ADR. Depending on what happened in ADR, players may thus react differently in litigation after breakdown. Formally, this means that the litigation game serves as a belief dependent outside option to ADR.

Chapter 3 generalizes the findings of chapter 2 by characterizing optimal conflict management for a general class of conflict games. Economically, conflicts are relevant in many settings and often the default resolution option provides unsatisfying results and externalities on other, not directly involved, agents. Therefore, conflict management has a long tradition in most areas were conflicts occur. In chapter 2 we have seen that players interpret breakdown of conflict management in light of the underlying game. In this chapter we show that this insight carries over to more generalized settings. Again, beliefs after breakdown play a crucial role. Further, we simplify the problem by showing that there are several equivalent versions of the problem that allow us to use different techniques for finding a solution. Moreover, we provide necessary and sufficient conditions for implementation of any solution that is derived in expected values. We show that if the players cannot commit to ignore publicly available information, the mechanism designer can always ensure full participation at the optimum by convexifying the players rejection payoffs via publicly available signals. We offer a general approach to limit the set of solution candidates to those belief choices that provide a concave expected payoff in the conflict game. Our findings provide a guideline for practitioners on how to design the conflict management mechanism as a function of the underlying game.

Chapter 4 contains a positive analysis of a model of project choice. I look at this in a merger context in which a firm can choose between two different merging partners. The merger serves as a project and the proposing firm can choose between two different mergers. The choice is made endogenously and I assume efficient bargaining among firms who then propose the merger to an antitrust authority that can accept or reject the proposed merger. An accepted merger is implemented. If a merger is rejected the status quo remains and the firm may propose a merger again in the subsequent period. I am interested in how the dynamics of the game shape the authoritie's understanding of the proposed mergers and find that in such a game there are essentially two types of equilibria. In one equilibrium type the authority screens by accepting some proposals with a probability smaller one which leads the firm to pander towards the authorities preferences. This equilibrium is called the pandering equilibrium. A second class is that of "waiting equilibria" in which the firm signals by waiting that she has high synergies in the larger merger. While the pandering equilibrium is the most desired equilibrium of the firm, the most desired equilibrium of the authority always lies in the class of waiting equilibria. The model offers a new insight on the dynamics of project proposal. Using the idea of delay as a signal, I provide a model that shows that, contrary to conventional wisdom, delay may beneficial for society. My results offer an explanation as to why firms push towards fast processes in merger control. At the same time, I show that ruling out waiting equilibria by limiting the time-horizon may harm society. More generally, my model offers a novel theory of project choice. Contrary to existing project choice models, I incorporate a notion of signalling by waiting into the problem. I show that if it is impossible for the sender to end the game after her first proposal, an uniformed decision maker may gain from additional time periods. Thus, even in the absence of verifiability, time can help a party to screen projects. In fact, the longer the time-horizon the better for the uninformed receiver.

2 Alternative Dispute Resolution in Contests

with Benjamin Balzer

2.1 Introduction

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

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

ADR breaks down sometimes to screen parties and to ensure truth-telling during ADR.

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

Thus, most jurisdictions encourage parties to engage in some form of ADR before starting the formal litigation process. The U.S. Alternative Dispute Resolution Act of 1998 states that courts should provide litigants with ADR-options in all civil cases. ADR is defined as “any process or procedure, other than an adjudication by a presiding judge, in which a neutral third party participates to assist in the resolution of issues in controversy” (Alternative Dispute Resolution Act, 1998). However, ADR supplements the “rule of law” rather than replacing it. Ultimately, each party has the right to return to formal litigation.[1] Hence, ADR indeed happens “in the shadow of the court:” whenever no settlement is achieved via ADR, litigants return to the traditional litigation path.

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

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

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

evidence choice and vice versa. Hence, litigation strategies after ADR-breakdown are a function of the players' *belief system*.

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

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

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

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

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

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

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

Outline. After discussing the literature in Section 2.2, we set up the model in Section 2.3 and derive the optimal mechanism in Section 2.4. Subsequently, we discuss the findings in Section 2.5 and several extensions in Section 2.6. Section 2.7 concludes.

2.2 Related Literature

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

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

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

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

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

The second strand of literature we relate to is that of mechanism design with endogenous outside options, i.e. mechanisms which cannot fully replace an underlying strategic game. Similar to Cramton and Palfrey (1995) and Celik and Peters (2011, 2013), we consider a mechanism that needs to be ratified by both parties. Without mutual consent, parties play the litigation game. In our model moreover, mediation sometimes breaks down and parties are referred to the underlying game. Breakdown is informative as in Cramton and Palfrey (1995) and Celik and Peters (2011). While Cramton and Palfrey (1995) are interested in finding worst off-path beliefs, Celik and Peters (2011) show that for some games it is optimal to design a mechanism without full participation. In our model, both channels are not present and full participation is optimal. Instead, we explore an additional channel: we ask how on-path references to the default game *by the mechanism* interact with the belief structure of the players after breakdown.

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

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

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

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

2.3 Model

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

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

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

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

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

and

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

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

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

Formally, the mediator is a mechanism designer who cannot enforce actions in the contest. In principle, we could allow the mediator to send players non-binding recommendations within the contest. It is without loss of generality to abstract from such recommendations as they would induce a communication equilibrium in the litigation game and all communication equilibria in all-pay contests are payoff equivalent to the unique Nash equilibrium.[8]

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

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

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

[8]See Pavlov (2013), especially Proposition 6 for details on this. Uniqueness of the Nash equilibrium is discussed in section 2.4.1.

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

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

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

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

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

Interim individual rationality of the players is assumed for the ease of notation, only. In Section 2.6 we show in line with the argument by Hörner, Morelli, and Squintani (2015) that assuming ex-post individual rationality changes results only arbitrarily little.

Finally, the assumption that the mediator aims to minimize breakdown is in line with the theoretical literature on conflict resolution. Courts have an enormous backlog in pending cases. Mainly because of the backlog, the time from filing to trial takes typically more than two years. Decreasing the number of court

[9] In Section 2.6 we show in an extension that assuming ex-post individual rationally can changes results arbitrarily little.

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

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

2.4 Analysis

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

We organize the remainder of this section as follows: for each step we first state its result and provide an intuition thereafter. Formal proofs are provided in Section 2.C.

2.4.1 Equilibrium Characterization of the Continuation Game

The continuation game after breakdown of mediation is an all-pay contest with type-dependent probabilities as defined in Section 2.3.

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

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

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

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

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

[11] Siegel (2014) shows that in principle little can be said if (M) is violated. In our setting, the mediator can only induces Bayes' plausible belief structures. Thus, it is actually possible to characterize the non-monotonic equilibria explicitly. We characterize them in the Section 2.C.

- *both players play mixed strategies with piecewise constant densities on at most three subintervals of* $(0, \Delta_{l,l}]$.

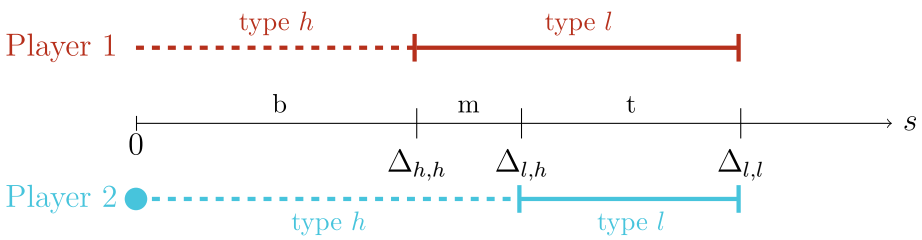

Figure 2.1: *Strategy support of player 1 and 2 with type-dependent priors.*

The Lemma is a direct application of Siegel (2014) to our setting. Figure 2.1 summarizes the equilibrium strategies. The horizontal axis depicts the score s. The dark-red and the light-blue line denote equilibrium strategy support for both players if player 1 is more likely to have low-cost. Player 1 (dark-red line at the top), type h (dashed part), is indifferent for all scores on the bottom interval b from 0 up to and including $\Delta_{h,h}$. This is the lower bound for the score of $1l$ (solid part) who is indifferent on all scores on intervals m and t up to and including $\Delta_{l,l}$ given the strategy of player 2. Player $2h$ (light-blue dashed line at the bottom) is indifferent between a score of 0 (indicated by the dot) and on intervals b and m up to and including $\Delta_{l,h}$. Player $2l$ is indifferent on interval t. If players become ex-ante symmetric, interval m vanishes, the mass point at 0 disappears, and strategies become fully symmetric.

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

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

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

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

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

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

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

Moreover, utilities are linear in beliefs, if beliefs are type-independent. If beliefs are symmetric, $F_{2,h}(0) = 0$.

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

2.4.2 Deviator Payoffs in the Continuation Game

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

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

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

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

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

Definition 2. *Deviation contest:* the contest is called deviation contest of player ik if player i, type k holds a belief $p_{-i}(l|\neg k)$ that is the same belief that player i, who

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

is *not* k, holds on-path. This belief is called the deviator's belief. Player $-i$, however, holds her on-path belief $p_i(l|k)$ about player i. Thus, generically, there is no common knowledge of beliefs in this contest.

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

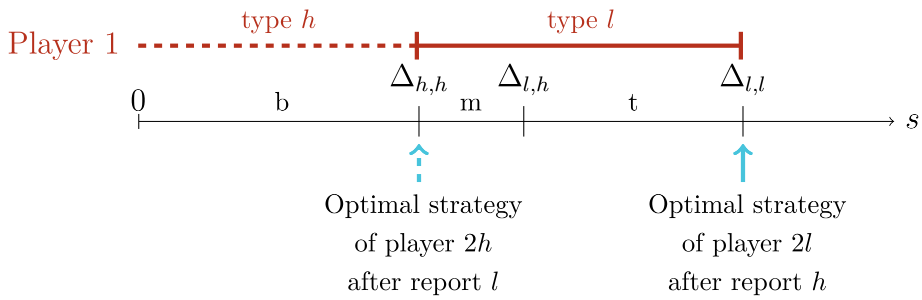

Figure 2.2: *Optimal behavior in the deviation contest of player 2 if $p_1(l|h) > p_1(l|l)$. Notice that the deviation strategies are conditional on $2l$ reporting h and $2h$ reporting l without player 1 noticing.*

Figure 2.2 illustrates the optimal strategies for player 2's deviation contest in case it is more likely that l-types appear after an h-report, i.e. $p_1(l|h) > p_1(l|l)$. The horizontal axis describes the scores, the dark-red line the strategy of player 1, which is the same as in equilibrium. The light-blue, dashed arrow points to the unique best response of player $2h$ who reported l, the solid arrow to that of player $2l$ who reported h.

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

Whenever reporting h increases the likelihood to meet an l-type opponent for player 2, reporting l must increase the likelihood to meet an h-type, i.e. $p_1(h|l) > p_1(h|h)$. Similar to the case of $2l$, a deviation by $2h$ makes her increase the score against an h-type (interval b in Figure 2.2), but decrease it against an l-type (interval m in Figure 2.2), since those occur less likely. Thus, her optimal response is $\Delta_{h,h}$ which leads to a win against all h-types. High-cost types occur with higher probability as $p_1(h|l) > p_1(h|h)$, and hence, $2h$ prefers the deviation contest to the on-path contest.

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

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

2.4.3 Rewriting the Problem

We now turn to the problem of the designer. Note that the problem is highly non-convex and standard techniques do not apply. To be able to characterize the solution we need to transform it to a tractable problem. We do so in several steps. As the transformation is a series of technical issues we proceed as follows. First, we state the proposition describing the reformulated problem. Second, we state the original problem. Third, we provide a brief, non-technical comment on each transformation step in the main text. We refer the interested reader to Section 2.A for the corresponding detailed description of the transformation including the intermediate lemmas.

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

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

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

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

The Original Problem of the Mediator. As the mechanism needs to pass a ratification stage it is not necessarily without loss of generality to assume full participation. Given the payoff structure of the litigation game, however, we can use a result of Celik and Peters (2011) to conclude that full participation is indeed optimal in our setting, the corresponding lemma stating this result is included in Section 2.A. Given full participation, the mediator's problem is

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

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

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

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

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

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

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

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

Whenever the value of vetoing is smaller than 1/2 for low-cost types, however, the mediator could offer parties a sharing rule of $(1/2, 1/2)$ for each type-realization and settle all cases. To make the problem interesting we make the following assumption.

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

Assumption 1 translates into the following condition on parameters: $\kappa > \frac{2-2p}{1-2p}$.

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

$$\Pi_i(k|m) = z_i(m) + \gamma_i(m)U_i(k|m), \tag{2.1}$$

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

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

of the breakdown probabilities conditional on the opponents type

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

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

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

and analogously for player 2. Equation (2.1) shows how optimal mediation relies on the litigation game. While the value of settlement, z_i, is similar to transfers in standard mechanism design, the utility of the contest continuation game is the screening device.

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

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

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

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

2.4.4 Optimal ADR-Mechanism

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Deviation utilities have a kink at type-independent beliefs by the all-pay contest assumption. The kink is a direct consequence of Lemma 3 as deviating high-cost players are only indifferent for type-independent beliefs. High-cost types score at the upper end of their on-path equilibrium strategy set for lower values of $p_1(l|l)$ and at the lower end for higher values of $p_1(l|l)$. Hence, for type-independent beliefs their utilities are non-differentiable and obtain a minimum. The left panel of Figure 2.3 plots the deviation utilities as a function of $p_1(l|l)$.

If we combine the effects on breakdown probabilities $\gamma_i(m)$ and contest utilities, we find that the minimum at type-independent beliefs prevails. The result can

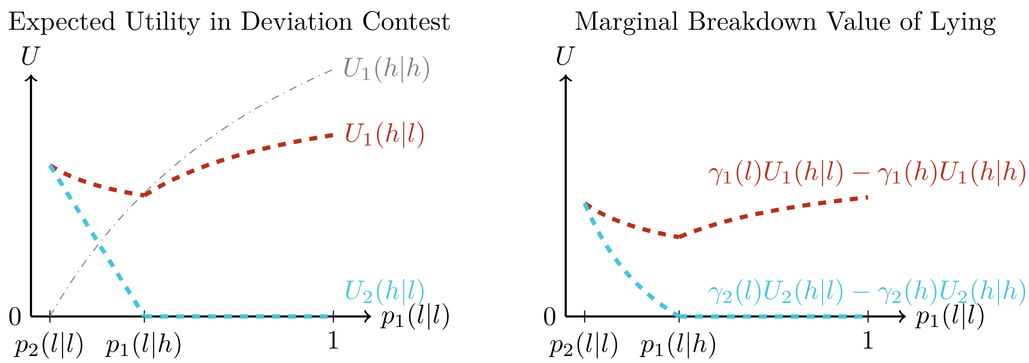

Figure 2.3: *The left panel depicts the high-types deviation utilities as a function of $p_1(l|l)$. The right panel depicts the marginal breakdown-value of lying. Red is for player 1, blue player 3. The gray line in the right panel is the on-path utility of the high-cost type of player 1.*

best be seen if we consider the marginal breakdown-value of lying. This breakdown value is the right-hand side of the following representation of the high-types incentive constraint,(IC_i^h),

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

The left hand side can be interpreted as the marginal settlement value of truth-telling which matches the right hand side being the marginal breakdown value of lying. The right panel of Figure 2.3 displays the marginal breakdown value of lying and illustrates how the minimum property prevails and type-independent beliefs are optimal. We can thus simplify notation and define ρ_i to be the probability that player i is the low-type post-mediation.

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

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

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

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

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

$$U_i(l|m) = (1 - \rho_2)\frac{\kappa - 1}{\kappa}, \qquad U_1(h|m) = (\rho_1 - \rho_2)\frac{\kappa - 1}{\kappa}. \tag{2.2}$$

These expressions are useful in the argument for asymmetry of the optimal mechanism which we turn to next. We discuss the general argument non-formally to provide a good understanding of the qualitative results. A more detailed and formal analysis is in Section 2.B.

The main argument for asymmetry lies in the structure of a contest. A symmetric contest is expected to be tight: parties expect to be matched with an opponent of similar strength and the marginal value of investment is high. By contrast, an asymmetric contest appears to be less tight, and the marginal value of investment is lower for both parties. This imposes an externality, especially for the high-cost type of the ex-ante stronger player. Her opponent's high-cost type is going to increase her investment but remains at a utility of 0 as she is the weakest of all player-types. Thus, the stronger player's h-type can reduce the investment and still has a reasonable chance to win the contest as the opponent believes she likely faces a low-cost type. This effect can be seen by inspecting equations (2.2). If we start in a symmetric setting and unilaterally increase the belief put on player 1, then l-types would not benefit in terms of expected utilities and neither would $2h$. However, player $1h$ actually achieves a positive utility in such a case which she would not under symmetry.

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

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

$$
\begin{aligned}
z_2 &= V(l) - \frac{1-\rho_2}{\rho_1}\frac{\kappa-1}{\kappa}p\gamma(l,l) \\
z_1 &= \underbrace{V(l) - \frac{1-\rho_1}{\rho_2}\frac{\kappa-1}{\kappa}p\gamma(l,l)}_{\text{symmetric part}} + \underbrace{(\frac{p}{\rho_1} - \frac{p}{\rho_2})\frac{\kappa-1}{\kappa}p\gamma(l,l)\kappa}_{\text{asymmetric part}}.
\end{aligned}
$$

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

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

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

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

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

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

$$\Big(\gamma_i(l) - \gamma_i(h)\Big)U_i(h|h) \leq \Big(\gamma_i(l) - \gamma_i(h)\Big)U(l|l). \tag{2.3}$$

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

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

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

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

Lemma 6 states that the probability of breakdown for low-types is larger than the probability of breakdown for high-types, i.e. $\gamma_i(l) \geq \gamma_i(h)$. In such a case one individual feasibility, (EPI), which is one of the two constraints coming from the reduced form, is always satisfied. Section 2.C provides details on this.

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

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

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

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

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

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

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

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

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

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

The cutoffs are given by:

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

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

As the prior p increases, the solution ρ_2 increases, too, and $\rho_2 \geq p$ does not bind anymore. The resource constraint, $\gamma^* \leq 1$, however, still does. If p is larger than p'', the solution of Lemma 5 can be implemented directly. For $p > 1/3$, on the other hand, low-cost types of player 2 have an incentive to misreport given the protocol from Lemma 5 which means that (IC_2^l) binds and the belief on player 2 is set to the prior, $\rho_2 = p$. The left panel of Figure 2.4 illustrates the findings. The dashed line plots the optimal protocol according to Lemma 5 whereas the solid line is the full model.

2.5 Discussion of the Results

Comparative Statics. Figure 2.4 depicts the probability of litigation under the optimal mechanism both as a function of the prior, p (left panel), and as a function of the distance between low and high cost, κ (right panel). The different colors indicate the different regimes as discussed in Lemma 7. Red and blue (for $p < p''$) denote the areas in which the resource constraint, $\gamma^* \leq 1$, binds; green (to the right of p''') is the area in which $2l$'s incentive constraint binds and black is the area

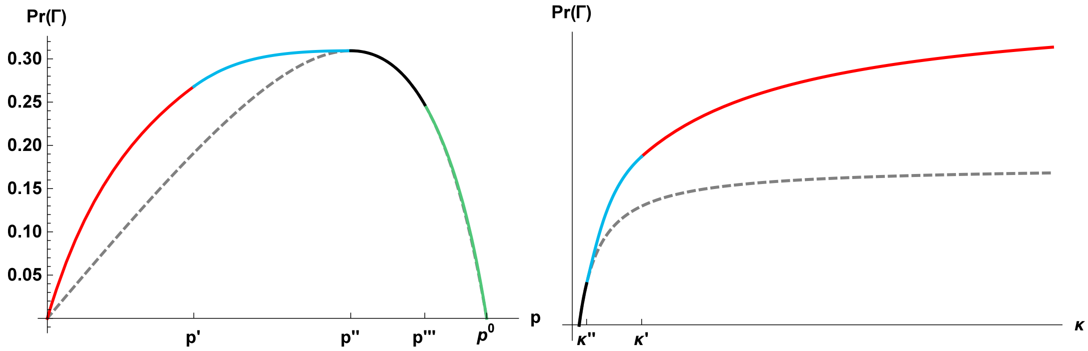

Figure 2.4: *Ex-ante probability of the contest as a function of* p *(left panel) and* κ *(right panel). The dashed line describes the situation of the unconstraint problem* $(P1')$ *as in Lemma 5. The green solid line corresponds to Lemma 6. All solid lines together display the result of Lemma 7.*

in which $(P1')$ is solved "unconditionally" as in Lemma 5. p^0 indicates the point at which Assumption 1 starts to fail and the mediator achieves full settlement for $p > p^0$. For comparison, the dotted line depicts the solution ignoring (IC_i^l) and $\gamma^* \leq 1$.

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

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

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

(a) Breakdown Probab. by Type (b) Expeceted Share by Type (c) Settlement Valuation by Type

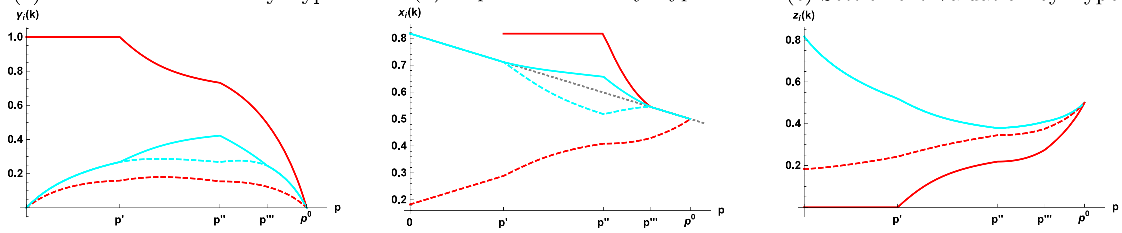

Figure 2.5: (a) *Expected Contest Probability,* (b) *Expected Share conditional on settlement taking place, and* (c) *Valuation of Settlement by player-type as a function of the ex-ante probability of being a low-cost type. Solid lines depict low-cost types, dashed lines depict high-cost types. Dark-red is player 1 and light-blue is player 2. The dotted gray line in* (b) *is the value of vetoing for low-cost type players. In* (c)*, player 2h has the same settlement value as 2l by incentive compatibility.*

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

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

Figure 2.5 illustrates the results of Proposition 3 as a function of the prior distribution. The left panel (a) describes breakdown utilities, the middle panel (b) expected shares conditional on settlement, $x_i(m) \equiv z_i(m)/(1-\gamma_i(m))$, and the right panel (c) the settlement valuation. Dark-red lines are for player 1 and light-blue lines for player 2. Dashed lines indicate high-cost types, solid lines indicate low-cost types. The linear, gray, dotted line in panel (b) denotes the value of vetoing for the l-type, $V(l)$.

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

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

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

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

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

2.6 Extensions

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

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

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

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

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

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

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

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

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

The mediation protocol developed in Section 2.4 does not directly carry over to PAIR and EPIR. In fact, given these commitment schemes, the mediator profits from the ability to communicate to parties even after ADR breaks down. If this is the case, the mediator can give parties non-binding recommendations for the play of the contest and by that restore the outcome under full-commitment. The modified game thus follows a slightly enhanced timeline:

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

Note that since the mediator first observes the behavior of the players with respect to the announced share she has the ability to detect a deviation in this stage (other than in the reporting stage). To restore the result of Section 2.4 the mediator uses the following slightly more sophisticated mechanism(s).

To find the optimal PAIR mechanism, we need to define a convex combination of the protocol derived in Section 2.4 and its mirror image switching roles of player 1 and 2. Define $\hat{\mathcal{X}}_\lambda$, a mediation protocol such that $\mathcal{X}_i$ applies with probability λ and $\mathcal{X}_{-i}$ with probability $(1-\lambda)$. $\mathcal{X}_i$ denotes a mediation protocol similar to the one discussed in Theorem 1. When mediation is successful, player i is treated as "player 1". To trigger litigation in this protocol, the mediator offers a share of 0 to at least one of the players. This share is going to be rejected such that parties move to the litigation game.

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

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

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

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

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

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

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

2.7 Conclusion

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

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

Trust in the mediator's discretion is an important driving force of the success of a mechanism.

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

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

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

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

Appendix

2.A Details on Rewriting the Problem

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

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

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

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

Given the constant outside option, the channels identified by Cramton and Palfrey (1995) and Celik and Peters (2011) are not present in our model as off path beliefs are less important.

Reduced Form Problem à la Border (2007). We reduce the problem by replacing the settlement shares, X_i, by the settlement values, z_i. For any given matrix of breakdown probabilities, G, this reduction is possible if and only if each settlement share is both individually feasible (condition (F), below) and ex-post implementable (condition (EPI), below). The following lemma states these conditions. With some abuse of notation, let $p(m)$ be the ex-ante probability that player i is of type m, that is $p(l) = p$ and $p(h) = 1-p$.

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

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

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

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

$$z_i(m) \leq 1 - \gamma_i(m) \tag{IF}$$

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

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

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

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

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

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

$$z_i(h) + \gamma_i(h)U_i(h|h) \geq z_i(l) + \gamma_i(l)U_i(h|l). \tag{IC_i^h}$$

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

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

$$z_i(l) + \gamma_i(l) U_i(l|l) \geq V(l). \tag{PC_i^l}$$

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

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

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

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

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

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

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

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

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

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

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

$$\underbrace{2V(l)-\gamma(l,l)Q(P)}_{\text{LHS of }(AF)} \le \underbrace{1-\gamma(l,l)R(P)}_{1-Pr(\Gamma)}, \tag{2.4}$$

where $\gamma(l,l)Q(P) := \sum_i \sum_m p(m)z_i(m) - 2V(l)$. Assumption 1 implies $Q(P) \geq R(P)$ and we can reformulate

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

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

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

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

2.B Forces of Asymmetry

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

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

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

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

consider the total resources distributed

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

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

Combining the two equations yields

$$\underbrace{2V(l) - 1}_{=\nu} = 2\gamma(l)U(l|l) - Pr(\Gamma). \tag{2.5}$$

Using Corollary 2 and (2.2) we can rewrite equation (2.5)

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

Solving for $Pr(\Gamma)$ yields

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

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

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

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

The first equality comes from (PC_2^l) and second from the results of Corollary 2 and the equations in (2.2).

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

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

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

$$\begin{aligned} pz_1(l)+(1-p)z_1(h) &= z_1(l)+(1-p)\Big(\gamma_1(l) - \gamma_1(h)\Big)U_1(h|h) \\ &= z_1(l) + \gamma_1(l)\left(1 - \frac{p}{\rho_1}\right)U_1(h|h) \end{aligned} \tag{2.6}$$

where the first uses (IC_1^h) and the second uses Corollary 2 to simplify. Simplifying this using (PC_1^l), (2.2), and $U_i(\cdot,\cdot)$ yields

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

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

2.C Proofs

Proof of Lemma 1

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

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

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

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

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

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

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

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

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

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

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

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

but also that

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

Thus, it must be the case that

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

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

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

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

which is ruled out by (M).

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

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

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

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

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

$$\begin{aligned} & U_i(h, s = s_i^h) \geq U_i(h, s = s_i^l) \\ \Leftrightarrow \quad & \sum_k p_i(k|h) F_{-i,k}(s_i^h) - \kappa c_l s_i^h \geq \sum_k p_i(k|h) F_{-i,k}(s_i^l) - \kappa c_l s_i^l \\ \Leftrightarrow \quad & \sum_k p_i(k|h)(F_{-i,k}(s_i^h) - F_{-i,k}(s_i^l)) \geq \kappa c_l (s_i^h - s_i^l). \end{aligned} \tag{2.7}$$

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

$$\sum_k p_i(k|l)(F_{-i,k}(s_i^h) - F_{-i,k}(s_i^l)) \leq c_l (s_i^h - s_i^l). \tag{2.8}$$

But, as $F_{-i,k}(\cdot)$ is always positive and $p_i(h|\cdot) = 1 - p_i(l|\cdot)$, inequalities (2.7) and (2.8) only hold if

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

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

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

and that of the middle interval $(\Delta_{h,h}, \Delta_{l,h}]$ is

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

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

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

Proof of Lemma 2

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

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

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

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

Proof of Lemma 3

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

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

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

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

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

The proof of the proposition is along the lines described in section 2.A.

