Managing A Conflict*

Benjamin Balzer[†] Johannes Schneider[‡]

February 2, 2017

Abstract

Two players conflict over a pie. They have the option to voluntarily participate in conflict management. In case conflict management cannot settle the conflict, it escalates to a costly Bayesian default game. Private information is only relevant in the default game which serves as both an endogenous outside option and a screening device. We show that optimal conflict management is equivalent to optimal post-escalation belief management. We characterize the set of feasible information structures post-escalation and link the mechanism design approach of eliciting information to the information design approach of processing information. We characterize the price of information revelation to the designer and show that additional public signals only play a minor role. Using two distinct examples we show how optimal conflict management links to the underlying games: while simple lotteries call for optimal sorting, contests advocate type-independent solutions.

1 Introduction

Conflict management is a tool used to solve disputes without letting them escalate to a costly fight. Its primary goal is to settle the conflict at little or no cost. However, conflict management is typically both voluntary and not always successful. If conflict management fails and the conflict escalates, players can use the information obtained about their opponent during conflict management when making further decisions. If optimal behaviour after escalation depends on players' information, players may use the information externality of conflict management strategically to extract information useful to them after escalation.

*We thank Eddie Dekel, Daniel Garrett, Susanne Goldluecke, Johannes Hoerner, Fabian Herweg, Benny Moldovanu, Volker Nocke, Martin Peitz, Klaus Schmidt, Till Stowasser, Thomas Troeger, Christoph Wolf, Asher Wolinsky, and seminar participants in Bonn, Mannheim, Munich, at the ESEM Geneva, the Jornadas in Palma, and the UECE meeting in Lisbon for comments and suggestions.

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

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

Thus, optimal conflict management must take behavioural effects of its information externality into account.

In this paper we study conflict management mechanisms when escalation is informative and triggers a non-cooperative Bayesian game. We consider escalation-minimizing conflict management when each player's strength in the post-escalation game is private information and unrelated to her preferences over settlement outcomes. Escalation is informative and players adjust their strategic behaviour post-escalation. These adjustments also influence the players' continuation utilities, since continuation-strategies depend on the beliefs players hold *after* learning that the conflict escalates. We show that the informational externality of escalation must not be ignored if players' optimal continuation strategies depend on their beliefs. In such cases escalation serves as an endogenous, belief-dependent outside option to the mechanism. We show generally that optimal conflict management is equivalent to optimally managing the beliefs players hold after escalation.

Examples of conflict management in which information is useful after escalation are abound. They include alternative dispute resolution escalating to litigation, mediated union-employer bargaining escalating to strikes, peace negotiations escalating to wars, or governed trade negotiations escalating to retaliations. All of the examples have three properties in common: (i) players can enforce the default game instead of conflict management unilaterally by constitution (litigation and strikes) or sovereignty (war and trade retaliation); (ii) conflict management aims at minimizing escalation because of a negative social externality on the legal system, the economy, global stability or global trade respectively, and (iii) after escalation, players play a given default game non-cooperatively. Further, parties typically hold private information about their individual strength within the default game which is unrelated to their preferences over settlement outcomes.

Taking into account the informational externality of optimal conflict management, we show that finding the optimal mechanism has a dual of finding the optimal beliefs players hold after escalation. We call the solution to the dual *optimal belief management*. This duality facilitates the problem significantly. Instead of solving a *mechanism design* problem with a complicated information externality we can instead focus on solving an *information design* problem that implies the solution to the mechanism design problem in a straight forward fashion. Optimal belief management provides an intuitive formulation of the economic problem the designer faces. Within the dual problem we identify the designer's two main motives. The information structure after escalation should ensure (i) discrimination (the *screening motive*), and (ii) little inefficiency (the *welfare motive*) in the continuation game. We establish additive-separable measures for both motives as objective of the dual.

The connection between mechanism design and information design is immediate. Players can use the rules of conflict management to Bayesian update their beliefs. The designer, in turn, can influence the updating procedure choosing these rules. *Belief management* refers to mechanisms set up to implement a particular information structure after escalation.

We connect optimal conflict management directly to the properties of the Bayesian default game via the belief management approach. This connection characterizes the general relation between the default way of conflict resolution and optimal conflict management. Further, the dual allows us to address the role of additional information the designer can release complementary to escalation. We provide simple sufficient conditions when such public signals are superfluous conditioning only on the set of binding constraints.

When a player forms her strategies for the continuation game, she uses all information she obtains during conflict management. In particular, she takes into account that the opponent updates his beliefs, too. Thus, the entire belief system influences the continuation value of escalation.

Conflict management makes the solution trivially less inefficient as settlement allows parties to save on the cost of fighting. However, given that parties need to mutually agree on conflict management either party can veto and enforce escalation. If strong parties expect to prevail after escalation they participate in conflict management only if they expect a large enough settlement valuation. A large settlement valuation attracts mimicking behaviour of weak types as strength is irrelevant under peaceful settlement. If pooling of *all* types is not affordable, escalation is the only screening device conflict management can use.

The fundamental challenge of the designer is to to keep the weak from mimicking the strong while making the strong participate. Each player's expected payoff from participation is the sum of the expected settlement value, financed by the efficiency gains of forgone fights, and the expected escalation value, determined by the expected utility after escalation. The more discriminatory the post-escalation game, the easier to deter mimicking behaviour. Simultaneously, the less inefficient the post-escalation game, the smaller the necessary settlement value to ensure participation. By choosing the belief system post-escalation, the designer influences the level of both discrimination and inefficiency.

Consider for example a standard war of attrition as the default game. Suppose both players expect to be of similar strength leading to large investment and small aggregate surplus. Compare this to a situation in which one player expects an opponent of similar strength, but the opponent instead expects the player to be (on average) weak. Given the logic of the war of attrition, the opponent invest less compared to the situation with similar expectations. As a response, the player, too, reduces her investment making the overall surplus larger than in the case of symmetric expectations. Not only strategies but also, expected individual utility and overall expected (utilitarian) welfare in the continuation game depend thus on the entire *belief system* and not only on the player's individual belief about her opponent's type.

The belief system also impacts the profitability of deviation within conflict management. Action choices after escalation can be adjusted after deviations, and provide an additional incentive for deviation. The reason is that any strategic adjustment a player makes subsequent to conflict management remains unresponded as the deviation is not detected. Suppose for example that a player expects to face a weaker opponent after a deviation. This deviating player adjusts her continuation-strategy to the expected strength of her opponent, but – different than on the equilibrium-path – *without* expecting her opponent to react. Thus, deviation becomes more attractive.

We use our approach and compare two examples of escalation games. Games in which the optimal continuation-strategy is independent of the beliefs (e.g. Hörner, Morelli, and Squintani (2015)) and games in which it is sensible to beliefs (e.g. an all-pay auction).

We show that conflict management in games with belief-independent continuation strategies induces *sorting* by the mechanism. That is, conflict management identifies easy to solve matches and promises them settlement while sending difficult matches to the default game.

Due to the belief-dependent behavioural adjustments the result differs if strategies react to beliefs, a case not considered much in the literature so far. We show for the example of the all-pay auction, that sorting would undermine the designer's screening motive. Instead, and to induce truth-telling, she promises each player the same information set independent of their behaviour. Further we show that, different to the first example, the solution generically involves an asymmetric type distribution in any continuation game even with ex-ante symmetric players.

1.1 Related Literature

The mechanism design literature on conflict preemption builds on the classical literature on trade mechanisms going back to Myerson and Satterthwaite (1983). Our general setup is related to Compte and Jehiel (2009) within that literature. We assumes a division of a pie as an outcome, a budget constraint mechanism, and an outside option that potentially depends on the information structure.

Different to Compte and Jehiel (2009), the outside option to conflict management depends not only on the player's type but also on that of the opponent, and escalation may not be avoided completely, despite being always inefficient. In our model, the information structure post-escalation *depends* on the design of the mechanism and *influences* the action choice in the continuation game via the entire belief system. The mechanism thus imposes an informational externality.[1]

[1]Interim dependence of the outside option on beliefs is also present in Fieseler, Kittsteiner, and Moldovanu (2003) and Jehiel and Moldovanu (2001) which allow for interdependent values. Inefficient escalation is a general feature of the conflict preemption literature (Bester and Wärneryd, 2006; Hörner,

Informational externalities of a mechanism are also present in bargaining models with interdependent valuations (Jehiel and Moldovanu, 2001), and auctions with resale as in Zheng (2002) who considers a sequence of mechanisms each proposed by a different agent.

Closer related to this paper are Philippon and Skreta (2012) and Tirole (2012) who study the informational externality of a bailout mechanism on future market behaviour. Similar to our approach they consider a model in which the design of the mechanism influences the interpretation of observed behaviour and thus subsequent behaviour in the market. A similar approach is taken by the literature on aftermarkets in auctions (Atakan and Ekmekci, 2014; Dworczak, 2016; Lauermann and Virág, 2012) that considers how informational externalities of the mechanism influence behaviour after the auction. The main difference to us is that the design of the auction affects the belief system in the aftermarket only one-sided and beliefs after the mechanism are type-independent by design. In our model, type-dependent beliefs are possible since the same players meet after escalation and the mechanism can influence the entire belief-system.

The literature on conflict preemption is evolving (Bester and Wärneryd, 2006; Hörner, Morelli, and Squintani, 2015; Spier, 1994). Contrary to that on aftermarkets, this literature typically involves type-dependent beliefs. Most models consider, however, only type-specific lotteries as continuation-games for tractability. In such games action choices post-escalation become irrelevant. We nest these models, but allow for general Bayesian continuation-games in which players' action choices are influenced by the mechanism and determine the outcome.[2] Our key contribution to this literature is to provide a general approach to conflict management applicable to a general class of continuation games. We identify the channel that links the properties of that (Bayesian) game to the optimal mechanism. We show that the properties of the game have a significant impact on the design problem.

We complement Zheng (2017) who studies necessary and sufficient conditions for full settlement for contest default games. We characterise the optimal mechanism when his conditions are violated. Meirowitz et al. (2015) study the effect of last-minute conflict management on early investment, while we focus on early stage conflict management that saves on this investment. Beliefs in Hörner, Morelli, and Squintani (2015) are only important under settlement due to limited commitment of the players. In our model, commitment is not an issue by design. Instead we focus on the role of beliefs in case of escalation.

Our model is an augmented information design problem (Bergemann and Morris, 2016a,b; Kamenica and Gentzkow, 2011; Taneva, 2016). The designer chooses two information struc-

Morelli, and Squintani, 2015; Spier, 1994). However, no post-escalation action choice is present in either of these models resulting in an linear outside options.

[2]As shown by example in Celik and Peters (2011), Bayesian default games can make full-participation non-optimal. Their channel is not present in the main part of our paper, but we provide an extension addressing these concerns, too.

tures (settlement and escalation) that combine to the initial information structure. The main difference to the pure information design literature is that the designer can directly determine the allocation under settlement, but only indirectly that under escalation.

Roadmap. The remainder of the paper is structured as follows. We describe the model and some preliminary simplifications 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 illustrates the power of our approach by comparing two classes of default games. We demonstrate how the properties map into significantly different optimal mechanism. In section 5 we discuss the robustness of our findings to relaxing of several of our assumptions. We conclude in section 6. All formal proofs are relegated to the appendix.

2 Setup

General Setup 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. Players can either mutually agree on a conflict management mechanism to solve the conflict or engage in a non-cooperative Bayesian default game. Independent of her valuation of the pie, each player i is endowed with a type θ_i independently drawn from $\Theta = \{1, 2, ..., K\}$. We say $\theta_i \in \Theta$ is player i's *strength in the default game.* The known ex-ante probability of being θ_i is $\rho^0(\theta_i) > 0$.

A basic outcome of our model is an element of the two-dimensional simplex, representing the distribution of (parts of) the pie.[3] We categorize 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. An element in $\mathcal{X}$ defines the shares attributed to each player under settlement. In the event of escalation the default game decides over the division of the pie. An element in $\mathcal{G}$ is the image of a fundamental 2-player Bayesian game Γ mapping from type profiles and an action pairs to outcomes.

