Managing A Conflict*

Benjamin Balzer† Johannes Schneider‡

March 23, 2017

Abstract

We investigate the potential of conflict management to settle disputes that otherwise escalate to a costly Bayesian game. Players possess private information relevant in the escalation game only. The threat of escalation serves both as an endogenous outside option and as a screening device. We provide an economically intuitive characterisation of optimal conflict management for a general class of two-player Bayesian escalation games. Two features are essential. First, belief management post-escalation is sufficient to solve the conflict-management problem. Second, optimal conflict management maximizes the sum of two simple functions of the information structure in the escalation game, measuring welfare and discrimination. Our results link the classical mechanism design problem of eliciting information to the information design problem of distributing that information. We illustrate our findings by comparing two common escalation games: simple lotteries that call for sorting mechanisms, and contests that advocate type-independence.

1 Introduction

Conflict management aims at preventing a dispute from escalating to a costly fight via settlement at little or no costs. Typically, conflict management is voluntary and not always successful. If conflict management fails and the conflict escalates, players may use information obtained about their opponent during conflict management when making further decisions. Examples abound. They include alternative dispute resolution (ADR) escalating to litigation, mediated union-employer bargaining escalating to strikes, peace negotiations escalating to wars, or governed trade negotiations escalating to tariffs and retaliations.

*We thank Eddie Dekel, Martin Dumav, Daniel Garrett, Johannes Hoerner, Benny Moldovanu, Volker Nocke, Martin Peitz, Thomas Troeger, Christoph Wolf, Asher Wolinsky, and seminar participants in Bonn, Mannheim, Madrid, Munich, Warwick, at the ESEM Geneva, the Jornadas in Palma, and the UECE meeting in Lisbon for comments and suggestions. Johannes Schneider gratefully acknowledges support from the Ministerio Economia y Competitividad, Maria de Maeztu grant (MDM 2014-0431).

†UT Sydney, benjamin.balzer@uts.edu.au

‡Carlos III de Madrid, jschneid@eco.uc3m.es

The examples share a set of properties: (i) failure of conflict management triggers a non-cooperative escalation game; (ii) conflict management is designed to minimize the probability of escalation to prevent negative social externalities on the legal system, the economy, civilians, or global trade respectively; (iii) the rules of the escalation game are exogenously determined and beyond the designer's control, either through constitution (litigation and strikes) or sovereignty (war and trade retaliation); (iv) players possess private information about their strength in the escalation game that is irrelevant under settlement; (v) if conflict management reveals information about the opponent's strength, players adjust beliefs and continuation strategies accordingly.

We study the optimal conflict management mechanism in this general setting. We provide a novel approach that is based on players' beliefs after escalation. Our *belief-management* approach characterises the optimal mechanism entirely via those beliefs and provides a simple, economically intuitive, and yet tractable solution to conflict management for a large class of Bayesian escalation games. Notably, and different to most existing models of conflict management, our analysis includes settings in which player's best-response functions in the continuation game depend on the information revealed by the mechanism.

Understanding the connection between the design of the mechanism and players' post-escalation behaviour is crucial if the mechanism has limited control over players' future interactions. Consider the example of legal disputes from above. If parties obtain valuable information about their opponents during an unsuccessful ADR-attempt, they are likely to use this information when preparing their litigation strategy. Expecting such behaviour leads to privacy concerns during ADR. Existing models of conflict management often abstract from the interdependency between the mechanism and the continuation strategies to retain tractability.

We show that including that interdependency into the designer's consideration significantly impacts results. However, it also complicates characterisation and interpretation of the solution using standard tools of mechanism design. Belief management overcomes this technical hurdle by treating the revealed information as a choice variable. Applying the belief-management approach, we identify and quantify the fundamental economic trade-off of conflict management: *balancing discrimination and welfare conditional on escalation.*

The fundamental challenge of the designer is to keep weak types from mimicking strong types while incentivising strong types to participate. Strong types require favourable settlements to participate, as they do not fear escalation much. Weaker types, in turn, fear escalation, but may mimic strong types to benefit from favourable settlements. A player's expected payoff from participation is the sum of her expected settlement value, financed by efficiency gains of forgone escalation, and her expected escalation value, determined by her expected continuation utility. The more discriminatory the game post-escalation, the easier to deter mimicking behaviour. Simultaneously, the less inefficient the game, the smaller the settlement value required for participation.

To solve the designer's problem, the information structure post-escalation therefore ensures the following in the continuation game: (i) discrimination to *screen* types, and (ii) little inefficiency to ensure *participation*. We establish additively separable measures for (i) and (ii) and show that maximizing their sum solves a dual problem to optimal conflict management. The duality simplifies the problem significantly. Instead of solving a *mechanism design* problem with a non-linear externality on the escalation game, we solve an *information design* problem that directly implies the solution to the mechanism design problem.

We use our solution approach to compare two examples. The first example generalises the existing theory of conflict preemption in which information revelation has no influence on escalation-game strategies (Bester and Wärneryd, 2006; Hörner, Morelli, and Squintani, 2015; Spier, 1994). It applies, for example, to last-minute peace negotiations. Using the belief-management approach the problem collapses to a simple, linear algorithm. Our algorithm identifies and settles easy matches and escalates difficult ones.

We show that the optimal mechanism changes drastically if players' escalation-game strategies respond to the information revelation. We use an all-pay auction – a classic model of legal conflicts – as escalation game to demonstrate the difference. Using the belief-management approach it becomes immediate that the sorting algorithm from last-minute peace negotiations undermines the *screening motive* of the designer. Optimal conflict management therefore promises that the reporting strategy has no influence on the learning possibilities. This way, players have no incentive to deviate to extract information. Additionally, and again different to the first example, the solution generically leads to asymmetries in beliefs *between players* to decrease inefficiencies post-escalation.

The screening motive interacts with post-escalation beliefs since deviations remain undetected post-escalation. Suppose, for example, that a player expects to face a weaker opponent after a deviation. This deviator chooses a continuation strategy best responding to that weaker opponent, but – different than on the equilibrium path – *without* expecting her opponent to, in turn, best respond to that choice. More generally, misreporting alters the expected distribution of opponent's types, but not the continuation strategy of each such type. Thus, a deviation provides an informational advantage to the deviator in the escalation game.

In addition to discrimination, beliefs influence on-path behaviour and thus the expected inefficiency of the continuation game. Suppose that both players are strong in the all-pay auction. If they symmetrically expect a strong opponent, both exert a lot of effort. Effort is costly and thus aggregate surplus small. To the contrary, if one strong player expects a strong opponent while the (strong) opponent expects an (on average) weaker player, *both players* exert less effort than in the symmetric case. The opponent expects a weaker player and reduces effort to save costs. The strong initial player anticipates this behaviour and responds reducing her effort, despite expecting a strong opponent.

In light of the examples, the connection between mechanism design and information design is immediate. Players use the rules of conflict management to calculate interim posteriors. The designer influences those posteriors by choosing the rules. We exploit that connection and show that a mechanism is identified by the information structure it induces. This finding motivates belief management. In addition, we disentangle the interim posterior induced via the escalation decision, and that induced via an additional (public) signal and provide a sufficient condition for settings in which additional signals are superfluous.

1.1 Related Literature

Conflict management builds on the classical mechanism design approach to bilateral trade initiated by Myerson and Satterthwaite (1983). Our general setup is related to Compte and Jehiel (2009) within that literature. We assume a division of a pie as an outcome, a budget constraint mechanism, and an outside option that depends on the information structure.

Different from Compte and Jehiel (2009), our model includes interdependent valuations in the sense of Jehiel and Moldovanu (2001). Similar to them, our mechanism induces an information externality and cannot achieve efficiency. Our focus is, however, different. We aim at characterising the second-best mechanism, and model interdependencies as (equilibrium) outcomes of the play of a costly Bayesian game serving as the players' outside option. If continuation strategies depend on the information structure, the valuation of the outside option in our model is generically non-linear in beliefs, a case mostly ignored in the literature.[1] Moreover, we depart from standard solution approaches and use belief management that allows us to simplify, characterise and interpret the optimal solution using the properties of the escalation game directly.

Close in spirit are Philippon and Skreta (2012) and Tirole (2012) who study the informational externality of a bailout mechanism on future market behaviour. In line with our approach, they consider a model in which the design of the mechanism influences the interpretation of observed behaviour and thus subsequent choices in the market. A similar approach is taken by the literature on aftermarkets (Atakan and Ekmekci, 2014; Dworczak, 2017; Lauermann and Virág, 2012; Zhang, 2014) that considers how informational externalities of a mechanism influence the players' behaviour in the aftermarket. Importantly,

[1] In addition to the above mentioned, Fieseler, Kittsteiner, and Moldovanu (2003) and the literature on second-best conflict preemption (Bester and Wärneryd, 2006; Hörner, Morelli, and Squintani, 2015; Spier, 1994) consider similar models. In all of those, the outside option is a linear function of the beliefs.

aftermarket competition is modelled as a game between the winner of the initial mechanism and new players that did not participate in the initial mechanism.[2] Therefore, the design of the mechanism affects the belief system in the aftermarket only one-sided and beliefs are type-independent by design. In our model, the same players participate in the mechanism and meet in the escalation game if no settlement is reached. Thus, the mechanism's design influences the information of all players and type-dependent beliefs are possible.

Contrary to the literature on aftermarkets, the conflict pre-emption literature (Bester and Wärneryd, 2006; Hörner, Morelli, and Squintani, 2015; Spier, 1994) typically involves type-dependent beliefs. For tractability however, most models exclude any effect of the mechanism on continuation strategies and post-escalation choices become irrelevant. We nest these models but allow for general Bayesian continuation games in which action choices are influenced by the mechanism.[3] Our key contribution to this literature is to provide a general approach to conflict management for a wide range of escalation games. We identify the channel linking the properties of the (Bayesian) game to the optimal mechanism and describe the interaction between escalation game and optimal mechanism.

We complement Zheng (2017) who studies necessary and sufficient conditions for full settlement in contest escalation games. We characterise the optimal mechanism when those conditions are violated. Contrary to Meirowitz et al. (2017), who study the effect of last-minute conflict management on early investment, we include early stage conflict management that saves such investment. In our model, players have full commitment power. Hence, we describe a channel complementary to Hörner, Morelli, and Squintani (2015) who study information effects under limited commitment.

The belief-management approach allows us to frequently apply techniques from the literature on information design (Bergemann and Morris, 2016a,b; Mathevet, Perego, and Taneva, 2017; Taneva, 2016). We solve the dual problem *as if* it was an information design problem with an omniscient designer. The optimal information structure we obtain, however, maps into the solution of a *mechanism design problem* in which the designer has limited control over the environment. Thus, we provide a natural connection between the two approaches not discussed in the literature thus far.

Roadmap. We first describe the model in Section 2. In Section 3, we derive our main result and establish the duality between conflict management and belief management. In Section 4, we illustrate our contribution comparing two classes of escalation games. In Section 5, we discuss a set of extensions. We conclude in Section 6. Formal proofs are in the appendix.

[2]The resale literature following Zheng (2002) takes a different road modelling a repeated mechanism design problem among the *same* set of bidders.

[3]Celik and Peters (2011) show that full participation may be suboptimal for Bayesian default games. We avoid the issue in the main model, but provide an extension addressing their concerns.

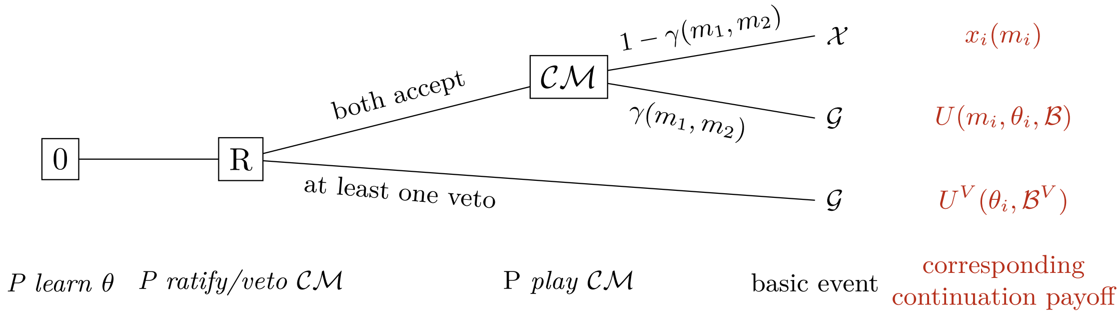

Figure 1: Timing of events and continuation payoffs. *The continuation payoff describes the expected value of the continuation game for player i conditional on escalation (after veto/$\mathcal{CM}$) or the expected settlement share. Private types influence this expected payoff only under escalation.*

2 Model

In this section, we outline the model. First, we describe the basic setup. Second, we describe the outcome of the continuation game as a function of the belief system that a particular information structure induces. We use this notation to construct the equivalence result in Section 3. Finally, we discuss the role of the crucial assumptions on the model.

Timing and Basic Events. Consider two ex-ante identical, risk-neutral players with linear preferences over basic outcomes. The players have a conflict over the distribution of a pie worth 1 to each player. Each player i is endowed with a private type θ_i independently drawn from $\Theta = \{1, 2, ..., K\}$. The known ex-ante probability of being θ_i is $\rho^0(\theta_i) > 0$. After privately learning their type in stage (0), players decide simultaneously whether they want to participate in the conflict management mechanism $\mathcal{CM}$ in stage (R). If they mutually agree on $\mathcal{CM}$ they select an action m_i. The mechanism results in either settlement or escalation. The probability of escalation is $\gamma(m_1, m_2)$ and depends on players' actions *within* the mechanism. Settlement leads to peaceful division of the pie, while escalation triggers a given escalation game with rules $\Gamma = (\Theta, A, \bar{u})$. After escalation is announced, players pick an action from the set, A, and the payoff function $\bar{u}$ maps the action and type profile, $A^2 \times \Theta^2$, into payoffs. If any player refuses to participate in stage (R), players learn about it and the conflict escalates immediately. The grand game summarized in Figure 1 and its rules are common knowledge. We assume that the designer has full commitment power, and players can commit to the mechanism at the ratification stage.

We characterize basic outcomes into two sets: the set $\mathcal{X}$ consists of all outcomes conditional on the event of settlement and the set $\mathcal{G}$ consists of all outcomes conditional on the event of escalation. A basic outcome in either of these sets is an element of the two-dimensional unit-simplex, representing the distribution of the (net) surplus.[4] An element

[4]Players are risk neutral, which is why we make no restrictions whether the distribution refers to an

in $\mathcal{X}$ defines the shares attributed to each player under settlement, (x_1, x_2), such that $x_1+x_2 \leq 1$. An element in $\mathcal{G}$ is the image of the payoff function, $\bar{u}$, mapping from type profiles in Θ^2 and action pairs in A^2 to outcomes. Players have type-independent preferences over settlement outcomes and thus their type is payoff relevant only when the outcome is in the set $\mathcal{G}$.

Conflict Management. Conflict management is a mechanism proposed by a non-strategic third-party, the designer, at the beginning of the game. An outcome of conflict management is either identified as settlement (i.e. in the set $\mathcal{X}$) or escalation (i.e. in the set $\mathcal{G}$). The revelation principle implies that it is without loss of generality to focus on direct revelation mechanisms. The set of (stochastic) conflict management mechanisms $\mathcal{CM}$ is defined as a mapping from the type space to the outcome space, that is,

$$\mathcal{CM}(\cdot) = (\gamma(\cdot), X(\cdot), \Sigma(\cdot)) : \Theta^2 \mapsto [0,1] \times \mathcal{X} \times \mathcal{S}. \tag{$\mathcal{CM}$}$$

The first component, the escalation rule $\gamma(\cdot)$, defines the probability with which the conflict escalates, the second component, the settlement rule $X = (x_1(\cdot), x_2(\cdot))$, defines the allocation conditional on settlement. In addition, the designer can commit to a public signal distribution $\Sigma(\cdot)$. A signal distribution is a random variable mapping the players' type reports to a stochastic, payoff-irrelevant realization $\sigma \in \mathcal{S}$. We are looking for conflict management mechanisms that minimize the ex-ante probability of escalation.