Proof of Lemma 8

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

Proof of Lemma 9

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

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

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

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

We want to implement **p** by the list **P** containing

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

where

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

Plugging in yields,

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

To state the conditions let in addition

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

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

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

- $\forall m \in \{h, l\}$ and $i = 1, 2$:
$$z_i(m) \leq 1 - \gamma_i(n) \tag{IF}$$

- $\forall i = 1, 2$
$$z_i(l)p(l) + z_i(h)p(h) \leq 1 - Pr(\Gamma) \tag{BC_2}$$
$$\sum_{i\in\{1,2\}} \sum_{k\in\{l,k\}} p(k)z_i(k) \leq 1 - Pr(\Gamma) \tag{AF}$$

- $\forall \{m, n\} \in \{h, l\}^2$ and $i = 1, 2$:
$$\sum_{k\in\{l,h\}} p_i(k)z_i(k) + pz_{-i}(n) \leq 1 - Pr(\Gamma). \tag{BC_4}$$

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

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

Hence,

$$z_i(h) \leq \gamma_i(l)U_i(l|l) - \gamma_i(h)U_i(l|h) + z_i(l) \leq (\gamma_i(l) - \gamma_i(h))U_i(l|l) + z_i(l), \tag{2.9}$$

where the last equality follows from Lemma 2.

If $\gamma_i(l) \geq \gamma_i(h)$ and $z_i(l) \leq 1 - \gamma_i(l)$ we can rewrite (2.9) to

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

which indeed is equation (IF). □

Proof of Lemma 10

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

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

Recall that

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

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

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

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

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

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

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

Proof of Lemma 11

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

$$p_1(l|l) = \frac{p\gamma(l,l)}{p\gamma(l,l) + (1-p)\gamma(h,l)}. \tag{2.10}$$

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

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

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

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

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

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

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

which after rearranging yields the desired

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

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

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

We proof the proposition in several steps. In line with the text, we first solve the "unconstrained problem" $(P1')$ which is also the proof of Lemma 5.[17] After that we introduce (IC_i^l) and proof Lemma 6 before finally introducing the remaining constraints with the proof of Lemma 7. Throughout this proof we make use of the following lemma

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

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

Proof of Lemma 4

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

$$p_2(l|h) = \frac{\Big(1 - p_1(l|l)\Big)\, p_2(l|l)p_1(l|h)}{p_2(l|l)p_1(l|h) - p_1(l|l)\Big(1 - p_2(l|l) - p_1(l|h)\Big)}. \tag{2.11}$$

$p_2(l|h) > p_2(l|l)$ if equation (2.11) divided by $p_2(l|l)$ is larger 1 that is

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

Rewriting yields,

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

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

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

Proof of Lemma 5

Proof. Notice that the unconstrained problem is $(P1')$ is a problem of three elements $P = (p_1(l|l), p_2(l|l), p_1(l|h)$ only, as the fourth is directly defined via consistency equation (C). We calculate the unconstrained optimum in several steps. First, we show that at the optimum the objective is not differentiable with respect to at least one of the three choices variables. Second, we show that if $p_1(l|l)$ is either $p_1(l|h)$ or $p_2(l|l)$m then it is $p_1(l|l) = p_1(l|h)$ and calculate this optimum. Finally, we show that a deviation to $p_1(l|l) = 1$ is not optimal.
Step 1: No optimum in the differentiable interior exists. To proof this claim we are going to proof that the objective $Obj(P) := R(P)v/(Q(P) - R(P))$ is locally concave at any critical point in $p_1(l|l)$ in what we call the "differentiable interior", meaning that such a critical point is in fact a local maximum in $p_1(l|l)$, which is sufficient to proof the claim. Let us begin with defining the differentiable interior.

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

Next, for the ease of notation define $\boldsymbol{\rho} = (\rho_1(l), \rho_2(l), \rho_1(h)) := \left(p_1^*(l|l), p_2^*(l|l), p_1^*(l|h)\right)$
Step 1a: Transform $R(P)$ and $Q(P)$.
Observe that

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

Defining the function

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

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

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

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

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

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

$$Obj''(\rho^{cp}) = \underbrace{f'(\rho^{cp})g(\rho^{cp})}_{=0 \text{ by equation (FOC)}} + f(\rho^{cp})g'(\rho^{cp}) \geq 0 \tag{2.12}$$

for every $\rho^{cp} \in \boldsymbol{\rho^{cp}}$. The first term is 0 by the standard envelope argument, such that (2.12) boils down to

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

By the linearity of $\tilde{R}$ and the observation that $\tilde{R} \geq 0$ by construction, a necessary and sufficient condition for (2.12) to hold is simply

$$\tilde{Q}''(\rho^{cp}) \leq 0 \tag{2.13}$$

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

To see this observe first by plugging in we can reduce $\gamma_i(l) = \gamma(l,l)p/\rho_{-i}(l)$ which in turn means that while $\tilde{\gamma}_2(l)$ is constant in $\rho_1(l)$, $\tilde{\gamma}_1(l)$ is linearly increasing in $\rho_1(l)$. In addition, we do not need to worry about $\gamma_2(h)$ as player $2h$ has no expected utility by Corollary 1. Further we can rewrite using Corollary 1 and

Lemma 11

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

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

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

it suffices to show that

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

is convex for every i.

For h_2, observe that by Lemmas 2 and 3, player $2h$ only gains from deviating if $\rho_1(h) > \rho_1(l)$. In such a case player $2h$, best post-deviation strategy is to play $\Delta_{h,h}$ with probability 1, which yields utility

$$U_2(h|l) = p_1(h|l) - \Delta_{h,h}\kappa c_l. \tag{2.14}$$

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

$$U_2(h|h) = p_1(h|h) - \Delta_{h,h}\kappa c_l = 0. \tag{2.15}$$

Subtracting equation (2.15) from equation (2.14) yields

$$p_1(h|l) - p_1(h|h) = \rho_1(h) - \rho_1(l) = U_2(h|l) \tag{2.16}$$

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

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

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

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

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

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

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

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

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

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

$$a\rho_1(h)^2 + b\rho_1(h) + c < 0 \tag{2.17}$$

where

$$a = (\kappa - 1 + \rho_2(l)^2) \tag{2.18}$$

$$b = 1 + 2\rho_2(l) - 2\Big(\rho_2(l)\Big)^2 + p(1 - \kappa) \tag{2.19}$$

$$c = \Big(\rho_2(l)\Big)^2\kappa - \rho_2(l)\Big((\kappa - 1)(1 - p) + \kappa\Big) + (\kappa - 1)(1 - p). \tag{2.20}$$

Note first, that (2.17) is decreasing in $\rho_2(l)$, second note that for $\rho_2(l) = \rho_1(h)$ condition (2.17) becomes

$$(\kappa - 1)\Big(1 - p - 2\rho_1(h)\Big) + \Big(\rho_1(h)\Big)^4 - 2\Big(\rho_1(h)\Big)^3 + \Big(\rho_1(h)\Big)^2(1 + 2\kappa) < 0. \tag{2.21}$$

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

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

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

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

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

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

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

In addition, $\gamma_1(l) = \gamma_2(l) = p/\rho(l)$ and therefore

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

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

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

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

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

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

$$\hat{R}'(\rho(l))\hat{Q} - \hat{Q}'(\hat{(l)})\hat{R}|_{\rho(l)<\rho(h)} = -2\rho(h)p\Bigg(\Big(U_i(l|l) - (1-p)U_i(h|l)\Big) - \hat{R}\Bigg). \quad (2.22)$$

Note that by construction Obj defines a probability and is thus in $[0,1]$. Whenever equation (2.22)=0, then $\hat{Q} - \hat{R} = (2\rho(h)p - 1)\hat{R}$ which can only be positive if $2\rho(h)p = 1$. As $p < 1/2$ this condition never holds. Therefore, we do not find an interior solution when $\rho(h) > \rho(l)$.

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

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

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

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

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

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

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

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

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

and

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

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

Notice that

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

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

Proof of Lemma 6

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

$$(\gamma_i(l) - \gamma_i(h))U_i(l|l) \geq (\gamma_i(l) - \gamma_i(h))U_i(h|h). \tag{2.23}$$

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

Calculating the difference yields

$$\gamma_i(l) - \gamma_i(h) = \frac{p}{\rho_{-i}}\frac{\rho_i - p}{1-p} \tag{2.24}$$

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

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

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

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

Plugging in and simplifying, we find that

$$\tilde{\gamma}_2(l) - \tilde{\gamma}_2(h) = \rho_1(h)\rho_2(l)(1-p) - \rho_1(l)p^2(1-\rho_2(l)) \tag{2.25}$$

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

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

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

$$\mathcal{L}(\lambda, \mu, \boldsymbol{\rho}) = Obj(\boldsymbol{\rho}) + \lambda g \boldsymbol{\rho}) + \mu(\rho_1(l) - \rho_1(h)). \tag{2.26}$$

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

$$\min \mathcal{L}(\cdot) \tag{2.27}$$

and

$$\lambda, \mu \geq 0. \tag{2.28}$$

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

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

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

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

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

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

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

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

As $\rho_1(l)$ has no extreme value on the interval $(\rho_2(l), \rho_1(h))$ by Step 1 in section 2.C, $\rho_1(l)$ must be the same in both optima.

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

Proof of Lemma 7

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

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

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

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

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

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

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

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

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

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

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

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

Proof of Proposition 2

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

$$Pr(\Gamma, \boldsymbol{\rho}^*) = \frac{(\kappa+1)p}{1 + 2(\kappa-1)p}, \tag{2.29}$$

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

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

$$4p\frac{\kappa p - (1-p)(\kappa-2)}{(\kappa-1)(7p^2 - 2p - 1) + 4p}, \tag{2.30}$$

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

$$\frac{\kappa - 4}{2\kappa - 5}, \tag{2.31}$$

which again is bounded by 1/2.

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

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

Proof of Proposition 3

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

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

Proof of Proposition 4

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

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

where b, m and t are the scoring ranges used in Figure 2.1 and the proof of Lemma 1. Further, $\underline{r}$ denotes the upper bound of range r and $\underline{r}$ denotes the lower bound of range r.

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

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

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

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

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

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

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

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

Proof of Proposition 5

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

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

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

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

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

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

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

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

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

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

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

which follows analogously to Lemma 11.

Plugging this into Sender l-types yields:

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

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

Proof of Proposition 6

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

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

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

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

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

Proof of Proposition 7

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

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

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

Suppose instead the mediator announces a mediation protocol $\mathcal{X}^{\epsilon}_{\lambda}$ in which reporting two l-types follows mediation breakdown with full information disclosure with probability ϵ and a protocol as that derived in Section 2.4 otherwise. As $\epsilon \to 0$, the result gets arbitrarily close to that of Theorem 1. However, the mediator can signal any l-type deviator that in fact the low-cost vs. low-cost litigation game is played, causing the l-type to also except ex-post shares. □

Proof of Lemma 12

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

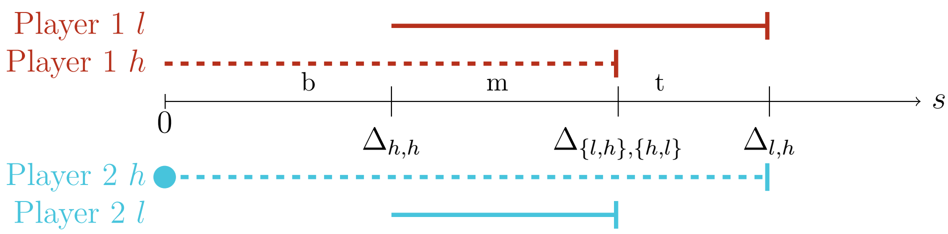

Figure 2.C.1: *Strategy support of player 1 and 2 if monotonicity fails.*

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

3 Conflict Management and Conflict Games

with Benjamin Balzer

3.1 Introduction

Many economic conflicts are by default solved via a costly, non-cooperative game. Examples include wars, strikes, legal disputes, and patent races. In many situations conflict management provides a way to settle the disputes without the need of the costly non-cooperative game. Participation in conflict management is typically voluntary and the default game serves as a last resort to conflict management. If parties possess private information relevant in the default game only, the conflict management mechanism might not be able to settle all conflicts. If, however, conflict management breaks down for whatever reason, players use information from conflict management to update their beliefs. Especially if the underlying game is Bayesian, updating effects their expected payoff at the point of breakdown. The knowledge of this relation, in turn, feeds back onto players' behaviour during conflict management.

In this paper, we study the optimal design of conflict management if the underlying game is Bayesian and private information is relevant only in the default game. We investigate the informational effect of breakdown of conflict management if parties are rational and use this information in the subsequent play of the default game. We are particularly interested in the following questions: What is the connection between the underlying default game and optimal conflict management? How much information is or can be released during the conflict management mechanism? how does the choice of mechanism influence the outcome of the default game in case conflict management fails?

The answers to these questions are of substantial economic importance. Conflict management helps to avoid not only costs borne by the fighting players, but also negative externalities on their surroundings: wars destabilizing the world economy, strikes affecting neighbouring firms, over-clocked courts limit access to the legal system, and patent wars holding off important substantial research are only a few examples of these externalities.

In reality there is indeed a large industry providing conflict management mechanisms for various disputes. Examples include peace conferences, alternative dispute resolution and standard setting organizations. A common feature of these mechanisms is that although they often lead to settlement, none of them guarantees settlement and the typical response to breakdown is to revert to the rules of the costly, initial mode of conflict resolution. Thus, the underlying default game can be seen as a realistic outside option to conflict management.

The relevance of private information in the default game is central to our findings and present in the applications mentioned above. The strength of an army, the financial resources, the access to evidence, or the current research of the R&D department may all be better known to a player herself and hard to verify without playing the default game to be avoided.

We find that for the third-party (call her an ombudsman) designing the conflict management mechanism, the players' private information is a double edged sword. On one hand, if the ombudsman possesses the private information of the parties, she can design a settlement contract that efficiently solves the conflict. On the other hand, the private information is irrelevant whenever the conflict is settled without playing the non-cooperative default game. If parties do not get punished via the play of the default game for making false claims during conflict management, then they also have little incentive to tell the truth to the ombudsman.

The ombudsman designs conflict management in a way that screens the players' types by inducing breakdown of the conflict management mechanism and thus invoking the play of the default game. If player's are rational, they are going to incorporate the information received through breakdown of conflict management before making decisions for the default game. If the default game is Bayesian, the players' updating directly effects their (and their opponent's) decision and thus the expected outcome of the default game. Obviously, the exact design of conflict management is going to influence the players' updating, and thus their action choices and expected payoffs. The underlying game therefore serves as an endogenous and belief dependent outside option to the conflict management mechanism, despite its fixed rules.

The aim of this paper is to study the design of optimal conflict management in the presence of a default game with asymmetric information. We show that the purpose of conflict management in such situations is to manage the beliefs of the parties. We highlight that the result of optimal conflict management can be entirely described by the beliefs that players hold in case no settlement can be found. We provide an intuitive formulation of the optimal mechanism and characterize its relation to properties of the underlying default game. In particular, we show that games in which the players' actions do not depend on her belief allow for "sure settlement" of certain pairs of types. This finding generalizes Hörner, Morelli, and Squintani (2015).

If this property is violated, that is, if the game has an action-relevant Bayesian component, "sure settlements" may not exist any more and it becomes more important for the ombudsman to increase the strong player's ability to separate. The ombudsman does so by increasing the *ability premium* that measures how much the fundamental characteristics of a strong player allow her to separate herself from a weaker player only by mimicking the strong player during conflict management and behaving optimally in the subsequent default game.

The ability premium is directly connected to the design of the conflict management mechanism in the following way: The design of and the players' behaviour within the mechanism determine the beliefs in case of breakdown. These beliefs influence the play of the subsequent continuation game. Mimicking another type implies that a player – in terms of beliefs – plays the continuation game *of the mimicked type* with two advantages: (i) she is not forced to behave as if she was the mimicked type in the default game, but can freely optimize, and (ii) her opponents do not know that a deviation occurred and play alongside the equilibrium path which induces a situation of superior knowledge of the deviator. Taking these two effects into account, the ombudsman aims to increase the ability to separate from mimicking types, by increasing the ability premium.

We begin our analysis by formulating a traditional conflict management problem with some underlying default game as an outside option. One of the key assumptions of our model is that the ombudsman has the ability to commit to a certain structure of her proposed conflict management mechanism. This common assumption in the mechanism design literature allows us to focus on the behaviour of the players having the conflict rather than on the commitment issues of the designer. We discuss the meaning of this assumption in the formulation of the model below.

Next, we show that we can re-express the problem of finding the right settlement contract as a search for conditional type distributions over players' types given breakdown. This allows us to relate the optimal contract directly to the expected performance within the default game. Alongside this central simplification we also formulate the problem in its reduced form using a tool from Border (2007). Further, we show that full-participation in conflict management problems is optimal.

After reformulation we can apply results from convex analysis to address the role of public signals sent by the ombudsman in case no settlement is found. Examining the relation between the expected posterior for any behaviour within conflict-management and the realised posterior after having learned the signal's realisation allows us to make statements about both the (limited) scope of signals in such problems and the choice of expected posteriors.

In the final part of the paper we then discuss several properties of the underlying default game and how they relate to optimal conflict management.

3.2 Related Literature

We contribute to the literature on conflict resolution by offering a general approach to solving problems of conflict management in the presence of a fixed default game. We take the mechanism design approach to conflict management proposed by Bester and Wärneryd (2006). Different to Bester and Wärneryd (2006) and the related work by Hörner, Morelli, and Squintani (2015) who study a particular mode of conflict, we allow for a general class conflict games. In particular and different to these papers we allow for games in which the choice of investment within the conflict game is made *after* the outcome of conflict management is known. The general formulation allows us to analyse conflict management at any stage of the conflict while previous literature has mainly focused on "last minute" resolution attempts. Economically, early stage conflict management has the advantage that settlement obtained at this early point generates the largest economic effect, as players do not engage in costly actions. Nonetheless, our approach is also valid for conflict management carried out at a later stage. In fact, we can generalize some of the findings of Hörner, Morelli, and Squintani (2015) to all conflict games in which a player's action choice is independent of her belief.

Another strand of literature related to ours is that of veto-constraint mechanisms. Similar to Cramton and Palfrey (1995) conflict management is only carried out if both players ratify the proposed mechanism. In our model, each player can enforce the play of the default game. Similar to these models our mechanism cannot influence any player's behaviour in case the default game is played. Different to them our designer can also send the players to the default game herself. Any of these breakdown situations is informative to the players as in Cramton and Palfrey (1995) and Celik and Peters (2011). Our scope is, however, different to these models. Cramton and Palfrey (1995) as well as Celik and Peters (2011) are interested in the set of implementable allocations by either finding the necessary off-path beliefs or restricting participation. We explore how the mechanism influences the post-breakdown behaviour of the parties and in particular how such behaviour feeds back on the design of the mechanism. Further, we show that full participation can be assumed without loss of generality.

Some of our results also connect to the literature on Bayesian persuasion such as Kamenica and Gentzkow (2011). Like them we are interested in a mechanism design problem without transfers and allow the mechanism to distort the processed information in a Bayesian manner. Similar to this literature we can use results from convex analysis to determine the scope of public signals. Yet, our setup is very different than that of Kamenica and Gentzkow (2011). The signals in our setup are in fact only supplementary and have, in general, only limited scope as the primary objective is to minimize the play of the game altogether.

3.3 Model

3.3.1 Setup

Default Game. Consider a model of two ex-ante identical, risk-neutral players that could potentially distribute a pie of size normalized to one via a non-cooperative game Γ. We refer to Γ as the *default game* or the *outside option*. Each player possesses a private signal θ_i independently drawn from a given distribution over a finite set, $\Theta = \{\theta^1, .., \theta^K\}$ of cardinality K. We refer to the signal as the player's *type*. The type measures the strength of the player when playing the default game and is payoff-irrelevant outside Γ. One interpretation of θ_i is to view it as an individual cost parameter determining how costly it is for agent i to exert effort within Γ. Each player also holds the same prior belief, $\beta_i = \beta_\emptyset > 0$ which is the ex-ante distribution of the signals. Beliefs are common knowledge, such that the belief system is given by $\boldsymbol{\beta}_\emptyset = (\beta_\emptyset, \beta_\emptyset)$. The k^{th} element $\beta_\emptyset^k$ describes the ex-ante probability of a player being type θ^k. Notice that any distribution β_i is pinned down by its $K-1$ elements. In what follows we are therefore going to treat β_i^K as an endogenous variable defined as $\beta_i^K = 1 - \sum\limits_{k \neq K} \beta_i^k$.

Assuming the existence of a Bayes' Nash equilibrium for every belief system $\boldsymbol{\beta}$, the expected payoff of player i, type θ_i is – up to the choice of equilibrium – given by $U_i(\theta_i, \boldsymbol{\beta})$. Throughout the paper we assume that the game is such that $U_i(\theta_i, \boldsymbol{\beta})$ is continuous in each element of $\boldsymbol{\beta}$. Similarly, fixing some equilibrium strategies, the reduced form payoff given the realisation of any type-profile can be denoted by $u_i(\theta_i, \theta_{-i}, \boldsymbol{\beta})$.[1]

We restrict attention to conflict games which are defined as follows.

Definition 4 (Conflict Game)**.** A game Γ is called a conflict game if it has the following properties:

i. *(non-productiveness.)* The sum of realised utilities in equilibrium is smaller than the contested pie, that is $\sum_i u_i(\theta_i, \theta_{-i}, \boldsymbol{\beta}) \leq 1$ for any $(\theta_i, \theta_{-i}, \boldsymbol{\beta})$.

[1]Observe that $u_i(\cdot)$ does not necessarily coincide with the realised payoff, since players may play mixed strategies even under full-information. Instead $u_i(\cdot)$ describes the expected payoff given the uncertainty about the players' types has been resolved.

ii. *(monotonicity in own type.)* The expected utilities $U_i(\theta_i^k, \boldsymbol{\beta})$ are non-increasing in k for any $(\theta_i^k, \boldsymbol{\beta})$.

iii. *(monotonicity in own belief.)* Take any belief β_i' that first-order stochastically dominates β_i. Then, for any type θ_i and a belief β_{-i} the expected utility $U_i(\theta_i, \beta_i, \beta_{-i})$ is not smaller than $U_i(\theta_i, \beta_i', \beta_{-i})$.

Non-productiveness guarantees that a conflict always reduces the amount of resources available in the economy and is therefore never desirable from a utilitarian point of view as a settlement is always weakly preferred. The two monotonicity properties ensure that the player's type is a sufficient statistic for the player's ability in the conflict game. While the first one states that higher ability results in higher expected utility, the second one ensures that it is always preferable to play against a weaker opponent.

Conflict management. Now assume that the parties can undertake a process of conflict management offered by an uninformed third party, the *ombudsman*, that seeks to settle the conflict without relying on the default game, if and only if they mutually agree to participate. A natural interpretation of such a settlement would be a contract offer conditional on the players' reports. The contract, if offered, then specifies the share of the pie assigned to each party. In case the ombudsman does not offer a solution, parties go back to the default game to resolve the conflict. We refer to this event as *breakdown* or $\mathit{\Gamma}$, since it invokes the play of Γ. Formally, this means that the ombudsman cannot directly influence the players' actions in the default game. The conflict management mechanism is a game leading to (a lottery over) allocations or the declaration of failure.

We are interested in conflict management that minimizes the probability that the default game is played, $Pr(\mathit{\Gamma})$ and satisfies ex-post weak budget balance, that is a mechanism that cannot grant any additional resources into the system.

Timing. At first the ombudsman proposes a conflict management mechanism and commits to it. Then players learn their private signal. Thereafter, both players decide simultaneously whether or not to participate in conflict management, and if so which action to play within conflict management. In case either (or both) of the players reject conflict management, both players learn about each others participation decision, update their beliefs and play the default game non-cooperatively.

If both players accept the mechanism, the mechanism is played. If the mechanism reaches a terminal node, payoffs realize. Alternatively, if the mechanism decides to send players to the default game, both players update their beliefs and play the default game non-cooperatively.

Fixing the equilibrium played in the default game for any belief system, we are looking for the conflict management mechanism that minimizes the ex-ante probability of playing the default game. Our solution concept is perfect Bayesian equilibrium.

Ratification and the Revelation Principle. As players can unilaterally and publicly reject any proposed mechanism (by sending an empty message), our problem is one of veto-constraint mechanism and it is therefore not obvious that the revelation principle holds. In particular it may be the case that full participation is not optimal. Nonetheless, we can, without loss of generality, restrict attention to direct revelation mechanisms. The mechanism maps player's type reports either into (a lottery over) shares or into a distribution of public signals of participating players. Any feasible mechanism induces a matrix of breakdown probabilities $\mathcal{G}$ and a matrix of shares $\mathcal{X}_i$ for each player. Each of these matrices have at most dimension $K \times K$. The element in the m^{th} row of the n^{th} column of $\mathcal{G}$ is denoted by $\gamma(\theta^m, \theta^n)$. With slight abuse of notation we say that $\gamma(\theta^m, \theta^n)$ describes the probability that the ombudsman enforces breakdown of settlement and sends parties back to the default game in case player i report type θ^m and player $-i$ reports θ^n. Similarly, $x_i(\theta^m, \theta^n)$ denotes the share player i receives if the report profile is (θ^m, θ^n). We assume budget balance in a sense that $x_1(\theta^m, \theta^n) + x_2(\theta^m, \theta^n) \leq 1$ for all $(\theta^m, \theta^n) \in \{\Theta\}^2$.

In addition the ombudsman may send a public signal $\sigma(\theta^m, \theta^n)$ before the default game is played. Importantly, we assume that no party can ex-ante commit to ignore public messages and that the ombudsman can send such a signal also if either of the parties decides to not participate in the mechanism. Thus, the matrix of all $\sigma(\cdot, \cdot)$, $\mathcal{S}$, is of dimension $K + 1 \times K + 1$. Observe that these signals induce a posterior that is "Bayes plausible" with respect to a given posterior which means they induce a mean preserving spread over that posterior. In particular the signals $\sigma(\emptyset, \theta^n)$ and $\sigma(\theta^m, \emptyset)$ induce a mean preserving spread over the prior $\beta_\emptyset$ such that $E[\beta_i(\emptyset, s)|\sigma(\emptyset, \theta^n)] = \beta_\emptyset$ with $\beta_i(\emptyset, s)$ being the posterior induced by signal realisation s.

Crucial Assumptions. As a benchmark we assume that all parties have full commitment power once the conflict management mechanism has been ratified. On the player's side this assumption is innocuous to a large extend. At the end we discuss an extension without full commitment on the players' side. The assumption of full commitment by the ombudsman on the other hand is crucial to our model. However, if we assume that the ombudsman repeatedly deals with conflicting situations, she has an incentive to commit by reputational concerns.

A second crucial assumption is that the players' private information is only payoff relevant *within the default game*. Our results can be seen as a benchmark on the possibilities of third party conflict management that cannot influence the behaviour of the players in the institutionalised default game. If each player's private information had a component that was valuable *beyond* the default game, the ombudsman could use this information to screen types without relying on the default game which would make it easier for her to find settlement. Our analysis also provides a benchmark on what third party conflict management can achieve by solely relying on (cheap-talk) messages of the parties. Such "pure-talking" conflict management is typically considerably cheaper to implement than more complicated modes of screening and thus often used as a first step. Hence, if we consider anything that follows as "default game", we can still use our approach to capture the problems first stage.

Our objective is to minimize the probability of the default game being played in equilibrium. This assumption is driven by the fact that we are mainly interested in applications in which invoking the default game comes at a large fix cost for uninvolved players, that is for example, civilians in case of a war, customers in case of a strike, or other cases in case of legal disputes.

The assumption that also non-participating players can receive messages from the ombudsman allows us to broaden the analysis without much further technical complication. If we were to remove this assumption our results would still be valid for games in which the expected payoff in Γ is convex in beliefs around the prior. However, if the expected payoff is for whatever reason non-convex, public signals allow us to convexify the payoff function in such a way that full participation is optimal. We are going to discuss this property next. Note that in many economic problems it is hard for any of the players to commit to ignoring publicly available information whenever it is interim beneficial to process it. Such situations are exactly those we are interested in.

3.3.2 Players' Behaviour and the Outside Option

By the revelation principle it is without loss of generality to restrict to truthful reporting for any participating player on the equilibrium path. To guarantee truthful reporting, the player must have no incentive to deviate from the truthful reporting strategy. In general a participating player's strategy consists of two additive components. Her expected utility in case the conflict settles, $z_i(m)$ and her expected utility in case the conflict advances to the default game Γ, $y_i(\theta, m)$. Observe that since types become only relevant within Γ, z_i is independent of the player's type. Thus, the ombudsman needs to rely on the default game to effectively screen types. Within the default game it is not only the player's own type that determines the payoff, but also the belief system after breakdown of conflict management.

The belief system has three different effects. First, it determines the distribution against which the player expects to play and by that influences expected payoffs. We call that effect the direct effect of beliefs. A direct effect of beliefs is typically also present even in games in which players cannot choose their action after breakdown but enter some form of distribution that potentially depends on the types of their opponent.

Second, as we assume a non-cooperative Bayesian game as the outside option to settlement, there is an indirect effect of beliefs. The underlying reason for the indirect effect of beliefs is the influence of belief system on the players' actions in the default game. For example, if a change in believes leads to one player playing more aggressively, the other player take this change of action into account and adjusts her action, too. As the continuation game is on-path played under common knowledge, all players are aware of the belief system of the other player. Thus, a change in any of the beliefs has a potential spill-over on the utility of any of the players.