Conflict Management. Conflict management is a mechanism proposed by a non-strategic third-party, the designer. 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 into the outcome space, that is

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

[3]Players are risk neutral, which is why we make no restrictions whether the distribution refers to an actual division of the pie, or whether players engage in a (fixed) lottery about the pie as a whole.

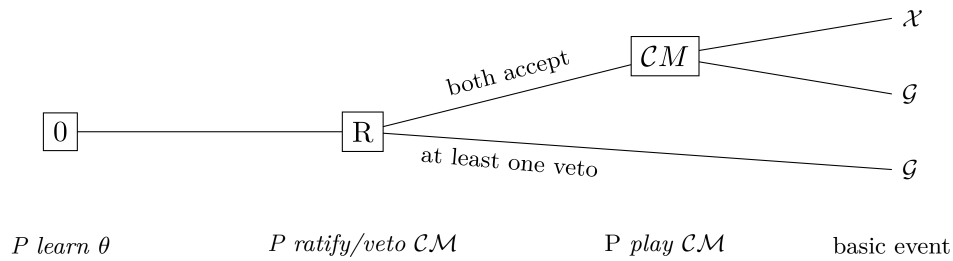

Figure 1: Timing of events.

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

Timing. At the initial stage (0) players privately learn their type. At the ratification stage (R) players decide simultaneously whether they want to participate in conflict management. If both agree on $\mathcal{CM}$, they report their types and $\mathcal{CM}$ results either in settlement or escalation. Under settlement, players receive a share and the game ends. Under escalation, players learn the signal realization σ, update their beliefs according to Bayes' rule and non-cooperatively decide on their action in the Bayesian game.[4] If at least one player vetoes conflict management at (R), the conflict escalates immediately to the Bayesian game. The rules of the grand game described in section 2 are common knowledge. We assume that the designer has full commitment power, and players can commit to accepting any outcome in $\mathcal{X}$ at the ratification stage.

Solution Concept and Beliefs. We use perfect Bayesian equilibrium as solution concept. Thus, whenever possible, players use all information available to update their beliefs according to Bayes' rule. Suppose escalation is announced signal σ realises. Now, player i who reported m during conflict management, assigns a conditional probability to each possible type of player $-i$. This conditional probability is denoted by $\beta_i(\theta_{-i}|m, \sigma)$, and the collection of conditional probabilities $\beta_i(\cdot|m, \sigma)$ is the *individual belief* of that player. 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, $\mathcal{B}(\sigma)$, the *realized belief system* given σ. A lottery over realized belief systems, $\mathcal{B}(\Sigma) := \{\mathcal{B}(\sigma), Pr(\sigma)\}_{\sigma\in\Sigma}$ is a

[4] In principle, signals could also realize if the outcome is $\mathcal{X}$. However, in the baseline model with full commitment signals have no effect on the final outcome in all settlement events.

collection of realized belief systems and their respective occurrence-probabilities. For notational simplicity we will distinguish between two individual beliefs $\beta_i(\cdot|\theta_i,\sigma)$ and $\beta_i'(\cdot|\theta_i,\sigma)$ only if player θ_i appears with positive probability given σ. A special case 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 $|supp(\Sigma)|=1$.

2.1 Preliminary simplifications

The On-Path Continuation Game. The fundamental game Γ consists of 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 into payoffs. Given escalation and the realization of σ, the (on-path) information structure is entirely determined by the belief system $\mathcal{B}(\sigma)$. Up to the choice of equilibrium, Γ and $\mathcal{B}(\sigma)$ are thus sufficient to determine the equilibrium outcome. In particular, take any fundamental game Γ, and any belief system $\mathcal{B}(\sigma)$ for which a unique equilibrium with some properties $(*)$ exists, then there is a function, $\mathbf{s}^*:\mathcal{B}(\sigma)\mapsto\Delta(A)^{2K}$, that fully describes the equilibrium (mixed-)strategies in that equilibrium. Given the equilibrium strategies, the von-Neumann-Morgenstern equilibrium utility of player θ_i that is matched with type θ_{-i} is denoted by $u_i(\theta_i,\theta_{-i},\mathcal{B}(\sigma)):=u_i(\theta_i,\theta_{-i},\mathbf{s}^*(\mathcal{B}(\sigma)))$ and the expected on-path utility is denoted by $U_i(\theta_i,\mathcal{B}(\sigma)):=\sum_{i=1}^K\beta_i(k|\theta_i,\sigma)u_i(\theta_i,k,\mathcal{B}(\sigma))$. We assume that an equilibrium exists for any realization of $\mathcal{B}(\sigma)$. Its selection is known by the designer.

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 to be type m as the limiting utility of what the deviator could obtain when choosing her strategy optimally, i.e.,

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

Observe that if $m=\theta_i$ the above equation describes the on path utility and $\mathbf{s}_i^*$ is determined by $\mathcal{B}(\sigma)$. 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 maximum continuation-utility of player θ_i, reporting to be type m after signal realization σ. The designer can neither influence Γ nor the choice of equilibrium. In what follows we treat the function $U_i(m_i,\theta_i,\mathcal{B}(\sigma))$ as a primitive to the optimal mechanism. We define the expected continuation utility given lottery $\mathcal{B}(\Sigma)$ by $\widehat{U}_i(m,\theta_i,\mathcal{B}(\Sigma)):=\sum_{\sigma\in\Sigma}Pr(\sigma)U_i(m,\theta_i,\mathcal{B}(\sigma))$.

Assumptions on the underlying game. To proceed we impose structure on U_i. Since U_i is entirely determined by an equilibrium of Γ, the assumptions we make are essentially assumptions on the default game. We assume upper hemi-continuity of U_i in $\mathcal{B}(\sigma)$ and an *anonymous conflict.* Define

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

Definition 1 (Anonymity)**.** The game Γ and the equilibrium choice rule satisfy *anonymity* if for any $\mathcal{B}(\sigma)$ and any type profile (θ_i, θ_{-i}) it holds that $u_i(\theta_i, \theta_{-i}, \mathcal{B}(\sigma)) = u_{-i}(\theta_i, \theta_{-i}, \mathcal{B}_-(\sigma))$.

Definition 2 (Conflict)**.** A game Γ and an equilibrium choice rule describe a *conflict* if for any $\mathcal{B}(\sigma)$ it holds that

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 in own type.)* $U_i(\theta_i, \mathcal{B}(\sigma))$ is non-increasing in θ_i.
iii. *(monotonicity in own belief.)* $U_i(\theta_i, \mathcal{B}(\sigma))$ increases if $\beta_i(\theta_i, \sigma)$ increases in the first order stochastic dominance sense.

Following Definition 2, we order types such that type 1 is *the strongest* and type K is the *weakest.* Non-productiveness guarantees that escalation always reduces the amount of resources available in the economy. Thus ex-ante conflict is never desirable from a utilitarian point of view compared to any settlement solution that distributes the entire pie. The two monotonicity properties ensure that the player's type is a sufficient statistic for the player's ability in the conflict game. While the first property states that higher ability results in higher expected utility, the second property ensures that player prefer weaker opponents.

Discussion of the Assumptions. We provide a detailed discussion of our assumptions in section 5. In this part we discuss the two assumptions most important to our model. First, we assume the designer has full commitment power. In reality, institutions designing conflict management tend to act repetitively and thus rely on reputations which gives them an (unmodelled) incentive to commit.[5]

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 via the mechanism. Our results can be seen as a benchmark on the possibilities of third party conflict management in two ways. (i) It is one of the limiting cases of a more general model in which the private information is about preferences *and* strength, the other limit being a classical trade mechanism with

[5]Our assumption of full commitment power of all participants 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. Evidence supporting our assumption of commitment on the designers' side can also be found in the online appendix of Hörner, Morelli, and Squintani (2015).

interdependent values á la Fieseler, Kittsteiner, and Moldovanu (2003) without transfers in which preferences and ability coincides. (ii) A second benchmark is on the technology of conflict management. We assume that the designer has no technology to test the players private information other than escalation. Thus, settlement relies only on soft information. In reality, such mechanisms provide the cheapest and fastest solution to the dispute. If known to the players when deciding on participation in conflict management, however, we can subsume any more complicated mechanism in the escalation stage making our model a benchmark on what can be achieved via the exchange of soft information only.

3 Analysis

In this section we develop our two main results. First, 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 (i.e., a screening measure), and welfare measure post-escalation. The two measures are entirely determined by the fundamental game Γ, the prior $\boldsymbol{\rho}^0$ and a belief system, $\mathcal{B}$. These measures directly link the post-escalation game to the optimal mechanism.

We proceed in steps. First, we reduce the choice set using the binding constraints. Then, we transform the problem into reduced-form, and define the set of consistent belief systems. Finally, we use the first-order approach to derive the dual, and state the equivalence result.

3.1 Binding Constraints

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 after a veto by player i and $v_i(\theta_i)$ the value of vetoing, that is, the expected utility from Γ of the vetoing player under the belief system $\mathcal{B}^V$. In particular, $\mathcal{B}^V$ contains the prior type distribution ρ^0 as belief for player i, the *individual veto belief* β^V which is the same for any type of player $-i$. We treat β^V and thus v_i as given and make an assumption that guarantees full participation.[6] For simplicity we assume that v_i is the same for both player and thus suppress the subscript.

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

The next assumption is necessary and sufficient for a non-trivial conflict management.

Assumption 2. $v(1) > 1/2$.

[6] Celik and Peters (2011) give an example in which full participation is not optimal, Cramton and Palfrey (1995) discuss the choice of β^V. We eliminate Assumption 1 by slightly changing the model in section 5.

A violation of Assumption 2 allows a 50/50 sharing leading to full settlement. Zheng (2017) constructs the conditions leading to Assumption 2 for general contest games.

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

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

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 $y(m_i, \theta_i)$. Note that the values $\gamma_i(m_i)$ and $x_i(m_i)$ correspond to θ_i's expected probability of escalation and her expected settlement share, and depend only on her report m_i.[7] Since preferences over outcomes are identical, the settlement value depends on the report only, but the escalation value depends on both the reported, and the actual strength in the default game. Observe, that the escalation value $y_i(m_i, \theta_i)$ is homogenous of degree 1 with respect to γ since $\mathcal{B}$ and thus $\hat{U}$ is homogenous of degree 0 with respect to γ.[8] By the revelation principle and Assumption 1 the set of participation constraints

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

and the set of incentive compatibility constraints

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

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. We proceed by identifying redundant constraints.

Lemma 1. *The following holds for the optimal mechanism*

i. all incentive compatibility constraints not concerning adjacent types are redundant,

ii. if both adjacent incentive compatibility constraints are redundant for type θ_i, then her participation constraint is satisfied with equality or $z_i(\theta_i) = 0$,

iii. the participation constraints for at least one type of every player is binding.

Although the set of binding constraints may depend on the exact location of the optimum, Lemma 1 provides enough structure to determine settlement values as a function of

[7]Note that the expected settlement share $x_i(m_i)$ depends also on the probability of escalation, γ, since shares are not independent of the escalation rule.

[8]We come back to this point in greater detail in section 3.2.

escalation values. This reduces the dimensionality of the choice set. We can specify $2K$ linear equations which uniquely identify the settlement values as a function of the settlement values with help of Lemma 1.