Solution Concept and Beliefs. Given a mechanism, $\mathcal{CM}$, the rules, Γ, and an equilibrium choice in the escalation game, we study the designer-preferred perfect Bayesian equilibrium of the grand game. Thus, whenever possible, players use all information available to update their beliefs according to Bayes' rule. Suppose escalation is announced and signal σ realises. Now, player i who reported m during conflict management updates her belief according to Bayes' rule given all her information. To calculate the probability that her opponent is of type θ_{-i}, she uses the fact that the elements of $\mathcal{CM}$ are common knowledge and that her opponent reports truthfully on the equilibrium path. Updating results in a conditional probability $\beta_i(\theta_{-i}|m, \sigma)$. The collection of conditional probabilities forms the player's *individual belief*, $\beta_i(\cdot|m, \sigma)$, a conditional probability mass function. By the revelation principle the on-path individual belief $\beta_i(\cdot|\theta_i, \sigma)$ of each player θ_i is common knowledge. We call the collection of all individual on-path beliefs for both players, $\mathcal{B}(\sigma) = \{\beta_i(\cdot|\theta_i, \sigma)\}^i_{\theta_i \in \Theta}$, the *realized belief system* given σ. Finally, given the signal distribution Σ, we define the lottery, $\mathcal{B}(\Sigma) := \{\mathcal{B}(\sigma), Pr(\sigma|\Sigma)\}_{\sigma \in \Sigma}$ as the collection of realized belief systems and their respective occurrence-probabilities. A special case of $\mathcal{B}(\Sigma)$ is the prior belief system $\mathcal{B}^0$ in which $\beta_i^0(\theta_{-i}|\theta_i, \sigma) = \rho^0(\theta_{-i})$ for any θ_i and Σ is degenerate. For notational simplicity we assign

actual division of the pie, or whether players engage in a (fixed) lottery about the pie as a whole.

an arbitrary individual belief to any player that does not occur with positive probability on the equilibrium path given realisation σ.

2.1 Structural Simplifications

The On-Path Continuation Game. The set of rules Γ of a game consist of the type space Θ, an action set A, and a function $\bar{u} : \Theta^2 \times A^2 \mapsto \mathbb{R}^2$ that maps from type and action profiles to payoffs. Given escalation and the realization σ, the (on-path) information structure is entirely determined by the belief system $\mathcal{B}(\sigma)$. The pair $(\Gamma, \mathcal{B}(\sigma))$ describes a Bayesian game and is, up to existence and choice of equilibrium, sufficient to determine the equilibrium outcome. In particular, take a set of rules Γ, a belief system $\mathcal{B}(\sigma)$, and an equilibrium of $(\Gamma, \mathcal{B}(\sigma))$. Then, there is a function, $\mathbf{s}^* : \mathcal{B}(\sigma) \mapsto \Delta(A)^{2K}$, fully describing the equilibrium (mixed-)strategies in that equilibrium. Under these equilibrium strategies, the Bernoulli equilibrium utility of player θ_i that is matched with type θ_{-i} is $u_i(\theta_i, \theta_{-i}, \mathcal{B}(\sigma)) := u_i(\theta_i, \theta_{-i}, \mathbf{s}^*(\mathcal{B}(\sigma)))$ and her von-Neumann-Morgenstern on-path utility is $U_i(\theta_i, \mathcal{B}(\sigma)) := \sum_{i=1}^{K} \beta_i(k|\theta_i, \sigma) u_i(\theta_i, k, \mathcal{B}(\sigma))$. We assume equilibrium existence and a given equilibrium choice rule for any realized game $(\Gamma, \mathcal{B}(\sigma))$.

The Off-Path Continuation Game. Similar to the on-path continuation game we can describe the off-path continuation game given Γ as a mapping from belief systems to outcomes. If a player deviates by misreporting her type, this deviation is undetected. Therefore, a deviation of player i during conflict management does not change the behaviour of the non-deviating player in case of escalation. Consequently, $-i$'s (continuation) strategy in the continuation game after deviation is a function of the on-path belief system $\mathcal{B}(\sigma)$. We define the continuation utility of a deviating type θ_i who reports being type m as the limiting utility of what the deviator could obtain when choosing her strategy optimally, that is,

$$U_i(m, \theta_i, \mathcal{B}(\sigma)) := \sup_{s_i} \sum_{i=1}^{K} \beta_i(k|m, \sigma) u_i(\theta_i, k, s_i, \mathbf{s}^*_{-i}(\mathcal{B}(\sigma))).$$

Observe that if $m = \theta_i$ the above equation describes the on-path utility and, given $\mathbf{s}^*_{-i}(\mathcal{B}(\sigma))$, s_i^* is determined by $\beta_i(\cdot|m{=}\theta_i, \sigma)$. Thus, we can simplify both on-path and off-path continuation utilities to the expression $U_i(m, \theta_i, \mathcal{B}(\sigma))$ which describes the utility player θ_i garners in the continuation game that starts after i reported to be type m and signal σ realized. As Γ is beyond the influence of the designer, we treat the function $U_i(m, \theta_i, \mathcal{B}(\sigma))$ as a primitive to the designer's problem. We define the expected continuation utility given lottery $\mathcal{B}(\Sigma)$ by $\widehat{U}_i(m, \theta_i, \mathcal{B}(\Sigma)) := \sum_{\sigma \in \Sigma} Pr(\sigma|m) U_i(m, \theta_i, \mathcal{B}(\sigma))$.

Assumptions on the escalation game. To proceed we impose structure on U_i. Since U_i is entirely determined by the choice of Γ, the assumptions we make are essentially assumptions on the rules of the escalation game. We assume upper hemi-continuity of U_i in $\mathcal{B}(\sigma)$, an *anonymous conflict*, and an adjusted version of the monotone likelihood ratio property (MLRP). Define the belief system with inverted identities as its reflection

$$Ref(\mathcal{B}(\sigma)) = \Big\{\tilde{\beta}_i(\cdot|\theta,\sigma) : \ \tilde{\beta}_i(\cdot|\theta,\sigma) = \beta_{-i}(\cdot|\theta,\sigma), \ \beta_{-i}(\cdot|m,\sigma) \in \mathcal{B}(\sigma), \theta \in \Theta, i = 1,2\Big\}.$$

Definition 1 (Anonymity)**.** The rules, Γ, and the equilibrium choice in the escalation game satisfy *anonymity* if $u_i(\theta,\theta',\mathcal{B}(\sigma)) = u_{-i}(\theta,\theta',Ref(\mathcal{B}(\sigma)))$ for any $\mathcal{B}(\sigma)$ and any type profile (θ,θ').

Definition 2 (Conflict)**.** The rules, Γ, and the equilibrium choice in the escalation game describe a *conflict* if for any $\mathcal{B}(\sigma)$ the following properties hold

i. *(non-productiveness).* $\sum_i u_i(\theta_i,\theta_{-i},\mathcal{B}(\sigma)) \leq 1$ for any $(\theta_i,\theta_{-i}) \in \Theta^2$.
ii. *(monotonicity).* $U_i(m_i,\theta_i,\mathcal{B}(\sigma))$ is non-increasing in θ_i.

Finally, we make an assumption on the interaction between belief and fundamental strength defining a version of MLRP for our game. Fix some $\mathcal{B}(\sigma)$ and define $\psi_i(k,\theta_i,\mathcal{B}(\sigma)) := U_i(k,\theta_i,\mathcal{B}(\sigma)) - U_i(k,\theta_i{+}1,\mathcal{B}(\sigma))$. Further let $\varphi_i(k,\theta_i,\mathcal{B}(\sigma)) := \frac{\psi(k,\theta_i,\mathcal{B}(\sigma))}{\psi(k+1,\theta_i,\mathcal{B}(\sigma))}$.

Definition 3 (MLRP)**.** The rules, Γ, and the equilibrium choice in the escalation game satisfy MLRP, if $\varphi_i(k,\theta_i,\mathcal{B}(\sigma))$ is non-decreasing in θ_i for any $\mathcal{B}(\sigma)$ and $k \in \Theta$.

Anonymity imposes a notion of symmetry on the game only to facilitates notation. Non-productiveness guarantees that escalation reduces the economy's aggregate resources. Monotonicity puts an order on types such that type 1 is *the strongest* and type K is the *weakest*. Finally, we impose a variant of MLRP *on the game*. MLRP is a sufficient condition to focus on local deviations during the analysis. Via MLRP we maintain connection to the mechanism design literature, although weaker sufficient conditions are possible.[5] We can interpret MLRP along the following lines: Fix the (realized) belief system and thus (on-path) strategies. Then, consider the decision problem of player i, holding belief $\beta_i(\cdot|m,\sigma) \in \mathcal{B}(\sigma)$. The function ψ_i describes the difference in payoffs *between two adjacent types*, θ_i and $\theta_i{+}1$, holding the *same belief* $\beta_i(\cdot|m,\sigma)$. Finally, we define the ratio of ψ_i for *of two adjacent beliefs*, $\beta_i(\cdot|k,\sigma)$, $\beta_i(\cdot|k{+}1,\sigma) \in \mathcal{B}(\sigma)$, by φ. We require this ratio to be monotone. Loosely speaking, MLRP states that becoming marginally stronger pays off most for strong players with weak beliefs. In Section 5 we discuss cases in which MLRP fails.

Discussion of the Assumptions. We discuss the robustness of our model in greater detail in Section 5. Here, we comment only the two most crucial assumptions. First, we

[5] We impose MLRP on *all* belief systems, although it suffices for the *optimal lottery.* In line with the guess-and-verify approach in the mechanism design literature, we suggest to *impose* MLRP first. If the solution satisfies MLRP, local incentive compatibility is sufficient.

assume the designer has full commitment power. This assumption is motivated by real-world observations. Institutions designing conflict management tend to act repetitively and thus rely on reputations that give them an (unmodelled) incentive to commit.[6]

Second, a key point to our model is that the player's strength under escalation is orthogonal to her preferences over outcomes. We use this assumption to motivate escalation after conflict management that is not replicable directly within the mechanism. Our results can be seen as a benchmark on the possibilities of third-party conflict management in two ways. First, it is one of the limiting cases of a more general model in which the private information is about preferences *and* strength, where the other limit is a mechanism with interdependent values and without transfers in which preferences and ability coincide. Second, our model benchmarks the designer's technology. We assume that the designer has no ability to test the players' private information other than escalation. Thus, settlement relies on soft information only. In reality, such mechanisms are the cheapest and fastest solution to the dispute. If known to the players when deciding on participation, however, we can subsume any more complicated mechanism in the escalation stage making our model a benchmark on the ability of soft information.

3 Analysis

In this section, we develop our two main results. First, the optimal conflict management mechanism is entirely determined by the choice of the optimal lottery $\mathcal{B}(\Sigma)$. Second, a dual to the escalation minimization problem exists. The dual consists of maximizing the combination of a measure of discrimination, and a measure of aggregate welfare. Both measures are defined on the level of the continuation game. Using the two measures and the first result we describe an economically intuitive characterisation of the optimal mechanism that links the rules Γ to the choice of the optimal belief structure $\mathcal{B}$.

We proceed in steps. First, we express the function X by the remaining choices γ and Σ using the binding constraints. We then transform the problem into reduced form and obtain an isomorphism between lotteries over belief systems and the mechanism's characteristics. Finally, we use the first-order approach to derive the dual, and state the equivalence result.

3.1 Preliminary Steps

Assumptions on the Value of Vetoing. Each party can trigger escalation unilaterally by vetoing $\mathcal{CM}$. Then, the conflict escalates immediately. Let $\mathcal{B}^V$ be the belief system

[6] Our assumption of full commitment power is in line with most of the mechanism design literature. A notable exemption is Bester and Strausz (2001) who consider a single agent limited commitment model. The online appendix of Hörner, Morelli, and Squintani (2015) provides evidence supporting our assumption.

after a veto by player i and $v_i(\theta_i)$ the value of vetoing, that is, the expected utility of θ_i from playing $(\Gamma, \mathcal{B}^V)$. In particular, $\mathcal{B}^V$ contains type-independent veto beliefs: the prior ρ^0 for player i, and an (in principle arbitrary) *veto belief* β^V for player $-i$. For simplicity, we assume symmetric β^V and drop the subscript in $v(\theta)$. For the main part, we focus on cases in which full participation is optimal and conflict management is non-trivial, that is no pooling mechanism exists that can guarantee full settlement irrespective of reports. Two technical assumption on $v(\theta)$ guarantee these properties. Assumption 1 implies full participation and Assumption 2 non-triviality. We relax Assumption 1 in Section 5. Given a particular escalation game, both assumptions define the set of possible off-path belief.[7]

Assumption 1 (Full participation)**.** $v(\theta)$ is convex with respect to ρ^0 given β^V for all θ.

Assumption 2 (Non-triviality)**.** $v(1) > 1/2$.

Relevant Constraints. Define the value from participation and the announcement m_i in a given mechanism as

$$\Pi_i(m_i, \theta_i) = \underbrace{\sum_{\theta_{-i}\in\Theta} \rho^0(\theta_{-i})(1-\gamma_i(m_i,\theta_{-i}))x_i(m_i,\theta_{-i})}_{=:z_i(m_i)(\text{settlement value})} + \underbrace{\gamma_i(m_i)\widehat{U}_i(m_i,\theta_i,\mathcal{B}(\Sigma))}_{=:y_i(m_i,\theta_i)(\text{escalation value})}, \tag{1}$$

that is, the (interim expected) utility of player θ_i, who participates in the mechanism, reports type m_i, and behaves optimally in the continuation game after escalation. We call the first part the settlement value, $z_i(m_i)$, and the second part the escalation value or $y(m_i, \theta_i)$. Each possible share $x_i(m_i, \theta_{-i})$ player i could receive is allocated to her with the probability that she is facing a type θ_{-i}, $\rho^0(\theta_{-i})$, and conflict escalates in that event, $\gamma(m_i, \theta_{-i})$. Preferences over outcomes are identical, and the settlement value depends on the report only. The escalation value, in turn, depends on both the reported and the actual strength in the escalation game. The expected escalation value for player m_i, $\gamma_i(m_i) = \sum_{\theta_{-i}} \rho^0(\theta_{-i})\gamma(m_i, \theta_{-i})$. Thus, $y_i(m_i, \theta_i)$ is homogeneous of degree 1 with respect to the escalation rule γ, since $\mathcal{B}$ and thus $\Sigma(\mathcal{B})$ and $\hat{U}$ are homogeneous of degree 0. By the revelation principle and Assumption 1 the set of participation constraints

$$\Pi_i(\theta_i, \theta_i) \geq v(\theta_i), \tag{PC}$$

and the set of incentive compatibility constraints

$$\Pi_i(\theta_i, \theta_i) \geq \Pi_i(m_i, \theta_i) \quad \forall m_i, \theta_i \in \Theta,\ i \in \{1, 2\}, \tag{IC}$$

[7] See Celik and Peters (2011), proposition 2, for sufficiency of Assumption 1 for full participation at the optimum, Zheng (2017) for a thorough discussion under which conditions a trivial solution exists in contests, and Cramton and Palfrey (1995) for a general discussion on off-path beliefs and veto mechanisms.

are satisfied at the optimum. Using (1) we interpret $y_i(m_i, \theta_i)$ as the screening parameter, and $z_i(m_i)$ as numeraire good.

To maintain simplicity we restrict attention to an environment in which the strongest type is sufficiently privileged from an ex-ante point of view. We relax the assumption in Section 5, showing that it does not affect the main results.

Assumption 3. $2\sum_{\theta\in Q}\rho^0(\theta)v(\theta) < \sum_{\theta\in Q}\rho^0(\theta)$, for any $Q \subseteq \Theta$ and $Q \neq 1$.

Using Assumption 3 we identify the set of binding constraints.

Lemma 1. *Suppose Assumption 1 to 3 hold. Then, the following is true at the optimum*

i. downward adjacent incentive compatibility constraints are satisfied with equality,

ii. all other downward incentive compatibility constraints are redundant,

iii. $z_i(\theta_i) > 0$ *for any* $\theta_i > 1$,

iv. only the participation constraint for the strongest type is binding.

Lemma 1 implies the existence of $2K$ linear equations defining optimal settlement values as a function of escalation values only.

Corollary 1. *In the optimal mechanism, the settlement values are*

$$z_i(\theta_i) = \begin{cases} z_i(\theta_i - 1) + y_i(\theta_i - 1, \theta_i) - y_i(\theta_i, \theta_i), & \text{if } \theta_i > 1 \\ v(1) - y_i(1,1), & \text{if } \theta_1 = 1 \end{cases}. \tag{Z}$$

Reduced Form Representation. Corollary 1 expresses settlement values as a function of escalation values and thus of γ. The tuple (z, γ) determines a reduced-form mechanism. Not all reduced-form mechanisms are implementable via X_i given γ. A reduced-form allocation is implementable if it satisfies the general implementation condition (GI) in Border (2007) adapted to our setting. Abusing notation, let the ex-ante probability of escalation be $Pr(\mathcal{G})$. Let $Q \subseteq \Theta$ be any subset of the type space. For given Q, define $Q_i := \{\theta_i | \exists \theta_{-i} : (\theta_i, \theta_{-i}) \in Q\}$, and $\bar{Q} := \{(\theta_1, \theta_2) \in \Theta^2 | \theta_i \notin Q_i \text{ for } i = 1, 2\}$ for each (Q_1, Q_2).