Finally, there is a third effect which is particularly relevant if we consider deviations of players within the mechanism. As discussed above, such deviations remain entirely undetected if the ombudsman can implement a settlement. However, even if conflict management breaks down, a deviation is not detected right away. Instead, the (non-deviating) opponent decides on her action as if there was no deviation. This gives the deviating player an informational advantage within the default game. Through the deviation she selects a different population to play against, which may change her optimal action compared to the play on path. This is in particular true since the opponent does not belief that there is a player of the deviating type expecting endowed with that belief. We call this the deviator's information advantage.

To illustrate the deviator's information advantage, consider the following extreme example. Suppose conflict management is such that breakdown occurs only if one player is the strongest and the other player is the weakest type. Suppose further that since both players know this, the weak type immediately gives in granting the strong type a high portion of the pie with little effort on her side, too. Now, suppose a weak type claiming to be a strong type. After breakdown the weak type faces another weak type who thinks she is facing a strong type. This weak type may immediately give in and the deviating weak type gains the good just by the 'reputation' of being a strong type. Thus, the player has an incentive to manipulate her own belief by deviation not only because it changes the population she plays against, but also because it allows her to gain an information advantage.

The three effects highlight the role beliefs play in the problem of the ombudsman and indeed, as we show in the following section, we can describe the solution entirely by considering how the ombudsman manages the beliefs of the players.

3.3.3 Simplifying the Problem

Prior to the in-depth analysis it is useful to simplify the problem already at this level of generality. The simplifying steps are largely technical in nature and we give a brief summary before laying out the different steps.

First, we show that full participation is always optimal. The reason for this is that the ombudsman can always 'punish' a non-participating deviator by releasing a public signal such that it is interim rational to incorporate that signal. By that the players payoff function is convexified which in turn provides sufficient slack on the participation constraint for everyone to participate. Further, we provide an assumption that guarantees that the ombudsman solves a non-trivial problem so that there is no solution that avoids the play of the default game altogether.

Next, we discuss the connection between the choice of breakdown probabilities and the realisation of a signal which can be seen as a mean-preserving spread over the expected posterior.

Thereafter, we use a result from reduced form auctions to rewrite our problem in expected terms. We show that the strong type's participation constraint and each type's upward adjacent incentive compatibility constraint always bind. The reason is simply that the strongest type is the most efficient within the conflict game and thus needs to have a sufficiently high value to participate in conflict management. Within the conflict management mechanism on the other hand, each type has an incentive to mimic the next stronger type as she expects a higher settlement result, but is similar enough when it comes to ability within the conflict game.

Finally, we use the two previous result to rewrite the entire problem in a way that highlights the ombudsman's sole objective to correctly manage the players' beliefs. We then provide several alternative formulations of the problem which are helpful both for understanding the ombudsman's considerations and the analysis of optimal conflict management.

Participation and Non-triviality

Lemma 13 (Full participation.)**.** *It is without loss of generality to assume that there is full participation at the optimum.*

The proof can, as all others not discussed in the text, be found in the appendix. It is due to Balzer and Schneider (2015) and relies on the fact that for any initial belief-system $\boldsymbol{\beta}_{\emptyset}$ we can find a Bayes plausible public signal $\sigma(\emptyset, n)$ released upon deviation of player 1 such that the value for player 1 of not participating in conflict management, $\tilde{V}_1(\theta_1)$, is convex in her belief. By symmetry the same is true for player 2. The second step is then to invoke proposition 2 of Celik and Peters (2011) stating that whenever the value of not participating is convex in a player's belief, it is without loss of generality to assume full participation. The value of non-participation, $\tilde{V}_i(\theta_i)$ corresponds to the expected payoff the player receives when rejecting the mechanism and behaving sequentially rational. Thus, $\tilde{V}_i(\theta_i)$ is the expected equilibrium payoff given the choice of public signals by the ombudsman. Formally, this means for player i, type θ_i

$$\tilde{V}_i(\theta_i) \equiv E_{\theta_{-i}}\left[E_{s(\theta_{-i})}\left[U_i\left(\theta_i, \beta_i\left(s\left(\theta_{-i}\right)\right), \beta_2^d\right)\right]\right],$$

where θ_{-i} player $-i$'s type and $s(\theta_{-i})$ is a possible signal realisation of the public signal $\sigma(\emptyset, \theta_{-i})$. Player i's, $\beta_i(s(\theta_{-i}))$, is the probability distribution over player $-i$'s types given signal realisation s. Similarly, β_{-i}^d is player $-i$'s off-path belief. Obviously $\tilde{V}_1(\theta_1)$ depends on the ombudsman's choice of public signals $\mathcal{S}$. The minimum value of non-participation that can be induced given any type θ_i is given by

$$V_i(\theta_i) \equiv \min_{\mathcal{S}} \tilde{V}_i(\theta_i).$$

Observe that if we assume passive beliefs, that is $\beta_2^s = \beta_\emptyset$, $V_i(\theta_i)$ is the value of the convex hull in β_i of $U(\theta_i, \boldsymbol{\beta})$ at $\beta_i = \beta_\emptyset$ by Bayes' plausibility of the signals. Therefore, the minimum value of non-participation can be easily computed using primitives only.

Definition 5 (Non-Triviality)**.** A problem is called non-trivial if there exists no conflict management mechanism such that settlement happens with probability 1.

Lemma 14 (Non-Triviality)**.** *The optimal conflict management mechanism is non-trivial if and only if the sum of minimum values of non-participation for the strongest type of each player is larger than* 1*.*

Assumption 2. (Non-Triviality.) The optimal conflict management mechanism is non-trivial.

Proof. If Assumption 2 was violated, the ombudsman's optimal solution is to offer the same sharing rule independent of the players' reports. If the ombudsman were to offer a trivial mechanism that satisfies the strongest player's participation constraint, then, whenever two strongest types meet, she cannot offer them more than 1 in total by weak budget balance. However, since types are irrelevant in case of settlement, she can only pool across types and thus must offer a sharing rule such that $x_1(m,n)+x_2(m,n) > 1$ for all type report to satisfy the incentive compatibility constraint. Thus, there is no trivial incentive compatible mechanism that satisfies the strongest players participation constraint and is budget balanced. $\square$

Breakdown Probabilities, Signals, and Beliefs

Due to monotonicity the strongest type always has the most attractive outside option for any signal realisation s. Thus, the ombudsman always wants to choose $\mathcal{S}$ such that the value of non-participation is minimized for the strongest type to limit the distortions via the participation constraint.

When the ombudsman offers a non-trivial solution, breakdown is valued differently by the different types. In particular, for any given believe the stronger types are less afraid of breakdown than the weaker types, as the default game is less costly to them. At the same time the beliefs players hold after breakdown are both a function of the mechanism *and* their report. If the mechanism specifies for example that the strongest type should never meet another strongest type after breakdown, that is $\gamma(\theta^1, \theta^1) = 0$, then any player type can prohibit herself from meeting the strongest type after breakdown by reporting θ^1. Thus, the belief system after breakdown plays an important role, as players could strategically misreport to induce a different belief system in case of breakdown. This makes the outside option not only type-dependent, but also endogenous to the design of the mechanism. For any report m and any realisation of the public signal s, player i holds a belief $\beta_i(m,s)$ conditional on breakdown. In other words, given a conflict management mechanism, any report m pins down a belief system and thus an expected payoff conditional on breakdown. Any public signal $\sigma(m,\cdot)$ then allows the player to adjust her action conditional on the signal realisation s she receives. To shorten notation we denote this expected payoff conditional on breakdown, given any signal realisation s and report m as

$$U_i(\theta_i|m,s) = E_{\theta_{-i}}[\beta_i^{\theta_{-i}}(m,s)u_i(\theta_i,\theta_{-i},\boldsymbol{\beta}(m,s))|m,s] \tag{3.1}$$

As the player does not know the signal realisation at the reporting state she forms an expectation over the signal realisations. Thus we can write

$$U_i(\theta_i|m) = E_s[U_i(\theta_i|m, s)], \tag{3.2}$$

and similarly

$$\beta_i(m) = E_s[\beta_i(m, s)]. \tag{3.3}$$

We refer to $\beta_i(m)$ as the *expected posterior* given m while $\beta_i(m, s)$ is called the *realized posterior* given m and s. Analogously to the matrix of breakdown probabilities $\mathcal{G}$, we can define a matrix of expected posterior distributions $\mathcal{B}$ which also has dimension $K \times K$.

Observation 2 (Reduced Form Definition of $\mathcal{B}$)**.** Each expected posterior $\beta_i(m)$ is uniquely determined by its $K-1$ first elements. Each matrix of expected posteriors $\mathcal{B}$ is uniquely determined by all expected posteriors of player 1, $\beta_1(m)$ and the expected posterior of player 2, reporting to be the strongest type, that is $\beta_2(\theta^1)$. We call these determinants the choice variables of $\mathcal{B}$.

The first part follows simply because posteriors are distributions and the second is a consequence of Bayes' rule and the following observation.

Observation 3 (Determinants of Posteriors.)**.** The expected posterior given m is entirely determined by the matrix of breakdown probabilities, $\mathcal{G}$, while the realised posterior given any expected posterior is determined by $\mathcal{S}$ only.

In what follows, we always refer to the reduced form definition of $\mathcal{B}$ and treat the remaining beliefs as endogenously determined functions of the $(K-1) \times (K+1)$ choice variable mentioned in Observation 2.

Using the assumption that the player's type is only relevant conditional on breakdown we can determine an additively separable term $y_i(m, \theta_i)$ which we call the *breakdown value* of reporting m and that is denoted by

$$y_i(m, \theta_i) = \gamma_i(m) U_i(\theta_i|m)$$

with $\gamma_i(m)$ being the expected breakdown probability given report m which is given by

$$\gamma_i(m) := \sum_{k=1}^{K} \beta_\emptyset^k \gamma(m, \theta_{-i}^k).$$

Likewise, we can determine the *settlement value*

$$z_i(m) = \sum_{k=1}^{K} \beta_\emptyset^k (1 - \gamma(m, \theta_{-i}^k)) x_i(m, \theta_{-i}^k).$$

Combining the two gives us the unconditional expected payoff of reporting m, that is

$$\Pi_i(\theta_i, m) = z_i(m) + y_i(m, \theta_i). \tag{3.4}$$

Observe that while the breakdown probabilities $\mathcal{G}$ influences both the settlement and the breakdown value $\mathcal{X}_i$ is only relevant for the settlement value and $\mathcal{S}$ only influences the breakdown value.

Binding Constraints and Border Transformation

Invoking the revelation principle we can restrict attention to direct mechanisms. Thus, the probability that player i sends message θ_i, $p_i(\theta_i^k)$, simply corresponds to the ex-ante probability that player i is of type θ_i^k, which is the k^{th} value of $\beta_\emptyset$. Now, similar to Border (2007) we can rewrite the problem in reduced form replacing $\mathcal{X}_i$ by the vector $\mathbf{z}_i$. Thus, we can solve the problem in expectations while applying Border (2007) provides a necessary and sufficient condition to implement the found solution also under ex-post budget balance.

Lemma 15 (Border Transformation)**.** *Fix some feasible $\mathcal{G}$ and $\mathbf{z}_i \geq 0$ for every i. Then there exists an ex-post feasible $\mathcal{X}_i$ that implements $\mathbf{z}_i$ if and only if the following constraints are satisfied for any $\{m, n\} \in K \times K$ and i:*

$$\beta_\emptyset^m(1-\gamma_i(\theta^m)-z_i(\theta^m))+\beta_\emptyset^n\left(1-\gamma_{-i}(\theta^n)-z_{-i}(\theta^n)\right) \geq (1-\gamma(\theta^m,\theta^n))\beta_\emptyset^m\beta_\emptyset^n \tag{EPI}$$

$$z_i(\theta^m) \leq 1-\gamma_i(\theta^m). \tag{IF}$$

To understand the intuition start with the individual-feasibility condition (IF) and observe that it provides a natural upper bound on the settlement value. Individual feasibility requires that no player can expect a settlement value that is higher than that in which she receives the whole pie whenever settlement is achieved. Using this upper bound we can interpret the ex-post feasibility condition (EPI). For a given type profile (θ^m, θ^n), the left-hand-side of (EPI) is the expected difference between offering each party precisely the upper bound of the settlement value and the actual settlement value. This difference then needs to be larger than the resources the parties *generate* by foregoing the default game which is equivalent to the right-hand-side of (EPI). This inequality needs to hold since the resources generated together can only be used once to settle for the given profile. Therefore, at most one of the parties can include these resources into her settlement value. Thus, the combined expected settlement values need to be smaller than the sum of their individual upper bounds by at least that amount. Finally, Lemma 15 implies that the sum of settlement values is ex-ante feasible, that is

$$\sum_i \sum_m \beta_i^m z_i(\theta^m) \leq 1 - Pr(\Gamma). \tag{AF}$$

Using Lemma 15 reduces the problem of choosing settlement values from 2 $K \times K$-dimensional matrices of $\mathcal{X}_i$ to two vectors each of dimension $K \times 1$ only. Further, as the agents make their choices in expected rather than realized values, we can use the binding constraints to reduce dimensionality even further.

Lemma 16 (Binding Constraints)**.** *It is without loss of generality to assume that strongest type's participation constraint as well as all upward adjacent incentive constraints hold with equality.*

The first part is an immediate consequence of the non-triviality assumption. Monotonicity of the default game on the other hand implies the second part. Further, we get two immediate corollaries to Lemma 16.

Corollary 3. *It is without loss of generality to ignore all participation constraints, but those of the strongest type.*

Corollary 4. *It is without loss of generality to consider upward deviations only locally.*

The linearity of the payoff function in the settlement value allows us to express all settlement values only in terms of breakdown values and the value of non-participation for the strongest type. To see this, consider the strongest type's binding participation constraint

$$V_i(\theta_i^1) = z_i(\theta_i^1) + y_i(\theta_i^1, \theta_i^1), \tag{PC}$$

and any type's upward adjacent incentive compatibility constraint

$$z_i(\theta_i^{k-1}) + y_i(\theta_i^{k-1}, \theta_i^k) = z_i(\theta_i^k) + y_i(\theta_i^k, \theta_i^k), \quad \forall k > 1. \tag{IC^+}$$

Rewriting Equation (PC) and Equation (IC^+) to

$$z_i(\theta_i^1) = V_i(\theta_i^1) - y_i(\theta_i^1, \theta_i^1) \tag{PC}$$
$$z_i(\theta_i^k) = z_i(\theta_i^{k-1}) + y_i(\theta_i^{k-1}, \theta_i^k) - y_i(\theta_i^k, \theta_i^k). \tag{IC^+}$$

Plugging all these constraints into each other we can fully determine $\mathbf{z}_i(\cdot)$ as an additive function of the strongest types value of non-participation and a set of breakdown valuations under truth-telling and deviation. Plugging all solutions into the ex-ante feasibility constraint (AF) yields

$$\sum_i \left(V_i(\theta_i^1) - \sum_m \sum_\theta q_i(m, \theta) y_i(m, \theta) \right) \leq 1 - Pr(\mathit{\Gamma}), \tag{AF}$$

where $q(m, \theta)$ is the weight put on $y_i(m, \theta)$ and is independent of the choice of mechanism.

Given the above considerations we can determine the solution to the problem entirely by the choice of $\mathcal{G}$ and $\mathcal{S}$ which, by Observation 3, induces an expected posterior via $\mathcal{G}$ and a mean-preserving spread over the expected posterior via $\mathcal{S}$. Taking this as given, we can check whether the remaining constraints are satisfied. All these constraints, too are a function of $\mathcal{G}$ and $\mathcal{S}$ only.

Three Alternative Formulations of the Problem

While any $\mathcal{G}$ induces a belief system $(\beta_i(m), \beta_{-i}(n))$ for any type profile (m, n), the reverse is not true as Bayes' rule gives us one additional degree of freedom in the choice of $\mathcal{G}$.

Observation 4 (Homogeneity)**.** Any expected posterior $\beta_i(m)$ is homogeneous of degree 0 with respect to the matrix of breakdown probabilities, $\mathcal{G}$.

Define $\{\mathcal{B}\}^+ := \{\mathcal{B}|\beta_i^1(\theta^1) > 0\}$ to be the set of all expected posterior such that the two strong types expect to meet conditional on breakdown with positive probability. Ignoring signals for a moment, we can combine the previous Lemmas and Observations to establish a duality result that allows us to solve the problem directly choosing expected posteriors.

Proposition 8 (Duality)**.** *Take any feasible matrix of expected posteriors, $\mathcal{B} \in \{\mathcal{B}\}^+$. Then, for any $\alpha \in (0, 1]$, $\mathcal{B}$ and α uniquely determine a feasible matrix of breakdown probabilities, $\mathcal{G}_\alpha$ in which the first element, $\gamma(\theta^1, \theta^1) = \alpha$ and all other elements $\gamma(m, n) = q(\mathcal{B})\alpha$. Conversely, any feasible $\mathcal{G}$ determines $\mathcal{B}$ and α.*

The duality of the two approaches greatly simplifies our analysis as it allows us to solve the whole problem by focusing on choice variables that directly connect to the rules of the underlying default game. Thus, we can directly connect the structure of the underlying game to the optimal mechanism. Finally, Proposition 8 nicely illustrates that the optimal choice of conflict management is directly depending on the role private information plays in the underlying default game. A direct consequence of Proposition 8 is that the ombudsman's objective is to *manage the beliefs* of the players in a sense that she, by choosing the optimal expected (and realized) posteriors, minimizes the likelihood of the default game being played.

Given Proposition 8, we can implement any matrix of breakdown probabilities $\mathcal{G}$ by the choice of expected posteriors and an appropriate α. More importantly, all elements of $\mathcal{G}$ are linear in α. This linearity implies that $\gamma_i(m)$ is linear in α. Since $U_i(\theta_i|m)$ is independent of α, we get that $y_i(\theta_i, m)$ is linear in α and so is the probability that the default game is played, $Pr(\Gamma)$. Thus, we can rewrite (AF) to

$$\sum_i V_i(\theta_i^1) - 1 \leq \alpha(Q(\mathcal{B}, \mathcal{S}) - R(\mathcal{B}, \mathcal{S})) \tag{AF}$$

with

$$
\begin{aligned}
Q(\mathcal{B},\mathcal{S}) &= \sum_i \sum_{m\in\Theta} \sum_{\theta\in\Theta} q_i(m,\theta)\frac{y_i(m,\theta)}{\alpha} \\
R(\mathcal{B},\mathcal{S}) &= \frac{Pr(\Gamma)}{\alpha},
\end{aligned}
$$

both depending on expected beliefs and a potential public signal only. Observe that the ombudsman seeks to minimize the probability that the default game is played, $\alpha R(\cdot)$ which is a linearly increasing function of α. Thus given feasibility, we would like to choose the smallest α possible which in turn means that it is without loss of generality to assume that (AF) holds with equality. Solving for α yields

$$
\alpha = \alpha^* := \frac{\sum_i V_i(\theta_i^1) - 1}{Q(\mathcal{B},\mathcal{S}) - R(\mathcal{B},\mathcal{S})}, \tag{3.5}
$$

if $\alpha^* \leq 1$ since we picked $\mathcal{G}$ such that $\gamma(\theta^1,\theta^1) = \alpha$.

Plugging everything back into the ombudsman's objective, we obtain an alternative formulation of the ombudsman's problem

$$
\min_{\mathcal{B},\mathcal{S}} \frac{R(\mathcal{B},\mathcal{S})}{Q(\mathcal{B},\mathcal{S}) - R(\mathcal{B},\mathcal{S})}\left(\sum_i V_i(\theta_i^1) - 1\right) \tag{P}
$$

subject to (EPI),(IF), $\alpha^* \leq 1$ and the set of downward incentive compatibility constraints (local and global) $(\mathcal{IC}^-)$.

While the set of choice variables is now sufficiently reduced, the economic interpretation of (P) is less clear. There is, however, another dual to our problem that provides an intuitive formulation of the ombudsman's problem.

A first step towards this problem is to observe that any solution to (P) subject to some constraints is also a solution to

$$
\max_{\mathcal{B},\mathcal{S}} \frac{Q(\mathcal{B},\mathcal{S})}{R(\mathcal{B},\mathcal{S})} = \sum_i \sum_m \sum_\theta \frac{q_i(m,\theta)y_i(m,\theta)}{Pr(\Gamma)} \tag{O_1}
$$

subject to the same set of constraints. Notice that

$$\frac{y_i(m,\theta)}{Pr(\mathit{\Gamma})} = \frac{\gamma_i(m)}{Pr(\mathit{\Gamma})} U_i(\theta|m) = \rho_i(m|\mathit{\Gamma}) U_i(\theta|m)$$

with $\rho_i(m|\mathit{\Gamma})$ denoting the ex-ante probability of seeing player i, type m conditional on breakdown, that is

$$\rho_i(m|\Gamma) := \frac{\beta_\emptyset^m \gamma_i(\theta^m)}{\beta_\emptyset^m \gamma_i(\theta^m) + \sum_{k \neq m} \beta_\emptyset^k \gamma_i(\theta^k)} = \frac{\beta_\emptyset^m \gamma_i(\theta^m)}{Pr(\mathit{\Gamma})} = \frac{g_i(m,\mathcal{B})}{f(\mathcal{B})},$$

where both $g_i(m)$ and f are linear functions of each element of $\mathcal{B}$ and more importantly are independent of $\boldsymbol{\beta}_\emptyset$. That is given any vector of expected ex-post posteriors $\boldsymbol{\beta}(m)$ the ex-ante expected probability of meeting a certain type after-breakdown is independent of the initial distribution. Further, the denominator f is the same for all type reports.[2]

Now, defining $\boldsymbol{\rho}_i$ to be the $1 \times K$ vector of all ex-ante expected conditional probabilities $\rho_i(m|\mathit{\Gamma})$ and $\mathcal{U}_i^m := \sum_\theta q_i(m,\theta) U_i(\theta_i|m)$ the m^{th} element of the $K \times 1$ vector $\mathcal{U}_i$ we can express (O_1) as

$$\max_{\mathcal{B},\mathcal{S}} \sum_i \boldsymbol{\rho}_i \cdot \mathcal{U}_i \tag{O_2}$$

Finally, spelling out $q_i(m,\theta)$ and slightly rearranging allows us to find a third equivalent expression,

$$\max_{\mathcal{B},\mathcal{S}} \sum_i E_\theta[D_i(\theta)|\mathit{\Gamma}] + E_\theta[U_i(\theta|\theta)|\mathit{\Gamma}], \tag{O_3}$$

with

$$E_m[D_i(\theta^m)|\mathit{\Gamma}] = \sum_\theta^m \rho_i(\theta^m) D_i(\theta^m),$$
$$D_i(\theta^m) = w(\theta^m)\psi(\theta^m),$$

and weighting function

$$w(\theta^m) = (1 - \sum_{k=1}^m \beta_\emptyset^k)(\beta_\emptyset^m)^{-1}$$

[2]Although notationally intense, these results are just due to iteratively applying definitions.

depending only on the prior, and

$$\psi_i(\theta^m) = \begin{cases} U_i(\theta^m|\theta^m) - U_i(\theta^{m+1}|\theta^m), & \text{if } m \neq K \\ 0, & \text{if } m = K. \end{cases}$$

We call ψ_i the *ability premium* and D_i the weighted ability premium.

Summarizing the results of the previous calculations gives us the following proposition

Proposition 9 (Different Formulations)**.** *The following problems are equivalent:*

$$\min_{\mathcal{B},\mathcal{S}} \frac{R}{Q-R} \tag{P}$$

$$\max_{\mathcal{B},\mathcal{S}} \frac{Q}{R} \tag{O_1}$$

$$\max_{\mathcal{B},\mathcal{S}} \sum_i \boldsymbol{\rho}_i \cdot \mathcal{U}_i \tag{O_2}$$

$$\max_{\mathcal{B},\mathcal{S}} \sum_i E_{\rho_i}[D_i] + E_{\rho_i}[U_i] \tag{O_3}$$

In particular problem (O_3) has an intuitive interpretation. It states the problem as the sum of two expected values. The latter is simply the expected utility of the on path conflict game, while the first part describes the expectations over what we call the weighted *ability premium*, D_i. To understand the weighted ability premium, let us first consider the un-weighted ability premium, $\psi_i(\theta^m)$. It describes how much a player of type θ^m can separate herself from the next strongest type via her true ability in the contest. It is the difference between the expected utility conditional on breakdown of a truthfully reporting type θ^m and the next best type's expected utility conditional on breakdown if she mimics θ^m during conflict management. This measure of distance between two adjacent types takes into account that the mimicking type has two advantages from misreporting in case of breakdown: she can adjust her behaviour and is the only one aware of a deviation. What remains is the premium that is due to the fundamental ability difference between the two types and is described by the ability premium. The actual ability premium is obviously determined by the mechanism as the mechanism defines the belief system of player i, reporting m. Nonetheless, for any mechanism that induces a certain belief structure we can straightforwardly compute the ability premium.

The ability premium enters the objective function in a slightly perturbed way, however. The perturbation comes from the fact that the ability premium is an inverse measure of the information rent that needs to be paid to the weaker type. The higher the ability premium, the lower the information rent for a player of type θ^{m+1}. By Lemma 16 the upward adjacent incentive constraints bind, however, and thus paying a higher information rent to type θ^{m+1} implies paying a higher information rent to all weaker types, too. Thus, the weight decreases if we increase the weight put on any type stronger than θ^m as we decrease the amount of players that request the information rent. Obviously, the information rent needs to be paid to any type not only those experiencing breakdown which leads to the prior being the relevant statistic for weighting. As downward incentive compatibility is not incorporated into (O_3), we can ignore the distribution of those types weaker than θ^m and care only about the weighting relative to θ^m which is why $w(\theta^m)$ is decreasing in $\beta_\emptyset^m$.

The expected utility conditional on ending up in the conflict game which forms the second part of (O_3) has an intuitive interpretation, too. By the feasibility constraint (AF) and the definition of α^* we see that the budget constraint is, both in expectations and realized values, always binding for the ombudsman. In turn, this means that the ombudsman has a valuation for decreasing the resources needed to satisfy the binding constraints (PC) and (IC^+). While the choice of the expected weighted ability premium is mostly driven by the attempt to relax (IC^+), the expected utility is mainly driven by (PC). If the ombudsman manages to increase the expected utility conditional on entering the default game, she makes breakdown less costly for the players. That in turn means she needs to pay less in case of settlement to achieve any expected payoff from taking part in conflict management, Π_i. The value of $\Pi_i(\theta^1)$ is pinned down by the strongest types' outside option $V_i(\theta^1)$ which then monotonically affects all other expected payoffs via incentive compatibility. For any fixed ability premium, making the play post-breakdown more efficient thus frees resources the ombudsman can use to settle more conflicts.

3.4 Optimal Conflict Management

To construct the optimal mechanism it is useful to first ignore any constraint on the reduced form problem. Therefore, we first restrict attention to the unconstraint, reduced form problem. Later, we comment on the remaining constraints $\alpha^* \leq 1$, (EPI), (IF) and $(\mathcal{IC}^-)$.

3.4.1 Relevance of Public Signals

By backward induction we solve for the optimal signal first as this can be done based on the expected posterior, only. That is, given any expected posterior we first ask: Is there a public signal that improves upon that expected posterior? The answer to this question is a direct application of Bayesian persuasion as in Kamenica and Gentzkow (2011). That is, whenever there are convexities to be exploited, a public signal improves the result. To see this consider the formulation (O_2) and assume we are at the following (hypothetical) interim state: The ombudsman has announced breakdown and players have updated their beliefs. The signals, however, have not yet realized. At this point all actions of all players (and the ombudsman) other than those when playing the default game are already bygone. Any public signal realization can now only affect actions of the default game.

Lemma 17. *The Border-constraints (EPI) and (IF) never mandate the use of public signals.*

The result is a direct consequence of the above mentioned hypothetical state. The Border-constraints only effect the expected posteriors and are thus independent of any signal. In fact, there is a straightforward notion of when signals indeed may have a positive effect ignoring any of the other constraints: If the (probability) weighted sum of the expected utility post-breakdown at the expected posterior is strictly convex, then introducing a signal is strictly beneficial for the ombudsman. This result follows directly from a Bayesian persuasion argument.

The argument outlined above is in particular helpful as we can answer the question whether signals have an effect with a clear "no" for the unconstraint problem.