Corollary 1. *Take any set of escalation values* $\{y_i(\theta_i,\theta_i), y_i(\theta_i+1,\theta_i), y_i(\theta_i-1,\theta_i)\}$ *for all* $\theta_i \in \Theta$*. Then there is a partition of* $P(\Theta) = \{\Theta^{z_i=0}, \Theta^{PC}, \Theta^{IC^+}, \Theta^{IC^-}\}$ *such that*

$$z_i(\theta_i) = z_i(\tilde{\theta}_i) + y_i(\tilde{\theta}_i, \theta_i) - y_i(\theta_i, \theta_i), \tag{Z}$$

with

$$\tilde{\theta}_i = \begin{cases} \theta_i + 1 & \text{if } \theta \in \Theta^{IC^+} \\ \theta_i - 1 & \text{if } \theta \in \Theta^{IC^-}. \end{cases}$$

Moreover $z_i(\theta_i) = 0$ *for* $\theta_i \in \Theta^{z_i=0}$*, and* $z_i(\theta_i) = v(\theta_i) - y_i(\theta_i,\theta_i)$ *for* $\theta_i \in \Theta^{PC}$*.*

Reduced Form Representation. Using Corollary 1 we can construct settlement values from escalation values using the set of binding constraints. Using z_i instead of the sharing rule X_i implies that we use a reduced-form approach. The reduced-form approach ignores implementation of the reduced-form z_i via a settlement rule X_i and the escalation rule γ_i. We use an adapted version of the general implementation condition (GI) in Border (2007) to state a necessary and sufficient condition when our reduced-form approach is indeed valid. Abusing notation slightly, we denote the ex-ante probability of escalation with $Pr(\mathcal{G})$. Le $Q_i \subseteq \Theta$ be any subset of the type space and define $\bar{Q} = \{(\theta_1,\theta_2) \in \Theta^2 | \theta_i \notin Q_i \text{ for } i = 1,2\}$ for each (Q_1, Q_2).

Lemma 2 (Sufficiency of Reduced-Form Mechanism)**.** *Fix a feasible* $\gamma(\cdot,\cdot)$ *and* $z_i(\cdot) \geq 0$*. An ex-post feasible* X *that implements the reduced form allocation* z_i *exists if and only if*

$$\sum_i \sum_{\theta_i \in Q_i} z_i(\theta)\rho^0(\theta_i) \leq 1 - Pr(\mathcal{G}) - \sum_{(\theta_1,\theta_2)\in\bar{Q}} (1-\gamma(\theta_1,\theta_2))\rho^0(\theta_1)\rho^0(\theta_2), \quad \forall Q_1, Q_2 \subseteq \Theta. \tag{GI}$$

Lemma 2 implies that the (expected) resource constraint holds at the optimum, that is,

$$\sum_i \sum_m \rho^0_m \, z_i(m) \leq 1 - Pr(\mathcal{G}). \tag{AF}$$

We treat the condition (GI) as an additional constraint of the designer.

3.2 Conflict Management and Belief Management

The option value of escalation depends on the belief system at the point in which players choose their actions in the continuation game. In this part we characterize the set of possible belief systems and show that given the optimal signal structure, finding the optimal belief system is isomorphic to finding the optimal mechanism. We establish an equivalence between conflict management and belief management on the level of realized belief systems. For simplicity, we focus on a degenerate signal distributions in the exposition, but state (and prove) the more general Theorem 1. We start with a straight-forward observation.

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

Recall that the escalation rule, $\gamma(\theta_1, \theta_2)$, is defined as the probability of conflict escalation conditional on the realization of type profile (θ_1, θ_2). Thus, whenever escalation occurs, each player θ_i uses γ to update the conditional probability of facing player θ_{-i}, $\beta_i(\theta_{-i}|\theta_i)$. The rule γ, the updating procedure of every player, and hence the belief system $\mathcal{B}$ at the point of escalation are common knowledge. The reverse statement to Observation 1 is not true for two reasons: First, the belief system is determined by *relative* escalation probabilities only. If γ implements $\mathcal{B}$, so does any $\alpha\gamma$ for any $\alpha \in (0, 1]$. Second, not every possible belief system is consistent with some γ. Any $\mathcal{B}$ implementable by some γ must 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 can, however, find necessary and sufficient conditions when a given belief system is consistent with some escalation rule. Intuitively, fix any $\gamma(1, 1) > 0$ and some belief system $\mathcal{B}$. Ignore the natural constrain $\gamma(\cdot, \cdot) \in [0, 1]$ for the moment. Our aim is now to construct all function values $\gamma(\theta_1, \theta_2)$ from $\mathcal{B}$ and the anchor $\gamma(1, 1)$. To construct $\gamma(\theta_1, \theta_2)$ we can use two paths. We can construct $\gamma(1, \theta_2)$ and from that $\gamma(\theta_1, \theta_2)$, or we can construct $\gamma(\theta_1, 1)$ and from that $\gamma(\theta_1, \theta_2)$. If and only if both path yield the same value for any $\gamma(\theta_1, \theta_2)$, then $\mathcal{B}$ is consistent with some γ. Constructing $\gamma(\theta_1, \theta_2)$ as sketched above 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 fixed $\gamma(1, 1) > 0$, 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 can use the continuity in the mapping from γ to $\mathcal{B}$.[9] 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*

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

[9]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. In addition there exists a function $g(\mathcal{B})$ such that any γ implementing $\mathcal{B}$ takes the form $\gamma = \alpha g(\mathcal{B})$ for some scalar α.

Lemma 3 provides $(K-1)^2$ independent equations which allow us to define any consistent $\mathcal{B}$ by choosing two objects: $\boldsymbol{\rho}^e$ and $\boldsymbol{\beta}_1^e$. The first, $\boldsymbol{\rho}^e = (\rho_1^e, \rho_2^e)$, describes the ex-ante expected distribution of types in the continuation game, where $\rho_i^e(\theta_i)$ is the likelihood of a player being type θ_i in the continuation game. The second object, $\tilde{\boldsymbol{\beta}}_1^e = \{\beta_1(\cdot|\theta)\}_{\theta<K}$, collects the individual beliefs of the $K-1$ strongest types of player 1, that is the conditional probability distributions over the opponent's type. We say $B = \boldsymbol{\rho}^e \cup \boldsymbol{\beta}_1^e$ is consistent if $\mathcal{B}$ implemented via B is consistent. One interpretation of $\boldsymbol{\rho}^e$ is that it determines the distribution of types while β^e determines how the different types are matched to realized continuation matches.

Lemma 4. *There exists a continuous bijection h such that $h(B) = \mathcal{B}$ for any consistent B and $\mathcal{B}$.*

Finally, we show that at any of the choice variables introduced above, i.e. $\mathcal{B}$ or B, entirely determines the optimal escalation rule.

Lemma 5. *Take any implementable mechanism with escalation rule γ and X determined via 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 $\mathcal{B}$.*

Lemma 6. *Any γ that does not satisfy* (AF) *with equality is not an optimal escalation rule.*

Lemmas 4 to 6 are the final results to establish the equivalence in choice sets. Any B determines γ only up to a scalar α. The probability of escalation $Pr(\mathcal{G})$ is monotone increasing in α while (AF) becomes tighter as α decreases. Thus, at the optimum (AF) holds with equality, and determines α^* and thus γ.

Remark. In what follows we are going to restrict attention to escalation rules with $\gamma(1,1) > 0$, and thus the set of possible mechanisms to $\mathcal{CM}^+ \subset \mathcal{CM}$. In what follows we are going to look for the infimum over $\mathcal{CM}^+$ which corresponds to the minimum over $\mathcal{CM}$ by continuity of the objective. This restriction is without loss of generality.[10]

Theorem 1. *There is a one-to-one mapping between the optimal mechanism $\mathcal{CM}$ and the lottery over realized belief systems, $\mathcal{B}(\Sigma)$.*

[10]Consider a convergent sequence of interior belief-systems $\mathcal{B}_n$ as in Lemma 3. $\gamma_n(\theta_1, \theta_2) = \frac{\beta^0(\theta_1)}{\beta^0(1)} \frac{\beta^0(\theta_2)}{\beta^0(1)} \frac{\beta_{2,n}(\theta_1|\theta_2)}{\beta_{2,n}(1|\theta_2)} \frac{\beta_{1,n}(\theta_1|1)}{\beta_{1,n}(1|1)} \gamma_n(1,1)\ \forall(\theta_1, \theta_2)$. By consistency, this is simply $\gamma_n(\theta_1, \theta_2) = \gamma_n(\theta_1, \theta_2)\ \forall n \Rightarrow \lim_{n\to\infty} \gamma_n(\theta_1, \theta_2) \in [0, 1]$ iff $\gamma_n(\theta_1, \theta_2) \in (0, 1)$ for n large.

Remark. For degenerate distributions the result follows directly from the lemmas discussed above. The extension to non-degenerated signals comes from the following observation: Each signal realization σ in the support of Σ induces an information structure. By iteratively combining all realized information structures in the sense of Bergemann and Morris (2016a) we can compute the information structure conditional only on escalation. That information structure is thus an expansion of any possible outcome of the lottery $\mathcal{B}(\Sigma)$. The prior, in turn, is the combination of the signal structure post escalation and that post settlement.

Theorem 1 states the isomorphism between conflict management and belief management, that is any mechanism that optimally manages the belief system post-escalation also optimally manages the conflict. The result has several immediate 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 and highlights how sensitive the optimal mechanism is to the default 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 link to literature on information design (Bergemann, Brooks, and Morris, 2016; Bergemann and Morris, 2016a,b; Taneva, 2016). We show that optimal conflict management is essentially determined by choosing information structures which then 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 simplify 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 3 (Minimization of Conflict Escalation)**.**

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

In what follows, we show that the characterization of the optimal mechanism corresponds to maximizing the sum of two measures on the continuation game: a measure on the degree of discriminating types in the continuation game (the (expected) ability premium) and a measure on the inefficiency in the continuation game (the (expected) welfare). To convey intuition we restrict attention to an environment in which the strongest type is sufficiently privileged from an ex-ante point of view. The intuition remains for other games but involves some case distinctions. We discuss it in section 5.

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 $Q \neq 1$.[11]

Using Assumption 3 we can state the following by combining Lemmas 1 and 5

Lemma 7. *Suppose Assumptions 1 to 3 hold. Then, the following is true at the optimum*

i. downward adjacent incentive compatibility constraints are satisfied with equality,
ii. $z_i(\theta_i) > 0$ *for any* $\theta_i > 1$,
iii. only the participation constraint for the strongest type is binding.

Under Assumption 3 the designer faces the following trade off: She must provide the strongest type a sufficient amount of expected value from $\mathcal{CM}$, while keeping weaker types from imitating the strongest. To solve this trade off, two motives are relevant: a *screening motive* and a *welfare motive*. Making the continuation game discriminatory and thus unattractive for weaker types deters imitation (the screening motive). Making the continuation game attractive for all types saves on resources by increasing its value in case of failure (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 rents have to be paid. A measure of the (relevant) discriminatory power is the *ability premium*.

Definition 4 (Ability Premium)**.** The ability premium, $\psi_i(\theta)$ is the difference in expected utility after escalation between a type θ and the a next strongest deviating type $\theta{+}1$. That is, $\psi_i(\theta_i, \sigma) = U_i(m{=}\theta_i, \theta_i, \mathcal{B}(\sigma)) - U_i(m{=}\theta_i, \theta_i{+}1, \mathcal{B}(\sigma)), \quad \forall \theta_i < K$.

Definition 5 (Weighted Ability Premium)**.** The *weighted ability premium* is $\Psi_i(\theta_i, \sigma) := w(\theta_i)\psi_i(\theta_i, \sigma)$, with $w(\theta_i) = (1 - \sum_{k=1}^{\theta}\rho^0(k))(\rho^0(\theta))^{-1}$ the inverse hazard rate of type θ given the prior.

The ability premium 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. Recall that 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

[11]The assumption imposes monotonicity on the problem. Examples for games in which Assumption 3 holds are given in section 4. Other examples include contests with costs increasing sufficiently convex in θ.

to the “closest deviator”, and thus measures the discriminatory power of the underlying game. Observe that monotonicity guarantees that the ability premium is weakly positive.

The larger the ability premium the less attractive mimicking behaviour, and the lower the pressure on (IC). Thus, all else equals, the designer desires a high ability premium. How important a particular player’s ability premium is depends on the prior distribution which weighs the ability premium with 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 ombudsman also 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 measure for this is expected welfare in the continuation game after escalation.

The set of binding constraints according to Lemma 7 consists of the downward adjacent incentive compatibility constraints and strongest types’ participation constraints. We call this set $C_R \subset \mathcal{C}$. Further, we denote $C_F := \mathcal{C} \setminus C_R$. Define $\mathbb{E}[f_i|\mathcal{G}] = \sum_{\theta\in\Theta} \rho_i(\theta) f_i(\theta)$ as the expected value of f_i conditional on escalation.