Lemma 2 (Sufficiency of Reduced-Form Mechanism)**.** *Fix an escalation rule* γ *and* $z_i(\cdot) \geq 0$. *An ex-post feasible* X_i *that implements* z_i *exists if and only if*

$$\sum_i \sum_{\theta_i \in Q_i} z_i(\theta)\rho^0(\theta_i) \leq 1 - Pr(\mathcal{G}) - \sum_{(\theta_1,\theta_2)\in\bar{Q}} (1 - \gamma(\theta_1, \theta_2))\rho^0(\theta_1)\rho^0(\theta_2), \quad \forall Q \subseteq \Theta. \tag{GI}$$

If $Q = \Theta$, then (GI) displays the aggregate resource constraint, which plays a prominent part in the course of the analysis,

$$\sum_i \sum_{\theta_i} \rho^0(\theta_i)\ z_i(\theta_i) \leq 1 - Pr(\mathcal{G}). \tag{RC}$$

3.2 Conflict Management and Belief Management

The option value of escalation depends on the belief system at the point at which players choose their actions in the continuation game. In this part, we characterise the set of possible belief systems and show that finding the optimal lottery over these belief systems is isomorphic to finding the optimal mechanism. For simplicity, we focus on degenerate signal distributions in the exposition, but state (and prove) the more general result as Theorem 1. We start with a straight-forward observation.

Observation 1. Any escalation rule γ determines some belief system $\mathcal{B}$.

The escalation rule, $\gamma(\theta_1, \theta_2)$, is defined as the probability of escalation conditional on the realization of type profile (θ_1, θ_2). The rules of the mechanism and the on-path behaviour are common knowledge. Each player, θ_i, uses γ to update her conditional probability $\beta_i(\theta_{-i}|\theta_i)$ of facing a player θ_{-i}. The reverse statement to Observation 1 is not true for two reasons: First, a belief system is determined by *relative* escalation probabilities only. If γ implements $\mathcal{B}$, so does $\alpha\gamma$ for any $\alpha \in (0, 1]$. Second, not every belief system is consistent with some γ. To be consistent with some γ, $\mathcal{B}$ has to be (i) internally consistent since $\gamma(\theta_1, \theta_2)$ influences both $\beta_1(\cdot|\theta_1)$ and $\beta_2(\cdot|\theta_2)$, and (ii) consistent with the prior since $\beta_i(\cdot|\theta_i)$ is a function of γ and the prior, ρ^0.

Using a network approach we characterise necessary and sufficient conditions on the set of belief systems consistent with some escalation rule. Intuitively, fix any $\gamma(1, 1) =$ *overlineg* > 0 and some belief system $\mathcal{B}$. Ignore the natural constraint $\gamma(\cdot, \cdot) \in [0, 1]$ for the moment. Our aim is to construct all function values $\gamma(\theta_1, \theta_2)$ from $\mathcal{B}$ and the anchor $\overline{g}$. Constructing $\gamma(\theta_1, \theta_2)$ can take several paths. We can construct $\gamma(\theta_1, \theta_2)$ via $\gamma(1, \theta_2)$ or via $\gamma(\theta_1, 1)$. If and only if both paths yield the same value for any $\gamma(\theta_1, \theta_2)$, then $\mathcal{B}$ is consistent with some γ. Constructing $\gamma(\theta_1, \theta_2)$ in this way may lead to some $\gamma(\theta_1, \theta_2) > 1$. However, since $\mathcal{B}$ determines γ only up to a constant, and we start with some *overlineg* > 0, we can always find an $\alpha \in (0, 1]$ such that all $\alpha\gamma(\theta_1, \theta_2) \leq 1$. More formally we can identify all belief systems consistent with some escalation rule by introducing a set of equations, (C). The construction outlined above is only valid if $\mathcal{B}$ is "sufficiently interior". To cover boundary cases we use continuity in the mapping from γ to $\mathcal{B}$.[8] We say that a belief system is *interior* if all individual beliefs have full support over the type space.

Lemma 3. *A belief system $\mathcal{B}$ is consistent with some γ, if and only if there exists a sequence of interior belief systems, $\mathcal{B}_n$, such that*

$$\beta_{2,n}(\theta_1|\theta_2)\beta_{1,n}(\theta_1|1)\beta_{1,n}(1|\theta_1)\beta_{2,n}(1|1) - \beta_{1,n}(\theta_2|\theta_1)\beta_{2,n}(\theta_2|1)\beta_{2,n}(1|\theta_2)\beta_{1,n}(1|1) = 0, \quad \text{(C)}$$

[8]The continuity follows directly from Bayes' rule.

for every type profile (θ_1, θ_2) *and* n*, where* $\lim_{n\to\infty} \mathcal{B}_n = \mathcal{B}$ *element-wise. For any consistent* $\mathcal{B}$ *with* $\beta_1(k|l) > 0$ *there exists a unique function* g *with the following properties:* $g(k,l) = 1$ *and any* γ *implementing* $\mathcal{B}$ *takes the form* αg *for some scalar* $\alpha \in (0,1]$*.*

Each consistent belief system $\mathcal{B}$ directly determines the following measures. An ex-ante distribution over types conditional only on entering the escalation game, ρ_1 and ρ_2, and the individual beliefs of all but the weakest type of player 1, $\{\beta_1(\cdot|\theta)\}_{\theta<K}$. One interpretation of $B = \{\rho_1, \rho_2\} \cup \{\beta_1(\cdot|\theta)\}_{\theta<K}$ is that ρ_i determines the post-escalation population and the individual beliefs determine the pairwise matching. We say the set B is consistent if there is a consistent $\mathcal{B}$ determining B. Equation (C) in Lemma 3 reduces the set of independent distributions to $K+1$. Hence, a consistent B also implies a consistent $\mathcal{B}$.

Lemma 4. *There is a continuous bijection* h *mapping from the set of consistent* B*'s into the set of consistent* $\mathcal{B}$ *such that* $h(B) = \mathcal{B}$*.*

Finally, we show that any choice of $\mathcal{B}$ (or B) determines a unique candidate for the optimal escalation rule. Given a consistent $\mathcal{B}$, Lemma 3 pins down the set of escalation rules that implement $\mathcal{B}$ as the product of a function g and a scalar α, i.e. $\gamma = \alpha g$. Using the conditions in Corollary 1 to reformulate the constraints from Section 3.1 in terms of g and α, we provide a lower bound on α, given g (and thus $\mathcal{B}$).

Lemma 5. *Take any implementable mechanism with escalation rule* γ *and* X *that satisfies the conditions stated in Corollary 1. Then there exists an* $\alpha^* \in (0,1]$ *such that an escalation rule* $\alpha\gamma \leq \gamma$ *is implementable if and only if* $\alpha^* \leq \alpha$*. Any such escalation rule* $\alpha\gamma$ *implements the same* $\mathcal{B}$*, and* α^* *is continuous in* $\mathcal{B}$*.*

The final step is to derive an upper bound on α that coincides with the lower bound α^*. For this result, we combine optimality and the resource constraint equation (RC) expressed as a function of α and g via Corollary 1. The escalation probability, $Pr(\mathcal{G})$ (linearly) decreases in α, while equation (RC) (linearly) tightens which pins down α^* given g.

Lemma 6. *The optimal escalation rule satisfies* (RC) *with equality.*

Combining Lemma 1 to 6 each consistent belief system $\mathcal{B}$ maps to a unique candidate for the optimal reduced-form mechanism. The result extends to non-degenerate lotteries and we obtain a function

$$CM : \mathbb{B} \mapsto (z, \alpha g), \tag{CM}$$

with $\mathbb{B}$ is the set of lotteries over consistent belief systems $\mathcal{B}(\sigma)$. The lottery $\mathcal{B}(\Sigma)$ is an *implementable candidate* only if $\alpha g(\theta_1, \theta_2) \leq 1$ for all (θ_1, θ_2), that is αg is an escalation rule.[9]

[9] CM uniquely defines the product αg, while the two elements separately may have multiple solutions due to homogeneity.

Theorem 1. *Let $\mathcal{B}(\Sigma)$ be the lottery over consistent belief systems in the post-escalation game of the optimal mechanism. Then, the reduced form of that mechanism, (z^*, γ^*), is equal to $CM(\mathcal{B}^*(\Sigma))$. That is $\mathcal{B}(\Sigma)$ is a sufficient statistic for (z^*, γ^*).*

To construct CM, express z in terms of αg via Corollary 1. For degenerate lotteries, we obtain αg from $\mathcal{B}$ using the construction above and verifying that $\alpha g(\theta_1, \theta_2) \leq 1$. The construction extends in a natural way to lotteries with multiple realizations. Following arguments from information design (Bergemann and Morris, 2016a), we construct $\mathcal{B}$ by combining the information structures of each signal realization σ in the support of Σ. This combination results in the information structure induced by the signal "escalation" alone. Finally, notice that monotonicity of (RC) in α provides a well-defined set $\mathbb{B}$.

Discussion of Theorem 1. Theorem 1 provides an isomorphism between conflict management and belief management. That is, any mechanism that optimally manages the beliefs also optimally manages the conflict. The result has several implications. First, it highlights the close link between the optimal belief system and the optimal mechanism. Understanding this link points towards the importance of understanding the continuation game for the optimal mechanism. It indicates sensitivity of the optimal mechanism to the escalation game. Second, the result shows that the main role of the mechanism is that of an informational gatekeeper. It is crucial for the success of conflict management that privacy of the players can be protected and at the same time information can be transmitted to the players. On a more abstract level Theorem 1 offers a direct connection to the literature on information design (Bergemann, Brooks, and Morris, 2016; Bergemann and Morris, 2016a,b). We show that optimal conflict management is essentially determined by choosing information structures which translate to (lotteries over) belief systems. Belief systems are a particular function of the information structure well studied for many games. We use it in the next section to characterise the solution to the mechanism design problem.

3.3 The Optimal Mechanism

In this part, we study the designer's problem. If $\mathcal{C}$ is the set of constraints, the primal problem of the designer is given by

Definition 4 (Minimization of Conflict Escalation)**.**

$$\min_{\mathcal{CM}} Pr(\mathcal{G}) \quad \text{s.t. } \mathcal{C}. \tag{P$_{min}$}$$

In what follows, we show that the characterisation of the optimal mechanism corresponds to maximizing the sum of two measures on the continuation game: a measure of the degree of discriminating types in the continuation game, the (expected) ability premium, and a measure of the inefficiency in the continuation game, the (expected) welfare. For simplicity, we focus again on degenerate signal distributions in large parts of the exposition, but state (and prove) the more general result as Theorem 2.

The designer faces the following trade-off: provide the strongest type a sufficient amount of expected value from $\mathcal{CM}$ while keeping weaker types from imitating the stronger. To solve this trade-off, two motives are relevant: a *screening motive* and a *welfare motive.* Making the continuation game discriminatory and unattractive *for weaker* types deters imitation (the screening motive). Making the continuation game attractive *for all* types saves on resources by increasing the escalation value (the welfare motive).

The classical mechanism design literature uses information rents to describe the designer's promises to weak types to deter imitation. In our setup, the discriminatory power of the Bayesian continuation game can be seen as an inverse to the information rent. Intuitively, the better the continuation game discriminates, the less information rent has to be paid. A measure of the relevant discriminatory power is the *ability premium.*

Definition 5 (Ability Premium)**.** The ability premium, $\psi_i(\theta_i, \theta_i, \mathcal{B})$, is the difference in expected utility after escalation between a type θ and the next strongest deviating type $\theta_i{+}1$. That is, $\psi_i(m_i = \theta_i, \theta_i, \mathcal{B}) = U_i(m{=}\theta_i, \theta_i, \mathcal{B}) - U_i(m{=}\theta_i, \theta_i{+}1, \mathcal{B})$.

Definition 6 (Weighted Ability Premium)**.** The *weighted ability premium* is $\Psi_i(\theta_i, \mathcal{B}) := w(\theta_i)\psi_i(\theta_i, \theta_i, \mathcal{B})$, with $w(\theta_i) = (1 - \sum_{k=1}^{\theta} \rho^0(k))(\rho^0(\theta))^{-1}$ the inverse hazard rate of type θ.

The ability premium, a measure already used in the definition of the MLRP, measures the distance between two adjacent types in the continuation game if the weaker type had pretended to be the stronger type during conflict management. Deviation has two effects. First, the deviator faces a different distribution of opponents after escalation than under truth-telling. Second, the deviator induces a situation of non-common knowledge as any adjustments she makes in the off-path continuation game remain undetected by the complying opponent. The ability premium of type θ_i captures the premium the continuation game pays to a complying player compared to the "closest deviator", and thus measures the discriminatory power of the underlying game. Monotonicity guarantees a positive ability premium.

The larger the ability premium, the less attractive mimicking behaviour, and the lower the pressure on (IC). Thus, all else equal, the designer desires a high ability premium. How important a particular player's ability premium is, depends on the prior distribution. The weighted ability premium takes this into account via the inverse hazard rate.

Optimizing over the (weighted) ability premium is, however, only one of two pillars of the optimal mechanism. Due to the welfare motive, the designer has an incentive to decrease inefficiencies in the continuation game. Intuitively, the lower such inefficiencies, the lower the compensation under settlement to reach any value from participation, Π. The relevant efficiency measure is expected welfare in the continuation game.

The set of unambiguously binding constraints according to Lemma 1 consists of the downward adjacent incentive compatibility constraints and the strongest types' participation constraints. We denote this set as $C_R \subset \mathcal{C}$ and the remaining set $C_F := \mathcal{C} \setminus C_R$. We define expectations of U_i and Ψ_i conditional on the event of escalation by $\mathbb{E}[\Psi_i|\mathcal{G}] := \sum_{\theta\in\Theta} \rho_i(\theta)\Psi_i(\theta,\mathcal{B})$, and $\mathbb{E}[U_i|\mathcal{G}] := \sum_{\theta\in\Theta} \rho_i(\theta)U_i(\theta,\theta,\mathcal{B})$.[10] Define the following maximization problem

$$\max_{\mathcal{B}} \sum_{i\in\{1,2\}} \mathbb{E}[\Psi_i|\mathcal{G}] + \mathbb{E}[U_i|\mathcal{G}] \quad s.t.\ C_F. \tag{P$^{\mathcal{B}}_{max}$}$$

Proposition 1 (Duality)**.** *Suppose Assumption 1 to 3 hold and fix the set of signal realizations to a singleton. A mechanism solves* (P$_{min}$) *if and only if its reduced form,* (z^*, γ^*)*, is equal to* $CM(\mathcal{B}^*)$*, and* $\mathcal{B}^*$ *solves* (P$^{\mathcal{B}}_{max}$)*.*

The equivalence in the choice set follows by Theorem 1, the transformation of the objective from comparing first order conditions. In appendix B we provide the full Lagrangian approach. Here, we summarise the argument under the restrictions of the proposition assuming $\gamma(1,1)>0$ at the optimum. The settlement value according to Lemma 1 is

$$z_i(\theta) = v(1) + \sum_{k=2}^{\theta} y_i(k-1,k) - \sum_{k=1}^{\theta} y_i(k,k).$$

Using Lemma 5, $\gamma_i(\theta_i) = \sum_{\theta_{-i}} g(\theta_i,\theta_{-i})\gamma(1,1)$, where g depends only on $\mathcal{B}$. Using the expected size of the settlement value, define $\mathcal{Q}(\mathcal{B})$ as the solution to $\sum_i \sum_\theta \rho^0(\theta)z_i(\theta) = 2v(1) + \gamma(1,1)\mathcal{Q}(\mathcal{B})$. Further, let $R(\mathcal{B}) := \sum_{\Theta^2} \rho^0(\theta_1)\rho^0(\theta_2)g(\theta_1,\theta_2)$ such that $Pr(\mathcal{G}) = \gamma(1,1)R(\mathcal{B})$. By Lemma 6, (RC) binds at the optimum and $2v(1) - \gamma(1,1)\mathcal{Q}(\mathcal{B}) = 1 - \gamma(1,1)R(\mathcal{B})$. Therefore, the optimum satisfies

$$Pr(\mathcal{G}) = \Big(\sum_i v(1) - 1\Big)\frac{R(\mathcal{B})}{\mathcal{Q}(\mathcal{B}) - R(\mathcal{B})} = \Big(\sum_i v(1) - 1\Big)\frac{1}{\frac{\mathcal{Q}(\mathcal{B})}{R(\mathcal{B})} - 1}.$$

Thus, any $\mathcal{B}$ that solves $\min \mathcal{P}r(\mathcal{G})$, solves $\sup \mathcal{Q}(\mathcal{B})/R(\mathcal{B})$. Finally, multiply the last term by $\gamma(1,1)/\gamma(1,1)$, substitute for $\gamma(1,1)\mathcal{Q}(\mathcal{B})$, $z_i(\theta_i)$ and $y_i(m_i,\theta_i)$, and notice that by Bayes' rule $\gamma(1,1)R(\mathcal{B}) = Pr(\mathcal{G})$ and $\gamma_i(\theta)/Pr(\mathcal{G}) = \rho_i(\theta)/\rho^0(\theta)$. Rearranging yields the objective of (P$^{\mathcal{B}}_{max}$).