Proposition 10 (No Signals). *It is without loss of generality to characterize the equilibrium outcome ignoring signals if neither the downward incentive compatibility constraints* $(\mathcal{IC})^-$ *nor the feasibility constraint* $\alpha^* \leq 1$ *are binding.*

The result follows from the fact that the ombudsman can implement any posterior she likes. For any convex combination of two points on the objective she can simply put all mass on the larger of these points, and thus improve the objective. Thus, signals never have an effect as the choice of expected posteriors is (up to Bayes' plausibility w.r.t. the prior) not limited. Using the Bayesian persuasion argument outlined above, we can derive a simple corollary to Proposition 10 excluding convexities in the expected utility of the default game.

Corollary 5 (No Convexities). *It is without loss of generality to exclude any point at which* $\sum_i \boldsymbol{\rho}_i \cdot \partial \mathcal{U}_i / \partial \beta_j^k > 0$ *for some* $k \in \{1, 2, ..., K\}$ *and* $j \in \{1, 2\}$.

Considering the formulation in expected values (O_3) we see that on-path expected utilities enter the objective positively while deviation utilities enter negatively. Thus, Corollary 5 states that on-path utilities need to be sufficiently concave at the optimum and off-path utilities need to be sufficiently convex in all of the beliefs to be a candidate for the optimal solution.

While Proposition 10 provides us with a clear statement as long as the ignored constraints are non-binding, the result is not true any more once those constraints become relevant. The reason is that the additional constraint exclude certain areas of the multi-dimensional belief-simplex. If the optimal solution lies within an excluded area, it is possible that signals may lead to a constraint that is satisfied in expectations, but not for all realizations. Although we cannot characterize the optimal signal, we can put some structure on that signal. In particular, if any of the downward incentive compatibility constraints bind, and the underlying conflict game (and its equilibrium) were chosen such that only types and beliefs, but not players matter, then the optimal signal is always going to be symmetrizing in a sense that at the hypothetical interim state discussed above, both players are still treated fully symmetric.

Definition 6 (Symmetric Game). The game is symmetric if $U_i(\theta_i, \beta_i, \beta_{-i})$ depends only on $(\theta_i, \beta_i, \beta_{-i})$, but not on i directly for any $(\theta_i, \beta_i, \beta_{-i})$.

Definition 7 (Symmetrizing Signals)**.** A *symmetrizing signal* is a signal that only publicly announces which of the players takes the role of player 1 and which player takes the role of player 2.

Lemma 18. *If in any symmetric game, the downward adjacent incentive compatibility mandates the use of a public signal than it is without loss of generality to assume that this signal is going to be a* symmetrizing signal*.*

The intuition for this Lemma is simple. Whenever only one of the incentive compatibility constraints bind, the ombudsman can always relax pressure on one of the constraints by throwing a fair coin right before the play of the default game which publicly determines which of the players acts as "player 1". The same holds true in essence even if both incentive constraints bind at the unconstraint optimum, but one has a larger shadow value in the Lagrangian sense. If both have the same shadow value the game is already symmetrized.

3.4.2 Relevance of the Ability Premium

In addition to its intuitive interpretation the formulation (O_3) has the advantage that we can easily connect characteristics of the optimal mechanism to the underlying game. Observe that within

$$\max_{\mathcal{B},\mathcal{S}} \sum_i E[w(\theta)\psi_i(\theta)] + E[U_i(\theta)] \qquad (O_3)$$

anything inside the expectations can directly be characterized for any belief, only knowing the parameters of the underlying game. Observe that $\psi_i(\theta^K) = 0$. Thus, the more weight is put on the weakest type the more important becomes the first part of (O_3). $w(\theta^k)$ decreases in the prior belief on any type $\beta^l_\emptyset$ with $l \leq k$ and is unchanged for any $l > k$. We can deduct that a shift in the distribution towards stronger types (in a FOSD-manner) decreases the relevance of the ability premium. Intuitively that means, the smaller the ex-ante probability of paying a high information rent, the less important it is to decrease this rent by offering a high ability premium. In turn, if the distribution is such that the strongest type becomes less and less frequent, her ability premium starts to dominate all other parts of the objective (O_3).

Lemma 19. *Fix any set of beliefs $\beta_\emptyset \setminus \beta_\emptyset^k$ with $k \neq K$. Then, if $\beta_\emptyset^k$ goes to 0, the optimal solution of the unconstraint problem gets arbitrarily close to the maximum of the ability premium of type θ^k, $\max_B \sum_i \psi_i(\theta^k)$.*

Recall that reduction in the ex-ante frequency $\beta_\emptyset^k$ corresponds to an increase in the frequency of the lowest type occurring by the endogenous definition of $\beta^K = 1 - \sum_i^{K-1} \beta_i^k$. The intuition behind this lemma follows directly from the ombudsman's accounting. Recall that the ability premium of type θ^k negatively affects the information rent paid to any type weaker than θ^k. That is, the ability premium is to be paid out to all types weaker than type θ^k. For the sake of the argument assume there are only two types. Then, whenever $\beta_\emptyset^1$ is close to 0, the ombudsman needs to pay the information rent almost always while any post-breakdown conflict game is likely to take place between two weak types and thus not very productive. Therefore, her aim is to minimize the information rent by maximizing the ability premium. If we now increase the type space a similar logic holds. As the ex-ante frequency of a certain type decreases, it becomes less important to separate this type from the stronger types which is done by the ability premia of types $\theta^l < \theta^k$, as the information rent paid to θ^k only becomes less important. However, separating weaker types, in particular the weakest becomes more important as her frequency increases by the endogenous definition of $\beta_\emptyset^K$. Thus, any mass that is shifted from an intermediate type θ^k to the weakest type corresponds to an increase in importance of the ability premium of θ^k.

3.5 Different Types of Conflict Games

Using the general results developed in Section 3.4, we can now apply these results to different types of games to characterize the solution for these type of games. We focus here on two types of games, namely "EPIC Games" and "Contests" which not only cover a large part of conflict games used in the literature and in practice, but also are very different in sensitivity of post-breakdown beliefs: While the actions in "EPIC Games" are invariant to beliefs, "Contests" are typically highly sensitive to the belief system.

3.5.1 EPIC conflict Games

Definition 8 (EPIC Conflicts)**.** A conflict game is called an ex-post incentive compatible conflict (EPIC conflict) if each player's action is invariant with respect to the belief system.

The class of EPIC conflicts includes games with dominant strategies, detail free games or games in which the conflict is inevitably resolved with full force. Examples of an EPIC conflict are those conflicts where the relevant cost of effort are paid prior to the dispute. Such models are for example studied by Hörner, Morelli, and Squintani (2015) in the context of international relations or Spier (1994) in the context of legal disputes. Formally, an EPIC conflict implies that the expected payoff takes the form

$$U_i(\theta_i, \boldsymbol{\beta}) = \sum \beta_i^k u(\theta_i, \theta_{-i}^k). \tag{3.6}$$

and $u(\theta_i, \theta_{-i})$ is independent of $\boldsymbol{\beta}$. Reports to the ombudsman in these conflict games have no strategic effect in case of failure. This has an important consequence for the design of the optimal conflict management namely that it is possible to establish "sure settlements".

Proposition 11 (Optimal Conflict Management with EPIC Conflicts)**.** *The solution to the reduced form optimal conflict management for EPIC conflict ensures that weakest types are guaranteed a settlement except when they meet the strongest type.*

This result generalizes the findings of Hörner, Morelli, and Squintani (2015) who derive a similar result with a particular EPIC default game played under a binary type assumption. The intuition for this is straight forward. The weakest player's settlement value is not "challenged from below" as there is no weaker type. Thus, the ombudsman does not need to send the weakest player into the conflict game for reasons of incentive compatibility. The only reason left for sending her to conflict is then to secure the strongest types' participation constraint which is why conflict games between strongest and weakest type may be possible. The result is also in line with conventional wisdom that the primary purpose for conflict management is to identify "sure settlements".

3.5.2 Contest Games

The property established in Proposition 11 is no longer true if the EPIC assumption does not hold any more. If the conflict game does not satisfy the EPIC assumption, the behaviour of the weakest type might depend on who she expects to meet in the conflict game. If the effort of the player decreases in the expected strength of her opponent, a strategy such as the one outlined in Proposition 11 may not be optimal any more.[3] In such a situation the weakest type expects to face the strongest type and decides not to invest too much within the game since she expects to lose anyway. This change in behaviour has a spillover on other types. If they claim to be the strongest type and enter the conflict game they may face an opponent that thinks she plays a strong opponent. Thus, conflict management may, for the sake of incentive compatibility, allow also for weakest types to meet within the game.

To understand the solution structure of non-EPIC conflict games consider again (O_3). In EPIC conflict games the ability premium of any type is a constant by definition. Thus, all the ombudsman can do is to change the post-breakdown distribution such that the type with the highest ability premium is most prominent after breakdown. Once we leave the class of EPIC games, this property is not true any more as the ability premium is defined endogenously and depends on the ombudsman's choice of mechanism.

One particular subclass in the class of non-EPIC games is the class of contest games. Indeed in reality many games are solved using some form of contest. Intuitively speaking, a contest is a type of game in which each player exerts some effort and a contest success function depending on the players' effort determines the outcome distribution. Formally, we define contest games as follows.

Definition 9. A conflict game is called a contest if, fixing any action, x_i, the expected utility takes the following form

$$U(\theta_i, \mathcal{B}) = F_{\mathcal{B}}(x_i) - c(x_i),$$

and $F_{\mathcal{B}}$ is a distribution function.

[3]Many conflict games with incomplete information feature such elements. Examples include contests, a war of attrition, or first price auctions.

Contest games have the advantage that the cost function is independent of the opponents behaviour and thus, given any action, independent of the beliefs, too. Using an envelop argument we can deduct that for any point at which x_i^* is differentiable,

$$\frac{\partial U(\theta_i, \mathcal{B})}{\partial \beta_i^k} = \frac{\partial F_{\mathcal{B}}(x_i)}{\partial \beta_i^k}|_{x_i = x_i^*},$$

where x_i^* is the optimal action given $\mathcal{B}$.

By the same argument, the second derivative is equal to the second derivative of F with respect to the belief β_i^k. Thus, we know that the second derivative of the ability premium in contest games can be displayed as

$$\frac{\partial^2 \psi(\theta_i^k, \mathcal{B})}{\partial \beta_i^k \partial \beta_i^k} = \frac{\partial^2 F_{\mathcal{B}}(x_i^*(\theta_i^k))}{\partial \beta_i \partial \beta_i^k} - \frac{\partial^2 F_{\mathcal{B}}(x_i^*(\theta_i^{k+1}))}{\partial \beta_i^k \partial \beta_i^k}, \tag{3.7}$$

for $x_i^*(k, m)$ being the optimal action in the contest of player i, type k, who claims to be type m. The ability premium is non-convex if and only if (3.7) is negative, or $F_{\mathcal{B}}$, is more convex at $x_i^*(\theta_i^{k+1}, \theta_i^k)$ than at $x_i^*(\theta_i^k, \theta_i^k)$. Combining now the results from Proposition 10 and lemma 19 we get that for any type θ^k for which the ability premium is (strictly) convex for some belief at some point, there is an ex-ante frequency $\bar{\beta}$ such that for any $\beta_\emptyset^k < \bar{\beta}$, the convexity of (3.7) dominates the curvature of the other types. Then, we can exclude this point from the set of hypothetical optima.

Although this convexity condition might seem weak at first glance, it is actually very useful in particular for contest with little noise. In an all-pay auction for example, a contest without any noise at all, the ability premium of the strongest type is convex in the belief of the strongest player, $\beta_2^1(\theta^1)$ for all $\mathcal{B}$ whenever the utility is differentiable. That means that given $\beta_2^1(\theta^1)$ is small we are guaranteed to get a corner solution. Since parties play a mixed strategy in an all-pay auction, the ability premium is differentiable whenever $\beta_2^1 \neq \beta_2^2$. Thus, given β_2^1 small enough we only need to check for $\beta_2^1(\theta^1) \in \{0, \beta_2^2(\theta^1), 1\}$ as we can exclude the entire interior via Proposition 10 and the rules of the game. It turns out that $\beta_2^1 = \beta_2^2$ becomes optimal in such a case. For the binary type case this directly implies that beliefs are report independent, that is $\beta_i^k(\theta^1) = \beta_i^k(\theta^2)$. In Chapter 2 we have shown that $\bar{\beta}$ in such a case lies beyond the no-triviality condition such that for any specification we get type independent beliefs. While adding noise to the contest changes the players' behaviour and by that (3.7) and $\bar{\beta}$, all functions are continuous in the contest success function such that adding some noise to the contest success function does not change the results qualitatively, as (3.7) remains to be positive for all $\mathcal{B}$.

Finally, the second result we have established in **??**, that is asymmetric beliefs between players, can also be seen here. Assuming type-independence and thus the maximization of the ability premium, the ombudsman wishes to maximize the on-path expected utility of the parties. Since asymmetric contests are in general less inefficient than symmetric ones, she has an incentive to introduce asymmetries between players. This has a negative effect on the ability premium, but a positive one on the expected contest utility. The ombudsman balances these two effects to maximize (O_3). Note again, that also here, as the ex-ante probability of a strong type goes to 0, the forces of asymmetry decrease and the solution converges towards symmetry.

3.6 Conclusion

In this paper we have generalized the findings of Chapter 2 and provided a general approach towards solving for optimal conflict management. We have shown that signals have no effect on the unconstraint optimal solution which allows us to exclude convex areas from the optimum. We have further shown, how the optimal solution may be affected when reintroducing the ignored constraints. We have highlighted the role of the ability premium and analysed two special cases of underlying games. We conclude the paper by briefly discussing two possible extensions to the model.

Ex-Post Participation. One of the main assumptions of our model is that once the mechanism is ratified by everybody any allocation of the pie can be enforced. In many settings with an underlying default game such commitment is not possible. Since many default games are institutionalized either through evolution or constitutions, it may be possible for any party to enforce the default game at any point in time. Such possibilities naturally constrain the ombudsman further. Typically, however this does not change the results too much.

One way to deal with this issue is to augment conflict management such that the ombudsman does not "order" on-path breakdown, but enforces it via a ridiculous offer to either of the parties which the party rejects on-path. If the party does so, the default game is invoked. The other party, however, has still not learned too much about her opponent's type. This can be used by the ombudsman to suppress deviations. Consider the following mechanism. With probability $1 - \epsilon$ a mechanism is played that enforces the same allocation as the one with full commitment. With probability ϵ the ombudsman announces breakdown and perfectly informs all players about the type of the others. Now, consider a player who declines an offer she should have accepted. In such a case the ombudsman knows about this deviation, but the opponent does not. If the ombudsman then commits to sending a signal to the opponent that induces a behaviour that is worst for the deviator, the deviator would only reject the offer if it was lower than this worst punishment. If no offer of the optimal mechanism is lower than the worst punishment the optimal mechanism with ex-post participation constraints is arbitrarily close to the performance of the optimal mechanism under full commitment since the two converge as ϵ goes to 0. This means, all mechanisms in which the lowest settlement share is higher than the worst expected payoff (as a function of the beliefs) performs at most ϵ-worse than the mechanism under full commitment. This is in line with e.g. Hörner, Morelli, and Squintani (2015) who show that the two may sometimes be completely equivalent.

Correlated Types. A second important extension is to consider what happens if types are correlated. Such correlations are beneficial to the ombudsman as she can elicit information of one player also from her opponent's report. In general, the logic of Crémer and McLean (1988) applies here where we can treat the shares like the classical model treats transfers. Budget balance, however, limits the use of Crémer-McLean type of mechanism. In particular, we can expect that the optimal mechanism for correlated types is likely to involve a binding Border-constraint. That is a deviator is punished as in Crémer and McLean (1988), but only to the extend that ex-post implementation is possible. Consequently, full efficiency cannot be achieved with any level of correlation.

Appendix

3.A Proofs

Proof of lemma 13

Proof. Using public signals we can lower the expected utility from not participating whenever the deviating player's utility is concave around the prior by using a signal consisting of two realizations $\hat{\beta}$ and $\check{\beta}$ such that $\lambda\hat{\beta} + (1 - \lambda)\check{\beta} = \beta_\emptyset$ by using the same argument as in Kamenica and Gentzkow (2011). Such signals relax the players participation constraint, such that the participation constraint. Since relaxing the participation is only to the benefit of the ombudsman, we can assume without loss of generality that she reduces the participation constraint as much as possible. That means, the value of non-participation is on the convex hull of the default games expected utility function with respect to the players belief. If this is the case, we can directly use proposition 2 of Celik and Peters (2011) to conclude that it is without loss of generality to assume full participation. □

Proof of lemma 15

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

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

An individual type pair in our setting is given by (θ^m, i) The general auction $\mathbf{p}$ in our setup is defined by a list

$$x_i(\theta^m, \theta^n).$$

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

$$X_i(m) := \sum_{k=1}^{K} x_i(\theta^m, \theta^k)\mu_i(k|m)$$

where

$$\mu_i(n|m) := \frac{\mu(m,n)}{\mu_i^{\bullet}(m)},$$
$$\mu(m,n) := \beta_\emptyset^n \beta_\emptyset^m \frac{1-\gamma(\theta^m,\theta^n)}{1-Pr(\mathit{\Gamma})},$$
$$\mu_i^{\bullet}(m) := \beta_\emptyset^m \frac{1-\gamma_i(\theta^m)}{1-Pr(\mathit{\Gamma})}.$$

Plugging in yields,

$$X_i(m) = \frac{\sum_{k=1}^K \beta_\emptyset^k (1-\gamma(m,\theta^k)) x_i(m,\theta^k)}{1-\gamma_i(\theta^m)} = x_i(\theta^m).$$

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

$$\beta_\emptyset^m z_i(\theta^m) + \beta_\emptyset^n z_{-i}(\theta^n) \leq \tag{BC}$$
$$1 - Pr(\Gamma) - \sum_{k \neq m} \sum_{l \neq n} (1-\gamma(\theta^k,\theta^l)) \beta_\emptyset^l \beta_\emptyset^k$$
$$z_i(\theta^m) \leq 1 - \gamma_i(\theta^m) \tag{BC_1}$$
$$\sum_{k=1}^K z_i(\theta^k)\beta_\emptyset^k \leq 1 - Pr(\Gamma) \tag{BC_2}$$
$$\sum_i \sum_{k=1}^K \beta_\emptyset^k z_i(\theta^k) \leq 1 - Pr(\Gamma) \tag{AF}$$

$$\sum_{k=1}^{K}\sum_{l=1}^{K}\beta_{\emptyset}^{k}z_i(k)+\beta_{\emptyset}^{l}z_i(l)\leq 1-Pr(\Gamma). \tag{BC_4}$$

Note that equation (IF) implies (BC_2) and equation (AF) which implies (BC_4). $\square$

Proof of lemma 16

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

$$z_1(\theta k+1)-z_1(\theta^k)>\gamma_1(\theta^{k+1})U_1(\theta^{k+1}|\theta^{k+1})-\gamma_1(\theta^k)U_1(\theta^{k+1}|\theta^k).$$

Recall that

$$z_1(\theta^{k+1})=\sum_{l=1}^{K}\beta_{\emptyset}^{l}(1-\gamma(\theta^{k+1},\theta^l))x_1(\theta^{k+1},\theta^l)$$

but then if $\mathcal{X}$ was feasible before, it remains feasible if we reduce $x_1(\theta^{k+1},1)$ such that (IC_1^h) holds with equality. Changing this has no effect on the right hand side of the inequality and (IC_i^-) gets relaxed as it is

$$z_1(\theta k+1)-z_1(\theta^k)\leq\gamma_1(\theta^{k+1})U_1(\theta^k|\theta^{k+1})-\gamma_1(\theta^k)U_1(\theta^k|\theta^k)$$

Similarly, suppose (PC_i^1) is not binding, then

$$z_i(\theta^1)>V_i(\theta^1)-\gamma_i(\theta^1)U_i(\theta^1|\theta^1).$$

Provided that $z_i(\theta^1)>0$ the ombudsman could reduce $z_i(\theta^1)$ such that the participation constraint binds. Then she can adjust all $z_i(\theta^k)$ such that the (IC^+) is binding and arrives at another $\mathcal{X}$ which gives the same value of the objective, is feasible and features binding constraints.

If $z_i(\theta^1)=0$, we cannot apply this logic. However, the ombudsman could use the homogeneity of degree 1 of $\gamma_i(k)$ and the homogeneity of degree 0 w.r.t. G to satisfy (PC_i^1) by multiplying all elements of G by $\alpha<1$. Then the same logic applies as above provided that all $z_i(\theta^k) > 0$. If one of them was 0 than multiplying G with $\alpha < 1$ has a positive effect on incentive compatibility.

It is a direct consequence of monotonicity of U, that local incentive compatibility holds globally. Similarly monotonicity of U also ensures that if the participation constraint holds for the strongest type it also holds for the weaker ones. □

Proof of proposition 8

Proof. By Bayes' rule the post-breakdown belief that player j is of type θ^1, given that player i reported θ^1 is

$$\beta_i^1(\theta^1) = \frac{\beta_\emptyset^1 \gamma(\theta^1, \theta^1)}{\sum_{k=1}^K \beta_\emptyset^k \gamma(\theta^1, \theta^1)} = \frac{\beta_\emptyset^k \gamma(\theta^1, \theta^k)}{\gamma_i(\theta^1)} \tag{3.8}$$

Solving this for $\gamma(\theta^1, \theta^n)$ yields

$$\gamma(\theta^1, \theta^n) = \frac{\beta_\emptyset^1}{\beta_\emptyset^n} \frac{1}{\beta_i^1(\theta^1)} \gamma(\theta^1, \theta^1) - \sum_{l \neq 1} \frac{\beta_\emptyset^l}{\beta_\emptyset^n} \gamma(\theta^1, \theta^l) \tag{3.9}$$

which is obviously linear all $\gamma(\theta^1, \theta^n)$ with a non-zero slope. Plugging all $\gamma(\theta^1, \theta^n)$ into each other provides us with $(K-1)$ linear equations which can be solved such that $\gamma(\theta^1, \theta^n)$ can be expressed as a linear function of $\gamma(\theta^1, \theta^1)$ with non-zero slope.

Now we can look at

$$\beta_{-i}^k(\theta^n) = \frac{\beta_\emptyset^k \gamma(\theta^k, \theta^n)}{\sum_{l=1}^K \beta_\emptyset^l \gamma(\theta^l, \theta^n)} = \frac{\beta_\emptyset^k \gamma(\theta^k, \theta^n)}{\gamma_{-i}(\theta^n)} \tag{3.10}$$

which we can then solve in a similar fashion such that all $\gamma(\theta^k, \theta^n)$ are linear functions of $\gamma(\theta^1, n)$ and thus of $\gamma(\theta^1, \theta^1)$ with non-zero slope. Moreover throughout the construction no constant entered the utility function such that we got the desired result for all $\gamma(m, n)$. Finally setting $\gamma(\theta^1, \theta^1) = \alpha$ yields the desired result. The reverse follows by simply plugging into (3.10).

□

Proof of proposition 9

Proof. The proof is to a large extend provided in the main text. Thus, we provide here only the missing steps.

The weighting function q is given by

$$q_i(m,\theta^k) = \begin{cases} (1 - \sum_{l=1}^{k-1} \beta^l_\emptyset) & \text{if } m = \theta^k \\ -(1 - \sum_{l=1}^{k-1} \beta^l_\emptyset) & \text{if } m = \theta^{k-1} \\ 0 & \text{else.} \end{cases}$$

Plugging everything into Q then yields (AF).

Equivalence of equation (P) and equation (O_2) follows from the first order conditions.

Suppose we are at an optimum in any variable x and denote $\partial f(x)/\partial x$ as f' for any function f. Then The first order condition of objective (P) are

$$\Big(\frac{R}{Q-R}(\sum_i V_i(\theta^1)) - 1)\Big)' = \frac{\sum_i V_i(\theta^1)}{(Q-R)^2}(R'Q - Q'R)$$

which, by construction, has the sign determining function $(R'Q - Q'R)$. If we consider the first order condition of Q/R we arrive at a sign determining function of $(Q'R - R'Q)$ such that the first order conditions of the problems

$$\min \frac{R}{Q-R}(\sum_i V_i(\theta^1) - 1),$$

and

$$\max \frac{Q}{R}$$

yield identical predictions in terms of first order conditions. For sufficiency we need to check second order derivatives two which are equivalent since (P) is minimized and (O_1) is maximized whenever

$$Q''R - R''Q > 0.$$

We show that $\rho(m|\mathit{\Gamma})$ is independent of $\beta_\emptyset$ for the case of $K = 2$. However, iteratively applying the same argument yields the result also for $K > 2$. If $K = 2$ we get

$$\gamma_i(\theta^1) = \beta_\emptyset^1\gamma(\theta^1,\theta^1) + (1-\beta_\emptyset^1)\gamma(\theta^1,\theta^2) = \beta_\emptyset^1\gamma(\theta^1,\theta^1) + (1-\beta_\emptyset^1)h_1(\mathcal{B})\gamma(l,l)\frac{\beta_\emptyset^1}{1-\beta_\emptyset^1}$$

where h_1 is independent of $\beta_\emptyset$

$$\beta_\emptyset^1\gamma_i(\theta^1) = (\beta_\emptyset^1)^2(1+h_i^1(\mathcal{B}))\gamma(\theta^1,\theta^1)$$

Spelling out in the same manner yields:

$$(1-\beta_\emptyset^1)\gamma_i(\theta^2) = p^2h_2(\mathcal{B}) + (1-p)\frac{p}{(1-p)}h_3(\mathcal{B})\frac{p}{(1-p)}h_1(\mathcal{B})$$

Such that we can write $(1 - Pr(\mathit{\Gamma}) = (\beta_\emptyset^1)^2 f(\mathcal{B})$ and $\beta_\emptyset^k\gamma_i(\theta^k) = (\beta_\emptyset^1)^2 g_i^k(\mathcal{B})$, where both H and h are independent of the prior.

The transformations to (O_2) and (O_3) follow from spelling out the arguments given in the text. □

Proof of lemma 17

Proof. The result directly follows from a combination of lemma 15 and Observation 3 □

Proof of proposition 10

Proof. Assume for a contradiction, that signals are optimal. By Bayes plausibility each signal implements a mean preserving spread over an expected posterior. Each of the realizations induces a realized posterior. By linearity of expectations the expected value for the ombudsman over the signals is never greater than the maximum over the realized values. Since any solution feasible including signals, i.e. any posterior that is Bayes plausible to the prior is feasible already in expected posteriors, it is always possible to choose the posterior that induces the highest value for the ombudsman directly, making public signals not beneficial for the unconstraint problem. □

Proof of lemma 18

Proof. Assume either of the two downward incentive compatibility constraints is non-binding and the signal is non-symmetrizing. Then, the ombudsman can relay the constraint by introducing a symmetrizing signal, since, by symmetry the players who's downward incentive compatibility binds for the initial signal now receives additional probability mass on states in which her incentive compatibility constraint is non-binding. This relaxes the pressure of the incentive compatibility constraint and is thus profitable for the ombudsman.

Next assume that both incentive compatibility constraints bind, but the shadow value is large for one of the two. By symmetry a symmetrizing signal then relaxes the pressure on the incentive compatibility constraint of the type with the larger shadow value making it profitable for the ombudsman.

□

Proof of lemma 19

Proof. Consider the representation (O_3) of the problem in the following sense

$$\sum_i (\rho_i(\theta^k|\Gamma)w(\theta^k)\psi(\theta^k)) + \sum_{l\neq k} \rho_i(\theta^l)D_i(\theta^l) + E_{\rho_i}[U_i]).$$

Now consider any changes in $\beta^k_\emptyset$. As $\beta^k_\emptyset \to 0$ $E[U_i]$ is not affected and all elements of $\sum_{l\neq k} \rho_i(\theta^l)D_i(\theta^l)$ are bounded, but $w(\theta^k) \to \infty$ such that $\psi(\theta^k)$ becomes arbitrarily important compared to the other parts of the objective. □

Proof of proposition 11

Proof. We show that $\beta^1_1(\theta^K) = 1$ to proof the claim, as this implies by Bayes plausibility that $\beta^1_2(\theta^K) = 1$.

Consider the representation of (P). Next, observe that R is (weakly) decreasing in $\beta_1^k(K)$. Thus it is sufficient to show that $Q-R$ is increasing in $\beta_1^1(K)$ to proof the proposition.

Observe that using the construction that leads to the result of Proposition 8, $\gamma_i(\theta^1)$ is independent of any $\beta_1^k(\theta^K)$. Moreover all $\gamma_1(\theta^l)$ are independent of $\beta_i^k(K)$ for $l \neq K$. Finally, by construction, all utilities $U_1(\theta^i|\theta^l)$ are independent of $\beta_i^k(\theta^K)$ for $l \neq K$ and so is $U_2(\theta^i|\theta^1)$.