Proposition 1 (Duality)**.** *Suppose Assumptions 1 to 3 hold and fix the set of signal realizations to a singleton. Then $\mathcal{CM}$ solves* (P_{min}) *if and only if it solves*

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

over the set of consistent $\mathcal{B}$.

The equivalence in the choice set has been established in theorem 1, the transformation in the objective can be obtained by comparing first order conditions. In appendix A we provide the full Lagrangian approach. Here, we summarize our the argument under the restrictions of the proposition.

Combining corollary 1 and lemma 7 we express the settlement value as

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

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 $\sum_i \sum_\theta \rho^0(\theta) z_i(\theta) = \sum_i \upsilon(1) + \gamma(1,1)\mathcal{Q}(\mathcal{B})$. 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, (AF) binds at the optimum and $\sum_i v_i(1) - \gamma(1,1)\mathcal{Q}(\mathcal{B}) = 1 -$

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

Ignoring public signals, Proposition 1 characterizes the solution linking it directly to properties of Γ. It provides a tractable solution approach to finding the optimal mechanism. Given the equilibrium characterization of Γ for any consistent $\mathcal{B}$, the only additional object we need is the off-path utility of a type θ_i+1. That is the utility of a type θ_i+1 endowed with the individual belief of type θ_i facing an opponent that expects equilibrium play according to $\mathcal{B}$. Those off-path utilities are easily calculated as they are the solution to a simple decision problem. The characterization 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, the characterization 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 solution to the problem of Proposition 1. Using our integrative approach we obtain a simple condition when the problem in proposition 1 is sufficient.

Corollary 2. *Consider the solution to* (P_{max})*. If it is least constraint, then the optimal signal structure is degenerate.*

Via Lemma 1 we incorporate the most relevant constraints directly in the objective of (P_{max}). If the remaining constraints are redundant, signals do not improve. The reason is in line with the Bayesian persuasion literature, e.g. Kamenica and Gentzkow (2011). Solving (P_{max}) without signals determines an initial information structure to a pure information design problem. Each (interior) solution is concave with respect to $\mathcal{B}$ by definition. Thus, the (information) designer cannot exploit convexities to increase the objective and signals are superfluous.

If the solution to the unconstraint problem in proposition 1 does not satisfy C_F, signals may be optimal. Different to the classical Bayesian persuasion literature we cannot rely on the concavification arguments of Aumann and Maschler (1995) to find the optimal signal. The on-path information design problem *after* escalation departs from that literature in two ways. First, the designer is –at this point– omniscient in the sense of Bergemann and Morris (2016a), and second, the expected continuation utility $\widehat{U}(m, \theta, \mathcal{B}(\Sigma))$ must satisfy the constraints in C_F. We follow a two step procedure. We fix some belief system $\mathcal{B}$ and determine the set of lotteries S with this mean. Then we look for the optimal, implementable lottery $\Sigma^* \in S$. Finally, we solve for the best $\mathcal{B}$ taking the optimal lottery for each $\mathcal{B}$ as given.[12]

Definition 6 (Admissible belief system)**.** We say that a belief system $\mathcal{B}$ is in the set of admissible belief systems $\mathcal{B}^a$ if (i) it is the mean of a lottery over belief-systems, $\mathcal{B}(\Sigma)$, with each outcome being consistent for some γ, and (ii) the constraints C_F hold with respect to $\mathcal{B}(\Sigma)$.[13]

Each admissible belief system has a set of implementable lotteries $S^a(\mathcal{B})$ that satisfy $\mathcal{C}_F$ and have mean $\mathcal{B}$. Given objective $\mathcal{O}$, we identify an element $\mathcal{B} \in \mathcal{B}^a$ by its value

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

Theorem 2 (Duality of problems)**.** *Suppose Assumptions 1 to 3 hold. Any mechanism* $\mathcal{CM}$ *solves equation* (P_{min}) *if and only if it also solves*

$$\sup_{\mathcal{B} \in \mathcal{B}^a} \mathcal{V}\left(\mathcal{B}, \sum_{i \in \{1,2\}} \mathbb{E}[\Psi_i | \mathcal{G}] + \mathbb{E}[U_i | \mathcal{G}]\right). \tag{P_{max}}$$

Remark. In contrast to Kamenica and Gentzkow (2011) the solution is not necessarily on the concave hull of the objective. The main reason is that we need to take the constraints C_F into account on the level of $\hat{U}$ that depends not only the mean $\mathcal{B}$ but on the entire lottery $\mathcal{B}(\Sigma)$.[14] Instead we exploit the fact that we can separate the pure information design problem given the mean $\mathcal{B}$ and the belief management problem of finding $\mathcal{B}$.

The economic interpretation of the optimization problem is straight forward. Mimicking behaviour has two effects on the continuation game: The deviator (i) inherits the posterior distribution over the opponents types from the mimicked, 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 adopt the strategy of the mimicked, but can freely choose her behaviour. Any such choice remains unresponded by the opponent who plays

[12] The procedure is an extension to that described in Bergemann and Morris (2016a,b). We treat each $\mathcal{B}$ as an initial information structure in the sense of Bergemann and Morris (2016a) and look for the optimal (public) Bayes correlated equilibrium given our objective. In the second step we then solve (P_{max}) with respect to $\mathcal{B}$ taking the solution to the information design problem as given.

[13] Note that $\mathcal{B}$ itself does not need to be a consistent belief system per se.

[14] In appendix A we show that the concavification approach is valid on the level of the Lagrangian functional which automatically takes all constraints into account.

as-if she is in the on-path game. The information advantage of the deviator reduces the discrimination of the continuation game and makes it particularly attractive to mimic seldom types. To account for that, seldom types receive a large ability premium and the choice of ability premium depends on the prior distribution. It is intuitive that the welfare motive is independent of the prior. It considers only on-path behaviour and is thus an aggregate measure of (in)efficiency post-escalation.

Theorem 2 provides several insights. First, although the designers is agnostic about the outcome once the conflict escalates, there is a direct incentive to reduce inefficiencies after escalation to save on resources for additional settlements. Further, optimal conflict management is entirely determined by the information structure in the continuation game.

Second, combining Theorems 1 and 2 caters to the understanding of the 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 specifies the role of information. It shows that the designer's choice set can be reduced to finding the right information structure post-escalation. This information structure is only based on the expected performance in the continuation game. In particular, it identifies, separates, and quantifies the screening motive and the welfare motive of the designer providing an intuitive and tractable notion of both motives.

Finally, Theorem 2 provides a solution algorithm to the general problem. First, we solve an information design problem à la Bergemann and Morris (2016a) for each admissible $\mathcal{B}$. Using this solution we can then find the optimal choice of $\mathcal{B}$. Our approach allows to solve all steps in the space of belief systems using the same objective. Further corollary 2 provides a sufficient condition ont the level of binding constraints that determines when the subsequent information design problem is redundant.

4 Examples

In this section we apply our results to two different underlying games. An exogenous simple lottery á la Hörner, Morelli, and Squintani (2015) and an all-pay-auction. The two examples highlight the link of the optimal mechanism to the underlying game. In each, we can directly identify the channel driving the results.

The first example focuses on the distribution channel. We show that games similar to that used in Hörner, Morelli, and Squintani (2015) result in a *sorting mechanism.* That is, conflict management identifies "easy to settle" matches and guarantees settlement for these matches while other types are referred back to the conflict game.

Contrasting this finding, the second example focuses on the informational advantage of a deviator. We show that if the conflict game is an all-pay-auction no sorting as in the lottery case takes places. Instead, the optimal mechanism is always *type independent.* That is, conflict management ensures that each type holds the same conditional distribution over the opponents types. Without type independence the mechanism would always provide an incentive to deviate to "steal" another types individual belief and use the information advantage in the off-path continuation-game. While beliefs are the same across types for a given player and, they differ across players for a given type.

4.1 Simple Lotteries

Consider a generalization of the default game in Hörner, Morelli, and Squintani (2015) which we call simple lotteries.

Definition 7 (Simple Lottery)**.** A game Γ is called a *simple lottery* if the von-Neuman Morgenstern utility of a match is independent of the belief system, that is $u_i(\theta, \theta', \mathcal{B}) = u_i(\theta, \theta')$.

Lotteries are in particular relevant for "last-minute" conflict management, that is settlement negotiations at the day of the trial, or peace negotiations at the verge of war. In such situations, the action choice post-escalation is limited and typically involves a dominant strategy such as showing all evidence collected or unleashing the troops. Thus, the expected outcome depends solely on the individual belief about the opponents type rathher than the entire belief system.

The expected utility for any belief system $\mathcal{B}(\sigma)$ is

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

The utility depends only on the individual belief and is entirely linear. Thus, we can abstract from public signals, and denote expected welfare as

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

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

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

with match ability premium and match aggregate utility[15]

$$A(\theta,k)=u(\theta,k)-u(\theta{+}1,k) \text{ and } W(\theta,k)=u(\theta,k)+u(k,\theta), \tag{5}$$

respectively. It includes the model of Hörner, Morelli, and Squintani (2015) as a special case with constants $W(\theta,k)=\bar{W}$ and $A(\theta,k)=\bar{A}$ for any type and player. Thus, the welfare motive is shut down and escalation always leads to the same aggregate payoff. In general, linearity directly implies that the optimal solution is entirely symmetric. Further, given Assumption 3 we maximize equation (4) subject to

$$\alpha R\left(\underbrace{\sum_i E[\Psi_i|\Gamma]+E[U_i|\Gamma]}_{=(4)}-1\right)\geq 2v(1)-1,\ \text{with } R=\frac{\rho^0(1)^2}{\rho(1,1)}, \tag{AF}$$

and

$$\sum_k^K\Big(\rho^0(2)\rho(1,k)-\rho^0(1)\rho(2,k)\Big)\,A(1,k)\geq 0. \tag{IC$_F$}$$

We assume that $\widetilde{VV}(\theta,k)$ is weakly decreasing.

Proposition 2. *Suppose Assumptions 1 to 3 hold and the default game is a simple lottery. The optimal reduced-form conflict management has the following characteristics*

- *two weak types settle for sure,*
- *two strong types escalate with positive probability,*
- *signals never improve on the "no-signals" solution.*

Moreover, there exists a $\bar{\rho}\in(0,1)$, such that whenever the prior $\rho^0(1)\leq\bar{\rho}$ only the two strongest types meet post-escalation.

The optimal solution in lotteries requires the mechanism to "sort" matches. In particular the mechanism can guarantee mutual weak matches settlement while mutual strong matches escalate with high probability. The consequence is that beliefs in case of escalation are generically not the same for each type. While the weakest type only expects to compete with stronger types if the conflict escalates, stronger types may meet the weakest even after escalation.

The main reason for this result is that higher order beliefs do not affect the expected outcome as players actions in the continuation game are independent of the prospects from it. Thus a potential deviator has no informational advantage, but at most distributional

[15] We assume that $A(K,k)$ is some finite, positive real number to avoid case distinction.

one. Using our framework, the linearity of the problem immediately identifies a "virtual valuation of matches" for the designer and a linear program for maximization. It also shows that the optimal mechanism in Hörner, Morelli, and Squintani (2015) is robust and depends mainly on the assumption that the default game is indeed a simple lottery.[16]

4.2 All-Pay Auction

Consider an all-pay auction with binary private information. The prize is normalized to 1 and players have constant marginal bidding cost $c_i \in \{1, \kappa\}$ with $\kappa > 2$. The probability that a player has type $c_i = 1$ is given by ρ^0. To simplify assume $\rho^0 := \rho^0(1) = \delta\bar{\rho}$ with $\bar{\rho} = (\kappa - 2)/(2\kappa - 2)$ and $\delta \in [0.7, 1]$.[17]

All-pay auctions have been frequently used to model situations of conflict, including legal disputes, international conflicts, strikes, and patent races.[18] In light of the equilibrium characterization 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 nature of the top-down algorithm in Siegel (2014) already suggests that not only the players individual belief, but the 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 default 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 each realization of the 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.*

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

[16] Proposition 2 nests Lemma 1 of Hörner, Morelli, and Squintani (2015) as a special case. If affordable, we put all mass on the highest $\widehat{VV}$ (part (4) in their Lemma 1). If we cannot gneerate enough ressources to support this program, we need to shift mass to asymmetric matches to ease (AF) (part (3) of that Lemma). In their setup however (IC_F) never binds, while this may be the case in the more general setup.