Ignoring public signals Proposition 1 characterises the solution and links it directly to properties of Γ. It provides a tractable solution approach to finding the optimal mechanism. Given the equilibrium characterisation of Γ for any consistent $\mathcal{B}$, the only additional object

[10] We define $\psi_i(\cdot, K, \mathcal{B}) \equiv 0$.

is the off-path utility of type θ_i+1. The off-path utility is the utility of a type $\theta_i + 1$ with the individual belief of type θ_i facing an opponent who expects equilibrium play according to $\mathcal{B}$. These off-path utilities are the solution to a simple decision problem.

The characterisation highlights the designer's fundamental motives: the welfare motive and the screening motive. The formulation provides an additively separable notion of both these motives. Finally, it determines the relative weights on both the screening motive and the welfare motive as a function of the prior distribution only. While the (weighted) ability premium depends on the prior distribution, the expected utility is independent of it.

Incorporating signals does not always improve upon the result of Proposition 1. Using our integrative approach we obtain a simple condition when solving the problem in Proposition 1 is sufficient.

Corollary 2. *Consider the solution to* ($\mathrm{P}^{\mathcal{B}}_{max}$) *ignoring* C_F. *If this solution does not violate* C_F, *the optimal signal structure is degenerate.*

Via Lemma 1 we incorporate the most relevant constraints directly in the objective of ($\mathrm{P}^{\mathcal{B}}_{max}$). If the maximum of the integrated objective satisfies all remaining constraints, signals do not improve. In that case, the designer can directly implement the post-escalation belief system that implies the highest value of the objective. By definition of a maximum, this leaves no room to exploit convexities in the sense of Aumann and Maschler (1995) and additional signals are superfluous.

If the solution to the unconstraint problem in Proposition 1 does not satisfy C_F, signals may be optimal, as the (unconstraint) optimum is not directly attainable without violating the constraints. Since beliefs may be type-dependent, we cannot directly use a concavified objective as in the Bayesian persuasion literature. Instead we use a two-step procedure. We say the mean of lottery $\mathcal{B}(\Sigma)$ is $\overline{\mathcal{B}} = \sum_{\sigma\in\Sigma} Pr(\sigma)B(\sigma)$.

Definition 7 (Admissible Means)**.** A belief system $\overline{\mathcal{B}}$ is in the set of admissible means $\overline{\mathcal{B}}^a$ if there is a lottery $\mathcal{B}(\Sigma)$ over consistent belief systems $\mathcal{B}(\sigma)$ that (i) satisfies the constraints C_F, (ii) has mean $\overline{\mathcal{B}}$.[11]

The set of signal structures Σ that induce a lottery with mean $\overline{\mathcal{B}}$ satisfying C_F is $S^a(\overline{\mathcal{B}})$. Given objective $\mathcal{O}$ we define the value of $\overline{\mathcal{B}} \in \overline{\mathcal{B}}^a$ as

$$\mathcal{V}(\overline{\mathcal{B}}, \mathcal{O}) := \max_{\Sigma(\mathcal{B})\in S^a(\overline{\mathcal{B}})} \sum_{\sigma} Pr(\sigma)\mathcal{O}(\mathcal{B}(\sigma)). \tag{2}$$

[11] $\overline{\mathcal{B}}^a$ can, in general, contain a non-consistent belief $\overline{\mathcal{B}}$, iff no type-independent lottery has mean $\overline{\mathcal{B}}$.

Using (2) we derive the general formulation of ($\mathrm{P}^{\mathcal{B}}_{max}$)

$$\max_{\overline{\mathcal{B}}\in\overline{\mathcal{B}}^a} \mathcal{V}\left(\overline{\mathcal{B}}, \sum_{i\in\{1,2\}} \mathbb{E}[\Psi_i|\mathcal{G}] + \mathbb{E}[U_i|\mathcal{G}]\right). \tag{$\mathrm{P}^{B(\Sigma)}_{max}$}$$

Theorem 2 (Duality of problems)**.** *Suppose Assumption 1 to 3 hold. A mechanism solves equation* (P_{min}) *if and only if its reduced form,* (z^*, γ^*)*, is equal to* $CM(\mathcal{B}(\Sigma)^*)$*, and* $\mathcal{B}(\Sigma)^*$ *solves* ($\mathrm{P}^{B(\Sigma)}_{max}$)*.*

Remark. The lottery-means do not have an economic meaning per se. This approach, however, exploits the linearity of (2) to represent the maximization problem in a concise manner. An equivalent, yet notationally more involved, approach is to consider the choice of the lottery as a (pure) information design problem that inherits its initial information structure from the escalation rule γ. Each γ implies an information structure $\mathcal{B}^\gamma$ by Observation 1. Any signal structure Σ leads to a (public) Bayes' correlated equilibrium (Bergemann and Morris, 2016a) in the game $(\Gamma, \mathcal{B}^\gamma)$. That is, an omniscient designer takes the belief system $\mathcal{B}^\gamma$ as given and chooses a set of public signals such that the information structures induced by a signal realization combine to the information structure of $\mathcal{B}^\gamma$. Similar to the mean-based approach we can formulate the value of an initial information structure as $\mathcal{V}^{BCE}(\mathcal{B}^\gamma, \mathcal{O})$ and then maximize over consistent $\mathcal{B}^\gamma$. Then, $\mathcal{B}^\gamma$ and Σ determine $\Sigma(\mathcal{B})^*$ and thus $\mathcal{CM}^*$. If the lottery $\mathcal{B}(\Sigma)$ leads to type-independent belief systems only, we have that $\overline{\mathcal{B}} = \mathcal{B}^\gamma$. That is, whenever $Pr(\sigma|\theta_i) = Pr(\sigma)$ for any θ_i the mean coincides with the $\mathcal{B}^\gamma$, a fact exploited in the Bayesian persuasion literature (Kamenica and Gentzkow, 2011).

The economic interpretation of the objective is intuitive. Mimicking behaviour has two effects on the continuation game: The deviator (i) inherits the posterior distribution over the opponent's types from the type she mimics, and (ii) gains an informational advantage as she is the only one aware of entering an off-path game. In the continuation game, the deviator is not forced to also adopt the strategy of the mimicked type, but can freely choose her behaviour. Any such choice remains unresponded by the opponent who plays *as if* she is in the on-path game. The information advantage of the deviator reduces discrimination in the continuation game and makes it particularly attractive to mimic seldom types. The weight of the ability premium depends on the prior distribution in the familiar fashion via the inverse hazard rate. The welfare motive is independent of the prior. It considers only on-path behaviour and is thus an aggregate measure of (in)efficiency conditional on escalation.

Discussion of Theorem 2. Our results offer several insights. First, to save resources for settlement the designer wants to reduce inefficiencies in the escalation game despite being agnostic about the outcome once the conflict escalates.

Second, Theorem 1 and 2 together fully describe the fundamental economic problem. Theorem 1 shows that the informational externality of the mechanism drives the results and the optimal mechanism depends on the role of information in the continuation game. Theorem 2 uses this insight and characterises the role of information. It shows that the designer's choice set can be reduced to finding the right post-escalation information structure. This information structure depends only on expected performance in the continuation game. In particular, it identifies, separates, and quantifies the screening motive and the welfare motive of the designer. It provides an intuitive and tractable notion of both motives.

Finally, Theorem 2 describes a solution algorithm to the general problem. First, we focus on the potential to relax incentive problems by pure information design techniques in the continuation game. Having solved this problem independent of the escalation rule, we then turn to the core mechanism design problem to determine the optimal escalation decision. Corollary 2 provides a sufficient condition when the first step is redundant.

4 Examples

In this section, we apply our characterisation to two different underlying sets of rules, Γ. A *simple lottery* and an *all-pay-auction*. The examples highlight the flexibility of our approach and show how drastically results change once the outside option is non-linear in beliefs.

Simple lotteries have an outside option that is linear in beliefs. These are games in the spirit of Hörner, Morelli, and Squintani (2015) often used to model last-minute peace negotiations. The distribution channel is most important resulting in a *sorting mechanism*. Conflict management identifies "easy to settle" matches and guarantees settlement for these matches while other types are referred back to the conflict game.

In the all-pay auction – a game frequently used to model for example legal conflicts – the informational advantage of a deviator is most important. Sorting as in simple lotteries undermines incentive compatibility. Instead, the optimal mechanism induces *type independent* beliefs. That is, conflict management ensures that a player's belief is independent of her previous behaviour. Absent this feature, any mechanism provides an incentive to deviate and "steal" another type's individual belief to obtain an informational advantage in the off-path continuation game. Beliefs are type-independent at the optimum, but differ between the two players, leading to a *generically asymmetric mechanism*.

4.1 Last-Minute Peace Negotiations (Simple Lotteries)

In this part, we discuss a general version of Hörner, Morelli, and Squintani (2015). Two countries are on the verge of war. Armies are fully prepared for the military conflict, and war immediately starts if peace negotiations fail. Parties hold private information about their strength in war, but cannot adjust strategies to information obtained during peace negotiations. Given a match, (θ_1, θ_2), the outcome of war is thus independent of beliefs. We call such games simple lotteries as the expected utility takes the form of a lottery.

Definition 8 (Simple Lottery)**.** A set of rules Γ is a *simple lottery* if the Bernoulli utility of a match is independent of $\mathcal{B}$, that is $u_i(\theta_i, \theta_{-i}, \mathcal{B}) = u_i(\theta_i, \theta_{-i})$.

The expected utility given belief system $\mathcal{B}(\sigma)$ depends only and linearly on the individual belief given m, that is

$$U_i(m, \theta_i, \mathcal{B}(\sigma)) = \sum_{\theta_{-i}\in\Theta} \beta_i(\theta_{-i}|m, \sigma) u_i(\theta_i, \theta_{-i}). \tag{U^L}$$

Thus, we can abstract from public signals. Expected welfare is

$$E[U_i|\mathcal{G}] = \sum_{\theta_i\in\Theta} \rho(\theta_i) \sum_{\theta_{-i}\in\Theta} \beta_i(\theta_{-i}|\theta_i) u(\theta_i, \theta_{-i}), \tag{3}$$

and the expected ability premium is calculated similarly. We define $\rho(\theta, k) := \rho_1(\theta)\beta_1(k|\theta)$, the ex-ante probability of a particular match using Bayes' rule with $\sum_{(\theta,k)} \rho(\theta, k) = 1$. We rewrite the maximization problem from Theorem 2,

$$\max_{\rho(\cdot,\cdot)} \sum_{(\theta,k)\in\Theta\setminus(K,K)} \rho(\theta, k) \underbrace{(\omega(\theta)A_1(\theta, k) + \omega(\theta)A_2(k, \theta) + W(\theta, k) - W(K, K))}_{\widetilde{VV}(\theta,k)} \tag{4}$$

with match ability premium and match aggregate utility[12]

$$A_i(\theta_i, \theta_{-i}) = u_i(\theta_i, \theta_{-i}) - u_i(\theta_i{+}1, \theta_{-i}) \text{ and } W(\theta, k) = u_1(\theta_1, \theta_2) + u_2(\theta_1, \theta_2), \tag{5}$$

respectively. This model nests that of Hörner, Morelli, and Squintani (2015) as a special case with constants $W(\theta_1, \theta_2) = \overline{W}$ and $A_i(\theta_i, \theta_{-i}) = \overline{A}$ for any type and player. In their model, the welfare motive is shut down and escalation yields a constant payoff. More generally, linearity directly implies that an entirely symmetric solution exists.

We assume that $\widetilde{VV}(\theta_1, \theta_2)$ is weakly decreasing in both arguments.[13] We guess that only the strongest type's participation constraint binds, and all downward adjacent incentive constraints hold with equality. Thus, we maximize equation (4) subject to

$$\mathcal{R}\Big(\underbrace{\textstyle\sum_i E[\Psi_i|\mathit{\Gamma}] + E[U_i|\mathit{\Gamma}]}_{=(4)} - 1\Big) \geq 2v(1) - 1, \text{ with } \mathcal{R} = \frac{\rho^0(1)^2}{\rho(1,1)}, \tag{6}$$

[12]We treat $A_i(K, \theta_{-i})$ as some finite, positive real number to avoid case distinction.

[13]Sufficient conditions that assure this property are for example: $\omega(\theta_i)$ is weakly decreasing in θ_i and $u_i(\theta_i, \theta_{-i}) - u_i(\theta_i{+}1, \theta_{-i})$ is weakly decreasing in θ_{-i} and θ_i.

which guarantees that the selected $\mathcal{B}$ satisfies (RC). That is, $\mathcal{B}$ can be implemented by an escalation rule with $\gamma(1,1) \in (0,1]$ that satisfies (RC) with equality. Given $\rho(1,1)$ satisfies equation (6) the condition $\rho(\theta_1,\theta_2) \leq \rho^0(\theta_1)\rho^0(\theta_2)(1/\mathcal{R})$ is necessary and sufficient for (RC).

Definition 9 (Top-Down Algortihm)**.** The top-down algorithm is described as

1. Increase $\rho(1,1)$ until $\rho(1,1) = 1$ or equation (6) binds. In the latter case, go to 2.
2. Take the next highest virtual valuation pair (θ_1,θ_2) and increase $\rho(\theta_1,\theta_2)$ until either $\sum_{k\in\Theta}\rho(k_1,k_2) = 1$, or $\rho(\theta_1,\theta_2) \leq \frac{\rho^0(\theta_1)\rho^0(\theta_2)}{\mathcal{R}}$ binds. In the latter case, repeat step 2.

Proposition 2. *Assume the game is a simple lottery, $\widetilde{VV}$ is weakly decreasing, and Assumption 2 holds. The Top-Down Algorithm solves the reduced-form problem. All but the downward adjacent incentive constraints are redundant. Moreover, all but the strongest type's participation constraints are redundant.*

The proposition illustrates that the problem of the optimal mechanism reduces to a simple algorithm when applying Theorem 1 and 2. The algorithm partitions, or *sorts* the possible matches, such that some matches are guaranteed settlements while others escalate. In particular, sorting follows a top down structure, and weak matches settle while strong matches escalate. A consequence of this mechanism is that post-escalation beliefs are not type-independent, but each type has a different individual belief.[14]

The result crucially relies on the fact that countries' action choices in war are independent of their information set. Thus, a ceteris paribus change in $\beta_i(\cdot|\theta_i)$ affects at most the expected utility of player θ_i. Further, a potential deviator has no informational advantage, but at most distributional one. Thus, all the designer guarantees a strong country is a sufficiently strong *match* in case of escalation. Technically, the problem reduces to a simple linear program using our approach. It highlights that the result in Hörner, Morelli, and Squintani (2015) is robust as long as the escalation game is a simple lottery.[15]

4.2 Third-Party Settlement in Legal Conflicts (All-pay Auction)

In this part, we consider two players in a legal dispute. Before they engage in formal litigation, the court offers a third-party settlement-mechanism. In case settlement negotiations fail, players enter the formal litigation process. However, parties can adjust their strategy in court to the information obtained during the negotiations. In line with, among others, Baye, Kovenock, and Vries (2005), Posner (1973), and Spier (2007), we model litigation as

[14] For brevity of the argument we ignore ex-post implementability. Due to linearity, inclusion is straightforward and does not change the economic insight.

[15] Proposition 2 nests Lemma 1 of Hörner, Morelli, and Squintani (2015) as a special case. If implementable, put all mass on the highest $\widetilde{VV}$ (part (4) in their Lemma 1). If this violates (6) we shift mass to asymmetric matches to generate sufficient resources (part (3) of that Lemma).

