Mediation and Costly Evidence *

Yi Chen[†] Pengfei Zhang[‡]

Oct. 23, 2021

Preliminary Draft

Please do not circulate

Abstract

This paper studies the efficient mediation procedure where disputants are asymmetrically informed and can acquire hard evidence with some probability at a cost. The model encompasses both facilitative mediation where the mediator only transmits information, and evaluative mediation where the mediator bases recommendation on acquired information. Our efficient mediation procedure is consistent with the practice and empirical facts of professional mediators, in which weak cases are settled by facilitation and strong cases are settled by evaluation if efficiency demands settlement of all cases. While facilitation is the exclusive focus of previous literature, our results suggest that evaluation is equally important for efficiency. Our findings discipline the use of evidence in mediation, which have implications for dispute resolution, two-sided online platform, and international relations.

Keywords: Mediation, Costly Evidence, Burden of Proof, Alternative Dispute Resolution, Mechanism Design, Information Design

JEL classification: C72, C78, D82, D83, K41

*We are grateful to Ben Chen, Sebastian Kodritsch, Henrik Lando, Barton Lipman, Bruno Strulovici, and Kathy Spier for valuable comments and suggestions. Pengfei owes special thanks to Kris Iyer for the encouragement and suggestion upon initiation of this project. For helpful comments and discussions, we thank seminar and conference participants at Cornell University, AEA 2021, ALEA 2021.

[†]Samuel Curtis Johnson Graduate School of Management, Cornell University, yc2535@cornell.edu

[‡]Corresponding author. Department of Economics, Cornell University, pz82@cornell.edu

1 Introduction

Information asymmetry is at the heart of many inefficient resource allocations, leading to bargaining impasse, and many times failure of dispute resolution. Mediation, a *self-enforcing* procedure, is increasingly adopted to alleviate such informational problems. It is used to resolve disputes in matters like breach of contract, online transaction, torts cases including medical malpractice, in labor issues like worker compensation and wrongful termination, family issues like divorce, and even international relation issues when the two parties are sovereignty.

Previous literature has exclusively focused on facilitative mediation where the mediator structures a communication process including holding private meetings with each party, asking them questions, understanding their concerns and passing on useful information to the other party. The role of the mediator is to transmit information.

However, equally important in the practice of mediation is evaluative mediation where the mediator assists parties in reaching a resolution by pointing out the weakness of their cases, and predicting what a likely outcome of the trial will be. An evaluative mediator would make evidence-based recommendations to the parties in relation to their outside option, in the hope that both parties can find a common ground for their agreement. Therefore the mediator also has a role to acquire information.

In this paper we analyze both facilitative mediation and evaluative mediation, and consider any hybrid of these two prominent forms of mediation along the spectrum. We ask ourselves, what is the efficient mediation procedure among all of them. In pursuit of that question, we are also able to answer a host of related questions: is the focus on facilitation without loss of generality? When should we use evaluation? What kinds of cases are settled by facilitation and vice versa?

To approach these questions, we apply mechanism design and information design techniques to a canonical bargaining model that captures the essence of the problem. There are three players: The mediator, the informed party which we call the plaintiff, the uninformed party which we call the defendant. The plaintiff and the defendant find themselves in a bargaining situation where the mediator tries to reach an agreement for the two parties.

Upon receiving a report from the plaintiff, the mediator commit whether to pay a cost to acquire evidence and collect some transfers if the parties reach an agreement. The transfers are here to finance the cost of evidence acquisition. The evaluation process is not always conclusive - it can lead to no evidence found with some probability. The mediator then converts the reports and the evidence to a recommended allocation and announces it to both parties, according to a pre-committed random mapping. The procedure is self-enforcing such that any party can opt out anytime. Rejection leads to the outside option that can mean going to trial, going on strike, or even a war. The mediator tries to maximize their total payoffs using information transmission and possibly acquisition subject to budget balance, the players' incentive compatibility for truth telling, and their obedience constraints such that when the mediator recommends a settlement, they will find it optimal to settle.

We solve this design problem in two steps. First, we solve a closely related problem of costly auditing, where the mediator can directly learn the truth if he pays the cost. We then show the mediation problem with costly evidence faces stronger constraints, but nevertheless achieves the same optimal value.

Our main findings are as follows: (i) Facilitation is always involved in efficient mediation, (ii) Evaluation is required by efficient mediation if and only if efficiency demands all cases to be settled, (iii) In any efficient mediation plan, weak cases are settled by facilitative mediation, whereas strong cases are settled by evaluative mediation if they are required by efficiency to be settled, (iv) Resolution for stronger case is based on more precise yet risky evidence.

To satisfy individual rationality, the mediator wants both players to get at least their outside option, such that settlement is always weakly better off. Incentive compatibility has two components, depending on whether the mediator acquires evidence. With hard evidence, the mediator can punish lying by refusing to mediate. Without evidence, the mediator has to ensure that a higher type always has a weekly higher payoff so that the plaintiff finds truth-telling optimal. Efficiency requires the mediator to settle as much as possible, and save cost of evidence in the meantime. Since facilitation does not cost anything, he will push that to the limit, which defines the threshold. Below the threshold, all cases are settled, they have exactly the same allocation, and no transfer is ever being paid. Above the threshold, all cases would have to proceed to trial if evaluation is not used. If evaluation is more cost-effective than trial, then it is worthwhile to acquire evidence, after which state of the world becomes common knowledge. The plaintiff and the defendant will realize they have a common interest to avoid the loss for both of them. Thus they are willing to settle, and they are willing to pay for the cost.