[17] The upper bound on ρ^0 guarantees that Assumption 2 is satisfied, while the lower bound is sufficient that the optimal belief system ignoring (AF) is implementable satisfying (AF). Cases in which the latter does not hold are discussed in Balzer and Schneider (2015b).

[18] See Konrad (2009) for a general discussion also on related games.

First, the result that all matches escalate with positive probability, is tightly connected with that of type-independent beliefs. The reason for this lies in the ability premium. Fix some signals and the ex-ante expected type distributions $\rho_1^e(\sigma)$ and $\rho_2^e(\sigma)$ after signal realization σ. The only variable left to determine $\mathcal{B}(\sigma)$ according to Lemma 4 is the individual belief of player $c_1 = 1$, that is $\tilde{\beta}(\sigma) := \beta_1(1|1,\sigma)$.

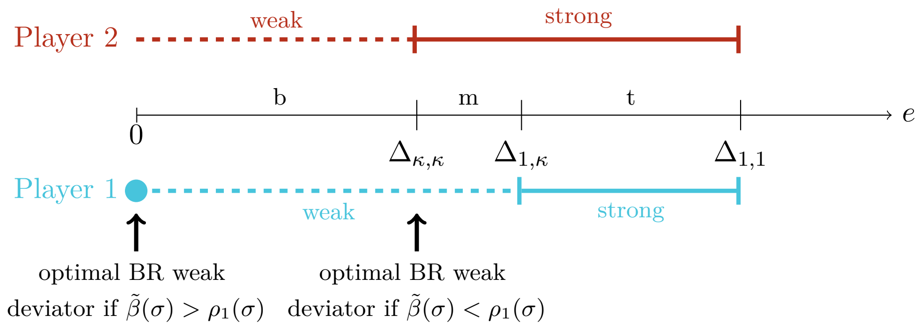

Figure 2: Equilibrium strategies in the all-pay auction assuming $\rho_1 > \rho_2$. Bold lines depict the suport for the storng player, dashed lines that of the weak. The dot indicates a mass point.

Changing $\tilde{\beta}(\sigma)$ has two effects on the ability premium. It reduces the stronger players expected utility in the continuation-game by monotonicity, but it also decreases the deviator's expected utility in the continuation-game as she, too, faces a stronger opponent. As depicted in 2 the deviator's best response changes at $\tilde{\beta}(\sigma) = \rho_1^e(\sigma)$. The reason is straight-forward. Whenever the deviator expects a stronger opponent than on the equilibrium path, that is $\tilde{\beta}(\sigma) > \rho_1^e(\sigma)$, her marginal gain from increasing her bid is reduced and vice versa. If $\tilde{\beta}(\sigma) = \rho_1^e(\sigma)$ she is indifferent as on the equilibrium path. This change in deviation strategies, provides a kink in the expected utility from the continuation game after deviation.

In the all-pay auction this kink is so strong that independent of the choice of $\rho_1^e(\sigma), \rho_2^e(\sigma)$, the ability premium increases for $\tilde{\beta}(\sigma) < \rho_1^e(\sigma)$ as the negative effect on $U_i(1,\kappa,\mathcal{B})$ is stronger than that on $U_i(1,1,\mathcal{B})$. The ability premium decreases for $\tilde{\beta}(\sigma) > \rho_1^e(\sigma)$ as the negative effect on $U_i(1,\kappa,\mathcal{B})$ is weaker than that on $U_i(1,1,\mathcal{B})$. Further, straightforward calculations show, that this property also survives if we include the welfare motive. Thus type-independency, $\tilde{\beta}(\sigma) = \rho_1^e(\sigma)$ is always desired.[19]

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 in our model and thus asymmetry decreases inefficiency and may thus be helpful. However, increased asymmetry comes at a cost of reducing

[19] 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 $\rho_i^e(\sigma)$ providing us immediately with the optimality of a corner solution. This emphasizes the tractability gained by theorems 1 and 2. Details are found in the appendix.

the ability premium. At the optimum the level of assymetry is determinde by that tradeoff. We find that asymmetry increases with the prior ρ^0 which reduces the (relative) weight of the ability premium compared to aggregate welfare.[20]

Finally, since the condition (GI) is always satisfied, signals can only improve if the strong type's (IC) holds with equality. None of the two binds with equality for any $\rho^0 < 1/3$, and they never bind simultaneously. 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 as in the case ignoring the strong type's (IC).

5 Discussion

In this section we relax several assumptions to address robustness of our findings.

Arbitrary sets of constraints binding. Using Assumption 3 we have focused on monotone, well-behaved problems. We can eliminate Assumption 3 by redefining the ability premium. Lemma 1, remains valid and the upward adjacent incentive constraint for some of the types may bind. In this case the ability premium is not only an upward measure, but instead involve a downward component, too. Moreover, more than one type could have a binding participation constraint. We can use an auxiliary linear minimization problem to determine the set of binding constraints for each consistent $\mathcal{B}$. This translates into step functions for the ability premium and the resource constraint. The ability premium for any type θ_i is given as

$$\widehat{\psi}_i(\theta_i) = \begin{cases} \widehat{U}(\theta_i, \theta_i, \mathcal{B}) - \widehat{U}(\theta_i, \theta_i{+}1, \mathcal{B}) & \text{if } (IC^{+}_{\theta_i+1}) \text{ binds,} \\ \widehat{U}(\theta_i, \theta_i, \mathcal{B}) - \widehat{U}(\theta_i{-}1, \theta_i, \mathcal{B}) & \text{if } (IC^{-}_{\theta_i}) \text{ binds,} \\ 0 & \text{else.} \end{cases}$$

The constant component in the resource constraint (AF) consists of $\sum_i \sum_{\theta \in \mathcal{A}(\mathcal{B})} f_\theta(\rho^0, \mathcal{B}) v_i(\theta)$ with $\mathcal{A}(\mathcal{B}) = \{\theta | \text{ (PC)}_\theta \text{ holds with equality given } \mathcal{B}\}$, and a set of functions f_θ s.t. $\sum_{\theta \in \mathcal{A}(\mathcal{B})} f_\theta(\rho^0, \mathcal{B}) = 1$. The set of binding constraints is a function of the belief system and we augment the problem accordingly. Everything else remains as in section 3. An approach to determine the set of binding constraints is discussed in Appendix A.

Non Convex Veto-Values. If the value of vetoing is not convex with respect to the prior given the veto off-path belief, we cannot guarantee that full-participation is optimal.[21] However, if we augment the nature of the conflict management game slightly, we can restore full participation. Assume the following changes to the model: Instead of sequentially

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

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

ratifying conflict management and then communicating their type, assume players do this simultaneously. Assume further that the mechanism has the possibility to send a public signal about the participating player and even a vetoing player cannot ex-ante commit to ignore this information. Then, using the tools from Kamenica and Gentzkow (2011) the designer can always find a set of signals such that the value of vetoing is convex with respect to the prior given the deviator expects to receive these signals in case of vetoing. See Balzer and Schneider (2015a) for details on this.

Private Signals sent by the Designer. In our main analysis we abstract from private signals sent by the designer. However, our observations made in the analysis of public signals shows how private signals could be added to the problem. We could merge the optimal private signals into the description of the default game (by a move of nature e.g.) such that $\widehat{U}$ constitutes the expected utility from participation before the private signals are realised. If the perturbed game is still an anonymous conflict, our analysis remains valid subject to optimal signalling thereafter. Thus, we provide a notion of the benefits a designer that can offer to circumvent the costly game has *on top* of choosing the optimal signal structure. Further we provide a way to separate the pure information design problem downstream in the default game and the hybrid mechanism. Unfortunately, we cannot provide a sensible, yet general characterization when such signals are redundant due to the complexity of communication equilibria.

Limited Commitment by the Players. If players cannot commit ex-ante to obey the mediators settlement decision, the design of the mechanism may change. If players learn enough from the settlement decision, they may expect an easy victory if they decide to enter the default game despite the decision of the mechanism. However, if such rejection cannot credibly be announced publicly, the mechanism has a set of tools to deter players from deviation. Observe that different to false reports a rejection of an offer is observed by the mechanism. We can thus, similar to Hörner, Morelli, and Squintani (2015), assume a different version of our model in which the mechanism triggers escalation by making an unacceptable settlement offer (e.g. 0) with a recommendation to reject to one of the players. If the mechanism communicates this information privately to the players, a deviating player cannot be sure that her deviation indeed triggered escalation. The mechanism can use this to make both players believe they play an on-path escalation game which makes them choose strategies accordingly and punishes the deviating player. Thus the idea of Hörner, Morelli, and Squintani (2015) carries over to some extend. A detailed example of such a situation is discussed in Balzer and Schneider (2015b).[22]

[22]This construction provides a particular example of private signals. In both examples in section 4 such a twist to the model solves the potential commitment problem (almost) entirely (cf. Balzer and Schneider (2015b) and Hörner, Morelli, and Squintani (2015)). This notion may, however, not be true in a more general class of games.

Different Objective. In our paper we assume that conflict management aims at minimizing escalation. While we belief this is the most sensible assumption in our context, we could extend our analysis to that of minimizing (aggregate) inefficiencies from the viewpoint of the parties. While the main motives for the designer remain, the objective of Theorem 2 is no longer valid. In particular the welfare motive receives larger weight. Notice that given settlement resources are scarce (condition (AF) is binding), any settlement solution is by definition the least inefficient outcome. However, the designer in the baseline model gives equal weight to any failed settlement attempt and cares for efficiency post-escalation only to increase resources. If we were to change our objective to a utilitarian maximization over parties ex-ante expected utility, the welfare component would receive higher weight, increasing the pressure on the designer to decrease inefficiencies there. Generically this would not give the same solution as the one carried out here, but the designer would sacrifice some efficient settlements to provide a less inefficient escalation game in many cases. Reformulating the (maximization) objective accordingly, give us a gradient that weighs the two components asymmetrically.[23]

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

Transfers and Correlation. So far, we assumed that utility is not directly transferable. While, the settlement value serves as a numeraire good in our analysis there are two additional constraints necessary to guarantee implementation. The first is that z_i can never be negative. The second is captured by (GI) since we need to find a sharing rule X that implements a particular z_i. If we instead allow utilities to be directly transferable we can ignore these additional constraints. Thus the reduced form mechanism is always implementable and z_i can take negative values.

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. In the baseline model this exploitation is however limited through the the (GI) constraints. Allowing for unlimited transfers (and giving up the ex-post budget constraint), however, would allow a first-best solution á la Crémer and McLean (1988).

[23]Note that in the case of 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.

6 Conclusion

The main contribution of this paper is to provide a tractable approach to optimal conflict management for a large class of Bayesian games. We offer an economically intuitive dual to the problem that directly links properties of the underlying game to the optimal mechanism. We show that optimal conflict management is completely characterized by the optimal belief system in the event of escalation. We postulate two measures on the continuation-game post-escalation. A measure of discrimination and one of aggregate welfare. We show that the optimal mechanism can be derived by simply maximizing the sum of the two.

We use our general results to characterize the optimal solution in two examples. The first are default games in the form of simple lotteries as often used in the literature. We show that optimal conflict management reduces to a sorting mechanism for the class of simple lotteries. The mechanism escalates conflicts with the highest virtual valuations only as players' continuation-strategies are invariant to changes in the information structure.

As a second example we study the all-pay auction. Contrary to lotteries, strategies depend heavily on the belief system in the all-pay auction. We show that in this case, there is an incentive to misreport during conflict management to extract information from the mechanism. The main channel is the informational advantage a deviator receives by deviating during conflict management and using the undetected deviation to adjust behaviour in the continuation-game without fearing a response. We show that the optimal solution always prevents such deviations by offering type independent beliefs such that the player receives the same information about her opponent independent of her behaviour in the mechanism.