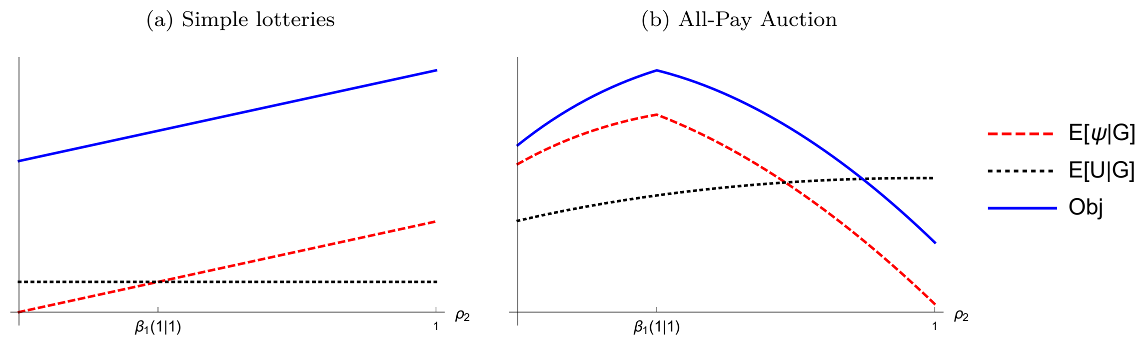

Figure 2: Comparing the optimum in both examples: *Expected ability premium (Dashed), expected welfare (Dotted) and value of the objective (Solid) in* ($\mathrm{P}^{B(\Sigma)}_{max}$) *in (a) the model of Hörner, Morelli, and Squintani (2015) and (b) the binary all-pay auction. All graphs plot functions of the probability that player 2 is the strong type conditional on escalation,* ρ_2. *In (a) the objective is maximized if* $\rho_2 = 1$. *The optimum in (b) is at* $\rho_2 = \beta_1(1|1)$, *the point with type-independent beliefs. The kink is due to the discontinuous change in the optimal continuation strategy post-deviation.*

an all-pay auction.[16] Both parties exert effort by collecting and presenting evidence, and whoever provides the most convincing evidence, wins the lawsuit. A victory has value 1 to both players. Players have private information about their cost of effort, $c_i \in \{1, \kappa\}$ with $\kappa > 2$. The probability that a player has type $c_i = 1$ is ρ^0. We assume $\rho^0 := \rho^0(1) = \delta\bar{\rho}$ with $\bar{\rho} = (\kappa-2)/(2\kappa-2)$ and $\delta \in [0.7, 1]$. Assumption 1 to 3 hold using this specification.[17]

Using the equilibrium characterisation in Rentschler and Turocy (2016) and Siegel (2014), it is immediate that –unlike the lottery case above – strategies in all-pay auctions are sensitive to beliefs. The algorithm in Siegel (2014) already suggests that not only the player's individual belief, but the entire belief system is relevant for the equilibrium strategies. Thus, the information advantage of a deviating player becomes relevant. We now state the result for optimal conflict management and discuss the intuition thereafter.

Proposition 3. *Suppose the escalation game is the all-pay auction above. Then optimal conflict management has the following characteristics*

- *all matches escalate with positive probability,*
- *the individual belief when entering the continuation game is independent of the player's behaviour during conflict management,*
- *in every continuation game, one player appears to be stronger than her opponent,*
- *signals improve on the optimal no-signal solution if and only if* $\rho^0 > 1/3$. *The optimal signal randomizes which player takes the role of "player 1" in the underlying game.*

[16]The all-pay auction is widely used to model conflict beyond the case of litigation. See Konrad (2009) for a general discussion also on related games. Our results apply also to conflict management in these cases.

[17]The upper bound guarantees that Assumption 2 is satisfied, the lower bound is sufficient for redundancy of the constraints in Lemma 2. Cases in which the (unconstraint) optimal $\mathcal{B}$ violates (RC) are discussed in Balzer and Schneider (2017).

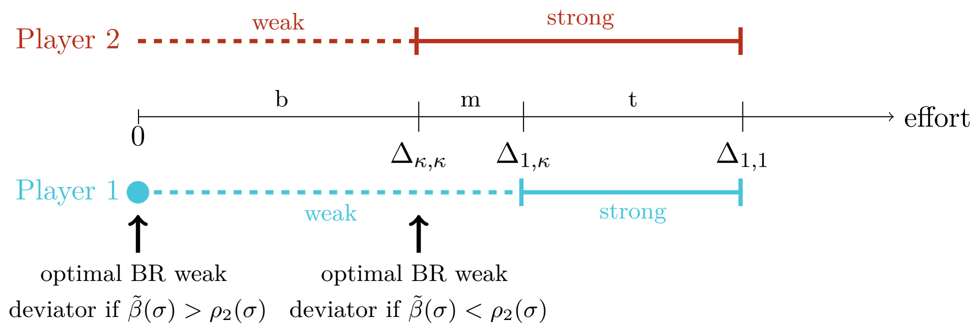

Figure 3: Equilibrium strategies in the all-pay auction assuming $\rho_1 < \rho_2$. Bold lines depict the support for the strong, dashed lines that of the weak. The dot indicates a mass point.

The difference between conflict management for the all-pay auction and that for a lottery is substantial in several dimensions.

The first two results, positive escalation probability for all matches and type-independent beliefs, are strongly connected. The reason lies in the ability premium. Fix some signals and type distributions $\rho_1(\sigma)$ and $\rho_2(\sigma)$ after escalation and signal realization σ. By Lemma 4, the only variable left to determine $\mathcal{B}(\sigma)$ is the individual belief of player $c_1 = 1$, that is $\tilde{\beta}(\sigma) := \beta_1(1|1, \sigma)$.

Changing $\tilde{\beta}(\sigma)$ has two effects on the ability premium. It reduces the strong player's expected utility in the continuation game, but also decreases the deviator's expected utility in the continuation game as she, too, faces a stronger opponent. As depicted in Figure 3 the deviator's best response changes at $\tilde{\beta}(\sigma) = \rho_2(\sigma)$. The reason is straight-forward. Whenever the deviator expects a stronger opponent than on the equilibrium path, that is $\tilde{\beta}(\sigma) > \rho_2(\sigma)$, her marginal gain from increasing her bid is reduced and vice versa. If beliefs are invariant to reports, that is $\tilde{\beta}(\sigma) = \rho_2(\sigma)$ the deviator is indifferent, too. The discrete change in deviation causes a kink in the expected continuation utility.

It turns out that this effect is of first order independent of the choice of $\rho_1(\sigma), \rho_2(\sigma)$. The ability premium increases for $\tilde{\beta}(\sigma) < \rho_2(\sigma)$ and decreases for $\tilde{\beta}(\sigma) > \rho_2(\sigma)$. Straightforward calculations show that type-independence prevails when considering the entire objective. Thus type-independence, $\tilde{\beta}(\sigma) = \rho_2(\sigma)$, is always desired.[18]

The third property comes directly from the welfare motive. Shared with many other competition formats, bidding behaviour in the all-pay auction is less aggressive under asymmetric type distributions. Bids are a loss and thus asymmetry decreases inefficiency and is beneficial. However, increased asymmetry comes at the cost of reducing the ability premium. At the optimum, the level of asymmetry is determined by that trade-off. We find

[18] A technical advantage of our approach is that the expected utilities in every realization of the escalation game are piecewise linear in $\beta_1(1|1, \sigma)$ for any $\rho_i(\sigma)$ providing us immediately with the optimality of a corner solution. This emphasizes the tractability gained by Theorem 1 and 2. See appendix C.7 for details.

that asymmetry increases with the prior ρ^0 which reduces the (relative) weight of the ability premium compared to aggregate welfare.[19]

Finally, since (GI) is always satisfied, signals can only improve if the strong type's upward (IC) holds with equality. Upward (IC)'s are redundant if $\rho^0 < 1/3$, and never bind for both strong players simultaneously at the candidate. Providing a symmetrizing signal, that is a coin-flip that then publicly announces who takes the role of player 1 is sufficient to achieve the same solution value as when ignoring the additional incentive constraint.

5 Discussion

In this section, we discuss robustness of our findings via several extensions to our model.

Arbitrary Sets of Constraints Binding. While Assumption 3 implies that downward incentive constraints are active at the optimum, MLRP implies that downward adjacent incentive constraints are sufficient for global downward incentive compatibility. If we relax Assumption 3downward incentive constraints may become redundant and MLRP has no bite. The following condition implies sufficiency of local incentive compatibility absent MLRP and Assumption 3.

Definition 10 (Linearity in Types)**.** The rules, Γ, and the equilibrium choice in the escalation game are linear in types if there is a pair of functions a, c such that $u_i(\theta_i, \theta_{-i}, s_i, \mathbf{s}^*_{-i}(\mathcal{B}(\sigma))) = a(\theta_{-i})\theta_i + c(\theta_{-i})$ for any $\mathcal{B}$ and $s^i = \arg\sup_{s_i} \sum_{i=1}^{K} \beta_i(k|m,\sigma) u_i(\theta_i, k, s_i, \mathbf{s}^*_{-i}(\mathcal{B}(\sigma)))$.

If the game is linear in types, the insights of Section 3 remain and the result in Theorem 2 changes only slightly. The ability premium for a particular type is defined upwards rather than downwards if only upward incentive constraints bind for that type. Then, the optimal mechanism aims at increasing discrimination between a type and the *next worst* type. All other arguments prevail.

Even if none of the conditions for sufficiency hold, the procedure in Section 3 is still a necessary first step to solve the problem. If the optimal solution obtained satisfies all global constraints, we have found an optimum. If the optimal solution violates any omitted constraints, we replace the objective with a Lagrangian objective incorporating the global constraints. The results from Theorem 1 and 2 remain under the adjusted objective.[20]

Non-Convex Veto-Values. If we give up Assumption 1, the value of vetoing may not be convex with respect to the prior given the off-path belief after a veto. In such a case

[19]In the binary example $E[\Psi_i|\mathcal{G}] = \rho_i \Psi_i(1) = (1-\rho^0)/\rho^0 \rho_i \psi_i(1)$.

[20]In Section 4.1 we use that guess-and-verify approach to global constraints to prove Proposition 2. A formal treatment of the general Lagrangian is in the supplementary material to this paper.

full participation may not be optimal.[21] Augmenting the problem slightly eliminates this issues. We make two changes to the model: First, instead of sequentially ratifying conflict management and then communicating their type, players do both simultaneously. Second, the designer can announce a public signal even after a veto and any vetoing player cannot ex-ante commit to ignore such information. Then, the designer can threaten to use a Bayesian persuasion mechanism a la Kamenica and Gentzkow (2011) on a vetoing player. These adjustments to the setup are sufficient to retain full participation at the optimum.[22]

Private Signals Sent by the Designer. We abstract from private signals sent by the designer mainly for tractability reasons. However, the link to the information design literature remains valid with private signals. We can use the same two-step procedure à la Bergemann and Morris (2016a) as in the discussion of Theorem 2. The only difference is that the set of signal structures includes private signals complicating the second stage.

Limited Commitment by the Players. Different to the second part in Hörner, Morelli, and Squintani (2015), we assume commitment by the players once they accept a certain mechanism. In some applications, this assumption appears to be too strong as pointed out by Hörner, Morelli, and Squintani (2015). They show that in their model, the designer can offer an equally efficient mechanism without commitment by not fully disclosing the realized match via the settlement offer. For general escalation games this approach is not valid. We can, however, apply an augmented version to our problem. Suppose instead of offering the settlement shares publicly, the designer can privately offer each party their share. Then, the designer can trigger escalation by making one party reject her settlement offer. In principle, the designer can use an announcement (e.g. $x_i = 0$) to trigger on-path rejection, as we allow her to dispose part of the surplus. Then, in case of an off-path rejection the designer is aware of the deviation and can commit to sending the worst possible signal. But then, the designer can, for any θ_i, find a worst $\mathcal{B}$ and induce it on-path with a small, but positive probability. If the utility given this worst $\mathcal{B}$ is smaller than any on-path settlement share x_i, the player accepts all shares on the equilibrium path. Thus, relaxing players' commitment imposes almost no threat on the optimal solution if offers are made privately.[23]

Different Objective. Throughout the paper, we assume that conflict management aims at minimizing escalation. While this assumption is motivated by reality in our context, the analysis extends to maximizing parties' expected welfare. While the main motives for the designer remain, the objective of Theorem 2 is no longer valid. In particular,

[21] An example when optimality of full participation fails in an otherwise different problem is discussed in Celik and Peters (2011).

[22] In Balzer and Schneider (2015) we discuss how the threat of Bayesian persuasion can discipline players in general mechanism design problems.

[23] In Balzer and Schneider (2017) we show that this condition holds for the binary all-pay auction and the result under limited commitment gets arbitrarily close to that under full commitment.

the welfare motive receives a higher weight. Notice that given settlement resources are scarce (condition (RC) is binding), any settlement solution is by definition an efficient outcome. In the main model, the designer values any failed settlement attempt equally and cares for efficiency post-escalation only for the purpose of increasing resources. If we change the objective to a utilitarian maximization over parties' ex-ante expected utility, the welfare component receives higher weight increasing pressure to decrease inefficiencies in the continuation game. Generically, this would not imply the same solution as the one we derive, but the designer may sacrifice some efficient settlements to provide a less inefficient escalation game. Reformulating the (maximization) objective accordingly, the updated first-order condition weighs the two components asymmetrically:[24]

$$\frac{\partial \sum_i E[\Psi_i|\mathcal{G}]}{\partial \mathcal{B}} h_\Psi(\mathcal{B}) + \frac{\partial \sum_i E[U_i|\mathcal{G}]}{\partial \mathcal{B}} h_U(\mathcal{B}) = 0, \quad \text{with } h_U \geq h_\Psi.$$

Transfers and Correlation. So far, we assumed that utility is not directly transferable. Although the settlement value, z_i, serves as a numeraire good in our analysis, we need to find a sharing rule X that implements a particular z_i. This is captured by the implementation constraints (GI). If we instead allow utilities to be directly transferable we can ignore these additional constraints and any reduced form mechanism is implementable.

Related to transferable utility is the case of correlated types. If types are correlated, the designer could use techniques similar to those in Crémer and McLean (1988) to exploit correlation and thus to achieve a higher settlement rate. Without transfers such exploitation is limited through the (GI) constraints. Correlated types, unlimited transfers, and no ex-post budget constraint allows a first-best solution à la Crémer and McLean (1988).

6 Conclusion

The main contribution of this paper is to provide a tractable approach to optimal conflict management for a large class of (Bayesian) escalation games. We propose an economically intuitive dual to the problem that directly links properties of the escalation game to the optimal mechanism. We show that optimal conflict management is completely characterised by the optimal belief system in the event of escalation. We formulate two measures on the post-escalation continuation game. A measure of discrimination and a measure of aggregate welfare. We show that the optimal mechanism maximizes simply the sum of these measures.

We use our general results to compare two classic models of conflict. The first considers

[24] In Hörner, Morelli, and Squintani (2015) the two objectives coincide since expected aggregate welfare after escalation is invariant to changes in the belief system. This is not the case in the all-pay auction.

last-minute peace negotiations where players cannot react to information received during conflict management. Optimal conflict management reduces to a sorting mechanism in such case. Conflicts with the highest virtual valuations escalate while others settle for sure.

If players, to the contrary, react to information from conflict management, such sorting may undermine incentive compatibility. We study the optimal mechanism in that case using a second classical model of conflict, the all-pay auction. There, the discrimination measure is non-monotone and deviators have a large informational advantage. To minimize this informational advantage, the optimal mechanism prohibits type-dependent learning. This leads to the surprising situation that, contrary to the first example, even a match of two weak players can escalate.

Our results suggest a number of directions for future research. First of all, our approach opens the field for research on how to compare different conflict games and their potential for (partially-)peaceful solutions. A second road is to extend the abilities of the designer to a set in which she can partially, but not fully control the rules of the continuation game. Such cases apply to existing conflict management mechanisms such as small claims courts or mini-trials for example. In our model, such methods could be implemented by allowing the designer to choose from a menu of screening devices each associated with some cost. Each of these devices would trigger a different information structure allowing the designer to further partition the matches. Our model is flexible enough to extend to such situations. Finally, our results lay the ground for extensions to a more dynamic setup. In early-stage conflict management players may expect exogenous news arrival during or after conflict management. Such a model suggests the use of a sequential approach to conflict management. A different trajectory is to consider repeated interaction between players with (partially) persistent types. That is, players can use some conflicts to learn something about their opponent. In such a setup, escalation and conflict management compete also in the degree of learning each provides to the players. We are confident that the belief-management approach offers a flexible framework for these more complex models.