Next, by Observation 2 note that the endogenously determined beliefs are $\beta_2(\theta^k)$ for $k \neq 1$ and $\beta_i(\theta^K)$. Moreover, by the definition of the utility function in EPIC games each $U_i(\theta^1|\theta^k)$ of a weighted sum of $\beta_i^l(\theta^k)$,

$$U_i(\theta^k|\theta^m) = \sum_{l=1}^{K} \beta_i^l(m) u_i(\theta^k, \theta^l).$$

Noticing, that, again by construction $\gamma_i(\theta^k)\beta_i^l(\theta^k)$ is independent of $\beta_1(\theta^K)$ for all, $k \neq K$ and thus $y_i(\theta^k, \theta)$ is independent of $\beta_1(\theta^K)$ for all $k \neq K$

We therefore can write $Q-R$ as

$$q_1(\theta^K, \theta^K)(y_1(\theta^K, \theta^K)) + q_1(\theta^K, \theta^K)(y_1(\theta^K, \theta^K) - \beta_\emptyset^K \gamma_1(\theta^K) + C$$

with C a constant independent of $\beta_1^1(\theta^K)$ and the second to last term, the part of R depending on $\beta_1^1(\theta^K)$.

Let us now assume the type space was only binary which yields $\gamma_i(\theta^K) = \frac{\xi}{\beta_1^1(\theta^2)} - \phi_i$, with $\phi_i \geq 0$ with equality for $i = 1$, and $\xi > 0$ the same for both i.

Thus we know that $\gamma_2(\theta^K) = \gamma_1(\theta^K) - \phi_2$ and therefore

$$\begin{aligned}
Q - R - C &= (1 - \beta_\emptyset^1)\left(\sum_i \gamma_i(\theta^2) U_i(\theta^2|\theta^2) - \gamma_1(\theta^2)\right) \\
&= (1 - \beta_\emptyset^1)\gamma_1(\theta^2)\left(U_1(\theta^2|\theta^2) + U_2(\theta^2|\theta^2) - 1\right) - \phi_2 U_2(\theta^2|\theta^2) \\
&= (1 - \beta_\emptyset^1)\gamma_1(\theta^2)\left(\sum_i \left(\beta_i^1(\theta^2)\left(u_i(\theta^2, \theta^1) - u_i(\theta^2, \theta^2)\right) + u_i(\theta^2, \theta^2)\right) - 1\right) \\
&\quad - \phi_2\left(\beta_2^1(\theta^2)\left(u_2(\theta^2, \theta^1) - u_2(\theta^2, \theta^2)\right) + u_2(\theta^2, \theta^2)\right)
\end{aligned}$$

Letting $\Delta_i := u_i(\theta^2, \theta^1) - u_i(\theta^2, \theta^2)$, slightly rearranging and dropping all θ^2 arguments for the sake of readability and taking the derivative with respect to $\beta_1^1(\theta^2)$

$$
\begin{aligned}
\frac{\partial \frac{Q-R-C}{1-\beta_\emptyset^1}}{\partial \beta_1^1(\cdot)} &= -\frac{\gamma_1(\cdot)}{\beta_1^1(\cdot)}\left(\sum_i \beta_i^1(\cdot)\Delta_i + \sum_i u_i(\cdot,\cdot) - 1\right) \\
&\quad + \gamma_1(\cdot)\left(\Delta_1 + \Delta_2 \frac{\partial \beta_2^1(\cdot)}{\partial \beta_1^1(\cdot)}\right) - \frac{\partial \beta_2^1(\cdot)}{\partial \beta_1^1(\cdot)}\phi\Delta_2 \\
&= \gamma_1(\cdot)\left(1 - \sum_i u_i(\cdot,\cdot)\right) + \left(\gamma_2(\cdot)\frac{\partial \beta_2^1(\cdot)}{\partial \beta_1^1(\cdot)} - \gamma_1(\cdot)\frac{\beta_2^1(\cdot)}{\beta_1^1(\cdot)}\right)\Delta_2,
\end{aligned}
$$

where the last term is 0 since the part in parenthesis is 0, and the first part is positive by non-productiveness.

If the type space was not binary, we could use the same steps at the cost of additional notation. However, at this point we had only established that $\beta_1^1(\theta^K)$ should be $1 - \sum_{k=2}^{K-2} \beta_1^k(\theta^K)$. Having established this, we would perform the same exercise given $\beta_1^1(\theta^K) = 1 - \sum_{k=2}^{K-2} \beta_1^k(\theta^K)$ showing that $1 - \sum_{k=2}^{K-2} \beta_1^k(\theta^K)$ should always be increased until $\beta_1^1(\theta^K) = 1$. Finally, $\beta_1^1(\theta^K) = 1$ never interferes with the feasibility constraint or the downward-adjacent incentive constraints which concludes the proof. $\square$

4 Pandering, Persuasion and Sequential Proposal

4.1 Introduction

An agent who implements a certain project often imposes an externality on the rest of the economy through implementation. Examples include mergers that change the market concentration, patent applications that limit competition, and research projects that require resources that otherwise are available for different purposes. To mitigate potential welfare losses caused by the project choices, agents are often required to apply to an authority who decides whether the proposed project should be implemented. In many cases, however, the authority is less informed than the agent about both, the project's realized quality itself, and the quality or existence of a potential alternative project. If private benefits of the project do not coincide with social benefits, the agent might have a strategic incentive to withhold a project for reasons beneficial only to her. Thus, an authority receiving a project proposal must therefore answer the following questions: How large is the likelihood that an even better project exists, but is not proposed by the agent? To what extend is it worth denying the current proposal in hope for a better proposal tomorrow?

The aim of this paper is to study dynamic project choice in a sender-receiver framework. The sender has multiple rounds to propose projects to the receiver, but only one project can be implemented in total. The receiver, in turn, can only implement projects if they have been proposed. Thus, all the receiver can use as a signal for the quality of both the proposed and the not proposed projects is the choice of the sender's proposal. That means the receiver might reject a proposal in hope for a better proposal tomorrow, even if she expects the current proposal to be beneficial.

I show that in principle two types of equilibria arise: the first is a mixed-strategy equilibrium in which the authority accepts the receiver's less preferred project only with some probability. Similar to a related, static model by Che, Dessein, and Kartik (2013) the sender panders towards the receiver-preferred project as she expects this to be implemented with a larger probability. However, the dynamic setting allows for a second type of equilibria, if the time horizon is large enough. These are equilibria in which the sender persuades by waiting to propose the sender-preferred project. The set of waiting equilibria is potentially large. The waiting equilibrium with the shortest waiting time yields the same payoff for the sender, but a lower payoff for the receiver than the mixed strategy equilibrium. The receiver's profits are highest in a waiting equilibrium with intermediate waiting time. The reason is that in such a case good realizations of the sender-preferred project are implemented, but all others are deterred by the waiting period. Starting from the optimum an increase in waiting time deters desirable projects. A reduction in the waiting time, to the contrary, attracts undesirable projects. Both is not beneficial to the receiver.

I contribute to the understanding of project choice problems in many environments by showing that the choice of an equilibrium with considerable delay is to the benefit of the receiver, while the sender never prefers any delay. This may be important in setting up institutions that determine the equilibrium choice. Authorities prefer to pick equilibria that involve considerable delay while private parties always prefer a quick solution.

As an application, I consider a firm (firm 0) within an oligopoly that proposes a merger with another firm to an antitrust authority. Upon seeing only the proposed merger, but not the realized synergies, the authority needs to decide whether to allow the merger or to block it. Even if the merger is preferred to the status quo, the authority faces the trade-off whether to allow the proposed merger or to block it in hope that firm 0 proposes an alternative merger with another firm in the next period. Different from Nocke and Whinston (2010), I consider non-disjoint mergers which is why a myopic policy is not optimal in this setting. I show that the implemented merger depends on the choice of equilibrium.

The analysis is motivated by two observations: first, a merger proposed by a given firm has typically two components: the merging partner and a realization of synergies. Second, if a merger is denied, a firm can offer an alternative merger in the subsequent period.

The two dimensions of a given merger proposal are particularly interesting as preferences are aligned in one dimension, but orthogonal in the other. Preferences about post-merger marginal cost are aligned: both the firm and the authority prefer less costly projects. Across projects, however, preferences differ: while the firm favors less competition post-merger, the authority wants to keep the level of competition as high as possible.

The dynamic component becomes particular interesting if synergies are non-verifiable. Then the merging firms can only commit on their identities, but not on the particular realization. The authority, however, observing the merging firms identities can only form beliefs about the merger's quality based on the merger choice by the proposing firm. Without commitment power this means that the authority updates, both her beliefs about the proposed merger as well as those about the non-proposed mergers before making a decision whether to accept or reject the merger.

Although both, firm and authority, discount future payoffs in the same way, increasing the time needed to implement the firm-preferred project is to the advantage of the authority. If the equilibrium is chosen such that decisions are always made early in the process, the firm is willing to propose her preferred project whenever synergies are not far worse than in the authority-preferred project. Different to that, an equilibrium that involves a longer waiting time for the firm-preferred merger results in the firm proposing the authority-preferred merger more often. As the authority-preferred merger can be allowed right at the beginning, the time loss is mitigated for the authority. However, any equilibrium with a very long waiting time is neither preferred by the authority: on the one hand, the firm would, of course, always propose the ex-ante authority-preferred merger if available. On the other hand, she would also do so in case the synergies in this merger are very low, but those in the ex-ante firm-preferred merger are very high. As authority and firm would prefer the same implementation in such a scenario scenario, both would prefer an equilibrium with intermediate waiting time.

The remainder of this paper is structured as follows: in Section 4.2, I give an overview on the related literature. Section 4.3 introduces the general model and the solution concept. In Section 4.4, I derive equilibria for the case of two possible projects, and analyze the differences. Section 4.5 concludes.

4.2 Literature Review

The paper by Che, Dessein, and Kartik (2013) is probably the closest to mine. Similar to my model, they are interested in a game between an informed sender and an uniformed receiver. In their model, however, the receiver can in principle decide freely what to do, even without the proposal of the sender. While this assumption is reasonable in case a decision maker hires an expert to tell her what to do, my model is more concerned about an authority that regulates what firms are allowed to do. In many situations authorities have only the power to block certain actions, but cannot enforce a particular action of the firm. The pandering equilibrium in their model has similar characteristics to the mixed strategy equilibrium in my model. I show, however, that adding a dynamic component also allows for a second class of equilibria, which are worthwhile to study, as the authority is actually better off in some equilibria within this class.

As proposals are needed, my model is not a model of cheap talk per se. However, given the multidimensionality of the project realizations, I relate to a model by Chakraborty and Harbaugh (2010) who introduce multidimensionality into the cheap talk literature. They show, that multidimensional cheap talk often leads to full revelation. The major difference to this model is that senders in Chakraborty and Harbaugh (2010) can choose in which dimension to communicate to the receiver, while the dimension is fixed in my model. Moreover, they, too, rely on the fact that the receiver has all the decision power and can implement whatever product she wants.

Different to the literature on sequential delegation, e.g. Kovác and Krähmer (2013), who show that sequential delegation supersedes if the differences in preference is small, I do not have gradual information arrival in my model, but a fixed multidimensional type space.

This model also contributes to the old literature on multidimensional signaling of Quinzii and Rochet (1985), Wilson (1985), and Milgrom and Roberts (1986). While Wilson (1985) and Milgrom and Roberts (1986) consider one dimensional types with multidimensional signals, my model has multidimensional types, as e.g. Quinzii and Rochet (1985), but only a one-dimensional signal. Moreover, I consider a dynamic environment which is why, even without explicit cost of signaling, the firm is equipped with a signaling motive.

Since producing the signal itself is costless, this paper also relates to dynamic bargaining models with one-sided incomplete information going back to Sobel and Takahashi (1983). In line with that Fudenberg and Tirole (1983) and Admati and Perry (1987) formulated a theory which uses time as a strategic variable. While Sobel and Takahashi (1983) and Fudenberg and Tirole (1983) have a similar negotiation structure to mine, that is one party makes an offer, the other can only accept or reject and the game continues upon rejection, Admati and Perry (1987) show that strategic delay can be used to signal one's type in two sided-bargaining problems. The main difference to the bargaining literature is that the signal space in this paper is reduced to the identity of the project and payments are not allowed. Firms want to signal that whatever they propose is "without alternative" which, if credible, would lead to immediate implementation. Thus, the firm tries not so much to signal anything about the value of the current project, but tries to discourage the authority's hope for a better project in the future.

In the application part, the model closest to mine is that of Nocke and Whinston (2013). They, model a delegated choice problem to derive optimal merger control. The difference is that is, despite their model being static, that the firm can verify her synergies in the proposed merger. Thus, the only private information the firm has is about the characteristics of the *not proposed* mergers. Further, Nocke and Whinston (2013) assume full commitment of the authority, which is not assumed in my model.

Asymmetric information between firm and authority has been studied also by Besanko and Spulber (1993) who were probably the first to model merger control explicitly as a game between firms and authorities.[1] Ottaviani and Wickelgren (2011) introduce a choice of timing on the side of the authority and derive conditions when ex-post merger control is better than ex-ante merger control. Dynamic

[1] A small literature on merger remedies also connects to this paper, as "projects" can also be seen as different types of remedies. Papers that asses remedies before are, e.g. Lyons and Medvedev (2007), Cosnita and Tropeano (2009), Vasconcelos (2010), and Cosnita-Langlais and Tropeano (2012).

merger review is studied by Nocke and Whinston (2010) and Sørgard (2009). Both papers look at disjoint mergers and how the decision on proposing the merger depends on the approval rule of the authority. Sørgard (2009) derives a rule that determines an optimal investigation probability, while Nocke and Whinston (2010) consider sequential proposal. They show that a myopic policy is optimal if mergers are disjoint. The main difference to my model is, that mergers are not disjoint in my model. Much to the contrary, I look at the different merger options by on single firm and ask how this effects the authorities decision. I find, different to the setup of Nocke and Whinston (2010) that a myopic policy is not optimal in such a case. In fact I show, that the authority can actually use the time dimension to screen merger realization.

4.3 The General Model

This section gives a formal description of the game to be played.

4.3.1 Setup

Consider a game of two players, a sender (the firm) and a receiver (the AA). The firm can choose one among a set of projects to recommend to the AA by sending message m, and the AA can accept or reject the proposed project and a proposed project only. The number of projects is N. The realisation of each project can be one of the following

1. with probability $1 > \lambda_i > 0$ project i is not available,[2] or
2. with probability $1 - \lambda_i$ project i is available. If project i is available, its realisation c_i follows a random variable. I assume each c_i to be independently

[2]Note: The non-availability probability subsumes not only the cases in which a certain project is not available in the usual sense (in a merger context this could be due to personal problems between CEOs, ...). Moreover non-availability may also occur if the project is neither profitable for the firm nor the AA (for example if synergies are too low). In such a case the project never gets proposed and may therefore be treated as if it was not available at all.

drawn from a distribution F_i with a density f_i that is continuous on its support $[\underline{c}, \overline{c}]$.[3]

The firm receives a payoff of $\delta^{t-1}\pi_i(c_i)$ if project i is implemented in period t, the AA receives a payoff of $\delta^{t-1}w_i(c_i)$.

I assume that the functions are ordered such that $\infty > \pi_1(c) > \pi_2(c) > ... > \pi_N(c) > 0$ and $w_1(c) < w_2(c) < .... < w_n(c) < \infty$ for any $c \in [\underline{c}, \overline{c}]$.[4] Each $\pi_i,\ w_i$ is twice continuously differentiable on $[\underline{c}, \overline{c}]$ and decreasing in its argument.

The game has a total number of T periods and both players discount with factor $\delta < 1$. The value of the status quo is normalised to 0.

For the sake of simplicity and to make the decisions non-trivial assume the following:

Assumption 3. Given it is the only project that is available, each project has, in expectation, a positive payoff to both players. In other words:

$$E[w_i] > 0 \quad \forall i.$$

Assumption 4. The probability distributions are such that

$$E[w_{n+1}] > E[w_n] \quad \forall n.$$

Assumption 5. Suppose the firm naively proposes the project that maximises its payoff under full acceptance. Then, the distributions are such that there exists at least one case in which the AAs best response is to wait for a better project in the next round. More formally, this means:

$$\exists j \neq i \in \{1, 2, ..., N\} : E[\delta w_j | \pi_i \in \max_{k \in N} \pi_k] > E[w_i | \pi_i \in \max_{k \in N} \pi_k].$$

[3] I assume common bounds for the ease of notation. None of the results depend on this if there is at least some overlap between the two.

[4] In the case of the merger framework, this reflects the firms preferences c.p. towards larger mergers (functions of lower cardinality) while the AA prefers smaller mergers (functions of higher cardinality). In addition, to reverse the preference order completely simplifies computations, but the general results do not depend too much on this particular preference order. All I need is that firm and AA do not share the exact same order of preferences.

Assumption 3 is only to facilitate computations. It serves to ensure that the threat of non-availability of other projects is present for any project i. Assumption 4, too, is for simplicity. It ensures that the AA not only pointwise prefers projects with higher cardinality, but prefers them in expectation, too. Both assumptions can be relaxed to milder versions, in which they only hold in conditional expectations. Finally, the last assumption is made to make the problem interesting. It is, thus, crucial to the problem. If assumption 5 was not fulfilled, preferences would be completely aligned, and the solution to the problem would be trivial.

All, but the realisations of each project is common knowledge.

The timing of the game is as follows:

1. Before the beginning of the first period, the firm privately observes the vector of realisations $\psi = (c_1, c_2, ..., c_N)$. I may refer to the vector also as the firm's type. The firm's type remains constant throughout the game.

   To deal with the non-availability of certain projects, assume that if a project was not available (which as noted above happens with probability λ_i), the entry in ψ is some number $\bar{\bar{c}} > \bar{c}$.

2. At the beginning of each period, the firm can propose
   a) nothing (by posting identity "0")
   b) exactly one project identity that
      i. is available, i.e. where $c_i \neq \bar{\bar{c}}$ and[5]
      ii. has not been proposed in any previous period.[6]

3. At the end of each period, after observing a non-"0"-proposal, the AA can

[5]Again this assumption is made to simplify the argument. An alternative way to model reduced form unavailability would be to assume that an unavailable project has a negative payoff to firm and authority. This way, if the AA accepts the proposal with positive probability, the firm would never propose it or withdraw her offer after it has been made, which is essentially the same as proposing 0. To avoid unnecessary notation, I simply assume such a proposal is not possible.

[6]This assumption reduces the set of possible equilibria. Allowing re-proposal might cause equilibria in which the firm persistently proposes the same project. All equilibria I derive in this setting here, easily survive in a world were re-proposal is possible. Moreover, the set of equilibria I derive here without re-proposal survives certain refinements, which they survive in the re-proposal setup as well.

a) accept the proposal: in this case we arrive at a terminal node, the proposed project gets implemented and payoffs realize

b) deny the proposal: in this case the game advances to the next period.

4. After the firm has proposed "0", the game continues independently of AA's action.[7]

I am looking for perfect Bayesian Nash equilibria in which no player plays a weakly dominated strategy.[8]

4.3.2 Action and Strategies

The firm's choice is a function that determines its proposal at each time t. This function is a mapping from the set of possible types ψ defined as Ψ to some message m_t. That is,

$$m_t : \Psi \mapsto M_t(\Psi, H_{t-1}).$$

By the rules of the game the set of available messages M_t depends both on the type of the firm (since non-available states are excluded) and the history (since re-proposal is excluded). Thus, at time t and for a given state ψ the set of available messages, $M_t(\psi)$, is defined as

$$\begin{aligned} M_1(\psi) &= \{i : c_i \neq \overline{\overline{c}}\} \cup 0 \\ M_{t+1}(\psi) &= M_t(\psi) \setminus m_t \cup 0 \quad \forall t \geq 0. \end{aligned}$$

Thus, the firm can only post an available project identity that has not been proposed before (i.e. a positive number) or "0" (the Null-Project) which means that it does not wish to implement anything.

The strategy of the AA is slightly more complex and involves a function that determines the acceptance probability and a function of beliefs over types.

[7] A variant in which this is not the case is studied in section 4.4.1.

[8] Since this is a two player game, this restriction is equivalent to using an strategic form trembling-hand refinement.

Before taking an action, the AA forms a belief about the type of the firm. The belief function takes the following form

$$\beta_t : H_{t-1} \mapsto \Delta(\Psi),$$

where $H_t = \{m_1, ..., m_t\}$ is an ordered set of received messages[9] and $\Delta(\Psi)$ is the set of probability distributions over the possible states.

The AA chooses further a function ρ_t in each period to determine the acceptance probability after having received proposal m_t. That is

$$\rho_t : M_0^{max} \setminus H_{t-1} \mapsto [0, 1].$$

where M_0^{max}, the maximal message space at the beginning of the game, is simply the vector $\{1, ..., N\}$ since in principle each project has a positive probability of being available and thus in the message set. After each round, the AA can remove all elements out of the message set that already have been proposed since we excluded re-proposal.

4.3.3 Equilibrium Description

A configuration $\left\{\{m^*\}_{t=1}^{T}\ ,\ \{\rho_t^*\}_{t=1}^{T}\ ,\ \{\beta_t^*\}_{t=1}^{T}\right\}$ is a trembling-hand perfect Bayesian Nash equilibrium if the following conditions hold:

- m_t^* is chosen out of a set of best responses to the equilibrium probability function of the AA. This set of best responses, $M^*(\psi, \rho_t^*, H_{t-1})$ is described as

$$M^*(\psi, \rho_t^*, H_{t-1}) = \{i \in M_t : V(i, \psi, \rho_t^*, H_{t-1}) \geq V(j, \psi, \rho_t^*, H_{t-1}) \quad \forall j \in M_t\} \tag{$*$}$$

[9] Formally, to be precise we need to define the order $\succeq^H$ as $m_i \succeq^H m_j \Leftrightarrow i > j$ and $\tilde{H}_t$ as the set of all messages send up to point t. $H_t(\tilde{H}_t, \succeq^H)$ is then the function that describes this ordered set. For the ease of notation I am going to suppress the arguments $\tilde{H}_t$ and $\succeq^H$.

where $V(i, \psi, \rho_t^*, H_{t-1})$ is the value of proposing project i in period t after having proposed history H_{t-1} in the past and acting optimally thereafter.[10]

To put some structure on $V(i, \psi, \rho^*, H_{t-1})$ it is useful to distinguish between two types of value functions:

i. $V(i, \cdot)$ denotes the *beginning of period value function*, that is the value of proposing project i *at the beginning of a period* and acting optimally thereafter.

ii. $\tilde{V}(i, \cdot)$ is the *end of period value function*, that is, it describes the value the choice of i has on all future periods given the firm acts optimally after that particular period.

Note that $\tilde{V}$ is only an auxiliary construction to properly describe the V, since for message m_t it holds that $\tilde{V}(m_t, \rho_{t+1}^*, \psi, H_{t-1}) = \delta V(m_{t+1}^*, \rho_{t+1}^*, \psi, H_{t-1} \cup m_t)$.

$\tilde{V}$ serves the simple purpose to spell out the *beginning of period value function:*[11]

$$V(i, \cdot) = \rho_t^*(i)\pi_i(c_i) + (1 - \rho_t^*(i))\,\tilde{V}(i, \cdot).$$

- The equilibrium probability function ρ_t^* is chosen such that, for all messages sent with a ex-ante positive probability, it holds that

$$\rho_t^*(m_t) \in \begin{cases} 0 & \text{if } E_t[w_{m_t}|H_t] < \delta\Upsilon\left(H_t\right) \\ 1 & \text{if } E_t[w_{m_t}|H_t] > \delta\Upsilon\left(H_t\right) \\ [0,1] & \text{else} \end{cases}. \qquad (**)$$

Where $E_t[\cdot|H_t]$ denotes the (rational) expectations of the AA in period t conditional on a history of messages H_t, that is already including m_t.

Further, $\Upsilon\left(H_t\right)$ is the expected value at the beginning of the next period conditional on rejecting this periods offer and acting optimally thereafter.

[10] Note that the time dependency of the decision is already incorporated in the variables ρ_t and H_{t-1} such that the value function it self can be modelled as time independent.

[11] For the ease of notation it is, without loss of generality, assumed that $\rho_t^*(0) := 0$.

- Beliefs β_t^* are consistent with Bayes' rule whenever possible.
- No player plays a weakly dominated strategy.

4.4 Analysis

For the sake of tractability, I am considering the case of only two projects here.

4.4.1 Equilibrium of the modified game

Before looking at the general game, I want to focus on a slightly modified version for the ease of exposition.

Definition 10. A game is called "modified" if it follows the setup as it was laid out in section 4.3 except for the additional rule that the game ends directly whenever the firm proposes 0, i.e. the "null-project".[12]

This modification is primarily of didactic use as it reduces the number of equilibria. Later, I am going to show that the result also survives in the general game even under quite demanding refinement criteria and discuss other equilibria (and there relation to this one) that exist only in the general game.

The modified game with only two possible projects has a simple type space, namely a vector $\psi = (c_1, c_2)$, and a message set in the first period that is at most $\{0, 1, 2\}$. Under assumption 3 and 5, proposing naively cannot result in an equilibrium. The AA would never accept project 1 and thus proposing 1 when 2 is also available is not optimal. If $\lambda_2 > 0$, rejecting 1 always and accepting 2 can also not be an equilibrium if we do not allow for weakly dominated strategies. Since we ruled out weakly dominated strategies, the firm would always propose project 1 if it is the only project available and under assumption 3 the AA should accept it. Moreover, it is even possible to state the following.

Lemma 20. *For the case of $N = 2$ and under assumption 3 and 5, no pure strategy trembling-hand perfect equilibrium exists in the modified game.*

[12] An equivalent formulation would be a setting in which the authority can shut down negotiations by formally "accepting" the null project and implementing it. Again, the equilibrium proposed in this section would survive such a game.

The proof of this lemma can be found in section 4.A, as can all others not in the text.

The intuition behind this lemma is straight forward. Proposing "0" is always weakly dominated if any other proposal is available since a proposal of "0" necessarily terminates the game and keeps the status quo, which is never preferred by the firm if any project is available. Thus the firm is always going to propose something. Assumption 5 rules out that the AA just waives whatever project is proposed. Since with some probability each project is not available, it cannot be optimal to unconditionally block a certain project either. The firm best responds to such a blocking strategy by proposing the project only if it is the only one available. With this strategy of the firm, however, unconditional blocking is not optimal since each project is profitable in isolation (assumption 3). Thus, pure strategies played by the AA can never be optimal.

Since every finite game has at least one trembling-hand-perfect Nash equilibrium,[13] the trembling-hand-perfect equilibria must involve at least one mixing party. In fact, since project realizations are continuously distributed, the AA is the only player that is going to mix on more than a probability zero set. To construct the equilibrium, first recall the equilibrium conditions from equation ($*$) and equation ($**$). For the equilibrium construction assume that the second project is the conditionally better looking one, i.e.

Assumption 6. $E[w_1|\pi_1 > \pi_2] < \delta E[w_2|\pi_1 > \pi_2]$.

Although it puts some restriction on the distributions, this assumption still allows for a wide range of distributions. Whenever both state variables c_1, c_2 are identically distributed or if c_2 does not have too much weight on high cost draws we always can find a $\delta < 1$ such that the assumption is fulfilled. Further assume that project 2 is always accepted.[14]

[13]The proof of this is due to Selten (1975) and can be found in various textbooks such as Fudenberg and Tirole (1991) and Mas-Colell, Whinston, and Green (1995).

[14]Later, I show that in equilibrium this is indeed the case.

Under these assumptions, the probability $q_\rho(\psi)$ of proposing project 1 for a given acceptance probability $\rho_1(1) = \rho$ is characterised by

$$q_\rho(\psi) = \begin{cases} 0 & \text{if} \quad \rho\pi_1(c_1) \leq \zeta\pi_2(c_2) \\ 1 & \text{if} \quad \rho\pi_1(c_1) > \zeta\pi_2(c_2) \end{cases} \tag{4.1}$$

with

$$\zeta = (1 - \delta + \rho\delta) \in [0, 1],$$

which accounts for both the alternative proposal and the second round. In equation (4.1) we may assume that the firm does not mix since indifference only happens on a probability zero set of ψ.