Our results suggest a number of directions for future research. The most obvious lies in the application to specific conflict management problems. As argued in the introduction most of this literature focuses on lotteries mainly for reasons of tractability. Our approach facilitates inclusion of more complicated Bayesian underlying games and thus a better understanding of optimal conflict management. As the optimal solution is characterized by the post-escalation continuation game only, our results provide an approach to link observations on post-escalation behaviour to the quality of the mechanism. A second road is to allow for a more dynamic game. In particular in early stage conflict management players may expect exogenous news arrival in case conflict management fails. Then, the question arises whether there is potential for sequential conflict management that takes this news evolution into account. A different trajectory when moving to a more dynamic game would be to consider repeated interaction between players with (partially) persistent types. That is, players can use some conflicts to learn something about their opponent which may be helpful for future conflicts. We are confident that our simple approach provides a valid starting point and benchmark for such complex models.

A General Problem

In this part of the appendix we construct a solution algorithm to general case. Using this general approach directly provides a proof for Theorem 2.

A.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 8. *The Lagrangian approach yields the global optimum.*

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

Define

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

with $G(\cdot)$ being the set of inequality constraints and $Pr(\Gamma)(cs)$ being a function from the choice variable in the probability of escalation.

Observe that $\mathrm{w}(t)$ is convex in t: Assume 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 $\alpha \mathrm{w}(t_1) + (1-\alpha)\mathrm{w}(t_2) < \mathrm{w}(t_0)$. Denote by cs_j the optimal solution, so that $Pr(\Gamma)(cs_j) = \mathrm{w}(t_j)$.

Then, consider the choice cs_0 so 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 $\gamma^{\sigma_j}(k_1,k_2) = \gamma_j(k_1,k_2)$.

By construction the constraints are satisfied and and the solution value is equals to the convex combination, that is,

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

Hence, $\mathrm{w}(t)$ is convex. By Luenberger, 1969 chapter 8 the Lagrangian approach finds a global minimum.

□

Recall that we minimize $Pr(\Gamma)$ subject to the following constraints, which are required to hold for all i and all k:

$$-(z_i(k) - z_{-i}(k-1)) - y_i(k,k) + y_i(k-1,k) \leq 0 \qquad (IC_i^-)$$

$$-(z_i(k) - z_{-i}(k-1)) - y_i(k,k) + y_i(k-1,k) \leq 0 \qquad (IC_i^+)$$

$$-z_i(k) - y_i(k,k) + \upsilon_i(k) \leq 0 \qquad (PC_i)$$

$$-1 + \sum_i \sum_k \rho^0(k) z_i(k) + Pr(\Gamma) \leq 0 \qquad (AF)$$

$$z_i(k) + \gamma_i(k) - 1 \leq 0 \qquad (IF_i)$$

$$-z_i(k) \leq 0 \qquad (EPI)$$

$$\gamma(k_1, k_2) - 1 \leq 0 \qquad (F)$$

In the following we derive the Lagrangian representation of the maximization problem. As first step, we introduce the complementary slackness conditions and introduce the Lagrangian multiplier of the respective constraint:

$$[z_i(k) - z_{-i}(k-1) + y_i(k,k) - y_i(k-1,k)]v^i_{k,k-1} = 0, \qquad v^i_{k,k-1} \geq 0$$

$$[z_i(k) - z_{-i}(k-1) + y_i(k,k) - y_i(k-1,k)]v^i_{k-1,k} = 0, \qquad v^i_{(k-1,k)} \geq 0$$

$$[z_i(k) + y_i(k,k) - \upsilon_i(k)]\lambda^i_k, \qquad \lambda^i_k \geq 0$$

$$[1 - \sum_i \sum_k \rho^0(k) z_i(k) - Pr(\Gamma)]\delta = 0, \qquad \delta \geq 0$$

$$[z_i(k) + \gamma_i(k) - 1]\xi^i_k = 0, \qquad \xi^i_k \geq 0$$

$$\left[(1 - \gamma(k_1,k_2))\rho^0(k_1)\rho^0(k_2) - \sum_i \rho^0(k_i)(1 - \gamma_i(k_i) - z_i(k_i))\right]\eta_{k_1,k_2} = 0, \quad \eta_{k_1,k_2} \geq 0$$

$$z_i(k)\zeta^i_k = 0, \qquad \zeta^i_k \geq 0$$

$$[\gamma(k_1,k_2) - 1]\mu_{k_1,k_2} = 0, \qquad \mu_{k_1,k_2} \geq 0,$$

For any Lagrangian multiplier, say t, we introduce the following notation $\tilde{t} \equiv \frac{t}{\delta}$.

Let $e^i(k) := \rho^0(k) \sum_{k_{-i}} \tilde{\eta}_{k,k_{-i}}$ and define

$$\tilde{\Lambda}^i(k) := \sum_{v=1}^{k} \tilde{\lambda}^i_v, \quad \tilde{\Xi}(k) := \sum_{v=1}^{k} \tilde{\xi}_k, \quad \tilde{E}^i(k) := \sum_{v=1}^{k} \tilde{e}^i(v), \quad \tilde{Z}^i(k) := \sum_{v=1}^{k} \tilde{\zeta}^i_k$$

We next state a characterization of our solution in terms of a Lagrangian objective. Because of Theorem 1 and **??** we have some degree of freedom in the choice of the maximizer. We use $\boldsymbol{\rho}$, the complete collection of all consistent type-profile probabilities, post escalation. Let $\boldsymbol{\rho}(\sigma)$ be the collection of those probabilities, conditional on escalation and signal realization σ. Moreover, note that there is a bijection between $\rho(\sigma)$ and $\mathcal{B}(\sigma)$. Consider the symmetrized problem.

Lemma 9. *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*

$$\begin{aligned}
\hat{\mathcal{L}}(\mathcal{B}(\sigma) &:= \sum_i \Bigg[\sum_{k=1}^{K} \rho_i(k)(\frac{\mathsf{m^i_k}}{\rho^0(k)})U_i(k, \mathcal{B}(\sigma)) \\
&+ \sum_{k=1}^{K-1} \frac{\mathsf{M^i(k)} - \tilde{v}^i_{k,k+1}}{\rho^0(k)} \rho_i(k)(U_i(k, k, \mathcal{B}(\sigma)) - U_i(k, k+1, \mathcal{B}(\sigma))) \\
&- \sum_{k=1}^{K} \frac{\mathsf{M^i(k)} - \tilde{v}^i_{k+1,k}}{\rho^0(k+1)} \rho_i(k+1)[U_i(k+1, k, \mathcal{B}(\sigma)) - U_i(k+1, k+1, \mathcal{B}(\sigma))]\Bigg] \\
&+ \mathcal{T}(\mathcal{B}(\sigma)),
\end{aligned} \tag{6}$$

where $\mathsf{m^i_k} := \rho^0(k) + \tilde{\xi}^i_k + \tilde{e}^i_k - \tilde{\zeta}^i_k$, *and* $\mathsf{M^i(k)} := \tilde{\Lambda}^i(k) - \sum_{v=1}^{v=k} \rho^0(v) - \tilde{\Xi}^i(k) - \tilde{E}^i(k) + Z^i(k)$, *and*

$$\begin{aligned}
\mathcal{T}(\mathcal{B}(\sigma)) &:= \\
&- \sum_i [\sum_{k=1}^{K} \frac{\rho_i(k,\sigma)}{\rho^0(k)} \tilde{\xi}^i_k] - \sum_{k_1 \times k_2} [-\rho(k_1)\beta_1(k_2|k_1,\sigma) + \rho_1(k_1,\sigma) + \rho_2(k_2,\sigma)]\tilde{\eta}_{k_1,k_2} \\
&- \sum_{k_1 \times k_2} \frac{\rho_1(k_1,\sigma)\beta_1(k_2|k_1,\sigma)}{\rho^0(k_1)\rho^0(k_2)} \tilde{\mu}_{k_1,k_2}
\end{aligned} \tag{7}$$

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 (AF) constraint is always binding, i.e., $\delta > 0$.*
- *If the Border constraints are redundant, then $\tilde{\xi}^i_k = \tilde{e}^i_k = \tilde{E}_i(k) = \tilde{\Xi}^i(k) = 0$*
- *If $\tilde{\Lambda}^i(k) + \tilde{Z}^i(k) - \sum_{v=1}^{v=k} \rho^0(v) - \tilde{\Xi}^i(k) - \tilde{E}^i(k) > 0$, then the downward adjacent incentive constraints are binding. If in this case the upward adjacent incentive constraints are redundant $\tilde{v}^i_{k,k+1} = 0$.*
- *$\tilde{\Lambda}^i(k) + \tilde{Z}^i(k) - \sum_{v=1}^{v=k} \rho^0(v) - \tilde{\Xi}^i(k) - \tilde{E}^i(k) < 0$, the upward adjacent incentive constraints are binding. If in this case the downward adjacent incentive constraints are redundant $\tilde{v}^i_{k+1,k} = 0$.*

Proof. The proof consists of two steps. First, we state and manipulate the Lagrangian objective, $\mathcal{L}$, until we arrive at a more tractable maximization problem. Second, we verify that we are at the concave hull of the objective.

$\mathcal{L}$ takes the following form:

$$\begin{aligned}
\mathcal{L} = &Pr(\Gamma) + \delta[-1 + \sum_i \sum_{k=0}^{K} \rho^0(k) z_i(k) + Pr(\Gamma)] \\
&+ \sum_i \sum_{k=1}^{K} [-z_i(k) - y_i(k,k) + v_i(k)]\lambda^i_k \\
&+ \sum_i \sum_{k=1}^{K-1} [z_i(k-1) - z_i(k) - y_i(k,k) + y_i(k-1,k)] v^i_{k,k-1} \\
&+ \sum_i \sum_{k=2}^{K} [z_i(k) - z_i(k-1) - y_i(k-1,k-1) + y_i(k,k-1)] v^i_{k-1,k} \\
&+ \sum_i \sum_{k=1}^{K} [z_i(k) + \gamma_i(k) - 1]\xi^i_k \\
&+ \sum_{k_1 \times k_2} \left[(1-\gamma(k_1,k_2))\rho^0(k_1)\rho^0(k_2) - \sum_i \rho^0(k_i)(1-\gamma_i(k_i) - z_i(k_i))\right] \eta_{k_1,k_2} \\
&+ \sum_{k_1 \times k_2} [\gamma(k_1,k_2) - 1]\mu_{k_1,k_2} - \sum_i \sum_k z_i(k)\zeta^i_k
\end{aligned} \tag{8}$$

Using Lemma 3 we optimize with respect to $\{z_i(\cdot), \gamma^\sigma(1,1), \boldsymbol{\rho}(\sigma)\}$, where $\gamma^\sigma(1,1) := Pr(\mathcal{G}, \sigma | 1, 1)$.

<u>Step 1: FOCs w.r.t. $z_i(\cdot)$</u>

We take the first order conditions w.r.t. $z_i(k)$:

Let $v^i_{K+1,K} := 0 =: v^i_{1,0}$, then

$$\rho^0(k)\delta - \lambda^i_k - v^i_{k,k-1} + v^i_{k-1,k} + v^i_{k+1,k} - v^i_{k,k+1} + \xi_k + \rho^0(k_i) \sum_{k_{-i}} \eta_{k_i,k_{-i}} - \zeta^i_k = 0 \tag{9}$$

Summing over the K conditions (those in (9)) yields

$$1 = \tilde{\Lambda}^i(K) - \tilde{\Xi}(K) - \tilde{E}^i(K) - \tilde{Z}^i(k) \tag{10}$$

Moreover, (9) is satisfied if and only if

$$\tilde{v}^i_{k-1,k} - \tilde{v}^i_{k,k-1} = \sum_{v=1}^{v=k-1} \rho^0(v) - \tilde{\Lambda}^i(k-1) + \tilde{\Xi}(k-1) + \tilde{E}^i(k-1) - Z^i(k-1) \tag{11}$$

<u>Step 2: Reformulation of the Lagrangian terms</u>

Given the above necessary conditions, we manipulate the Lagrangian objective in order to derive a more tractable maximization problem. As first step consider all terms that involve $z_i(\cdot)$. The first order conditions with respect to $z_i(\cdot)$, i.e., (9), imply that these terms cancel out.