Appendix

Organisation: Appendix A provides the missing steps towards Theorem 1. In Appendix B we prove Theorem 2. Appendix C gives the remaining formal arguments for claims in both main text and Appendices.

A Conflict Management and Belief Management

A.1 Conflict Management and Belief Management

In this part we provide the steps that lead to Lemma 3 to 6. We first focus on interior belief systems and then extend to boundary cases.

We first derive a representation that links the belief system to the escalation rule. We interpret every type profile as a node in a network. Node (θ_1, θ_2) has the value $\gamma(\theta_1, \theta_2)$. The (values of the) nodes are linked using a *transition* function $q_i(\theta_1, \theta_2)$.

Observation 2. Consider nodes (θ_i, θ'_{-i}), (θ_i, θ_{-i}) and the *transition* function $q_i(\theta'_{-i}, \theta_{-i}|\theta_i) = \frac{\rho^0(\theta_{-i})}{\rho^0(\theta'_{-i})} \frac{\beta_i(\theta'_{-i}|\theta_i)}{\beta_i(\theta_{-i}|\theta_i)}$. Then,

$$\gamma(\theta'_1, \theta_2) = q_2(\theta'_1, \theta_1|\theta_2)\gamma(\theta_1, \theta_2),$$

$$\gamma(\theta_1, \theta'_2) = q_1(\theta'_2, \theta_2|\theta_1)\gamma(\theta_1, \theta_2).$$

The result follows using Bayes' rule as

$$q_i(\theta'_{-i}, \theta_{-i}|\theta_i) = \frac{\rho^0(\theta_{-i})}{\rho^0(\theta'_{-i})} \frac{\gamma(\theta_i, \theta'_{-i})}{\gamma(\theta_i, \theta_{-i})} \frac{\rho^0(\theta'_{-i})}{\rho^0(\theta_{-i})}.$$

Fix two nodes in the network, say (θ_1, θ_2) and (k_1, k_2). There are several paths that connect the two nodes. Starting from (θ_1, θ_2) we can go to (k_1, θ_2) and then to (k_1, k_2). Equivalently we can approach (k_1, k_2) through (θ_1, k_2). Bayes' rule implies that both paths have the same length, or, the values of the nodes are the same. Using Observation 2 we have that

$$\gamma(k_1, k_2) = q_1(k_2, \theta_2|k_1)q_2(k_1, \theta_1|\theta_2)\gamma(\theta_1, \theta_2),$$

$$\gamma(k_1, k_2) = q_2(k_1, \theta_1|k_2)q_1(k_2, \theta_2|\theta_1)\gamma(\theta_1, \theta_2).$$

Definition 11 (Bayes' consistency)**.** A belief-system is Bayes' consistent if for every (θ_1, θ_2), (k_1, k_2)

$$q_1(k_2, \theta_2|k_1)q_2(k_1, \theta_1|\theta_2) = q_2(k_1, \theta_1|k_2)q_1(k_2, \theta_2|\theta_1). \tag{7}$$

Definition 12 ($(1,1)$-Consistent)**.** A belief-system is $(1,1)$-consistent if for every (θ_1, θ_2),

$$q_1(\theta_2, 1|\theta_1)q_2(\theta_1, 1|1) = q_2(\theta_1, 1|\theta_2)q_1(\theta_2, 1|1). \tag{8}$$

Lemma 7. *An interior belief system is Bayes' consistent if and only if it is* $(1,1)$*-consistent.*

Proof. Bayes' consistency trivially implies (1,1)-consistency. For the reverse see C.8. □

Using Lemma 7 consistency is reduced to (1,1)-consistency. Thus, any consistent belief system is implemented by some escalation rule only if every node (θ_1, θ_2) has a value weakly below 1. The value of a node is given by the length of the path connecting the node with the initial node, that is

$$\gamma(\theta_1, \theta_2) = q_1(\theta_2, 1|\theta_1)q_2(\theta_1, 1|1)\gamma(1, 1). \tag{9}$$

Finally, note that the above exposition together with Lemma 3 implies the following:

Observation 3. An interior belief-system is feasible if and only if it is consistent.

Lemma 8. *Let $\mathcal{O}$ be a continuous function defined on the domain of $\gamma \in [0,1] \cap C$ where C consists of those γ's that satisfy a given set of weak inequality constraints, each of which is continuous in γ. Then,* $\arg \max_{\gamma \in [0,1] | \gamma \in C} \mathcal{O}(\gamma) = \arg \sup_{\gamma \in (0,1) | \gamma \in C} \mathcal{O}(\gamma)$.

Proof. Without loss of generality suppose the argument that maximizes $\mathcal{O}$, γ^*, gives rise to a non-interior belief system, $\mathcal{B}^*$. Then, Lemma 3 implies that we can approximate $\mathcal{B}^*$ by a convergent sequence of consistent interior belief-systems. Because $\mathcal{O}$ is continuous in γ (and through Lemma 3 continuous in $\mathcal{B}$), it follows that $\lim_{n \to \infty} \mathcal{O}(\mathcal{B}_n) = \mathcal{O}(\mathcal{B}^*)$. Moreover, because the constraints are inequality constraints and continuous in γ (and $\mathcal{B}$), there is n' such that every element $\mathcal{B}_n$ with $n > n'$ satisfies the constraints. Therefore, $\max_{\gamma \in [0,1] | \gamma \in C} \mathcal{O}(\gamma) = \sup_{\gamma \in (0,1) | \gamma \in C} \mathcal{O}(\gamma)$ and $B^* = \lim_{n \to \infty} \mathcal{B}_n$. Using Lemma 3 we note that for every $\mathcal{B}_n$ there is γ_n so that $\lim_{n \to \infty} \gamma_n = \gamma^*$. □

A.2 Proof of Theorem 1

Proof. If the signal structure is a singleton, equivalence follows from the construction of (CM) via Lemma 3 to 6 and appendix A.1. Now suppose the signal structure has multiple realizations. What remains is to construct a relationship between the belief system $\mathcal{B}$ implemented by γ and the optimal lottery $\mathcal{B}(\Sigma)$ to obtain (CM).

Any realized belief system is entirely determined by the escalation rule, the signal distribution and the realized signal. Observation 1 shows that the escalation rule already implements a consistent belief system $\mathcal{B}$. Given the escalation rule, the signal distribution leads to a lottery $\mathcal{B}(\Sigma)$ over realized belief-systems $\mathcal{B}(\sigma)$. Any such realized belief-system is consistent for some escalation rule, say γ^σ, that satisfies $\gamma^\sigma(\theta_1, \theta_2) = Pr(\mathcal{G}, \sigma|\theta_1, \theta_2)$. The lottery $\mathcal{B}(\Sigma)$ is a sufficient statistics for both the "intermediate" $\mathcal{B}$ and all realizations $\mathcal{B}(\sigma)$. Thus, any $\mathcal{B}(\Sigma)$ determines $\mathcal{B}$, and $\hat{U}_i(m_i, \theta_i, \mathcal{B}(\Sigma))$ and thus (CM).[25] □

[25] An alternative way is to argue that the lottery $\mathcal{B}(\Sigma)$ implements a (public) Bayes correlated equilibrium (Bergemann and Morris, 2016a) on the "interim" information structure described by $\mathcal{B}$. Using the techniques of Bergemann and Morris (2016a) we can then get from $\mathcal{B}(\Sigma)$ to $\mathcal{B}$ by combining the information structures in all realizations $\mathcal{B}(\sigma)$.

B General Problem

In this part of the appendix we construct a solution algorithm to the general case. The general approach nests that discussed in Section 3 and proves Theorem 2.

Remark. Given $\mathcal{B}$, the value of α^* depends on the choice of g, but the resulting γ does not. To facilitiate notation we restrict ourselves to g with $g(1,1) = 1$ only in this appendix. Thus, we implicitly consider only $\mathcal{CM}^+ = \{\mathcal{CM} \text{ s.t } \gamma(1,1) > 0\}$ as possible mechanisms. Both the designer's objective and the constraints are continuous in γ. Therefore, continuity of $\mathcal{B}$ in γ implies that solving for the infimum over $\mathcal{CM}^+$ does not compromise generality. Recall Lemma 8, which provides a formal proof to this remark.[26]

B.1 The general Problem

We first use the arguments of Luenberger (1969) chapter 8 theorem 1 to show that the Lagrangian methodology can be applied to solve the general problem.

Our choice variables are a finite set of signals Σ (together with realization probabilities), the $\gamma(\cdot,\cdot)$ and z. Let the choice set be CS, with element cs.

Lemma 9. *The Lagrangian approach yields the global optimum.*

Proof. Let T be the space of Lagrangian multiplier, with element t. Define

$$\mathrm{w}(t) := inf\{Pr(\mathcal{G})|cs = (\gamma, z, \Sigma) \in CS, G(cs) \leq t\},$$

with $G(\cdot)$ the set of inequality constraints and $Pr(\mathcal{G})(cs)$ a function from the choice variable in the probability of escalation. By Luenberger (1969) chapter 8 the Lagrangian approach finds a global optimum if $\mathrm{w}(t)$ is convex. Assume for a contradiction that $\mathrm{w}(t_0)$ is not convex at t_0. Then, there is t_1 and t_2 with $\alpha t_1 + (1-\alpha)t_2 = t_0$ so that $\alpha\mathrm{w}(t_1) + (1-\alpha)\mathrm{w}(t_2) < \mathrm{w}(t_0)$. Let cs_j describe the optimal solution, such that $Pr(\mathcal{G})(cs_j) = \mathrm{w}(t_j)$. Then, consider the choice cs_0 such that $z^0(\cdot) = \alpha z^1(\cdot) + (1-\alpha)z^2(\cdot)$, $\gamma^0(\cdot,\cdot) = \alpha\gamma^1(\cdot,\cdot) + (1-\alpha)\gamma^2(\cdot,\cdot)$ and $\Sigma = \{1,2\}$, with $Pr(\sigma_1) = \alpha$ and $\gamma^{\sigma_j}(\theta_1,\theta_2) = \gamma_j(\theta_1,\theta_2)$ for all type profiles. By construction constraints are satisfied and and the solution value equals that of the convex combination

$$\mathrm{w}(t_0) = Pr(\mathcal{G})(cs_0) = \alpha \sum_{\sigma_j} Pr(\mathcal{G}, \sigma_j) = \alpha\mathrm{w}(t_1) + (1-\alpha)\mathrm{w}(t_1)$$

A contradiction. □

[26] Whenever the infimum corresponds to the minimum, results are equivalent. Whenever the two differ, the optimal mechanism is on the boundary with $\gamma(1,1) = 0$.

We state the Lagrangian objective in the context of Lemma 1. A complete statement of the general problem, a derivation of the Lagrangian objective and a complete definition of the Lagrangian multiplier is in the supplementary material, appendix E.

Corollary 3. *The lottery $\{Pr(\sigma), \boldsymbol{\rho}(\sigma)\}_\sigma$ is an optimal solution to the designers problem if and only if there are Lagrangian multipliers that satisfy complementary slackness given the lottery and the lottery includes every $\boldsymbol{\rho}(\sigma)$ that maximizes*

$$\begin{aligned}\widehat{\mathcal{L}}(\mathcal{B}(\sigma)) := \sum_i \Bigg[& \sum_{\theta=1}^{K} \rho_i(\theta)(\frac{\mathsf{m}^{\mathsf{i}}_\theta}{\rho^0(\theta)})U_i(\theta, \mathcal{B}(\sigma)) \\ & + \sum_{\theta=1}^{K-1} \frac{\mathsf{M}^{\mathsf{i}}(\theta) + \nu^i_{\theta+1,\theta} - \nu^i(\theta)}{\rho^0(\theta)} \rho_i(\theta)(U_i(\theta,\theta,\mathcal{B}(\sigma)) - U_i(\theta,\theta{+}1,\mathcal{B}(\sigma))) \\ & - \sum_{\theta=1}^{K}\sum_{k=1}^{\theta-1} \frac{\mathsf{M}^{\mathsf{i}}(\theta) + \nu^i_{k,\theta} - \nu^i(\theta)}{\rho^0(\theta)} \rho_i(\theta)[U_i(\theta,k,\mathcal{B}(\sigma)) - U_i(\theta,\theta,\mathcal{B}(\sigma))]\Bigg] \\ & + \mathcal{T}(\mathcal{B}(\sigma)),\end{aligned} \tag{10}$$

where $\mathsf{m}^{\mathsf{i}}_\theta := \rho^0(\theta) + \tilde{e}^i_\theta - \tilde{\zeta}^i_\theta$, $\mathsf{M}^{\mathsf{i}}(\theta) := \tilde{\Lambda}^i(\theta) - \sum_{k=1}^{k=\theta} \rho^0(k) - \tilde{E}^i(\theta) + Z^i(\theta)$, $\nu^i(\theta) := \sum_{k=1}^{\theta}\{[\sum_{v=\theta+1}^{K} \tilde{\nu}^i_{k,v}] - \tilde{\nu}^i_{k+1,k}\}$ *and*

$$\begin{aligned}\mathcal{T}(\mathcal{B}(\sigma)) := & \sum_{Q\in Q^2} \sum_{(\theta_1,\theta_2)\in\tilde{Q}} [\rho(\theta_1)\beta_1(\theta_2|\theta_1,\sigma)]\tilde{\eta}_Q \\ & - \sum_{\theta_1\times\theta_2} \frac{\rho_1(\theta_1,\sigma)\beta_1(\theta_2|\theta_1,\sigma)}{\rho^0(\theta_1)\rho^0(\theta_2)} \tilde{\mu}_{\theta_1,\theta_2}.\end{aligned} \tag{11}$$

Hence, $\boldsymbol{\rho} = \sum_\sigma Pr(\sigma)\boldsymbol{\rho}(\sigma)$ is a maximizer of the concave hull of the above function. Moreover, the following is true at the optimum:

- *The* (RC) *constraint is always binding, i.e., $\delta > 0$. $\tilde{e}^i_\theta = \tilde{E}_i(\theta) = 0 = \tilde{Z}^i(\theta) = \tilde{\zeta}^i_\theta$.*
- *If $\tilde{\Lambda}^i(\theta) + \tilde{Z}^i(\theta) - \sum_{v=1}^{v=\theta} \rho^0(v) - \tilde{E}^i(\theta) > 0$, then the* downward adjacent incentive constraints *are* binding. *If in addition the upward incentive constraints are redundant, then $\tilde{v}^i_{\theta,k} = 0$ for all $k \geq \theta$.*
- *If $\tilde{\Lambda}^i(\theta) + \tilde{Z}^i(\theta) - \sum_{v=1}^{v=\theta} \rho^0(v) - \tilde{E}^i(\theta) < 0$, the* upward incentive *constraints are* binding.
- *If local incentive constraints are sufficient, then $\nu^i(\theta) = \tilde{\nu}^i_{\theta+1,\theta} - \tilde{\nu}^i_{\theta,\theta+1}$. In this case, $\tilde{\nu}^i_{k,\theta} - \nu^i(\theta) = -M^i(\theta)$ for any k such that $k > \theta + 1$ or $k < \theta + 1$.*

Proof. The proof follows from applying the more general statement Lemma 12 from the supplementary material appendix E with $\mu^i_{\theta,\theta'} = 0$ for any $\theta' < \theta{-}1$. ☐

B.2 Proof of Proposition 1

Proof. Whenever the designer has access to signals, i.e., she can implement spreads over consistent post-escalation belief systems, then (i) the Lagrangian approach yields the global maximum and (ii) the optimal solution lies on the concave hull of the Lagrangian function, where the Lagrangian function is defined on the domain of consistent post-escalation belief systems. Assume the designer has no access to signals. Hence, the differences are that (i) a critical point of the Lagrangian objective is only necessary but not sufficient for global optimality, (ii) every optimal solution must be a local maximum of the Lagrangian objective (but not of its concave hull), and (iii) constraints have to hold for the ex-post realized belief system (rather than for the lottery over realized belief-systems). Honouring these differences, we still can use the form of the Lagrangian function stated in Corollary 3. Using complementary slackness, i.e., dropping redundant constraint identified in Lemma 1, the form of the maximization problem stated in Proposition 1 follows. □