With the help of equation (4.1), it is possible to construct a cut-off function $\tilde{c}_\rho(c_2)$ that assigns, given ρ and c_2, a unique value $\tilde{c}$ such that for all $c_1 < \tilde{c}$ the firm prefers to propose project 1 while she is going to propose project 2 whenever $c_1 > \tilde{c}$. The function takes the following form:

$$\tilde{c}_\rho(c_2) := \begin{cases} \max\{c_1 : \rho\pi_1(c_1) \geq \zeta\pi_2(c_2)\} & \text{if it is not } \emptyset \\ \overline{c} & \text{if } c_2 = \overline{\overline{c}} \\ \underline{c} & \text{else.} \end{cases} \tag{$*'$}$$

Note that the function $(*')$ is defined for each $\rho > 0$ over $c_2 \in [\underline{c}, \overline{c}] \cup \overline{\overline{c}}$ and has the following properties:

i. it is continuous in ρ and c_2, since π_2 is continuous;

ii. it is weakly increasing in c_2 since $\pi_1(c_1)$ and $\pi_2(c_2)$ are both decreasing in their arguments;

iii. it is weakly increasing in ρ since the left-hand side of the inequality increases in ρ and $\pi_1(c)$ decreases in c;

iv. for fixed ρ it is weakly increasing in δ since the right-hand side decreases in δ, giving the inequality (weakly) more slack at all points.

v. as $\rho \to 0$, $\tilde{c}_\rho \to \underline{c}$ for all c_2. Since $\pi_2(\overline{c}) > 0$ for all c_2 there exists an $\underline{\varepsilon}_{c_2} > 0$ such that $\forall\ \rho < \underline{\varepsilon}_{c_2} \Rightarrow \tilde{c}_\rho(c_2) = \underline{c}$. Monotonicity of $\pi_2(c_2)$ ensures that $\underline{\varepsilon}_{c_2}$ is decreasing in c_2;

vi. since $\pi_1 > \pi_2$ by definition, it holds that there also exists an $\overline{\rho} < 1$ such that for all $\rho > \overline{\rho}$ we have $\tilde{c}_\rho(c_2) \geq c_2$;

vii. as a consequence of the monotonicity in both arguments, it also holds that the smallest c_2 at which the highest $\tilde{c}_\rho(c_2)$ is reached[15] (weakly) decreases in ρ. $\tilde{c}(\underline{c})$, on the other hand, increases in ρ.

To define the equilibrium action of the AA, recall equation $(**)$. By Lemma 20 we have excluded pure strategy equilibria, thus, $1 > \rho > 0$.

To re-write condition $(**)$, define 1_2 to be the event in which project 1 is recommended although project 2 is available, too, and define 1_0 to be the event at which 1 is sent and it is the only message other than 0 that is feasible. Since condition $(*')$ defines the optimal behaviour of the firm for any ρ and therefore determines the probabilities and expectations of events 1_2 and 1_0, it is sufficient to rewrite condition $(**)$ as follows

$$\begin{aligned} & E[w_1|m_1 = 1] = E[\delta w_2|m_1 = 1] \\ \Leftrightarrow \quad & (1-\tilde{\lambda}_2)E[w_1|1_2] + \tilde{\lambda}_2 E[w_1|1_0] = (1-\tilde{\lambda}_2)E[\delta w_2|1_2] \\ \Leftrightarrow \quad & (1-\tilde{\lambda}_2)E[w_1|1_2] + \tilde{\lambda}_2 E[w_1] = (1-\tilde{\lambda}_2)E[\delta w_2|1_2] \\ \Leftrightarrow \quad & (1-\tilde{\lambda}_2)\left(E[w_1 - \delta w_2|1_2]\right) + \tilde{\lambda}_2 E[w_1] = 0, \end{aligned} \tag{$**'$}$$

where $\tilde{\lambda}_2 = \frac{\lambda_2}{\lambda_2 + (1-\lambda_2)P(m=1|1_2)}$.

The first step divides expectations into two subsets. The second then makes use of the fact that, given that project 1, but not project 2, is available, the firm would always propose 1 in order to not play a weakly dominated strategy. Noticing that w_2 and w_1 in the first term are conditioning on the same event gives the third step of equation $(**')$.

Lemma 21. *If assumption 3 and 6 hold, there always exists a ρ such that $1 > \rho > 0$ and equation $(**')$ are fulfilled.*

[15] Typically this highest $\tilde{c}_\rho(c_2) = \overline{c}$, but one may choose a ρ small enough that even $\tilde{c}_\rho(\overline{c}) < \overline{c}$.

Intuitively, this lemma says that it is always possible to choose a ρ so low that whenever both projects are available, the firm chooses the second project. In turn, this choice means that the firm only (weakly) prefers to offer project 1 if nothing else is available. If it did so whenever only 1 was available, the AA would want to accept that proposal since it is welfare-increasing in expectations. Since gradual changes in ρ lead at most to a gradual change in both expectations, the difference of the two is continuous. If the difference is continuous, the indifference condition must, in line with the intermediate value theorem, eventually be satisfied for some $1 > \rho > 0$.

In fact, it is possible to derive an even stronger statement

Lemma 22. *If the conditions for Lemma 21 hold, then whatever ρ solves equation ($**'$) is unique.*

The intuition to this lemma is, again, quite simple. As ρ increases, the firm increases the number of states in which it proposes project 1. For any given c_2, this increase leads to additional states that are all worse than the worst state under the original regime. Therefore, the expected payoff the AA earns from implementing project 1 can only fall in ρ for proposals of 1. The expectations of postponing implementation, on the other hand, increase for the same reasons, since there is a larger set of "desirable" c_2 that lead to a proposal of 1. Thus, increasing ρ decreases the expected payoff when implementing 1 and increases the expected payoff of waiting for any ρ. Consequently equation ($**'$) can at most hold for one $\rho \in (0, 1)$.

To describe the equilibrium, it remains to show that it is optimal to always accept project 2, i.e. $E[w_2|m = 2] \geq \delta E[w_1|m = 2]$ needs to hold under the equilibrium ρ. This is guaranteed by assumption 4, as the following lemma shows.

Lemma 23. *The AA prefers project 2 over project 1 whenever 2 is proposed.*

The intuition underlying this lemma is that, in expectation, whenever the firm proposes a certain project, this project needs to be "better" than a certain threshold. In turn, whenever the firm does not propose a project it must have a worse realization than this threshold. Since in equilibrium the AA is indifferent under the proposal of 1, the expected payoff from project 2's realization must be higher for the authority if the firm proposes 2. Expectations of waiting after the proposal of project 2 must be smaller by the same argument. Thus always accepting project 2 is optimal in equilibrium.

Combining Lemma 21 and 23 provides existence of a mixed strategy equilibrium as the following lemma shows

Lemma 24. *Suppose assumption 3, 4 and 6 hold. Then, there exists a unique mixed strategy trembling hand perfect Baysian Nash equilibrium in the modified game with the following properties:*

- *The second project is accepted whenever it is proposed.*
- *The first project is accepted with probability $\rho < 1$.*
- *If the second stage is reached, proposal project 2 in the second stage gets accepted.*
- *Firms propose project one whenever ψ is such that $c_1 < \tilde{c}_\rho(c_2)$.*

4.4.2 Comparison to the original game

All the results of the previous section have been derived under the assumption that players play the modified game. In this section, I am going to discuss how these results carry over to the original game.

Theorem 2. *There exists an equilibrium of the original game that has the following properties:*

- *It is outcome equivalent to the unique equilibrium of the modified game*
- *On the equilibrium path actions are identical to the ones of the modified game*
- *The equilibrium does not fail the universal divinity, if any of the two conditions holds*
  - *i. the acceptance probability ρ^* is larger than the discount factor δ, or*
  - *ii. the worst project is smaller than the outside option $w_1(\bar{c}) < 0$.*

The crucial aspect for survival of the refinement is to find beliefs such that the AA wishes to deny off equilibrium proposals. For the second proposal this is not that hard since the firm never wishes to deviate since that only would incur time costs. Thus, any type is equally likely to deviate. The story is quite different for project 1. There might be a reason to wait for the firm if it beliefs that in later periods the acceptance probability was large enough. For sure, any off-path acceptance probability that attracts at least some type $\hat{\psi}$ must also attract all types that cannot propose project 2 because it is unavailable. However, discriminating between those is not possible, since they are all only interested in implementing project 1. This way, if any type has, for some off-path acceptance probability any incentive to go off-path, also the type $\underline{\psi} = (\bar{c}, \bar{\bar{c}})$ has an incentive to do so. This is, by definition the worst type in terms of payoff for the AA. Thus, if there is at least one state such that the AA prefers the outside option, she prefers the outside option to implementing project 1 under $\underline{\psi}$. Since this type engages in all off-path activities, it cannot be excluded under universal divinity. Thus, if the AA beliefs any deviator is of type $\underline{\psi}$ she has reason to deny it and the equilibrium sustains.

The crucial aspect for the sustainability of the equilibrium is the fact that there exists a positive probability that the better looking project does not exist.

Together with the second restriction, namely that $w_1(\bar{c}) < 0$, the non-availability guarantees universal divinity for any length of the game. Thus, the existence of the equilibrium of the modified game carries over to the general game even if we were to restrict our self to a rather narrow set of equilibria.

Uniqueness, however, does not necessarily survive. In particular, the two dimensional type space of the firm allows for a wide range of equilibria if we consider the generalized game. Together with the possibility to sent the “0” message as often as possible a wide range of justifiable off-equilibrium beliefs can be consistent even under strict refinements such as universal divinity. To understand this recall that the equilibrium of the modified game does not fail universal divinity in the general game simply because “all” types for which project 2 was not available survive the iterated D2 criterion. That is, if the AA sees a proposal of project 1 anywhere off the equilibrium path it can, even with strong restrictions be of any type (regarding project 1’s cost function).

If negotiations go on for sufficiently many rounds, that is if T is large enough, a second set of equilibria which I am going to call “waiting equilibria” would arises. In such a “waiting equilibrium” the firm proposes the null-project for sufficiently many rounds in order to signal that the ex-ante AA-preferred project is of poor quality or not available. In the traditional bargaining literature these are sometimes the only equilibria that survive even mild refinements and they are typically called “signalling equilibria”.[16] In the present context signalling also is an issue in the “ρ^*-equilibrium” described above, which is why I differ in terminology.

The decision whether to accept or reject an offer made by the firm in those type of equilibria is in fact very similar too the decision rule in the ρ^*-equilibrium. Whenever the AA believes that the project offered in the current round is at least as good as the discounted value of what she can expect in the future, she would accept the proposed project, whenever this is not the case she denies approval. Given this, the firm might delay its proposal strategically under some circumstance while it may prefer to propose something that gets accepted right away instead of suffering the cost of waiting. More precisely, suppose the AA always accepts project 2 and accepts project 1 only at time $t \geq \tau + 1$.[17] Then for each c_2 it is possible to construct a function $\hat{c}_\tau(c_2)$ such that the firm chooses to propose project 1 in period τ whenever $c_1 < \hat{c}_\tau(c_2)$ and project 2 (if possible) in all other cases. More formally this is

$$\hat{c}_\rho(c_2) := \begin{cases} \max\{c_1 : \delta^\tau \pi_1(c_1) \geq \pi_2(c_2)\} & \text{if it is not } \emptyset \\ \overline{c} & \text{if } c_2 = \overline{\overline{c}} \\ \underline{c} & \text{else.} \end{cases} \tag{$*'$}$$

This function almost perfectly corresponds to $\tilde{c}^*_\rho(c_2)$ defined in the beginning of this section. Consequently, this, too, provides a cutoff value for the firm given any τ. As in the other equilibrium the AA optimally chooses τ such that whenever the firm proposes project 1 in period $\tau + 1$ it holds that the AA only accepts the proposal if $E_{\tau+1}[w_1|H_{\tau+1}] \geq \delta E_{\tau+1}[w_2|H_{\tau+1}]$. Before describing the possible “waiting equilibria” let us first properly define the term.

[16]Examples would be Admati and Perry (1987) or Cramton (1992)

[17]To understand why this is $\tau + 1$,recall that a proposal in t=2 is discounted with δ^1 since participants ex-ante payoffs are given in period one values.

Definition 11. An equilibrium is called a "waiting equilibrium at τ" if it is an equilibrium of the game and the authority uses the following strategy:

- Project 2 is accepted in any period
- Project 1 is rejected in any period $t \leq \tau$ and accepted in all other periods.

Next, it is straightforward to use the cutoff function $\hat{c}_\tau$ derived above to describe the equilibrium that is closest related to the ρ^*-equilibrium from above.

Lemma 25. *Consider a certain specification of the game and fix the equilibrium at ρ^*. Ignoring the integer constraint in t there exists a "waiting equilibrium at $\underline{\tau}$" where*

$$\underline{\tau} = \frac{\ln(\rho^*) - \ln(1 - \delta(1 - \rho^*))}{\ln(\delta)}$$

that corresponds in the firms decision rule to the ρ^*-equilibrium, that is $\hat{c}_{\underline{\tau}}(c_2) = \tilde{c}_{\rho^*}(c_2)$ for all c_2. This equilibrium may be considered as the "shortest possible" waiting equilibrium in a sense that all other waiting equilibria require a waiting time $\tau \geq \underline{\tau}$. Moreover, if and only if $\rho^* > \frac{\delta}{\delta+1}$, then $\underline{\tau} < 1$.

The intuition behind this becomes quite clear if you think about what happens if the AA accepts proposals of project 1 "too early". Then the firm would propose them in a way, such that the AA better denies it in hope of a better outcome on the other project. From the construction of the ρ^*-equilibrium it is already clear that whenever the firm decides via cutoff rules, then there exists a unique decision function such that the AA is indifferent in expectations between the two projects. While this function was enforced in the ρ^*-equilibrium with the acceptance probability, it is now done by appropriately choosing τ. Since the two correspond we just need to look for the parameter in which the functions are identical. This is what describes the smallest τ.

In the model, of course, time is discrete and thus the above described point is generically never an integer. However, as the following proposition shows, for all $\tau > \underline{\tau}$ a waiting equilibrium at τ exists.

Theorem 3. *If $T - 1 > \underline{\tau}$, there exists a set of waiting equilibria at τ for all τ such that $\underline{\tau} \leq \tau \leq T - 1$. All waiting equilibria survive under the refinement of universal divinity if $w_1(\bar{c}) < 0$*

While each equilibrium on its own is in fact independent of the maximum duration of the game, there is a one to one mapping between numbers of proposal rounds played after $\underline{\tau}$ and the number of waiting equilibria that exist.

4.4.3 Welfare effects

To compare the different equilibria it is useful to think about how the equilibrium payoffs in the different equilibria behave and to exercise some comparative statics. Following the order in which the equilibria were characterized, I start by examining the ρ^*-equilibrium and thereafter looking at the waiting-equilibria.

Besides comparing the equilibria with each other, a natural benchmark would also be the (ex ante) equilibrium payoffs if the AA had the ability to commit ex-ante to accepting only project 2. This benchmark case is going to be called a "conservative AA" throughout the rest of the paper.

To simplify notation denote $E_a[w|\rho^*]$ to be the ex-ante payoff expectation of the AA under equilibrium acceptance probability ρ^* as in the equilibrium of section 4.4.1 and let $E_a[w|0]$ be the ex-ante expectations if the AA was to deny project 1 always, i.e. behaving conservatively.

Proposition 12. *The ρ^*-equilibrium always assigns a lower payoff to the AA than she gets when committing to only accepting the second project.*

Proof. The mixed strategy equilibrium increases the AA's payoff only if $E_a[w|\rho^*] > E_a[w|0]$. The following argues that this is not true.

Spelling out the first yields

$$E_a[w|\rho^*] =(1-\lambda_2)(1-\lambda_1)\int\limits_{\underline{c}}^{\overline{c}} w_2(c_2)\left(1-F_1(\tilde{c}_{\rho^*}(c_2))\right)\mathrm{d}F_2(c_2) \tag{4.2}$$

$$+(1-\lambda_2)\lambda_1 E[w_2] \tag{4.3}$$

$$+(1-\lambda_2)(1-\lambda_1)\int\limits_{\underline{c}}^{\overline{c}} \delta w_2(c_2)F_1(\tilde{c}_{\rho^*}(c_2))\mathrm{d}F_2(c_2), \tag{4.4}$$

where the first is the expected value in case both projects are available and project 2 is proposed given ρ^*, the second is the part in which project 2 is proposed since it is the only one available, and the last part covers the expectations whenever 1 is proposed, using the indifference condition $(**')$.

Further we may decompose

$$
\begin{aligned}
E_a[w|0] =&(1-\lambda_2)E[w_2] \\
=&(1-\lambda_2)\Big(\lambda_1 E[w_2] + (1-\lambda_1)E_[w_2]\Big) \\
=&(1-\lambda_2)\lambda_1 E[w_2] + (1-\lambda_2)(1-\lambda_1)\int\limits_{\underline{c}}^{\overline{c}} w_2(c_2) dF_2(c_2) \\
=&(1-\lambda_2)\lambda_1 E[w_2] \\
&+ (1-\lambda_2)(1-\lambda_1)\int\limits_{\underline{c}}^{\overline{c}} w_2(c_2)\left(F_1(\tilde{c}_{\rho^*}(c_2) + (1-F_1(\tilde{c}_{\rho^*}(c_2)))\right) dF_2(c_2) \\
=&E_a[w|\rho^*] + \underbrace{(1-\delta)(1-\lambda_2)(1-\lambda_1)\int\limits_{\underline{c}}^{\overline{c}} w_2(c_2)F_1(\tilde{c}_{\rho^*}(c_2))\mathrm{d}F_2(c_2)}_{\geq 0}. \qquad (4.5)
\end{aligned}
$$

The first step just spells out the conditional expectations, the second divides it into the events in which project one is or is not available. The third transforms them into integral form, the fourth multiplies with 1 to back out the ex-ante expectations of the equilibrium in the last step. Since expectations in equilibrium are positive when 1 is proposed, the leftover term is positive. Thus, from an ex-ante point of view the AA prefers accepting only project two if both were available. □

Intuitively, this is rather obvious if one thinks about the different channels at work in the model.

If the AA was able to allow only project 2 the firm can never exploit the game in her preferred way if opinions about optimal actions differ. Thus, the persuasion channel is completely shut down. Pandering, to the contrary, is driven to its maximum.[18] Finally, the sequentiality channel is shut down as well, since in

[18] In fact ex-ante pandering is optimal from the point of view of the AA. However, on an interim stage there might actually be over-pandering, that is the firm proposes a project that both firm and AA prefer less than the other one (interim), but that was not preferred ex-ante by the AA.

equilibrium there is no second period. This sequentiality works, as we saw in the discussion of the equilibrium, as an insurance for the firm to “try out” her preferred project. That way, commitment to pure strategies in fact must work in favour of the AA since it only shuts down the channels that benefit the firm. The AA therefore earns something like a “conservative commitment markup”. To facilitate later comparison it is useful to think of this markup as relative to the expected payoff if only project 2 was was proposed. To do so, define

$$\phi(\rho^*) := \frac{E[w|0] - E[w|\rho^*]}{E[w|0]}$$

and rewrite equation (4.5) as

$$E_a[w|0] = (1 - \phi(\rho^*))E_a[w|0] + \phi(\rho^*)E_a[w|0]. \tag{4.6}$$

In this context ϕ can be interpreted as the “relative conservative commitment markup” (RCCM).

With this reformulations it is easier to tell how much (in relative terms) the AA looses by being unable to ignore any proposals of 1 in equilibrium.

Proposition 13. *Consider the ρ^*-equilibrium. An increase in the time preference parameter δ leads to*

- *a decrease in the acceptance probability ρ^*,*
- *a decrease in the ex-ante likelihood that project 1 is being proposed and*
- *a decrease of the RCCM (and a higher absolute payoff for the AA).*

While an increase in δ decreases the cost of waiting one more round for both the firm and the AA, the increased expected welfare in the second period leads the AA to reduce its probability of accepting the first project. That means that the firm reduces the states at which it proposes the first project and panders more towards the second. This way it needs to give up some of the gains from the increased δ. Further, the “commitment markup” goes down and thus, the AA loses less from not being able to commit to a pure strategy.[19]

[19] In section 4.4.4 I discuss why commitment to a pure strategy is not optimal either and why this leads to lower credibility of a commitment assumption on the AA’s side.

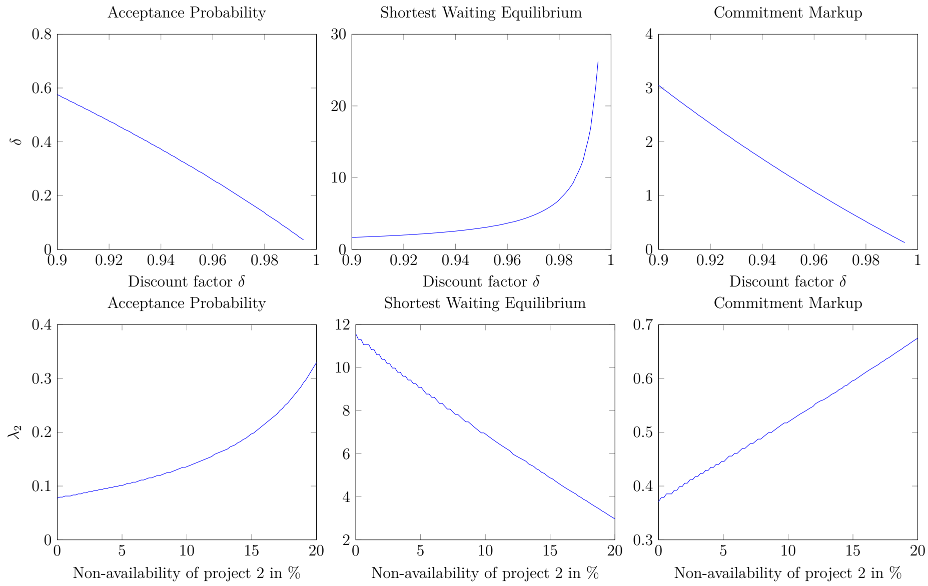

Figure 4.1: *Comparative statics with respect to availability of the second project λ_2 and discount factor δ. The upper panels deal depict effects on the ρ^*-equilibrium and the lower panels that on the shortest waiting equilibrium.*

Another interesting comparative statics in this model is the effect of non-availability of the second project. This sheds light on to the point of how much the threat of not having an alternative projects works in favour of the firm.

Lemma 26. *Consider the mixed strategy equilibrium ρ^*. An increase in the non availability probability of project 2, λ_2, leads to*

- *an increase in the acceptance probability ρ^* in equilibrium*
- *an increase in the ex-ante likelihood that project 1 is being proposed and*
- *an increase of the RCCM.*

This result is, of course, not very surprising. As it becomes more likely that the second project is not available at all, the firm creates, by proposing project 1, a larger threat to the AA that this is the only project it has in fact. Thus, given project 1 has been proposed, there is a higher chance that it is the only one around (in which case the AA would want to accept) and the interim expectations of the AA are driven down. She reacts by increasing the acceptance probability.

In terms of the AAs payoffs under no commitment, to the contrary, the story is not so clear since λ_2 effects both the ex-ante expectations of a conservative strategy as well as those under the equilibrium strategy.

The relative mark-up increases nevertheless, as the interim pressure the firm can put on the AA increases with a higher probability of project one being the only one available.

Turning to the waiting equilibria, it might at first be interesting to see how the set of possible waiting equilibria changes with the parameters. While the largest waiting equilibrium (if any such equilibrium exists) is always defined by T, as Theorem 3 shows, the shortest waiting equilibrium might change as we change the parameters of the model. With help of Lemma 25 it is obvious that all parameters except the discount factor δ effect $\underline{\tau}$ only through ρ^*. Observe, that

$$\frac{\partial \underline{\tau}}{\partial \rho^*} = \frac{1}{ln(\delta)\rho^*} \frac{1-\delta}{1-\delta(1-\rho^*)} < 0.$$

Thus, as ρ^* increases $\underline{\tau}$ decreases. The reason can easily be found in the equivalence of the decisions rules of the firm in both equilibria. Therefore, a higher effective acceptance probability needs to result in a smaller minimum waiting period.[20]

The only parameter, that effects $\underline{\tau}$ also in another way than through ρ^*, is the discount factor δ. Its effect on $\underline{\tau}$ is not all that obvious since the direct effect could go either way. Nonetheless, as the following lemma shows, the overall effect of a change in δ on $\underline{\tau}$ has the opposite sign than a change of δ has on ρ^*.

Lemma 27. *Consider a waiting equilbrium at $\underline{\tau}$. An increase in the discount factor δ leads to an increase in $\underline{\tau}$.*

Even if the effect though ρ^* always supersedes the direct effect of δ, the additional effect that a change in the discount factor has already points towards the presence of some difference in the analysis of the waiting equilibria compared to that of the

[20] In fact, in terms of elasticity, one can show that the movements do not correspond one-to-one but that $\underline{\tau}$ shrinks less than proportional compared to the increase in ρ^* if $\delta > \rho^*$ since that implies

$$0 > \frac{\partial \underline{\tau}}{\partial \rho^*} \frac{\rho^*}{\underline{\tau}} > -1.$$

ρ^*-equilibrium. These differences are what is going to be considered next. To do so, I especially focus on two effects not present in the ρ^*-equilibrium, that may, in addition, help to understand the change in payoff between the different waiting equilibria.

First, each round of waiting decreases the AAs payoff since the agreement is reached later in time. The AA is impatient and therefore suffers "cost of delay". Second, and quite different is what I call the "benefit of the doubt" effect which works in the other direction. If in equilibrium there is a longer waiting period, this does not only lead to pandering, but also to a smaller set of states for which the firm decides to wait. That means the more costly it is to wait for the firm, the more often she proposes the AA-preferred project already in the first round and thus the (potential) delay is to the benefit of the AA.

Which effect dominates, if we were to move from the waiting equilibrium at τ to the one at $\tau + 1$ is hard to say without further parameter restrictions. Ignoring again, the integer constraint in t and assuming that T is large enough, the "AA-most-preferred" equilibrium (in an ex-ante sense) lies in the interior of the interval $(\underline{\tau}, T-1)$. In fact the shortest waiting equilibria, i.e. that at $\underline{\tau}$ is never the most preferred one as it is always dominated by the ρ^*-equilibrium. For T sufficiently large, the most preferred equilibrium is in addition also neither the ρ^*-equilibrium nor an equilibrium in which for all possible states at most project 2 gets implemented as the following proposition shows:

Proposition 14. *Consider a game in which*

$$T - 1 > \overline{\tau} = \frac{\ln \pi_2(\overline{c}) - \ln \pi_1(\underline{c})}{\ln \delta}.$$

Then, ignoring the integer constraint in t, the "AA-most-preferred equilibrium" is a waiting equilibrium at τ^* with $\tau^* \in (\underline{\tau}, T-1)$. Moreover, independent of the length of the game, the "AA-most-preferred equilibrium" is never at $\underline{\tau}$.

Further description of the AA-most-preferred waiting equilibrium is, however, hard to accomplish since there are counteracting effects in the derivative of the expected payoff of the AA. To see this, it may be helpful to disentangle these expectations:

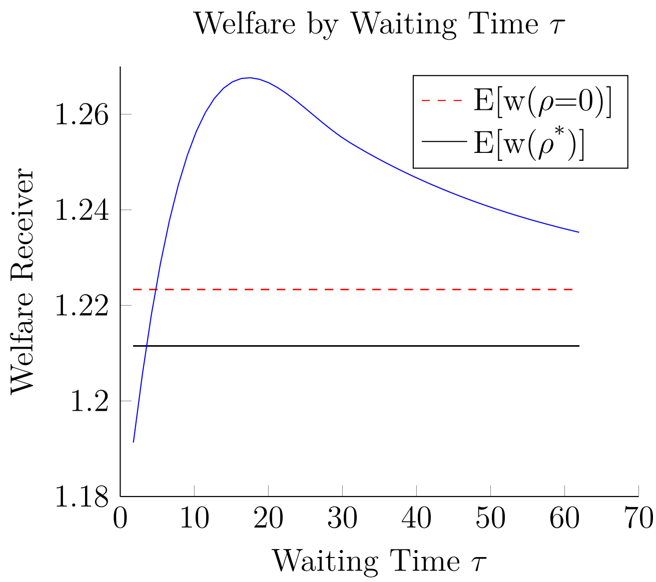

Figure 4.2: *Welfare effects in a linear demand Cournot economy for mergers from three firms to two. The blue line is the (ex-ante) payoff of the AA in a waiting equilibrium at τ. It always starts below the black solid line (by assumption 6) and is above the dashed line and decreasing to the right of $\overline{\tau}$. For comparison the black solid line is the (ex-ante)payoff of the ρ^* equilibrium and the red dashed line denotes the (ex-ante) payoff to the AA if she could (again ex-ante) commit to reject any proposal of the larger merger. Although single-peakness of the blue line is assured in simple linear-demand Cournot models, it cannot be generalized to more complex settings.*

$$
\begin{aligned}
E[w|\tau] = & \underbrace{(1-\lambda_1)(1-\lambda_2)\int_{\underline{c}}^{\overline{c}} w_2(c_2)\left(1-F_1\left(\hat{c}_\tau(c_2)\right)\right)\mathrm{d}F_2(c_2)}_{A} \\
& + \underbrace{\lambda_1(1-\lambda_2)E[w_2]}_{B} \\
& + \underbrace{(1-\lambda_1)(1-\lambda_2)\delta^\tau \int_{\underline{c}}^{\overline{c}}\int_{\underline{c}}^{\hat{c}_\tau(c_2)} w_1(c_1)\mathrm{d}F_1(c_1)\mathrm{d}F_2(c_2)}_{C} \\
& + \underbrace{\lambda_1(1-\lambda_2)\delta^\tau E[w_1]}_{D}.
\end{aligned}
$$

Terms A and B describe the expectations in cases in which the firm decides to propose project 2 right away. C and D describe them in cases in which project 1 is proposed after the waiting time. Within the two blocks, the first term (i.e. A and C) describes the situation in which both projects are available, the second (B and D) the case in which the proposed project is the only one available.