In the next step, we apply algebra manipulation, use the first order conditions and Bayes' rule to show that (8) admits the following representation

$$\mathcal{L} = Pr(\Gamma)(1+\delta) - \delta(C) - \delta \sum_{\sigma} \Pr(\Gamma, \sigma)\hat{\mathcal{L}}(\mathcal{B}(\sigma), \tag{12}$$

where C is a constant that is independent of the choices,

$$C := 1 - \sum_{i}\sum_{k} \tilde{\lambda}_k \upsilon_i(k) - \sum_{k_1 \times k_2} \tilde{\mu}_{k_1,k_2} - \sum_{i} \tilde{\Xi}^i(K) + \sum_{k_1 \times k_2} [\rho^0(k_1)\rho^0(k_2) - \rho^0(k_1) - \rho^0(k_2)]\tilde{\eta}_{k_1,k_2} < 0$$

Define $\gamma^{\sigma}(\theta_1, \theta_2) := Pr(\mathcal{G}, \sigma|\theta_1, \theta_2)$. It follows from lemma 3 that $\gamma^{\sigma}(\theta_1, \theta_2)$ obeys the following representation: $\gamma^{\sigma}(\theta_1, \theta_2) = f(\mathcal{B}(\sigma), \theta_1, \theta_2)\gamma^{\sigma}(1, 1)$, where $f(\mathcal{B}(\sigma), \theta_1, \theta_2)$ is a positive real number. Thus, $Pr(\mathcal{G}, \sigma) = [\sum_{\theta_1,\theta_2} \rho^0(\theta_1)\rho^0(\theta_2)f(\mathcal{B}(\sigma), \theta_1, \theta_2)]\gamma^{\sigma}(1, 1)$.

Define $R(\mathcal{B}(\sigma)) := [\sum_{\theta_1,\theta_2} \rho^0(\theta_1)\rho^0(\theta_2)f(\mathcal{B}(\sigma), \theta_1, \theta_2)]$, and note that (12) becomes

$$\mathcal{L} = \sum_{\sigma} \gamma^{\sigma}(1,1)R(\mathcal{B}(\sigma))(1+\delta) - \delta C - \delta \sum_{\sigma} \gamma^{\sigma}(1,1)R(\mathcal{B}(\sigma))\hat{\mathcal{L}}(\mathcal{B}(\sigma)) \tag{13}$$

The derivative of (13) w.r.t. $\gamma^{\sigma}(1, 1)$ satisfies

$$(1+\delta)R(\mathcal{B}(\sigma)) - \delta R(\mathcal{B}(\sigma))\hat{\mathcal{L}}(\mathcal{B}(\sigma)) = 0, \tag{14}$$

for each signal. Therefore, (i) every signal must give rise to the same value, (ii) $\delta = \frac{1}{\hat{\mathcal{L}}(\mathcal{B}(\sigma))-1}$ and (iii) the Lagrangian objective becomes:

$$\frac{-C}{\hat{\mathcal{L}}(\mathcal{B}(\sigma)) - 1} \tag{15}$$

Because every signal must give rise to the same value, δ, every $\mathcal{B}(\sigma)$ must maximize $\hat{\mathcal{L}}$.

Therefore, for the optimal multipliers one constructs the concave hull of $\hat{\mathcal{L}}$ by taking spreads over those $\mathcal{B}(\sigma)$ that are a global maximum of $\hat{\mathcal{L}}$. If there are multipliers and a unique maximizer, $\mathcal{B}(\sigma)$,that satisfies the complementary slackness conditions, then signals do not improve. Conversely, if there are multiple global optima and there is a spread that satisfies the complementary slackness conditions, signals do improve. □

A.2 Proof of Proposition 1

Proof. In Appendix A we argued that whenever the designer has access to signals, i.e., he 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. Here, we assume that 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 and (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). Honoring these differences, we still can use the definition of the Lagrangian function established in the proof of lemma 9. Doing so and using complementary slackness, i.e., dropping redundant constraint identified in lemma 7, the form of the maximization problem stated in Proposition 1 follows. □

A.3 Proof of Theorem 2

Proof. Proposition 1, shows that maximizing equation (P_{max}) 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 selects a state $\mathcal{B}(\sigma)$ and communicates this to the players. Then players play Γ 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 $\widehat{U}(m, \theta, \Sigma(\mathcal{B}))$ satisfies the constraints it is implementable and has the expected belief system $\widehat{\mathcal{B}}$. Further it leads to some lottery over both the expected ability premium and the expected welfare which we denote by $\mathbb{E}[\widehat{\Psi}|\mathcal{G}] := Pr(\sigma)\mathbb{E}[\Psi|\mathcal{G}, \sigma]$ and $\mathbb{E}[\widehat{U}|\mathcal{G}] := Pr(\sigma)\mathbb{E}[U|\mathcal{G}, \sigma]$. Each $\widehat{\mathcal{B}}$ may have many possible lotteries that support it and are feasible. We select the maximum among them. The $\widehat{\mathcal{B}}$ with the highest value $\mathcal{V}$ also solves the problem of lemma 9. □

B Conflict Management and Belief Management

B.1 Conflict Management and Belief Management

Fix am interior belief-system. We first derive a representation that links the belief system to the underlying feasible 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 through a *transition* function in the following way:

Observation 2. Consider the type profiles (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,

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

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

Proof. Applying Bayes' rule it follows that

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

□

Fix two nodes in the network, say (θ_1,θ_2) and (k_1,k_2). There are many paths that connect the two nodes. For example, starting from (θ_1,θ_2) we can go to (k_1,θ_2) and then to (θ_1,θ_2). Or, starting from (k_1,k_2) we can approach (θ_1,θ_2) through (θ_1,k_2). Bayes' rule implies that both paths give rise to the same length, or, the resultant values of the nodes are the same. That is, 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 8 (Bayes' consistency)**.** A belief-system is Bayes' consistent if for every (θ_1,θ_2) and (k_1,k_2) the following holds

$$q_1(k_2,\theta_2|k_1)q_2(k_1,\theta_1|\theta_2) = q_2(k_1,\theta_1|k_2)q_1(k_2,\theta_2|\theta_1) \tag{16}$$

A more tractable condition follows if we impose a more structure on the way we move in the network. Assume $\gamma(1,1)>0$, and consider $(1,1)$ as the origin of the network. Any node (θ_i,θ_{-i}) is reached using a path that starts at that origin. We can restate Definition 8.

Definition 9 ($(1,1)$-Consistent)**.** A belief-system is $(1,1)$-consistent if and only if for every (θ_1,θ_2) the following is true:

$$q_1(\theta_2,1|\theta_1)q_2(\theta_1,1|1) = q_2(\theta_1,1|\theta_2)q_1(\theta_2,1|1) \tag{17}$$

Lemma 10. *A belief-system with $\beta_1(1|1)>0$ is Bayes' consistent if and only if it is $(1,1)$-consistent.*

Proof. The proof considers interior belief-systems only to shorten notation. The same arguments work without that notion. Take any $\gamma(k_1,k_2)$ and $\gamma(\theta_1,\theta_2)$. We want to show that consistency implies

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

that is,

$$\frac{\beta_1(k_2|k_1)}{\beta_1(\theta_2|k_1)}\frac{\beta_2(k_1|\theta_2)}{\beta_2(\theta_1|\theta_2)} = \frac{\beta_2(k_1|k_2)}{\beta_2(\theta_1|k_2)}\frac{\beta_1(k_2|\theta_1)}{\beta_1(\theta_2|\theta_1)},$$

or

$$\frac{\beta_1(k_2|k_1)}{\beta_2(k_1|k_2)}\frac{\beta_1(\theta_2|\theta_1)}{\beta_2(\theta_1|\theta_2)} = \frac{\beta_1(\theta_2|k_1)}{\beta_2(k_1|\theta_2)}\frac{\beta_1(k_2|\theta_1)}{\beta_2(\theta_1|k_2)}. \tag{18}$$

Observe that consistency for $\gamma(k_1, \theta_2)$ implies

$$q_1(\theta_2, 1|k_1)q_2(k_1, 1|1) = q_2(k_1, 1|\theta_2)q_1(\theta_2, 1|1),$$

i.e.,

$$\frac{\beta_1(\theta_2|k_1)}{\beta_2(k_1|\theta_2)} = \frac{\beta_1(1|k_1)}{\beta_2(1|\theta_2)}\frac{\beta_1(\theta_2|1)}{\beta_1(1|1)}\frac{\beta_2(1|1)}{\beta_2(k_1|1)}. \tag{19}$$

Consistency for $\gamma(\theta_1, k_2)$ implies

$$q_1(k_2, 1|\theta_1)q_2(\theta_1, 1|1) = q_2(\theta_1, 1|k_2)q_1(k_2, 1|1),$$

i.e.,

$$\frac{\beta_1(k_2|\theta_1)}{\beta_2(\theta_1|k_2)} = \frac{\beta_1(1|\theta_1)}{\beta_2(1|k_2)}\frac{\beta_1(k_2|1)}{\beta_1(1|1)}\frac{\beta_2(1|1)}{\beta_2(\theta_1|1)}. \tag{20}$$

Consistency for $\gamma(\theta_1, \theta_2)$ implies

$$q_1(\theta_2, 1|\theta_1)q_2(\theta_1, 1|1) = q_2(\theta_1, 1|\theta_2)q_1(\theta_2, 1|1),$$

i.e.,

$$\frac{\beta_1(\theta_2|\theta_1)}{\beta_2(\theta_1|\theta_2)} = \frac{\beta_1(1|\theta_1)}{\beta_2(1|\theta_2)}\frac{\beta_1(\theta_2|1)}{\beta_1(1|1)}\frac{\beta_2(1|1)}{\beta_2(\theta_1|1)}. \tag{21}$$

Consistency for $\gamma(k_1, k_2)$ implies

$$q_1(k_2, 1|k_1)q_2(k_1, 1|1) = q_2(k_1, 1|k_2)q_1(k_2, 1|1),$$

i.e.,

$$\frac{\beta_1(k_2|k_1)}{\beta_2(k_1|k_2)} = \frac{\beta_1(1|k_1)}{\beta_2(1|k_2)}\frac{\beta_1(k_2|1)}{\beta_1(1|1)}\frac{\beta_2(1|1)}{\beta_2(k_1|1)}. \tag{22}$$

Equation (18) is satisfied because the right-hand side of Equation (22) times the right-hand side of Equation (21) is equal to the right-hand side of equation Equation (19) times the right-hand side of Equation (20).

□

By Lemma 10 a consistent belief system can be implemented by some escalation rule if and only if the escalation rule is feasible, i.e., every node (θ_1, θ_2) has a value weakly below 1. Using our network interpretation, the value of a node is given by the length of the path connecting the node with the origin, i.e.,

$$\gamma(\theta_1, \theta_2) = q_1(\theta_2, 1|\theta_1)q_2(\theta_1, 1|1)\gamma(1, 1) \tag{23}$$

Definition 10. A consistent belief system is feasible if there is an escalation rule that implements it.

Lemma 11. *A consistent belief-system with* $\beta_1(1|1) > 0$ *is feasibly if and only if there exists* $\gamma(1, 1) \in (0, 1)$ *such that for all* (θ_1, θ_2) *the following is true:*

$$\frac{\rho^0(1)}{\rho^0(\theta_2)}\frac{\rho^0(1)}{\rho^0(\theta_1)}\beta_1(\theta_2|\theta_1)\beta_2(\theta_1|1)\gamma(1, 1) \leq \beta_1(1|\theta_1)\beta_2(1|1). \tag{24}$$

Proof. Equation (23) implies that feasibility is satisfied if and only if

$$\frac{\rho^0(1)}{\rho^0(\theta_2)}\frac{\beta_1(\theta_2|\theta_1)}{\beta_1(1|\theta_1)}\frac{\rho^0(1)}{\rho^0(\theta_1)}\frac{\beta_2(\theta_1|1)}{\beta_2(1|1)}\gamma(1, 1) \leq 1,$$

which in turn implies Equation (24). □

Observation 3. A interior belief-system can be implemented by some feasible escalation rule γ if and only if it is *implementable*, that is, (i) feasible and (ii) consistent.

Proof. The only if part follows trivially from the above exposition.