B.3 Proof of Corollary 2

Proof. By Corollary 3 the solution to the considered optimization problem, say $\mathcal{B}^*$, maximizes $\hat{\mathcal{L}}$, that is (10). By hypothesis, $\mathcal{B}^*$ is in the set of least constraint solutions. Thus, the constraints in $\mathcal{C}_F$ are redundant. Thus, $\mathcal{B}^*$ maximizes

$$\sum_i \left[\sum_{\theta=1}^{K} \rho_i(\theta) U_i(\theta, \mathcal{B}(\sigma)) + \sum_{\theta=1}^{K-1} \frac{1 - \sum_{k=1}^{k=\theta} \rho^0(k)}{\rho^0(\theta)} \rho_i(\theta)(U_i(\theta, \theta, \mathcal{B}(\sigma)) - U_i(\theta, \theta{+}1, \mathcal{B}(\sigma))) \right],$$

which is (10) when setting the Lagrangian multipliers of those constraints in $\mathcal{C}_F$ to zero. Moreover, $\mathcal{B}^*$ satisfies complementary slackness by construction. Therefore, signals do not improve. □

B.4 Proof of Theorem 2

Proof. The proof of Proposition 1 establishes through Corollary 3 that maximizing ($\mathrm{P}_{max}^{B(\Sigma)}$) subject to $\mathcal{C}_F$ minimizes the probability of escalation. Now assume that instead of the continuation game Γ, an alternative continuation game $\widehat{\Gamma}$ is played. $\widehat{\Gamma}$ differs from Γ in that nature first draws a belief system $\mathcal{B}(\sigma)$ from the lottery $\Sigma(\mathcal{B})$ and communicates this to the players. Then, Γ is played under $\mathcal{B}(\sigma)$. The lottery $\Sigma(\mathcal{B})$ is sufficient to describe this augmentation. $\widehat{\Gamma}$ results in the continuation utilities $\widehat{U}(m, \theta, \Sigma(\mathcal{B}))$. If $\Sigma(\mathcal{B})$ satisfies the constraints, the lottery is implementable. Furthermore, $\Sigma(\mathcal{B})$ leads to some lottery over both the expected ability premium and the expected welfare which we denote by $\mathbb{E}[\widehat{\Psi}|\mathcal{G}] := \sum_\sigma Pr(\sigma)\mathbb{E}[\Psi|\mathcal{G}, \sigma]$ and $\mathbb{E}[\widehat{U}|\mathcal{G}] := \sum_\sigma Pr(\sigma)\mathbb{E}[U|\mathcal{G}, \sigma]$. Each $\overline{\mathcal{B}}$ may have many possible lotteries that support it and are feasible. We select the maximum among them. The $\overline{\mathcal{B}}$ with the highest value $\mathcal{V}$ also solves the problem of Corollary 3. $\square$

C Remaining Proofs

Lemma 10. *Let Θ_i^B be the set of types so that either the participation constraint is binding or $z_i(\theta) = 0$. Take any $\theta_i \in \Theta_i^B$ and let $\theta_i' := \min\{\theta \in \Theta_i^B | \theta > \theta_i\}$. Define $\Theta_i^I(\theta_i) := \{\tilde{\theta}_i \in \Theta_i | \theta_i' > \tilde{\theta}_i \leq \theta_i'\}$. Assumption 2 implies that the optimal mechanism features*

$$\sum_i \sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k) \left[\mathbb{1}_{PC}(\theta_i) \upsilon(\theta_i)\right] - 1 > 0. \tag{12}$$

Proof. From incentive compatibility we get $z_i(\theta_i - 1) - z_i(\theta_i) \leq y_i(\theta_i, \theta_i) - y_i(\theta_i - 1, \theta_i) =: \overline{\zeta}_i(\theta_i)$. Thus, for any $\tilde{\theta}_i \in \Theta_i^I(\theta_i)$

$$z_i(\tilde{\theta}_i) \leq \mathbb{1}_{PC}(\theta_i)[\upsilon_i(\theta_i) - y_i(\theta_i, \theta_i)] + \underbrace{\sum_{k=\theta_i+1}^{\tilde{\theta}_i} \overline{\zeta}_i(k)}_{=:\tilde{h}_{\theta_i}(\tilde{\theta})},$$

with $\mathbb{1}_{PC}(\theta_i)$ an indicator function with value 1 if $\theta_i \in \Theta^{PC}$. An upper bound on the expected sum of player i's incentive compatible shares is

$$\sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k) \left[\mathbb{1}_{PC}(\theta_i)[\upsilon(\theta_i) - y_i(\theta_i, \theta_i)] + \tilde{h}_i(k)\right].$$

Define

$$h(\gamma) := \sum_i \sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k) \left[\mathbb{1}_{PC}(\theta_i) y_i(\theta_i, \theta_i) - \tilde{h}_{\theta_i}(k)\right] - Pr(\mathcal{G}),$$

and observe that h_{θ_i} is homogeneous of degree 1 in γ, since y_i, $\tilde{h}_i$ and $Pr(\mathcal{G})$ are. In particular, $h(\alpha\gamma)$ converges to 0 if α is sufficiently small. Observe that $\alpha\gamma \to 0$ is the full settlement solution. Note that the constraint (RC) is satisfied if

$$\sum_i \sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k) \left[\mathbb{1}_{PC}(\theta_i) \upsilon(\theta_i)\right] - 1 \leq h(\gamma). \tag{13}$$

When LHS is negative escalation can be fully avoided, a contradiction to Assumption 2. $\square$

C.1 Proof of Lemma 1

Proof. Monotonicity and the MLRP imply that local downward incentive compatibility is sufficient for global downward incentive compatibility.[27] Second, assume there is a type for which both incentive constraints are redundant. Then, it is possible to reduce the players settlement share z_i at no cost for the designer until either an incentive constraint or the participation constraint starts to bind, or $z_i = 0$.

Lemma 10 implies that there must be one type and player with (PC) binding. Otherwise the LHS of (13) is negative contradicting Lemma 10. If there is exactly one type of one player with (PC) binding, the designer can offer an alternative mechanism: The mechanism determines at random who is assigned the role of player i and who that of $-i$ *after* players have submitted their report. Each of the two realizations satisfies the constraints and players are symmetric, and so does the combination. Under the alternative mechanism, no participation constraint is binding. A contradiction.

To arrive at the exact solution, consider again Lemma 10 and suppose for a contradiction that $\Theta_i^B \neq \{1\}$. There is an upper bound for this LHS of (13) given by

$$\sum_i \sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k) \left[\mathbb{1}_B(\theta_i)\upsilon(\theta_i)\right] - 1 \leq \sum_i \sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k)\upsilon(\theta_i) - 1, \tag{14}$$

Assumption 3 implies that the upper bound is negative if $\Theta_i^B \neq \{1\}$ and we have a contradiction to Assumption 2. If $\Theta_i^{PC} = \{1\}$, it is without loss of generality to focus on mechanisms in which the downward adjacent incentive constraints are satisfied with equality: Suppose $\theta_i > 1$'s downward incentive constraint were satisfied with strict inequality. Then, the designer could reduce $z_i(\theta_i)$ (and potentially burn the share) share without violation of any other constraint. □

C.2 Proof of Lemma 2

Proof. The proof follows diretcly from Border (2007) theorem 3. □

C.3 Proof of Lemma 3

Definition 13. Fix a feasible escalation policy, γ. A belief system, $\mathcal{B}$, is γ-feasible if it is consistent through Bayes' rule with γ. A belief system $\mathcal{B}$ is feasible, if there is a feasible γ, such that $\mathcal{B}$ is γ-feasible. Similarly, if $\mathcal{B}$ is γ-feasible, we say γ is $\mathcal{B}$-feasible.

Proof. Take a sequence of consistent $\mathcal{B}_n \to \mathcal{B}$. $\mathcal{B}_n$ is consistent, thus Observation 3 implies some function $f : \mathcal{B} \to [0,1]^{K\times K}$, such that $f(\mathcal{B}_n) = \gamma_n$ with γ_n $\mathcal{B}_n$-feasible. Since f is

[27] The proof is along the lines of more standard MLRPs in the literature. A detailed version is in the supplementary material.

continuous, $\lim_{n\to\infty} f(\mathcal{B}_n) = f(\lim_{n\to\infty} \mathcal{B}_n) = \gamma$. Equation (C) can be rewritten as

$$g_L(\mathcal{B}) = g_R(\mathcal{B}), \tag{15}$$

where both g_L and g_R are continuous functions from belief systems to the reals. Because $g_L(\mathcal{B}_n) - g_H(\mathcal{B}_n) = 0$, we can conclude that $g_L(\mathcal{B}) - g_R(\mathcal{B}) = \lim_{n\to\infty}[g_L(\mathcal{B}_n) - g_R(\mathcal{B}_n)] = 0$ and $\mathcal{B}$ satisfies equation (C).

Conversely, take any $\mathcal{B}$ being γ-feasible for some γ. We show that we can find a sequence of interior belief-systems that are consistent and converge to $\mathcal{B}$: Call γ^* the escalation rule so that $\mathcal{B}$ is γ^+-feasible. Choose a sequence of escalation rules in the interior that converges to γ^*. By Bayes' rule every element of the sequence, γ_n, corresponds to some belief system $\mathcal{B}_n$. Moreover, there is a continuous function, say $f^{-1} : [0,1]^{K\times K} \to [0,1]^{K\times K}$, such that $f^{-1}(\gamma_n) = \mathcal{B}_n . \mathcal{B}_n$ is γ_n-feasible and Observation 3 implies the system satisfies equation (C).

For the last part, take any $\mathcal{B}$ with $\beta_1(1|1) > 0$, fix $\gamma(1,1) = \alpha$, with $\alpha > 0$ and use equation (9) to construct all other $\gamma(\theta_1, \theta_2) \neq \gamma(1,1)$. Define the function g by $\gamma/\alpha =: g$. Thus, the set of all α s.t. $\alpha \leq 1/\max_{(\theta_1,\theta_2)} 1/g(\theta_1,\theta_2)$ determines all escalation rules which are $\mathcal{B}$-feasible. For any $\mathcal{B}$ such that $\beta_1(\theta_2|\theta_1) > 0$ perform the same steps fixing $\gamma(\theta_1,\theta_2) = 1$.

Finally we want to show that γ necessarily takes the form $\alpha g(\mathcal{B})$. First, we show that an interior and consistent belief system satisfies

$$\begin{aligned} \gamma_1(\theta_1)\beta_1(\theta_2|\theta_1) &= \rho^0(\theta_2)\gamma(\theta_1,\theta_2) \\ \gamma(\theta_1,\theta_2) &= \frac{\beta_1(\theta_2|\theta_1)}{\rho^0(\theta_2)(1-\beta_2(\theta_2|\theta_1))} \sum_{k\neq\theta_2} \gamma(\theta_1,k)\rho^0(k), \end{aligned} \tag{16}$$

for all θ_1, θ_2. Using equation (9) and substituting into equation (16) we get

$$\begin{aligned} q_1(\theta_2,1|\theta_1)q_2(\theta_1,1|1) &= \frac{\beta_1(\theta_2|\theta_1)}{\rho^0(\theta_2)(1-\beta_1(\theta_2|\theta_1))} \sum_{k\neq\theta_2} q_1(k,1|\theta_1)q_2(\theta_1,1|1)\rho^0(k), \\ \Leftrightarrow \qquad q_1(\theta_2,1|\theta_1) &= \frac{\beta_1(\theta_2|\theta_1)}{\rho^0(\theta_2)(1-\beta_1(\theta_2|\theta_1))} \sum_{k\neq\theta_2} q_1(k,1|\theta_1)\rho^0(k) \end{aligned} \tag{17}$$

Plugging in for $q_1(\cdot,\cdot|\theta_1)$ and rearranging yields

$$\frac{\rho^0(1)}{\rho^0(\theta_2)} \frac{\beta_1(\theta_2|\theta_1)}{\beta_1(1|\theta_1)} = \frac{\beta_1(\theta_2|\theta_1)}{\rho^0(\theta_2)(1-\beta_1(\theta_2|\theta_1))} \sum_{k\neq\theta_2} \frac{\beta_1(k|\theta_1)}{\beta_1(1|\theta_1)} \rho^0(1).$$

Using algebra we see that

$$1 = \frac{\sum_{k\neq\theta_2} \beta_1(k|\theta_1)}{1-\beta_1(\theta_2|\theta_1)}, \tag{18}$$

which holds since $\beta_1(\cdot|\theta_1)$ is a (conditional) probability. Therefore, an escalation rule is consistent with equation (16) if and only if it has the form $\alpha g(\mathcal{B})$. $\square$

C.4 Proof of Lemma 4

Proof. First observe that $\mathcal{B}$ determines the type-profile probability post-escalation, $\rho(\cdot,\cdot)$, by Bayes' rule. $\rho(\cdot,\cdot)$ together with $\mathcal{B}$ is sufficient for B.

Conversely, if we have given B we can, for every θ_2, compute

$$\rho_2(\theta_2) = \sum_{\theta_1 \in \Theta_1} \rho_1(\theta_1)\beta_1(\theta_2|\theta_1) = \sum_{\theta_1 \in \Theta_1} \rho(\theta_1, \theta_2). \tag{19}$$

Note that B and ρ determine (i) $\rho_1(K)$, $\rho_2(K)$ (through the requirement that $\sum_K \rho_i(k) = 1$), and (ii) $\beta_1(\cdot|K)$ through Equation (19). Hence, we have backed out the type-profile probability post-escalation which is sufficient for $\mathcal{B}$ through Bayes' rule. $\square$

C.5 Proof of Lemma 5 and Lemma 6

Proof. By Lemma 3 any escalation rule $\alpha\gamma$ implements the same $\mathcal{B}$ as γ. If γ is feasible it satisfies (RC). Substituting from Corollary 1, and rearranging, (RC) takes the form

$$\sum_i V_i(\upsilon(\Theta)) - 1 \leq \sum_i H_i(\gamma) - Pr(\mathcal{G}) \tag{RC'}$$

where $V_i(\upsilon(\Theta))$ is the probability weighted sum of those veto values for which (PC) binds, and $H_i(\gamma)$ is a weighted sum of escalation values y_i. The LHS is positive by Assumption 2.

$H_i(\gamma) \geq Pr(\mathcal{G})$ because γ is part of an implementable mechanism. Moreover, $H_i(\gamma)$ is homogenous of degree 1 w.r.t. γ, since y_i is. For any $\alpha\gamma$, increasing α leaves the LHS of (RC') constant, while the RHS increases. By linearity of the RHS there exists an $\alpha^* \in (0, 1]$ such that (RC') holds with equality. Lemma 1 ensures that no $\alpha \leq \alpha^*$ is implementable, but any $\alpha \in [\alpha^*, 1]$ is. This proves Lemma 5. Lemma 6 is true because $Pr(\mathcal{G})$ is monotone increasing in α and hence (RC) holds with equality at the optimum. $\square$

C.6 Proof of Proposition 2

Proof. By construction the problem is point-wise maximized, subject to the constraint in equation (6). We show monotonicity, i.e. $\rho^0(\theta_i)\rho(\theta_i{-}1, \theta_{-i}) \geq \rho^0(\theta_i{-}1)\rho(\theta_i, \theta_{-i})$, for all θ_i, θ_{-i}. We then use this feature to verify that ignored constraints are redundant.