Observe that term B is in fact completely independent of τ and D decreases in τ. The other two terms are not that easy to determine, so let us look at the derivatives a bit closer

$$
\frac{\partial A}{\partial \tau} = (1-\lambda_1)(1-\lambda_2)\int_{\underline{c}}^{\overline{c}} w_2(c_2)\left(1-f_1\left(\hat{c}_\tau(c_2)\right)\right)\frac{\partial \hat{c}_\tau(c_2)}{\partial \tau}\mathrm{d}F_2(c_2) \tag{4.7}
$$

$$
\begin{aligned}
\frac{\partial C}{\partial \tau} = & (1-\lambda_1)(1-\lambda_2)\ln(\delta)\delta^\tau \int_{\underline{c}}^{\overline{c}}\int_{\underline{c}}^{\hat{c}_\tau(c_2)} w_1(c_1)\mathrm{d}F_1(c_1)\mathrm{d}F_2(c_2) \\
& + (1-\lambda_1)(1-\lambda_2)\delta^\tau \int_{\underline{c}}^{\overline{c}} w_1(\hat{c}_\tau(c_2))f_1(\hat{c}_\tau(c_2))\frac{\partial \hat{c}_\tau(c_2)}{\partial \tau}\mathrm{d}F_2(c_2)
\end{aligned} \tag{4.8}
$$

While the sign of the first derivative and the second part of the second derivative remain unidentified due to the fact that w_i can take on different values, the first part of the second derivative is clearly negative. However, the results of Proposition 14 indicate that at some point $\frac{\partial A}{\partial \tau}$ in combination with the second part of $\frac{\partial C}{\partial \tau}$ must overturn the negative effects in the rest of the terms. Without further parameter restrictions, it is not possible to pin down when this overturning is going to take place or even to decide whether the $E[w|\tau]$ function is single peaked in τ between $\underline{\tau}$ and $\overline{\tau}$.

4.4.4 The case of commitment

In many application it is assumed that authorities such as an AA have the possibility to ex-ante commit themselves to a certain probability rule. This is often justified by the idea that an authority is a long lasting institution that cares for reputation and therefore has the opportunity to credibly commit itself to a certain action. As shown above, a commitment to a pure strategy is ex-ante preferred by the AA if she finds herself in the ρ^*-equilibrium. In this part I am going to argue that such a commitment cannot be optimal. Thus, the AA would need to commit to a mixed strategy. This, in contrary is in practice again hard to implement.[21]

In terms of the structure of the game, notice that allowing for commitment changes the following to aspects:

(i) With commitment the game changes in a way such that the AA is now a first mover.

(ii) With commitment the AA can only gain compared to the case without commitment.

The second statement is not only a corollary to Proposition 12, but in such settings an even more general point. If commitment is an option the AA could always replicate the outcome of the game as if it was played without commitment by simply committing to some ρ or τ as needed in any equilibrium. Since it has no private information at all, the firm would, given the probabilities, propose exactly as in the case without commitment. This way any commitment device can only be to the benefit of the AA.

Another important fact, that is useful for constructing a solution to the commitment game is the following

[21]One reason might be, that (real) randomization across proposals might violate legal statues. However, if the AA needs to justify denials, she might have an incentive not to stick to the commited randomization rule and thus acts in an interim optimal way as in my model.

Remark. A commitment game can be modelled w.l.o.g. ignoring any beliefs either of the players have.

To see this, observe that the only private information present here is the state vector ψ. If we let nature move after the AA has chosen their probability rule (which by definition only depends on common knowledge parameters, i.e. the prior), then there is never any uncertainty about the state in this game and we can restrict ourselves to (trembling-hand) subgame perfect equilibria.

Suppose now, the AA would commit to a rule where she accepts project 1 with probability 0 and project 2 with probability 1. In such a case the expected welfare was

$$(1 - \lambda_2)E[w_2], \tag{4.9}$$

since the firm always proposes the second project whenever it is available.

However, a pure strategy as described above is not optimal in the commitment case as the next proposition shows.

Proposition 15. *It cannot be profitable for the AA to commit to a rule of pure time-invariant strategies in which she always accepts one of the projects and always rejects the other (in each period). it is neither optimal to accept both projects in the first period nor to accept none.*

This proposition is strongly connected to Proposition 14. There it is already stated that the best waiting equilibrium yields a higher payoff than commitment to a pure time-invariant decision rule would.

One might now wonder whether commitment to a pure strategy is possible in which the firm accepts a certain project only after some time period much like the waiting equilibria. In fact this can also never be optimal as the following proposition shows:

Proposition 16. *In a game with commitment it is never optimal for the AA to commit to a strategy where she accepts one project right away and another after a certain time t has past.*

Proof. Recall the decision of the firm in any waiting equilibrium. For any τ it is possible to find a $\rho < 1$ such that

$$\delta^\tau = \frac{\rho}{1-\delta(1-\rho)}.$$

Under such ρ it holds that $\hat{c}_\tau = \tilde{c}_\rho$ by definition. But whenever the firm proposes 1 under τ she does so under ρ as well, but earlier than under ρ. Thus, the cost of delay shrinks, yielding a higher payoff to the AA. □

In other words, pure strategies are never optimal with commitment.[22] But that might lead to some "strategic trembling" as for example to accept project 1 once it has been proposed, since it might be the only one available. Afterwards, to maintain credibility, an institution would of course claim this only happened by "mistake" or has an even easier job by blaming the randomization device if we allow for commitment to mixed strategies. Thus, avoiding interim information updates of the AA might be a hard job even for an authority and not too easy to assume in this setting.

To sum up, the commitment option might serve as a benchmark, but should not be taken too seriously when it comes to implementation. If at all, the AA could impose a fixed waiting time to achieve at least second best and to use this as a coordination device to pick her most preferred equilibrium.

4.5 Concluding Remarks

I consider a dynamic sender-receiver game in which an informed firm can propose a project to an uniformed authority who decides whether to implement the project or to block it. A proposal is required to implement a project and the firm may propose an alternative project in case her proposal is rejected. Overall, only one project can be implemented.

In principle there a two types of equilibria: There is a mixed-strategy equilibrium in which the authority implements her less-preferred project with a certain probability already in the first period, and a set of waiting equilibria in which the

[22]Note that we ruled out trivial examples with assumption 6.

authority implements it only after a certain period of waiting. A firm that postpones her proposal signals a high quality of the authority's less-preferred project. The authority's most-preferred project, on the other hand, is always implemented right away in both types of equilibria.

I find that both firm and authority prefer the mixed-strategy equilibrium over a short waiting equilibrium to save on the cost of delay. However, as the time of waiting increases, intermediate realizations of the authority's less-preferred projects are deterred from being proposed. This has a positive effect on the authorities ex-ante expected payoff. If the waiting time becomes too long, however, all realization of the authority's less-preferred project are deterred which decreases the authorities payoff. The sender never prefers a waiting equilibrium over the mixed-strategy equilibrium.

My findings show that multiple periods allow authorities to learn from the proposal of the firm and yields a better solution than the static model from the authority's point of view. The reason is that the firm is in competition with her future self. The improvement materializes either through *pandering* by the firm such that she sometimes proposes a less-preferred project to avoid it being blocked. Alternatively, the firm may *persuade* the authority by delaying her proposal to signal a good realization of the firm-preferred project.

I show that firms always prefer a quick implementation of a project, while authorities prefer to delay the implementation of an ex-ante less-preferred project for an intermediate length. Delaying the implementation for too long is not beneficial for the authority, but still preferred to mixing at the initial period. A delay time that is too short on the other hand is not preferred by either of the two players.

The findings contribute to the discussion on how approval processes should be designed in a dynamic setting. I show that delaying certain approval decisions can be more effective in screening projects than probabilistic acceptance rules. Such waiting games are, however, not to the benefit of the firm who has an incentive to lobby for an equilibrium that leads to instantaneous proposals at any point. Thus, I provide a theory that suggests that delays in the merger review process might not entirely be determined by technical constraints. Instead, there may be a strategic component for delay depending on the choice of equilibrium. Moreover, although possible, an authority would not want to switch to an equilibrium with shorter waiting time as this harms her ex-ante expected payoff.

The insights gained in this analysis provide several interesting directions for future research. While I focus on a game theoretic approach in this paper, a natural follow-up is to consider a mechanism design approach instead. Close to the analysis here would be a mechanism without commitment power. That is a third-party that collects the multi-dimensional private value of the firm, i.e. the realizations of all projects available, and offers a recommendation to both firm and authority on how to proceed. This may overcome parts of the coordination failure in the mixed-strategy equilibrium and makes this equilibrium more attractive for the authority. A second possible direction is to allow the authority to costly investigate the proposal of the firm. If she had some technology to test any efficiency claims, e.g. via a mean-preserving spread as in Kamenica and Gentzkow (2011), at an interim stage, she may have an incentive to do so if the alternative project is likely to be similar in terms of quality. On the other hand, if she expects the alternative to be considerably different, she might not want to bear these cost. Finally, results may be affected if there was competition on the firms side, that is if two firms compete about whose project is going to be implemented. In such a case firms have a stronger incentive to pander towards the authority preferred project, as they do not compete only against their future selves, but also against a competitor within each period.

Appendix

4.A Proofs

Proof of Lemma 20

Proof. The proof proceeds by contradiction.

In general, in any pure strategy equilibrium, it needs to hold for the acceptance vector $\boldsymbol{\rho}_1 = (\rho_t^1, \rho_t^2)$ that $\rho_t^i \in \{0, 1\}$ for all i and t. In this case, both i and t need only to be considered within $\{1, 2\}$.[23]

In a first step, I show that if least one project is profitable in the second stage, there is no pure strategy equilibrium. The second step considers then the remaining case.

First Step

Define $E_t[w_i|j]$ the expectations of the authority at time t about project i if project j has been proposed in t.[24]

For an equilibrium in pure strategies in the subgame of the second period, it must be the case that $\rho_2(i) = 1$ if $E_1[w_i|j] > 0$ for $j \neq i$. Suppose this was the case.

Then, for a pure strategy equilibrium of the whole game, $\rho_1(j) = 1$ if $E_1[w_j|j] > \delta E_1[w_i|j] > 0$. Again suppose this was the case.

If all this was true, the Firm for sure proposes j if $\pi_j > \pi_i$ and whenever only project j is available. Assumption 5 implies that $\rho_1(i) = 0$, since otherwise the firm would propose i whenever it maximises its profits.

[23] Note that if the game comes to a third period, then the firm must propose 0 along the way which ends the modified game immediately.

[24] Note again by the rules of the modified game and $N = 2$ the tuple (j, t) suffices for a full history of proposals.

This implies that for all cases in which both $c_i \neq \bar{\bar{c}}$ and $c_j \neq \bar{\bar{c}}$, the firm proposes j.

If $c_j = \bar{\bar{c}}$ but $c_i \neq \bar{\bar{c}}$, the firm must propose i because proposing 0 is a weakly dominated strategy. In fact this is the only state at which the firm proposes i, but then, by assumption 3, $E_1[w_i|i] > 0 = E_1[w_2|i] \Rightarrow \rho_1^i = 1$, a contradiction.

Thus, whenever assumption 3 and 5 hold and $E_1[w_i|j] > 0$, it cannot be that $\rho_1(k) = 1$ for $k \in \{i, j\}$.

This means the only possibility left in this case is to deny all proposals in the first period.

With this, it is not possible that $\rho_2(k) = 1$ for both k since that would lead the firm to propose j whenever i is more profitable and the other way around. But by assumption 5 the AA would then accept at least one proposal. This violates the assumption that $\rho_1(k) = 0$ for both k and is a contradiction.

In turn, this means at least one merger needs to be denied even in the second round. Let this be merger j. As a result, whenever j is available, j is proposed and gets denied in the first stage for trembling-hand-perfection. But then again, if only i was available, this would be proposed for sure in period 1 due to the refinement and the assumption of the modified game. As discussed above the AA would then need to accept such a proposal since she knows that j does not exist. This violates $\rho_1(i) = 0$ and is, again, a contradiction.

Second Step

To complete the proof we need to consider the case in which $E[w_i|j] < 0$ for both $i = 1, 2$. Suppose this was the case, then we know that $\rho_2(k) = 0$ for any k. As we know from the previous discussion $\rho_1 = 1$ cannot be true, thus at least one merger needs to be declined in the first stage.

If only one gets denied in the first period, the same arguments as in step 1 hold. If the firm knows merger j gets rejected in both stages, while i gets accepted only in the first stage, it proposes i whenever possible in the first period. However, if only j is around it is going to be proposed. But then j must be accepted by the AA.

The remaining pure strategy equilibrium candidate is now that no project gets accepted at any stage. Thus, $E_t[w_i|i] < 0$ for all i and t. This again cannot be achieved if assumption 3 holds and the firm never chooses a weakly dominated strategy.

□

Proof of Lemma 21

Proof. Notice first that equation ($**'$) is continuous in ρ since $\tilde{c}_\rho$ is continuous. Notice further that by assumption 6 for $\rho = 1$ it holds that $E[w_1|m_1 = 1] < \delta E[w_2|m = 1]$ conditional on the firm best responding. Now pick some $\rho' > 0$ such that

$$q_{\rho'}(\psi) = 0 \quad \forall \psi : c_1, c_2 \neq \overline{\overline{c}},$$

which exists since $\pi_1(\overline{c}),\ \pi_2(\underline{c}) > 0$. Choosing such a ρ' is equivalent to setting $\tilde{c}_{\rho'} = \underline{c}$. The first part of the last line in equation ($**'$) reduces to 0 with $\tilde{c}_{\rho'}(c_2)$. The latter part, however, does not depend on it and is, by assumption 3, positive. Thus, the left hand side of the last line of equation ($**'$) can be both positive and negative within the relevant range and due to continuity and the intermediate value theorem, this means that equation ($**'$) also holds for at least one $1 > \rho > 0$. □

Proof of Lemma 22

Proof. To prove uniqueness first observer from the second line of equation ($**'$) that $\lambda_2 E[w_1]$ is constant over all $0 < \rho < 1$. To show that ρ is unique it suffices thus to show that $h(\rho) := E[w_1|1_2] - \delta E[w_2|1_2]$ is decreasing in ρ. I do so in two steps. The first proofs that $E[w_1|1_2]$ is in fact decreasing in ρ and the second is to show that $E[w_2|1_2]$ is increasing in ρ. This suffices to conclude that $h(\rho)$ is decreasing in ρ and thus, ρ is unique.

First Step To see that $E[w_1|1_2]$ is decreasing in ρ fix some $1 > \rho > 0$ and c_2.

Given this observe that the firm proposes project 1 if $c_1 \leq \tilde{c}_\rho(c_2)$. Suppose now that ρ was increased to $1 \geq \rho' > \rho > 0$. By the properties of $\tilde{c}_\rho$ we know that $\tilde{c}_{\rho'} \geq \tilde{c}_\rho$. This in turn means that there exists a (possibly empty) interval $[\tilde{c}_\rho, \tilde{c}_{\rho'}]$ for which the firm proposes project 1 under ρ' but not under

ρ. Whenever project 1 was proposed under ρ, it is also proposed under ρ'. Since w_1 is a decreasing function in c this means that $E[w_1|c_1 \leq \tilde{c}_{\rho'}, c_2] \leq E[w_1|c_1 \leq \tilde{c}_\rho, c_2]$ for all c_2. Integrating over all c_2 this leads to the following statement:

$$E[w_1|c_1 \leq \tilde{c}_{\rho'}] \leq E[w_1|c_1 \leq \tilde{c}_\rho],$$

and thus $E[w_1|1_2]$ is decreasing in ρ.

Second Step To see now that $E[w_2|1_2]$ is increasing in ρ fix again some $1 > \rho > 0$ and in addition c_1.

Define now

$$\hat{c}_2 := \begin{cases} \min[\tilde{c}^{-1}(c_1)], & \text{if it exists} \\ \underline{c}, & \text{otherwise.} \end{cases} \tag{4.10}$$

Then the firm proposes 1 for every draw $c_2 \geq \hat{c}_2$. If we now pick $1 \geq \rho' > \rho > 0$, the firm would propose 1 under the new regime whenever it has been proposed under the old regime. In addition due to the $\tilde{c}_\rho$ being increasing in ρ there exists an additional (possibly empty) interval $[\check{c}_2, \hat{c}_2]$ for which the firm proposes project 1 as well. Since $\check{c}_2 \leq \hat{c}_2$ this leads to (weakly) increasing expectations in the decreasing function w_2. Thus,

$$E[w_2|c_2 \geq \hat{c}_2] \geq E[w_2|c_2 \geq \check{c}_2].$$

Since this holds for a generic c_1 we can similar to step one draw the conclusion that $E[w_2|1_2]$ increases in ρ, thus $h(\rho)$ is decreasing in ρ which completes the proof.

$\square$

Proof of Lemma 23

Proof. The AA strictly prefers project 2 to project 1, whenever project 2 is proposed if and only if the following condition holds

$$E[w_2|m = 2] > \delta E[w_1|m = 2].$$

Claim. Suppose the firm plays its equilibrium strategy given any $0 < \rho < 1$. Then $E[w_1|m^* = 1] \geq E[w_1|m^* = 2]$.

To see this, fix some c_2 and ρ.

Recall that w_1 is a monotone decreasing function and observe that thus

$$E[w_1(c_1)|c_2, c_1 < \tilde{c}_\rho(c_2)] \geq E[w_1(c_1)|c_2, c_1 > \tilde{c}_\rho(c_2)] \tag{4.11}$$

since

$$\inf\{w_1(c_1) : c_1 < \tilde{c}_\rho(c_2)\} = w_1(\tilde{c}_\rho(c_2)) = \sup\{w_1(c_1) : c_1 > \tilde{c}_\rho(c_2)\}.$$

Since the relation in equation (4.11) holds for any c_2 it also holds that integrating out c_2 we have the following relation

$$\begin{aligned} & E[w_1(c_1)|c_1 < \tilde{c}_\rho(c_2)] \geq E[w_1(c_1)|c_1 > \tilde{c}_\rho(c_2)] \\ \Leftrightarrow \quad & E[w_1(c_1)|m^* = 1] \qquad \geq E[w_1(c_1)|m^* = 2] \end{aligned}$$

Note that for this we only used the decision rule of the firm and thus can by the same argument (just redefining the cut-off as a function of c_1) state that

$$E[w_2|m^* = 2] \geq E[w_2|m^* = 1].$$

Together with the equilibrium condition from equation ($**'$) this implies the following

$$E[w_2|m^* = 2] \geq E[w_2|m^* = 1] > \delta E[w_2|m^* = 1] = E[w_1|m^* = 1] \geq E[w_1|m^* = 2]$$

and thus,

$$E[w_2|m^* = 2] > \delta E[w_1|m^* = 2]. \tag{4.12}$$

□

Proof of Lemma 24

Proof. The first two properties directly come from Lemma 21 and 23 except for the part that $E[w_2|m = 2] > 0$. Uniqueness follows from Lemma 22.

For the remaining property, namely that project 2 gets accepted in the second period, observe that this only happens if project 1 is proposed in the first period. A proposal of 1 in $t = 1$ occurs if either 1 is the only available project or 1 and 2 are available, but 1 is more profitable to the firm given the AAs strategy. In the first event, the expected payoff is simply $E[w_1]$. For cases where both 1 and 2 are available, notice that this event implies that for every c_2 there exists a $\tilde{c}_\rho(c_2) \leq \bar{c}$. The firm only proposes 1 if $c_1 < \tilde{c}_\rho(c_2)$. Given any c_2 this implies $E[w_1|m = 1, c_2] \geq E[w_1] > 0$. Integrating over all c_2 can therefore not change the sign. Thus, $E[w_1|m = 1] = \delta E[w_2|m = 1] > E[w_1] > 0$ and the proposal in the second stage, project 2 is always accepted.

What is missing is now to show that $E[w_2|m = 2] \geq 0$. This holds since we already know from the above that $E[w_2|m = 1] > 0$. If this is the case, we also know that (similar to the discussion about $\tilde{c}_\rho(c_2)$) there exists a cut-off $\hat{c}(c_1)$ for each c_1 such that message project 2 is sent whenever $c_2 < \hat{c}(c_1)$. Since w_2 is decreasing in c it must be the case that $E[w_2|m = 2] \geq E[w_2|m = 1] \geq 0$. Thus, project 2 is not only preferred to project 1 if project 2 is proposed but it is also preferred to not accepting.

Trembling hand perfection of the overall game is easily checked, recognizing that players play no weakly dominated strategies if indifferent. Neither accepting nor rejecting is weakly dominated and so is no mixture of the two. The firm is only indifferent on a probability zero event, and even then there is no strategy that dominates the other. For all other cases, trembling of the AA would not change the result. For the outcome irrelevant subgame after 2 has been proposed and denied (which never occurs in equilibrium), we can take any of the trembling hand perfect equilibria in the subgame, which exist for sure.

The last part that is missing is to show that expectations are positive along the path. This is ensured by assumption 3 and the observation that project 1 needs to be sufficiently good to be proposed, thus its expectations cannot be negative. Since they equal the expected value of waiting this cannot be negative either. Lemma 23 then provides that the expectations of project 2 cannot be negative either. □

Proof of Theorem 2

Proof.

On the equilibrium path: First, observer that on the equilibrium path of the modified game, the AA accepts any proposal in the second round. Suppose now, that players play the same strategies in the original game, then there is no reason why in the second period either off the players should have an incentive to deviate from the strategies they played in the modified game. Taken this as given, the reduced form game of the first period leaves the same options for both players as the modified game. Except for the case in which the firm proposes 0, the game is of no difference from the perspective of the players as the reduced form game of the first period of the modified game. Thus, again there is no incentive to deviate from the equilibrium strategies of the modified game. Since 0 is only proposed if the firm has no project available in the modified game, the same is going to happen in the original game. This way, on the equilibrium path it is trivial that the firm would (for any history of 0) have only 0 in its message set and would thus propose 0 in every period until T. Thus, the actions of both players on the equilibrium path are identical to those of the modified game and the outcome is the same as that of the modified game. The same holds if the firm only has project 1 available and gets rejected. Then the firm proposes 0 in the second period and forever after until the termination period T.

Off the equilibrium path:

1. First consider the off equilibrium action in which the AA observes the firm proposing project 1 in the second period after a proposal of 0 in the first period. These are in terms of off equilibrium beliefs the most problematic ones.

   The proof of this part is of constructive nature and defines beliefs that are consistent with the notion of universal divinity[25] In each step I am going to describe which beliefs actually are in line with the criterion itself to thereafter argue why iterative application would not fail the criterion either.

   i. Assume $\rho^* > \delta$. Suppose further project 1 is available. This means that whenever the firm was to propose 1 on the equilibrium path, a necessary condition for a deviation with positive probability under the intuitive criterion (and therefore also necessary under D2) is

   $$\rho^* \pi_1 + (1 - \rho^*)\delta \max\{\pi_2, 0\} \leq \delta \pi_1.$$

   This can be rearranged to

   $$(\rho^* - \delta)\pi_1 + (1 - \rho^*)\delta \max\{\pi_2, 0\} \leq 0.$$

   The above cannot hold if $\rho^* > \delta$ and therefore whenever project 1 was proposed in the first round on the equilibrium path, it has a 0 probability to be proposed after a 0 is observed off the equilibrium path.

   If 2 was proposed in the first round it gets accepted for sure, thus a deviation into the second round can only be profitable for types where $\rho^* < \delta \rho_2'$, with $\rho_2' < 1$ being the second round acceptance probability for project 1. This necessity contradicts the assumption that $\rho^* > \delta$. Thus, also those types are eliminated under the intuitive criterion. Since no type survives the intuitive criterion, none thus survive D2. Thus, the equilibrium is robust to those refinements.

[25] This has been introduced by Banks and Sobel (1987).

Iteration in this case is not in question, since all types are eliminated already after the first round.

ii. Assume that $w(\overline{c}) < 0$.

For what follows, it is necessary to define the set of best responses the AA has to such an off equilibrium deviation. In principle the AA can respond either with 0 or 1 and both responses can be supported by at least one belief system. Due to the continuous density, there also exists at least one convex combination of beliefs that justify a response of 0 and 1 respectively which makes the AA indifferent. thus, all (mixed) actions of the AA are part of the best response.

Next, for a given state $\psi = (c_1, c_2)$ with $c_i \neq \overline{\overline{c}}$ and the equilibrium ρ^* a deviation for the firm as proposed can only be profitable if the following condition holds:[26]

$$\max\{\pi_2(c_2), \rho^*\pi_1(c_1) + (1-\rho^*)\delta\pi_2(c_2)\} \leq \rho'\delta\pi_1(c_1) + (1-\rho')\delta^2\pi_2(c_2) \tag{4.13}$$

This can be rearranged to

$$\max\{(1-\delta^2+\rho'\delta^2)\pi_2(c_2), \rho^*\pi_1(c_1)+(1-\rho^*-\delta(1-\rho'))\delta\pi_2(c_2)\} \leq \rho'\delta\pi_1(c_1) \tag{4.14}$$

Fixing some arbitrary $c_1 = c$ and ρ' for which the firm wants to deviate at least in some states and propose 0 in the first and 1 in the second period. With this, the left hand side of condition 4.14 is increasing in π_2. For the first element this is trivial, since $(1 - \delta + \rho'\delta) > 0$ by definition. For the second part, observe that the second element is increasing in π_2

[26]Note that I am assuming here, that project 2 gets accepted in the subsequent period. Assuming any kind of different acceptance probability for project 2 other than 0 would not change the result. Since this effect is independent of the parameter c_1 those types may not be excluded under certain belief systems. However, results (and passing the divinity criterion) does not depend on this at all.

if and only if

$$1 - \rho^* - \delta(1 - \rho') > 0.$$

Suppose now, this was not the case, i.e.

$$1 - \rho^* \leq \delta(1 - \rho').$$

Then, since $\delta < 1$, it must hold that $\rho' < \rho^*$. If that was the case, then a deviation is never profitable for any state since waiting for the second period actually would yield a worse outcome even ignoring any time cost δ. Thus, the ρ' that has been chosen does not fulfil the requirement that at least someone wants to deviate.

However, for $\rho' > \rho^*$ the left hand side of condition (4.14), is increasing in π_2.

That, in turn implies, that if condition (4.14) would not hold for any $\psi = \{c, c^x\}$ it does not hold for any $\psi\{c, c_2 > c^x\}$ either.

Finally, consider the case where $\rho' = \frac{\rho^*}{\delta}$. Under this regime it holds that

$$\rho'\delta\pi_1(c_1) < \rho^*\pi_1(c_1) + (1 - \rho^* - \delta(1 - \rho')\delta\pi_2(c_2),$$

but at the same time if project 2 was not available, it is true that

$$\rho'\delta\pi_1(c_1) = \rho^*\pi_1(c_1) \tag{4.15}$$

and thus, a deviation is (weakly) profitable.

Last, observe that in equation (4.15) the value of π_1 does not matter at all, i.e. whenever it is profitable to deviate for some state $\psi(\hat{c}, \overline{\overline{c}})$ deviations for all other states $\psi(c_1, \overline{\overline{c}})$, given $c_1 \neq \overline{\overline{c}}$ are also profitable for the agent.

With this, D2 eliminates all types accept for those where $c_2 = \overline{\overline{c}}$, and since this contemplates to the complete range of $c_1 \in [\underline{c}, \overline{c}]$ it is in fact possible for the firm to have any sort of believe that is consistent with

D2, that is any believe that excludes project 2 from being existent. A belief of $\beta(m_1 = 1) = B(\psi)$, where $B(\psi)$ is a probability function with

$$B(\psi) = \begin{cases} 1 & \text{if } \psi \geq (\overline{c}, \overline{\overline{c}}) \\ 0 & \text{else,} \end{cases}$$

would be of such kind and justifies a rejection in the second period, whenever 0 was observed in the first period as long as $w_1(\overline{c}) < 0$.

To complete universal divinity for this type of deviation, notice, that the set of best responses does not change, as long as $w_1(\overline{c}) < 0$. Thus, no further elimination takes place and universal divinity is fulfilled for this type of deviation.