Suppose that an interior belief-system satisfies (i) and (ii). Consider the following set of equations:

$$\gamma_1(\theta_2)\beta_1(\theta_2|\theta_1) = \rho^0(\theta_2)\gamma(\theta_1, \theta_2)$$

$$\gamma(\theta_1, \theta_2) = \frac{\beta_1(\theta_2|\theta_1)}{\rho^0(\theta_2)(1 - \beta_2(\theta_2|\theta_1))}\sum_{k \neq \theta_2}\gamma(\theta_1, k)\rho^0(k) \tag{25}$$

This defines a system of equations.

We show that a consistent belief system satisfies these equations. By definition we know that a consistent belief-system satisfies Equation (23). Substituting into Equation (25) we get

$$q_1(\theta_2,1|\theta_1)q_2(\theta_1,1|1)=\frac{\beta_1(\theta_2|\theta_1)}{\rho^0(\theta_2)(1-\beta_1(\theta_2|\theta_1))}\sum_{k\neq\theta_2}q_1(k,1|\theta_1)q_2(\theta_1,1|1)\rho^0(k),$$

which can be simplified to

$$q_1(\theta_2,1|\theta_1)=\frac{\beta_1(\theta_2|\theta_1)}{\rho^0(\theta_2)(1-\beta_1(\theta_2|\theta_1))}\sum_{k\neq\theta_2}q_1(k,1|\theta_1)\rho^0(k) \tag{26}$$

Using the definition of $q_1(\cdot,\cdot|\theta_1)$ Equation (26) becomes

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

Using algebra we see that

$$1=\frac{\sum_{k\neq\theta_2}\beta_1(k|\theta_1)}{1-\beta_1(\theta_2|\theta_1)}, \tag{27}$$

which is correct because $\beta_1(\cdot|\theta_1)$ is a belief. Therefore, a consistent belief system uniquely determines the escalation rule up to $\gamma(1,1)$. Because the belief-system is feasible, we can find a $\gamma(1,1)$ so that every $\gamma(\cdot,\cdot)$ is a real number between 0 and 1. □

Given Observation 3 we prove Lemma 3.

B.2 Proof of Lemma 3

Proof. Suppose there exist a implementable sequence $\mathcal{B}_n\to\mathcal{B}$. Because $\mathcal{B}_n$ is consistent, Observation 3 implies the existence of some continuous function, say $f:\mathcal{B}\to[0,1]^{K\times K}$, such that $f(\mathcal{B}_n)=\gamma_n$ with γ_n being feasible and implementing $\mathcal{B}_n$. Because $f(\cdot)$ is continuous it follows that $\lim_{n\to\infty}f(\mathcal{B}_n)=f(\lim_{n\to\infty}\mathcal{B}_n)=\gamma$. Moreover, $\mathcal{B}$ satisfies Equation (C): Equation (C) can be rewritten as

$$g_L(\mathcal{B})=g_R(\mathcal{B}), \tag{28}$$

where both $g_L(\cdot)$ and $g_R(\cdot)$ are continuous functions that map the belief system into a real number. Because $g_L(\mathcal{B}_n)-g_H(\mathcal{B}_n)=0$, we can conclude that $g_L(\mathcal{B})-g_R(\mathcal{B})=\lim_{n\to\infty}[g_L(\mathcal{B}_n)-g_R(\mathcal{B}_n)]=0$.

Conversely, take any $\mathcal{B}$ being implemented by some γ. We show that we can find a sequence of interior belief-systems that are implementable and converge to $\mathcal{B}$: Call γ^* the escalation rule that implements $\mathcal{B}$. Choose a sequence of escalation rules that lies in the interior and converges to γ^*. By Bayes' rule every element of the above sequence, say γ_n, implements 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 $f^{-1}(\gamma_n) = \mathcal{B}_n$. Because $\mathcal{B}_n$ is implemented by some γ_n Observation 3 implies that the system satisfies Equation (C). For the reverse, take any $\mathcal{B}$ such that $\beta_1(1|1) > 0$, fix $\gamma(1,1) = 1$ and use equation (23) to construct all other $\gamma(\theta_1, \theta_2) \neq \gamma(1,1)$. The function g can be obtained by setting the constructed $\gamma = g$. Finally the set of all α s.t. $g \leq 1^{K\times K}$ determines all escalation rules consistent with $\mathcal{B}$. For any $\mathcal{B}$ such that $\beta_1(\theta_2|\theta_2) > 0$ and perform the same steps fixing $\gamma(\theta_1, \theta_2) = 1$. □

B.3 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 compute

$$\rho_2(\theta_2) = \sum_{\theta_1 \in \Theta_1} \rho_1(\theta_1)\beta_1(\theta_2|\theta_1) = \sum_{\theta_1 \in \Theta_1} \rho(\theta_1, \theta_2), \tag{29}$$

for all θ_2.

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 (29). Hence, we have backed out the type-profile probability post-escalation which is sufficient for $\mathcal{B}$ through Bayes' rule. □

B.4 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 must satisfy (AF). Substituting the result from Corollary 1, and rearranging, (AF) takes the form

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

where $V_i(v(\Theta))$ 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 (AF') constant, while the RHS increases. By linearity of the RHS there exists an $\alpha^* \in (0,1]$ such that (AF') 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 (AF) holds with equality at the optimum. □

B.5 Proof of Theorem 1

Proof. The proof follows from Lemmas 3 to 6. □

C Remaining Proofs

Lemma 12. *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*

$$\sum_i \sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k) \left[\mathbb{1}_{PC}(\theta_i) \upsilon_i(\theta_i)\right] - 1 > 0. \tag{30}$$

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 $\tilde{\theta}_i \in \Theta_i^I(\theta_i)$t

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

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 (AF) is satisfied if

$$\sum_i \sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k) \left[\mathbb{1}_{PC}(\theta_i) \upsilon(\theta_i)\right] - 1 \leq h(\gamma). \tag{31}$$

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

C.1 Proof of Lemma 1

Proof. Monotonicity implies that local incentive compatibility constraints are sufficient. Second, assume there is a type for which both incentive constraints are redundant. In such a case it is always 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 $z_i = 0$.

We prove the final claim using lemma 12: There must be one type and player with (PC) binding. Otherwise the LHS of (31) is negative contradicting Lemma 12. Observe that 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. □

C.2 Proof of Lemma 7

Proof. Suppose by contradiction that $\Theta_i^B \neq \{1\}$. Consider the lower bound on the expected sum of shares of the optimal mechanism derived in lemma 12, term (31). This bound can be bounded by above the following way:

$$\sum_i \sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k) \left[\mathbb{1}_{PC}(\theta_i) \upsilon_i(\theta_i)\right] - 1 \leq \sum_i \sum_{\theta_i \in \Theta_i^B} \sum_{k \in \Theta_i^I(\theta_i)} \rho^0(k) \upsilon(\theta_i) - 1, \tag{32}$$

Assumption 3 implies that this upper bound is negative whenever $\Theta_i^B \neq \{1\}$.

Finally observe that because only the strongest types participation constraints are binding, 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 while if anything slacking any other constraint that is active. □

C.3 Proof of Lemma 2

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

C.4 Proof of Corollary 2

Proof. The solution to the problem is in the set of least constraint solutions. Thus, the constraints in $\mathcal{C}_F$ are redundant. Consider any $\mathcal{B}^*(\sigma^*)$ that solves equation (15). The degenerated signal structure, $\Sigma = \{\sigma^*\}$, satisfies all constraints by hypothesis. Hence, a non-degenerated signal structure cannot improve the solution value of the problem. $\square$

C.5 Proof of Proposition 2

Proof. The first claim is a direct consequence of (4) and monotonicity. Monotonicity ensures that $\widehat{VV} > 0$ such that the designer distributes all mass on matches (θ, k) such that $\theta, k < K$. Monotonicity in $\widehat{VV}$ guarantees positive mass on the type profile $(1, 1)$ and linearity of (4),(AF),(IC$_F$) ensures non-convexity at any solution making signals superfluous.

Finally if ρ^0 is small enough, both (AF) and (IC$_F$) are redundant. Thus, all mass is put on the strongest types and $\rho(1, 1) = 1$. $\square$

C.6 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 (2015b) we provide a more constructive version. We also omit the proof that the unique equilibrium in the all-pay auction given for the optimal belief system is monotonic. It is however, straight-forward to proof using the equilibrium descriptions of Rentschler and Turocy (2016) and Siegel (2014). We assume without loss of generality that $\rho_1 \geq \rho_2$ and proceed in several steps and make frequent use of the upper bound of interval b in Figure 2 above, $\Delta_{\kappa,\kappa}$. For large parts of the proof we consider the least constraint problem, invoke Proposition 1 such that signals are superfluous.

Part A (Linearity of the Objective).
Step 1: Linearity of individual beliefs. Given ρ_1, ρ_2 use Lemma 4 to describe each $\beta_i(\theta_{-i}|\beta_i) \in \mathcal{B}$ as a linear function of $\tilde{\beta}$.
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 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 2. 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 determinde by

$$\begin{aligned} F_{2_k}(\Delta_{\kappa,\kappa}) :&= \underbrace{F_{2_\kappa}(0)}_{\text{independent of } \tilde{\beta}} + \Delta_{\kappa,\kappa} \underbrace{\frac{\kappa}{\beta_1(\kappa|\kappa)}}_{\text{density of } 2_\kappa}. \\ &= F_{2_\kappa}(0) + \frac{1-\rho_2}{1-\rho_1}, \end{aligned} \tag{33}$$

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

Step 4: Linearity of equilibrium utilities. Given step 2 and 3 it is sufficient to show that quilibrium 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 and 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 which 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 $\tilde{\beta} = \rho_1 = (1+\rho^0)/2$ and $\rho_2 = (1-\rho^0)/2$.

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

$$\begin{aligned} \rho^0\left[E[\Psi(\gamma)|\mathcal{G}] + E[U|\mathcal{G}]\right]\Big|_{\tilde{\beta}<\rho_2} &= F_{2_\kappa}(\Delta_{\kappa,\kappa})\left(\beta_1(\kappa|1)\rho_2 + \rho^0(1-\rho_2)\right) \\ &\quad - \beta_2(\kappa|1)\left((1-\rho^0)\rho_2 F_{1_\kappa}(\Delta_{\kappa,\kappa})\right) \\ &\quad + \underbrace{\Delta_{\kappa,\kappa}}_{=\beta_2(\kappa|\kappa)/\kappa}\left((\rho_1+\rho_2)(\kappa-1) - (1+\rho_2)\kappa\rho^0\right). \end{aligned} \tag{34}$$

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

$$\lim_{\tilde{\beta}\to^{-}\rho_1} \left.\frac{\partial\rho^0\left[E[\Psi(\gamma)|\mathcal{G}]+E[U|\mathcal{G}]\right]}{\partial\tilde{\beta}}\right|_{\text{cand}} = \frac{\kappa(1-(\rho^0)^2)-(1-(\rho^0)^2)}{\kappa(1+\rho^0)}$$

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

$$\lim_{\tilde{\beta}\to^{+}\rho_1} \left.\frac{\partial\rho^0\left[E[\Psi(\gamma)|\mathcal{G}]+E[U|\mathcal{G}]\right]}{\partial\tilde{\beta}}\right|_{\text{cand}} = -\frac{\kappa(1-(\rho^0)^2)-1-(\rho^0)^2}{\kappa(1+\rho^0)},$$

Step 2: Type distribution. Taking the derivative of (34) with respect to ρ_i and evaluating at $\tilde{\beta} = \rho_2$ directly establishes the critical point irrespective of the choice of ρ_{-i}.[24] 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 (AF) with equality.
Step 3: Potential for signals. A sufficient condition for incentive compatibility of the 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

$$(\gamma_i(1) \geq \gamma_i(\kappa)) \Leftrightarrow \rho_i \geq \rho^0. \tag{35}$$

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 (35) holds by Assumption 2 as it becomes

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

$\square$

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.

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

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

— (2015b). “Optimal Alternative Dispute Resolution”. *mimeo.*

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. (2016). “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”.

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.*

Meirowitz, A., M. Morelli, K. W. Ramsay, and F. Squintani (2015). “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.

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.

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.

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

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