Monotonicity trivially holds if $\gamma(1, 1) \neq 1$, as it implies an optimum at $\rho(1, 1) = 1$. Now, assume $\gamma(1,1)=1$. By Bayes' rule

$$\gamma(\theta_1,\theta_2)=\frac{\rho(\theta_1,\theta_2)}{\rho(\theta_1)\rho(\theta_2)}Pr(\mathcal{G}).$$

When $\gamma(\theta_1,\theta_2)>0$, then $\gamma(\theta_1-1,\theta_2)=1$ and

$$\begin{aligned}Pr(\mathcal{G})\rho(\theta_1-1,\theta_1)&=\rho^0(\theta_1-1)\rho^0(\theta_2),\\ Pr(\mathcal{G})\rho(\theta_1,\theta_1)&\leq\rho^0(\theta_1)\rho^0(\theta_2).\end{aligned}$$

Thus, monotonicity holds since

$$\rho^0(\theta_1)\rho(\theta_1-1,\theta_1)\geq\rho^0(\theta_1-1)\rho(\theta_1,\theta_1).$$

Monotonicty implies that all but the local downward adjacent incentive constraints are redundant as

$$\sum_{\theta_{-i}=1}^{K}\Big(\rho^0(\theta_i')\rho(\theta_i,\theta_{-i})-\rho^0(\theta_i)\rho(\theta_i',\theta_{-i})\Big)\left[u_i(\theta_i,\theta_{-i})-u_i(\theta_i',\theta_{-i})\right]\geq 0, \tag{20}$$

for all θ_i and $\theta_i'>\theta_i$. Moreover, $z_i(\theta_i)-z_i(\theta_i-1)=y_i(\theta_i,\theta_i-1)-y_i(\theta_i,\theta_i)\geq 0$ which implies that $z_i(\theta_i)\geq 0$. Finally, we claim that only the lowest types' participation constraint binds at the optimum. It is straightforward to verify that this implies that downward adjacent local incentive constraints hold with equality. We verify the claim by induction. We first show that $\Pi_i(2,2)\geq\upsilon(2)$. By local incentive compatibility, and $\Pi_i(1,1)\geq\upsilon(1)$,[28] we know that $\Pi_i(2,2)=\upsilon(1)-y_i(1,1)+y_i(1,2)$. We want to show that $\upsilon(1)-\upsilon(2)\geq y_i(1,1)-y_i(1,2)$. The game is a simple lottery, so the condition the reduces to

$$\sum_{\theta_{-i}}[\rho^0(\theta_{-i})-\gamma_i(1)\beta_i(\theta_{-i}|1)][u_i(1,\theta_{-i})-u_i(2,\theta_{-i})]\geq 0.$$

Moreover,

$$[\rho^0(\theta_{-i})-\gamma_i(1)\beta_i(\theta_{-i}|1)]=\frac{1}{\rho^0(1)}[Pr(1,\theta_{-i})-Pr(1,\theta_{-i},\mathcal{G})]\geq 0.$$

Assume that $\Pi_i(\theta_i,\theta_i)\geq\upsilon(\theta_i)$ for all $\theta_i<K$. Using a similar argument verifies that $\upsilon(K-1)-\upsilon(K)\geq y_i(K-1,K-1)-y_i(K-1,K)$.

Hence, ignoring all participation constraints, but those of type 1 is without loss. As $z_i(\theta_i)$ is weakly increasing, downward adjacent incentive constraints are binding. $\square$

[28]In the optimal mechanism it holds that $\Pi_i(1,1)=\upsilon(1)$.

C.7 Proof of Proposition 3

Proof. **Structure of the proof.** We use a guess and verify approach to proof Proposition 3. A constructive proof is possible, too, but notationally intense. In a companion paper Balzer and Schneider (2017) we provide a more constructive version. We also omit the proof that the unique equilibrium in the all-pay auction given the optimal belief system is monotonic.[29] We abuse notation by denoting a player i with cost $c_i = \kappa$ as player i_κ. We assume without loss of generality that $\rho_2 \geq \rho_1$ and proceed in several steps. We make frequent use of the upper bound of interval b, $\Delta_{\kappa,\kappa}$, in Figure 3 on page 24. For large parts of the proof none of the ignored constraint binds and signals are superfluous by Proposition 1.

Part A (Linearity of the Objective).

Step 1: Linearity of individual beliefs. Given ρ_1, ρ_2 use Lemma 4 to describe each $\beta_i(\theta_{-i}|\theta_i) \in \mathcal{B}$ as a linear function of $\tilde{\beta}_1(1|1) =: \tilde{\beta}$.

Step 2: Linearity of $\Delta_{\kappa,\kappa}$. The upper bound of $\Delta_{\kappa,\kappa}$ is determined by the mass put on player 1_κ and her belief. Let $F_{i_\theta}(a_i)$ denote the equilibrium probability of i_θ choosing an action smaller than a_i according to Siegel (2014)'s algorithm, then player 2_κ's equilibrium support includes a_2 if and only if it maximizes

$$\beta_2(\kappa|\kappa)F_{1_\kappa}(a_2) + \beta_2(1|\kappa)F_{1_1}(a_2) - a_2\kappa.$$

Player 2_κ is indifferent on the interval b in Figure 3. Thus player 1_κ uses a mixed strategy with constant density $f_{1_\kappa}(x) = k/\beta_2(\kappa|\kappa)$ for $x \in b$. Second, again by construction, player 1_κ uses her entire mass on that interval. The length of the interval is linear in $\tilde{\beta}$ since

$$\Delta_{\kappa,\kappa} = \frac{\beta_2(\kappa|\kappa)}{\kappa} = \frac{1}{\kappa(1-\rho_2)}\left(1 - \rho_1 - \rho_2 + \tilde{\beta}\rho_1\right).$$

Step 3: Linearity of winning probability at $\Delta_{\kappa,\kappa}$. By construction, player 2's winning probability is the same as her belief that player 1 is type 1_κ, i.e. $\beta_2(\kappa|m)$, which is linear in $\tilde{\beta}$ by step 1. The probability that player 2_κ chooses an action below $\Delta_{\kappa,\kappa}$ is determined by

$$\begin{aligned} F_{2_k}(\Delta_{\kappa,\kappa}) &:= \underbrace{F_{2_\kappa}(0)}_{\text{independent of } \tilde{\beta}} + \Delta_{\kappa,\kappa} \underbrace{\frac{\kappa}{\beta_1(\kappa|\kappa)}}_{\text{density of } 2_\kappa}. \\ &= F_{2_\kappa}(0) + \frac{1-\rho_2}{1-\rho_1}, \end{aligned} \tag{21}$$

which is independent of $\tilde{\beta}$. The winning probability of player 1_θ is linear in $\tilde{\beta}$ by step 1.

[29] Again see Balzer and Schneider (2017) for a treatment of this case.

Step 4: Linearity of equilibrium utilities. Given steps 2 and 3 it is sufficient to show that equilibrium utilities can be expressed in the form $F_i(\Delta_{\kappa,\kappa}) - \Delta_{\kappa,\kappa}c_i$. This follows from the construction of $U_1(1,1,\mathcal{B})$, $U_2(\kappa,\kappa,\mathcal{B})$. Further, $\Delta_{\kappa,\kappa}$ is the supremum of the equilibrium support of 1_κ. As there are no mass points other than at 0, $U_1(\kappa,\kappa,\mathcal{B})$ has the desired structure. Finally, $U_2(1,1,\mathcal{B}) = U_1(1,1,\mathcal{B})$ by the common upper bound.

Step 5: Linearity of the Objective. It remains to show (piecewise) linearity in the deviation utilities of κ-types. A deviating κ-type always has either $\Delta_{\kappa,\kappa}$ or 0 in her best response set. Second, such a deviator is only indifferent between actions if $\tilde{\beta} = \rho_2$. In this case beliefs are type-independent, and the deviator expects the same distribution (and thus utility) as a non-deviating player. If $\tilde{\beta} < (>)\rho_2$ her best response is a singleton at $\Delta_{\kappa,\kappa}$ (0). $\Delta_{\kappa,\kappa}$ and 0 are both linear in $\tilde{\beta}$ and the winning probability is, too. Thus deviating utilities are linear in $\tilde{\beta}$ and have a kink at $\tilde{\beta} = \rho_2$.

Part B (Optimality). We guess the solution at $\tilde{\beta} = \rho_2 = (1+\rho^0)/2$ and $\rho_1 = (1-\rho^0)/2$.

Step 1: Type-independency. Assume, to the contrary that that $\tilde{\beta} < \rho_2$ at the optimum, and rewrite

$$\begin{aligned}
\rho^0 \left[E[\Psi(\gamma)|\mathcal{G}] + E[U|\mathcal{G}]\right]\Big|_{\tilde{\beta}<\rho_2} &= F_{2_\kappa}(\Delta_{\kappa,\kappa})\left(\beta_1(\kappa|1)\rho_2 + \rho^0(1-\rho_2)\right) \\
&\quad - \beta_2(\kappa|1)\left((1-\rho^0)\rho_2 F_{1_\kappa}(\Delta_{\kappa,\kappa})\right) \\
&\quad + \underbrace{\Delta_{\kappa,\kappa}}_{=\beta_2(\kappa|\kappa)/\kappa}\left((\rho_1+\rho_2)(\kappa-1) - (1+\rho_2)\kappa\rho^0\right).
\end{aligned} \tag{22}$$

The derivative w.r.t. $\tilde{\beta}$ at the candidate solution for ρ_i, say ρ_i^c, is positive since

$$\lim_{\tilde{\beta}\to^-\rho_2} \frac{\partial\rho^0\left[E[\Psi(\gamma)|\mathcal{G}] + E[U|\mathcal{G}]\right]}{\partial\tilde{\beta}}\Bigg|_{\rho_i=\rho_i^c} = \frac{\kappa(1-(\rho^0)^2) - (1-(\rho^0)^2)}{\kappa(1+\rho^0)}$$

Instead assume $\tilde{\beta} > \rho_2$. Then the same derivative is negative since

$$\lim_{\tilde{\beta}\to^+\rho_2} \frac{\partial\rho^0\left[E[\Psi(\gamma)|\mathcal{G}] + E[U|\mathcal{G}]\right]}{\partial\tilde{\beta}}\Bigg|_{\rho_i=\rho_i^c} = -\frac{\kappa(1-(\rho^0)^2) - 1 - (\rho^0)^2}{\kappa(1+\rho^0)},$$

Step 2: Type distribution. Taking the derivative of (22) with respect to ρ_i and evaluating at $\tilde{\beta} = \rho_2$ directly establishes the critical point irrespective of the choice of ρ_{-i}.[30] Second order conditions are satisfied at the desired point and we can conclude that a local optimum exist in case we face a least constraint problem. Due to our assumptions on ρ^0, there always exists an $\alpha \leq 1$ such that the optimal solution satisfies (RC) with equality.

Step 3: Potential for signals. A sufficient condition for incentive compatibility of the

[30] By continuity of the objective the same holds true if we took the objective given $\tilde{\beta} \geq \rho_i$ instead.

candidate is obtained by directly plugging into the incentive constraint using $U(1,1,\mathcal{B}) = U(\kappa,1,\mathcal{B}) \geq U(m,\kappa,\mathcal{B})$. That is

$$\gamma_i(1) \geq \gamma_i(\kappa) \Leftrightarrow \rho_i \geq \rho^0. \tag{23}$$

This always holds for player 1_1, but not for player 2_1 if $\rho^0 > 1/3$. Now consider the following mechanism with public signals. There are two realizations σ_1 and σ_2 both equally likely. Under realization σ_1 the mechanism proceeds as above, under realization σ_2 proceeds as above but flips players identities. By ex-ante symmetry, the value of the value of the problem remains and condition (23) holds by Assumption 2 as it becomes

$$\frac{1}{2}\left(\gamma_i(1,\sigma_1)+\gamma_i(1,\sigma_2)\right) \geq \frac{1}{2}\left(\gamma_i(\kappa,\sigma_1)+\gamma_i(\kappa,\sigma_2)\right) \Leftrightarrow \frac{1}{2} \geq \rho^0.$$

□

C.8 Proof of Lemma 7

Proof. Take any $\gamma(k_1,k_2)$ and $\gamma(\theta_1,\theta_2)$. We want to show that (1,1)-consistency implies

$$q_1(k_2,\theta_2|k_1)q_2(k_1,\theta_1|\theta_2) = q_2(k_1,\theta_1|k_2)q_1(k_2,\theta_2|\theta_1).$$

By (1,1)-consistency we know that

$$q_1(\theta_2,1|k_1)q_2(k_1,1|1) = q_2(k_1,1|\theta_2)q_1(\theta_2,1|1), \tag{24}$$

$$q_1(k_2,1|\theta_1)q_2(\theta_1,1|1) = q_2(\theta_1,1|k_2)q_1(k_2,1|1), \tag{25}$$

$$q_1(\theta_2,1|\theta_1)q_2(\theta_1,1|1) = q_2(\theta_1,1|\theta_2)q_1(\theta_2,1|1), \tag{26}$$

$$q_1(k_2,1|k_1)q_2(k_1,1|1) = q_2(k_1,1|k_2)q_1(k_2,1|1). \tag{27}$$

Plugging into $q_i(\theta_{-i},1|\theta_i)q_{-i}(\theta_i,1|1) = q_{-i}(\theta_i,1|\theta_{-i})q_i(\theta_{-i},1|1)$, and rearranging yields

$$\frac{\beta_i(\theta_{-i}|\theta_i)}{\beta_{-i}(\theta_i|\theta_{-i})} = \frac{\beta_i(1|\theta_i)}{\beta_{-i}(1|\theta_{-i})}\frac{\beta_i(\theta_{-i}|1)}{\beta_i(1|1)}\frac{\beta_{-i}(1|1)}{\beta_{-i}(\theta_i|1)}. \tag{28}$$

Arrange all equations (24) to (27) according to (28). Observe on these transformations that the RHS of (24) times that of (25) is the same as the RHS of (26) times that of (27). Using the respective LHS of the equations yields

$$\begin{aligned} & \frac{\beta_1(k_2|k_1)}{\beta_1(\theta_2|k_1)}\frac{\beta_2(k_1|\theta_2)}{\beta_2(\theta_1|\theta_2)} = \frac{\beta_2(k_1|k_2)}{\beta_2(\theta_1|k_2)}\frac{\beta_1(k_2|\theta_1)}{\beta_1(\theta_2|\theta_1)} \\ \Leftrightarrow \quad & q_1(k_2,\theta_2|k_1)q_2(k_1,\theta_1|\theta_2) = q_2(k_1,\theta_1|k_2)q_1(k_2,\theta_2|\theta_1). \end{aligned}$$

□

References

Atakan, A. E. and M. Ekmekci (2014). “Auctions, actions, and the failure of information aggregation”. *American Economic Review* 104, pp. 2014–2048.

Aumann, R. J. and Maschler (1995). *Repeated games with incomplete information*. Cambridge, MA: MIT press.

Balzer, B. and J. Schneider (2015). “Mediation and Informational Punishment”. *mimeo*.

— (2017). “Optimal Alternative Dispute Resolution”. *mimeo*.

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

Bergemann, D., B. A. Brooks, and S. Morris (2016). “First price auctions with general information structures: Implications for bidding and revenue”. *Econometrica*, forthcoming.

Bergemann, D. and S. Morris (2016a). “Bayes correlated equilibrium and the comparison of information structures”. *Theoretical Economics*, pp. 487–522.

— (2016b). “Information Design, Bayesian Persuasion, and Bayes Correlated Equilibrium”. *American Economic Review* 106, pp. 586–91.

Bester, H. and R. Strausz (2001). “Contracting with imperfect commitment and the revelation principle: the single agent case”. *Econometrica* 69, pp. 1077–1098.

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.

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

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

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.

Dworczak, P. (2017). “Mechanism Design with Aftermarkets: Cutoff Mechanisms”. *mimeo*.

Fieseler, K., T. Kittsteiner, and B. Moldovanu (2003). “Partnerships, lemons, and efficient trade”. *Journal of Economic Theory* 113, pp. 223–234.

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

Jehiel, P. and B. Moldovanu (2001). “Efficient design with interdependent valuations”. *Econometrica* 69, pp. 1237–1259.

Kamenica, E. and M. Gentzkow (2011). “Bayesian Persuasion”. *American Economic Review* 101, pp. 2590–2615.

Konrad, K. A. (2009). *Strategy and Dynamics in Contests*. Oxford Univ. Press.

Lauermann, S. and G. Virág (2012). “Auctions in Markets: Common Outside Options and the Continuation Value Effect”. *American Economic Journal: Microeconomics* 4, pp. 107–130.

Luenberger, D. G. (1969). *Optimization by Vector Space Methods*.

Mathevet, L., J. Perego, and I. Taneva (2017). “On Information Design in Games”. *mimeo*.

Meirowitz, A., M. Morelli, K. W. Ramsay, and F. Squintani (2017). “Dispute Resolution Institutions and Strategic Militarization”. *mimeo*.

Myerson, R. B. and M. A. Satterthwaite (1983). “Efficient mechanisms for bilateral trading”. *Journal of Economic Theory* 29, pp. 265–281.

Philippon, T. and V. Skreta (2012). “Optimal Interventions in Markets with Adverse Selection”. *American Economic Review* 102, pp. 1–28.

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

Rentschler, L. and T. L. Turocy (2016). “Two-bidder all-pay auctions with interdependent valuations, including the highly competitive case”. *Journal of Economic Theory* 163, pp. 435–466.

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

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

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

Taneva, I. A. (2016). “Information design”. *mimeo*.

Tirole, J. (2012). “Overcoming Adverse Selection: How Public Intervention Can Restore Market Functioning”. *American Economic Review* 102, pp. 29–59.

Zhang, J. (2014). “Optimal Mechanism Design with Aftermarket Interactions”. *mimeo.*

Zheng, C. (2002). “Optimal Auction with Resale”. *Econometrica* 70, pp. 2197–2224.

— (2017). “A necessary and sufficient condition for Peace”. *mimeo.*