2. Second, consider deviations in which the firm after $l < T$ periods of proposing 0 suddenly proposes project 1.

   Since a deviation in case 1. fulfils the criteria of universal divinity, it mus also hold that any later period fulfils the criteria since the only thing that changes is the discount factor which is now δ^l. Since $0 < \delta^l < 1$ the same as in case 1. applies.[27]

3. Third, consider deviation in which the firm proposes 2 off the equilibrium path.

   This deviation is (due to the discount factor) never profitable for any state (given that the AA rejects 1 with probability 1 at all times except $t = 1$). Thus, the AA cannot infer any beliefs from the observed deviation and the equilibrium action is arbitrary and we can (for the sake of simplicity) assume that 2 is always accepted after l rounds of 0s.

4. Finally, consider deviations in which the firm proposes 0 in period 2 after having proposed 1 in period 1.

   Then again, since on the equilibrium path the firm accepts 2 always, there is no possible gain from deviation and the AA cannot infer anything from

[27] An exception might be the terminal period. If $w_1(c_1) < 0$, there exists a reasonable belief that justifies $\rho_T = 0$ even in the terminal period and the argument goes through.

it. Thus, we might as well assume whatever we want, e.g. that 2 is always accepted.

□

Proof of Lemma 25

Proof. The proof of existence is a straightforward adaptation of the results of the ρ^*-equilibrium. Observe that due to the two decision rules whenever

$$\delta^{\tau} = \frac{\rho}{1-\delta(1-\rho)}, \tag{4.16}$$

the function $\hat{c}_\tau$ is in fact identical to $\tilde{c}_{\rho^*}$. Thus, the decision rule of the firm is the same. Applying the ln on both sides of equation (4.16) yields the expression for $\underline{\tau}$. This constitutes an equilibrium only if the AA acts optimally, too. On path, this holds, since we know already from the ρ-equilibrium that $E[w_2|2] > \delta E[w_1|2]$ in the first period. Since the firm is going to propose 2 in the first period whenever $c_1 > \hat{c}_{\underline{\tau}}(c_2)$, all subsequent periods must be due to some trembling error the firm has made. It is always possible to find a totally mixed strategy sequence such that for $\epsilon \to 0$ the AA prefers to accept 2 whenever proposed.

If project 1 is proposed in any period prior to (and including) $\underline{\tau}$ this is an off path action and can be rejected by the AA if we assign beliefs that such early proposal of project 1 are often enough[28] made by types for which the second project is more beneficial to the firm. After period $\underline{\tau}$ the firm nonetheless proposes project 1 whenever $c_1 < \hat{c}_{\underline{\tau}}(c_2)$. As in the description of the ρ^*-equilibrium, this decision rule makes the AA indifferent at time $\underline{\tau}$ between accepting or rejecting the proposal. Thus, there is no incentive to deviate from the proposed strategy. Finally, for all subsequent periods anything gets accepted for example, if we assign beliefs to the AA that only "good enough" types make the mistake of proposing very late in the game.

To see that no equilibrium with less waiting time exists, fix the set $\tilde{\boldsymbol{\psi}}$ of possible states ψ in which the firm proposes project 1 in the first period of the ρ^*-equilibrium. This set is identical to the set $\hat{\boldsymbol{\psi}}_{\underline{\tau}}$ of states in which the firm chooses

[28] In a sense that it is very likely that such off path behaviour comes from this type of firms.

to wait. If the waiting need was reduced to $\tau < \underline{\tau}$, the set $\hat{\boldsymbol{\psi}}_{\tau} \supset \hat{\boldsymbol{\psi}}_{\underline{\tau}}$. Since $\delta < 0$ there exist some $0 < \rho' < 1$ such that

$$\delta^{\tau} = \frac{\rho'}{1 - \delta(1 - \rho')}$$

Since the left hand side decreases in τ and the right hand side increases in ρ, the corresponding $\rho' > \rho^*$. By the monotonicity arguments used in the proof of the ρ^*-equilibrium, this means that the AA is not indifferent any more between accepting the offer and delaying it another period. In fact, she prefers to wait. Thus, there exists an incentive to deviate on path and thus, there does not exist any shorter waiting equilibrium than $\underline{\tau}$.

The last part follows by simple algebra. $\underline{\tau} \leq 1$ is equivalent to

$$\delta \leq \frac{\rho^*}{1 - \delta(1 - \rho^*)}.$$

This statement is in turn equivalent to

$$\rho^* \geq \frac{\delta}{\delta + 1}$$

□

Proof of Theorem 3

Proof. Recall from Lemma 25 that under the decision rule that is described there, the AA is indifferent between accepting and rejecting the offer of the firm in $\underline{\tau}$. If the AA increases τ, monotonicity in $\hat{c}_{\tau}$ shrinks the set of states in which the firm chooses to wait. As in the proof of Lemma 22, "worse" states of project 1 are dropped earlier than "better" ones. Thus, the AA is not indifferent but in fact strictly prefers acceptance in period $\tau + 1$ under the firms decision rule $\hat{c}_{\tau}$ whenever $\tau > \underline{\tau}$. By the same reasoning, the firm accepts project 2 whenever it is proposed. By asuumption $E[w_2] > 0$. However, only if $\pi_2(c)$ is "bad enough" the firm chooses to wait. Thus, expectations are higher than $E[w_2]$ and positive whenever project 2 is proposed. Finally, for the same reason it never pays off for the AA to wait for a proposal of project 1 whenever the firm offers project 2. Thus, the AA has no incentive to deviate. The firm has neither, since she follows her optimal decision rule $\hat{c}_\tau$.

For universal divinity consider first the following equilibrium situation. The firm has some type $\psi = \{c_1, c_2\}$ for which it is optimal to wait for τ periods to propose project 1 in the $(\tau+1)^{th}$ period. Consider now a deviation in which the firm proposes project 1 at an earlier stage, that is e.g. τ. The firm would benefit from this type of deviation whenever the acceptance probability of project 1 in period τ is such that $\delta^{\tau-1}\rho_\tau(1)\pi_1(c_1) > \delta^\tau \pi_1(c_1)$. Since this condition is in fact independent of c_1, all such types are equally likely to deviate and non can be excluded by universal divinity. The same reasoning obviously holds for any $\tau' < \tau$

Second, consider a type ψ' in which the firm chooses to propose project 2 in period 1 in equilibrium. When, would such a type choose to deviate and to propose project 1 in τ? The condition for this is a bit more involved and requires that $\delta^{\tau-1}\rho_\tau(1)\pi_1(c_1) > \pi_2(c_2)$. This is now of course not independent of the type any more, but due to the fact that the firm chose to propose project 2 in equilibrium, it must hold that $\delta^\tau\pi_1(c_1) < \pi_2(c_2)$. Combining the two, requires thus that $\delta^{\tau-1}\rho_\tau(1)\pi_1(c_1) > \delta^\tau\pi_1(c_1)$, but then no one has a greater desire to deviate to τ than all types that propose 1 at $t = \tau + 1$ in equilibrium and we can exclude all types that propose 2 in equilibrium. Since there is no further distinction possible, all surviving types survive any further iteration of this reasoning and cannot be separated. Thus, any type that chooses to wait in equilibrium can be part of the belief of the AA. If we, as in Theorem 2, assign a degenerate belief that the firms type is $\underline{\psi} = \{\overline{c}, \overline{\overline{c}}\}$ whenever a deviation is observed, then the firm has even under universal divinity no incentive to react any different than staying on the equilibrium path. With this belief, the AA denies all earlier off-equilibrium proposals and the equilibrium survives universal divinity. □

Proof of Proposition 13

Proof. First, recall again the equilibrium condition as used in the proof of Lemma 22

$$\begin{aligned} \lambda_2 E[w_1] + (1-\lambda_2)h(\rho^*) &= 0 \\ \Leftrightarrow \lambda_2 \left(E[w_1] - h(\rho^*)\right) + h(\rho^*) &= 0 \end{aligned} \tag{4.17}$$

where

$$h(\rho^*) = E[w_1|1_2] - \delta E[w_2|1_2] = \frac{\int\limits_{\underline{c}}^{\bar{c}} \int\limits_{\underline{c}}^{\tilde{c}_{\rho^*}(c_2)} [w_1(c_1) - \delta w_2(c_2)] dF_1(c_1) dF_2(c_2)}{\int\limits_{\underline{c}}^{\bar{c}} F_1(\tilde{c}_{\rho^*}(c_2)) \mathrm{d}F_2(c_2)}$$

where the first term is the expected value whenever only project one is available which is independent of δ and ρ. The $h(\rho)$ function, however, depends on δ both directly and indirectly via ρ^*.

Since equation (4.17) exists and holds for all $0 < \delta < 1$ and $0 < \lambda_2 < 1$ and by assumption 3 (unconditional expectations are positive), it holds that $h(\rho^*) < 0$.

Using that equation (4.18) must hold as δ changes, we can derive the following

$$\frac{\partial \lambda_2 E[w_1] + (1-\lambda_2)h(\rho^*)}{\partial \delta} = \tag{4.18}$$
$$(1-\lambda_2)\left[\underbrace{\frac{\partial h(\rho^*)}{\partial \delta}}_{<0} + \underbrace{\frac{\partial h(\rho^*)}{\partial \tilde{c}_{\rho^*}}}_{<0}\left(\underbrace{\frac{\partial \tilde{c}_{\rho^*}}{\partial \delta}}_{\geq 0} + \underbrace{\frac{\partial \tilde{c}_{\rho^*}}{\partial \rho^*}}_{\geq 0}\frac{\partial \rho^*}{\partial \delta}\right)\right] = 0.$$

The sign of the direct derivative can be seen by inspection, that of the derivative with respect to ρ has been shown in the proof of Lemma 22 and the discussion of $\tilde{c}_\rho(c_2)$. Recall further that for all c_2 for which the equality that describes $\tilde{c}_{\rho^*}$ is not binding a marginal change in ρ^* has no effect, while for all c_2 such that

$$\tilde{c}_{\rho^*}(c_2) = \max\left\{c_1 : \frac{\rho^*}{\zeta}\pi_1(c_1) = \pi_2(c_2)\right\}$$

$\tilde{c}^*_\rho$ is increasing in ρ^*. Since the acceptance probability is independent of c_2, we may conclude that

$$\frac{\mathrm{d}\tilde{c}_{\rho^*}(c_2)}{d\delta} = \frac{\partial\tilde{c}_{\rho^*}}{\partial\delta} + \frac{\partial\tilde{c}_{\rho^*}}{\partial\rho^*}\frac{\partial\rho^*}{\partial\delta} \leq 0$$

is necessary for equation (4.18) to hold and requires

$$\frac{\partial\rho^*}{\partial\delta} < 0.$$

Second, observe that the ex-ante probability that project 1 gets proposed is

$$(1-\lambda_1)\left(\lambda_2 + (1-\lambda_2)\int_{\underline{c}}^{\bar{c}} F_1(\tilde{c}_2)\ \mathrm{d}F_2(c_2)\right).$$

Taking the derivative with respect to δ yields

$$(1-\lambda_1)(1-\lambda_2)\int_{\underline{c}}^{\bar{c}} f_1(\tilde{c}_{\rho^*})\frac{\mathrm{d}\tilde{c}_\rho(c_2)}{d\delta}\ \mathrm{d}F_2(c_2)\ \leq 0,$$

since $f_1 \geq 0$ by definition and $\frac{\mathrm{d}\tilde{c}_\rho(c_2)}{d\delta}$ is negative for reasons given above.

Finally, recall from equation (4.5) that

$$\phi E[w|0] = (1-\delta)(1-\lambda_2)(1-\lambda_1)\int_{\underline{c}}^{\bar{c}} w_2(c_2)F_1(\tilde{c}_{\rho^*}(c_2))\mathrm{d}F_2(c_2).$$

The RHS decreases in δ both directly and indirectly via $\tilde{c}_{\rho^*}$.

Thus the RCCM goes down. Since $E[w|0]$ is independent of δ, $E[w|\rho^*]$ increases in absolute terms, too. □

Proof of Lemma 26

Proof. Similar to the proof of Proposition 13, we take the derivative of equation (4.17) and take the derivative with respect to λ_2, that is

$$\begin{aligned}\frac{\partial\lambda_2 E[w_1]+(1-\lambda_2)h(\rho^*)}{\partial\lambda_2} &= \\ E[w_1]-h(\rho^*)+(1-\lambda_2)\left[\frac{\partial h(\rho^*)}{\partial\tilde{c}_{\rho^*}}\frac{\partial\tilde{c}_{\rho^*}}{\partial\rho^*}\frac{\partial\rho^*}{\partial\lambda_2}\right] &= 0\end{aligned} \tag{4.19}$$

Now recall from equation (4.17) that, since $h(\rho^*)<0$, it must hold that $E[w_1]-h(\rho)>0$ Thus, the second term in equation (4.19) must be negative. Since the first derivative is negative, the second is positive and $\lambda_2<1$ it must hold that

$$\frac{\partial\rho^*}{\partial\lambda_2}>0$$

The second part results from taking the derivative of the ex-ante probability of proposing 1 with respect to λ_2 which is:

$$(1-\lambda_1)\left[\left(1-\int_{\underline{c}}^{\bar{c}}F_1(\tilde{c}_2)\,\mathrm{d}F_2(c_2)\right)+\int_{\underline{c}}^{\bar{c}}f_1(\tilde{c}_{\rho^*})\left(\frac{\partial\tilde{c}_{\rho^*}}{\partial\rho^*}\frac{\partial\rho^*}{\partial\lambda_2}\right)\,\mathrm{d}F_2(c_2)\right]\geq 0. \tag{4.20}$$

Within the parenthesis the first term is (weakly) positive since the conditional probability cannot be greater than 1. The second part is positive in all derivatives which makes the whole equation positive. Thus, the probability of proposing project 1 increases in λ_2.

To see the effects of a change in λ_2 on the RCCM, recall first the last term of equation (4.5) and observe that we may write this part as

$$(1-\lambda_2)g(\rho^*):=(1-\delta)(1-\lambda_2)(1-\lambda_1)\int_{\underline{c}}^{\bar{c}}w_2(c_2)F_1(\tilde{c}_{\rho^*}(c_2))\mathrm{d}F_2(c_2),$$

where $g(\rho^*)$ only depends indirectly (via ρ^*) on λ_2.
using $\phi(\rho^*)$ as defined above this enables us to state the following

$$\phi(\rho^*)E[w|0]=(1-\lambda_2)g(\rho^*)$$

dividing by $1-\lambda_2$ this straightforwardly yields

$$g(\rho^*) = \phi(\rho^*)\frac{E[w|0]}{(1-\lambda_2)}$$

Observe now that $\frac{E[w|0]}{(1-\lambda_2)}$ is constant over both λ_2 and ρ_2 and ϕ is thus proportional to g w.r.t. ρ^* (and λ_2).

Knowing that an increase λ_2 leads to an increase in ρ^* and thus an increase in $\tilde{c}_{\rho^*}(\cdot)$ also $g(\rho^*)$, and therefore $\phi(\rho^*)$ increases in ρ^*.

$\square$

Proof of Lemma 27

Proof. First, recall from Lemma 25 that in the decision rule of the firm about which project to propose (in general) is the same in both the ρ^*-equilibrium and the waiting equilibrium at $\underline{\tau}$. Hence, the function $\tilde{c}_{\rho^*} = \hat{c}_{\underline{\tau}}$. Further, recall, that for any c_2 such that $\hat{c}_{\underline{\tau}}(c_2) \neq \{\overline{c}, \underline{c}\}$ the following equation holds

$$g\left(\delta, \underline{\tau}, \hat{c}_{\underline{\tau}}(c_2)\right) := \delta^{\underline{\tau}}\ \pi_1\left(\hat{c}_{\underline{\tau}}(c_2)\right) - \pi_2(c_2) = 0$$

Since expectations are positive, there is at least one state c_2 for which $g(\delta, \underline{\tau}, \hat{c}_{\underline{\tau}}(c_2)) = 0$.

Next, totally differentiate $g(\cdot)$ to get

$$\begin{aligned}
\mathrm{d}g(\cdot) &= 0 \\
\Leftrightarrow \quad 0 \quad &= (\underline{\tau}\ \delta^{\underline{\tau}-1}\ \pi_1\left(\hat{c}_{\underline{\tau}}(c_2)\right) \qquad \mathrm{d}\delta \\
&\quad + \delta^{\underline{\tau}}\ \ln(\delta)\ \pi_1\left(\hat{c}_{\underline{\tau}}(c_2)\right)\ \mathrm{d}\underline{\tau} \\
&\quad + \delta^{\underline{\tau}}\ \pi_1'\left(\hat{c}_{\underline{\tau}}(c_2)\right) \qquad \mathrm{d}\hat{c}_{\underline{\tau}}(c_2) \\
\Leftrightarrow \quad \frac{\mathrm{d}\underline{\tau}}{\mathrm{d}\delta} &= \frac{-1}{\ln(d)}\left[\frac{\underline{\tau}}{\delta} + \frac{\pi_1'\left(\hat{c}_{\underline{\tau}}\right)}{\pi_1\left(\hat{c}_{\underline{\tau}}\right)}\frac{\mathrm{d}\hat{c}_{\underline{\tau}}}{\mathrm{d}\delta}\right]
\end{aligned}$$

Since $\hat{c}_{\underline{\tau}} = \tilde{c}_{\rho^*}$ for all δ the following also holds

$$\frac{\mathrm{d}\hat{c}_{\underline{\tau}}}{\mathrm{d}\delta} = \frac{\mathrm{d}\tilde{c}_{\rho^*}}{\mathrm{d}\delta} \leq 0$$

as discussed in the proof of Proposition 13. Since $\pi_1' < 0$ by definition the term in square brackets is positive, and so is the one outside the brackets as $\delta < 1$. Thus,

$$\frac{\mathrm{d}\underline{\tau}}{\mathrm{d}\delta} > 0.$$

□

Proof of Proposition 14

Proof. The proof proceeds in three steps. First I show that the AA has a lower payoff in the waiting equilibrium at $\underline{\tau}$ than in the ρ^*-equilibrium. Second I show that if $T-1 \geq \overline{\tau}$, the waiting equilibrium at $T-1$ yields a higher payoff for the AA than the ρ^*-equilibrium and that the AA's waiting equilibrium payoffs are continuous in τ. Finally, I show that the derivative of the AA's payoff at the waiting equilibrium at $\overline{\tau}$ is in fact negative (and remains negative for all $T-1 \geq \overline{\tau}$) and thus, the "AA-most-preferred" equilibrium lies in the interior if T> $\overline{\tau}$.

First Step The claim in this part is, that $E_a[w|\rho^*] > E_a[w|\tau = \underline{\tau}]$, where $E[w|\tau]$ is the ex-ante payoff the AA expects in an waiting equilibrium at τ. The derivation of this claim is pretty straightforward. Recall that the decision rule in the two equilibria is nearly the same, only that in all cases in which the firm does not propose project 2 in both equilibria the firm chooses to wait in the waiting equilibrium while she directly proposes project 1 in the ρ^* equilibrium. In those cases the decision by the authority is again the same, once project 1 has been proposed. She is (interim) indifferent. Thus whenever the the AA sees a proposal of project 1 she knows at time t when message one is sent that

$$E_t[w_1|m_t = 1] = \delta E_t[w_2|m_t = 1].$$

However, in the waiting equilibrium we have $t = \underline{\tau}$ while in the ρ^*-equilibrium $t = 1$. Thus, ex-ante expected payoffs for the AA are

$$E_a[w|\rho^*] = E_a[w|0] - (1-\delta) \int_{\underline{c}}^{\overline{c}} \tilde{r}_{\rho^*}(c_2)\, \mathrm{d}F_2(c_2) \tag{4.21}$$

$$E_a[w|\tau = \underline{\tau}] = E_a[w|0] - (1-\delta^{\underline{\tau}+1}) \int_{\underline{c}}^{\overline{c}} \hat{r}_{\underline{\tau}}(c_2) \mathrm{d}F_2(c_2) \tag{4.22}$$

where

$$\tilde{r}_{\rho^*}(c_2) = (1-\lambda_2)(1-\lambda_1) \int_{\underline{c}}^{\overline{c}} w_2(c_2) F_1(\tilde{c}_{\rho^*}(c_2)) \mathrm{d}F_2(c_2)$$

and $\hat{r}_{\underline{\tau}}(c_2)$ is defined respectively. Since $\tilde{c}_{\rho^*} = \hat{c}_{\underline{\tau}}$, the two expressions are identical except for the factor in front of the last term. Since $\underline{\tau} > 0$ by construction, in particular by assumption 6, (4.21)> (4.22) and the waiting equilibrium at $\underline{\tau}$ can never be optimal.

Second Step To see the second part, observe that when $T-1 \geq \overline{\tau}$, it must be the case that under a waiting equilibrium at $\overline{\tau}$, the decision rule $\hat{c}_{\overline{\tau}}(c_2) = \underline{c}$ for all $c_2 \neq \overline{\overline{c}}$. In other words, whenever project 2 is available, the firm proposes it already in the first period. Thus, the AA gains profits of at least $E[w_2]$. If project 2 is not available, the firm still can make some profit by proposing project 1 after a waiting time of $T-1$ periods. Thus, the AA's ex-ante payoff is

$$E[w|\tau = \overline{\tau}] = (1-\lambda_2)E[w_2] + \delta^{\overline{\tau}}\lambda_2(1-\lambda_1)E[w_1]. \tag{4.23}$$

This way, it is clear, that $E[w|\rho^*] < (1-\lambda_2)E[w_2]$.

Continuity can be shown, by observing that $\hat{c}_{\tau}(c_2)$ is continuous in τ for all c_2 since δ^{τ} is. Then, each part in the additive form of the ex-ante expectations of the AAs welfare is a continuous in τ and by that the expectations itself are as well.

Third Step Finally, I show that the derivative of the expected welfare for the AA at $\overline{\tau}$ is negative and remains negative as τ increases.

The latter is easy to see by inspecting equation (4.23). The first term is independent of τ while the second decreases in τ. As a consequence, each waiting equilibrium at $t \geq \overline{\tau}$ yields lower welfare than the waiting equilibrium at $\overline{\tau}$. If projects are such that $\pi_1(\underline{c}) \geq \pi_2(\overline{c}$, then it is even possible to find a neighbourhood of size ϵ around τ such that the "AA-most-preferred" waiting equilibrium is in fact shorter than $\overline{\tau}$.

□

Proof of Proposition 15

Proof. To begin with, observe that, if the AA would only play pure strategies, delaying a decision is not optimal. I first prove that a rule at which only project two is accepted cannot be optimal. The proof can w.l.o.g. be applied to a situation where the candidate is the other way around, that is 1 gets always accepted and 2 never. In the second part I show that an acceptance vector of (1,1) cannot be optimal either.

First Part In this part, I show that it is never optimal for the AA to commit to a strategy in which project 2 is always accepted and project 1 is always rejected.

First, recall that in such an equilibrium the ex-ante expectation of the AA were $(1 - \lambda_2)E[w_2]$ since the firm always proposes 2 if it is available and something else in all other cases, but only two gets accepted.

Second, suppose now that the AA accepts project 1 with a probability $\rho_1^1 = \epsilon$, where ϵ is chosen such that

$$\epsilon\pi_1(\underline{c}) + (1 - \epsilon)\delta\pi_2(\overline{c}) < \pi_2(\overline{c}). \tag{4.24}$$

From the discussion in section 4.4 we know that such an ϵ exists (since $\pi_i(\overline{c}) > 0$) and the firm would under this probability rule still always propose

2 whenever it is available. However, if only 1 was available the firm would actually strictly prefer to propose 1.

The ex-ante expected value of such a probability rule for the AA would be

$$(1-\lambda_2)E[w_2] + \epsilon\lambda_2(1-\lambda_1)E[w_1] > (1-\lambda_2)E[w_2]. \tag{4.25}$$

Thus, it cannot be optimal to accept 2 always and never 1.

Second Part In this part I, show that a "accept all" policy is even worse than what we derived in the first part. This follows, essentially by definition and is repeated here only for the sake of completeness. Suppose the firm accepts all proposals in the first round. If the AA were to choose an acceptance probability of 1 for both projects, the firm would propose naively. Thus the ex-ante expected payoff for the AA was

$$E[w]_{(1,1)} := P_{\pi_1>\pi_2}E[w_1|\pi_1 > \pi_2] + P_{\pi_2>\pi_2}E[w_2|\pi_2 > \pi_1], \tag{4.26}$$

where $P_{\pi_i>\pi_j}$ denotes the probability that $\pi_i > \pi_j$.

Now recall the payoff the AA earns when only commiting to project two, that is

$$(1-\lambda_2)E[w_2], \tag{4.27}$$

By assumption 6 it holds that

$$E[w]_{(1,1)} < P_{\pi_1>\pi_2}\delta E[w_2|\pi_1 > \pi_2] + P_{\pi_2>\pi_2}E[w_2|\pi_2 > \pi_1].$$

But then (4.26)<(4.27) or in other words, "accept all" is ex-ante less profitable for the AA than "accept only 2".

Finally, accepting no project at all is not optimal since it has no influence on the behaviour of the firm, but leads only to waiting costs for both parties.

□

Bibliography

Admati, A. R. and M. Perry (1987). "Strategic delay in bargaining". *The Review of Economic Studies* 54, pp. 345–364.

Akerlof, G. (1970). "The market for" lemons": Quality uncertainty and the market mechanism". *The quarterly journal of economics*, pp. 488–500.

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

Balzer, B. and J. Schneider (2015). "Mediation and Informational Punishment". *mimeo.*

Banks, J. S. and J. Sobel (1987). "Equilibrium Selection in Signaling Games". *Econometrica* 55, pp. 647–61.

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

— (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.

Besanko, D. and D. F. Spulber (1993). "Contested mergers and equilibrium antitrust policy". *Journal of Law, Economics, and Organization* 9, pp. 1–29.

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

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

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

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

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

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

Chakraborty, A. and R. Harbaugh (2010). "Persuasion by Cheap Talk". *The American Economic Review* 100, pp. 2361–2382.

Che, Y.-K., W. Dessein, and N. Kartik (2013). "Pandering to Persuade". *American Economic Review* 103, pp. 47–79.

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

Cosnita, A. and J.-P. Tropeano (2009). "Negotiating remedies: revealing the merger efficiency gains". *International Journal of Industrial Organization* 27, pp. 188–196.

Cosnita-Langlais, A. and J.-P. Tropeano (2012). "Do remedies affect the efficiency defense? An optimal merger-control analysis". *International Journal of Industrial Organization* 30, pp. 58–66.

Cramton, P. C. (1992). "Strategic Delay in Bargaining with Two-Sided Uncertainty". *The Review of Economic Studies* 59, pp. 205–225.

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

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.

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

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

Fudenberg, D. and J. Tirole (1983). “Sequential Bargaining with Incomplete Information”. *The Review of Economic Studies* 50, pp. 221–247.

— (1988). “Perfect Bayesian and Sequential Equilibria - A clarifying Note”. *mimeo.*

— (1991). *Game Theory.* MIT Press.

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

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

Kamenica, E. and M. Gentzkow (2011). “Bayesian Persuasion”. *American Economic Review* 101, pp. 2590–2615.

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

Kovác, E. and D. Krähmer (2013). “Optimal Sequential Delegation”. *Mimeo.*

Lyons, B. and A. Medvedev (2007). “Bargaining over remedies in merger regulation”. *Mimeo.*

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. Oxford University Press.

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

Milgrom, P. and J. Roberts (1986). "Price and advertising signals of product quality". *The Journal of Political Economy*, pp. 796–821.

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

Nocke, V. and M. D. Whinston (2010). "Dynamic Merger Review". *Journal of Political Economy* 118, pp. 1201–1251.

— (2013). "Merger Policy with Merger Choice". *American Economic Review* 103, pp. 1006–33.

Ottaviani, M. and A. L. Wickelgren (2011). "Ex ante or ex post competition policy? A progress report". *International Journal of Industrial Organization* 29, pp. 356–359.

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

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

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

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

Quinzii, M. and J.-C. Rochet (1985). "Multidimensional signalling". *Journal of Mathematical Economics* 14, pp. 261–284.

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

Selten, R. (1975). "Reexamination of the perfectness concept for equilibrium points in extensive games". *International Journal of Game Theory* 4, pp. 25–55.

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

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

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

Sobel, J. and I. Takahashi (1983). "A Multistage Model of Bargaining". *The Review of Economic Studies* 50, pp. 411–426.

Sørgard, L. (2009). "Optimal Merger Policy: Enforcement vs. Deterrence". *The Journal of Industrial Economics* 57, pp. 438–456.

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

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

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

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

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

Vasconcelos, H. (2010). "Efficiency gains and structural remedies in merger control". *The Journal of Industrial Economics* 58, pp. 742–766.

Wilson, R. (1985). “Multi-dimensional signalling”. *Economics Letters* 19, pp. 17–21.

Curriculum Vitae

2011-2016	Ph.D. student at University of Mannheim
2008-2009	Visiting student at University of Bologna
2006-2011	Undergraduate studies in Economics at Albert-Ludwig-University Freiburg