We then consider an alternative model where it is the sender who can present an evidence at a cost. We find the two models are equivalent in terms of information transmission, i.e., who bears the burden of costly proof is irrelevant for efficiency.

We conclude this paper by relating it to the literature that no unmediated negotiation procedures can achieve the same mediated result. Mediation has a strict benefit. We also advocate the policy of mediation default to resolve costly disputes, especially in developing countries where the legal costs are high (see evidence in Djankov et al. (2003)), and with the advancement in information technology, the cost of evaluation is getting ever lower.

1.1 Related Literature

Works on mediation and alternative dispute resolution have been relatively little. Myerson (1991) pioneered in showing that mediation between parties with misaligned interests can improve the efficiency of the communication even though parties are restricted to unverifiable messages (aka Crawford and Sobel (1982)). Myerson's insight is that by adding noise, mediator actually makes the communication more informative. In a similar sender-receiver game, Goltsman et al. (2009) compares the ability of mediation, arbitration, and unmediated negotiation (finite rounds of cheap talk) to maximize the receivers payoff. The optimal mediation filters the senders private information and adds noise. Arbitration is (generically) more effective than mediation, while mediation is only sometimes more effective than unmediated negotiation. In settings similar to bilateral bargaining when the preferences of the two parties are completely conflicting (i.e., split a pie), the mediator recommends a split of the pie based on agents' reports of their types, and if either party opts out, a default division of a reduced pie is implemented. In this setting, Fey and Ramsay (2010) shows that mediation cannot improve on unmediated communication if uncertainty only concerns agents private costs of fighting. When symmetric agents share of the (reduced) pie from conflict is determined by privately known strengths, however, Hrner, Morelli, and Squintani (2015) shows that arbitration and mediation are equally effective at minimizing conflict. Both outperform unmediated communication when the intensity of conflict is high, or asymmetric information is substantial. Extending this game, Meirowitz et al. (2019) shows that unmediated peace talks increase the incentive to militarize and so increase eventual conflict, but mediated peace talks reduce militarization and conflict. As mentioned in the introduction, previous theoretical works has exclusively focused on facilitative mediation (Myerson, 1991; Goltsman et al., 2009; Horner et al., 2015; Fanning, 2021). Exciting empirical research on mediation is emerging, highlighting the importance of evaluation (McDermott & Obar, 2004; Klerman & Klerman, 2015). In a very interesting paper, Balzer & Schneider (2020) also studies mediation and evidence, where evidentiary hearing is an outside option to the mediation process. By contrast, this paper highlights the use of evidence *in* the mediation process.

This paper contributes to the recent literature on evidence and mechanism design (Hart, Kremer & Perry, 2017; Ben-Porath, Dekel & Lipman, 2020). There are two approach to incorporate evidence: (i) verifiable disclosure of the informed party, and (ii) costly verification of the uninformed party. The early insight of this literature is that skepticism of the receiver can force the sender to voluntarily disclose and unravels to full revelation of truth (Grossman and Hart (1980), Grossman (1981), Milgrom (1981)). A large body of literature followed and can be categorized into two strands: one maintains the GM assumption and extends the conclusion, the other questions its robustness by bringing in other elements. In the first category, Okuno-Fujiwara et al. (1990) extends the unraveling result to more general games. Recently, Hart et al. (2017) shows that the uninformed party's commitment to transfer policy makes no difference to the outcome in a verifiable disclosure setting. Ben-Porath et al. (2019) further generalizes that result to multi-player mechanism design setting. In the second category, Jovanovic (1982) and **?** show that some information will be withheld if there is cost associated with disclosure [1]. Following their approach, we pays particular attention to the cost of evidence, and we consider all possible communication protocols by studying the information design problem. Moreover, we compare the disclosure setting to our main setup where it is the receiver who can request evidence. As such, we relax the assumption that off equilibrium message has to be truthful. There is a literature on contracting with costly state verification (Townsend (1979), Border and Sobel (1987), Mookherjee and Png (1989), Glazer and Rubinstein (2004), Ben-Porath et al. (2014)). The main differences are three-folds: i) In a model of costly verification, the uninformed party can directly learn the true type at the cost. We instead explicitly specify the process of evaluation on top of the evidence structure, and thus allow for risk consideration of information acquisition. ii) our central concern in this paper is the communication process instead of the transfer policy, iii) verification is committed in this literature, while evaluation of evidence has to be obedient for the receiver in our paper.

This paper relates to the burgeoning literature on litigation and pretrial Bargaining (see Spier (2007) and Daughety and Reinganum (2017) for excellent overviews of the literature). Most models follow earlier bilateral bargaining models by Bebchuk (1984), Reinganum and Wilde (1986), and Spier (1992). Spier (1994) and Klement & Neeman (2005) study dispute resolution from a mechanism design perspective. In both models, information revealed determines incentives to (re-)negotiate. Farmer & Pecorino (2005) study how costly evidence affects pretrial bargaining.

This paper relates to the literature on Bayesian communication and information design (Kamenica and Gentzkow (2011), Bergemann and Morris (2016)). In particular, it is most closely related to the mediation literature initiated by Myerson (1991) where the information designer has no informational advantages over the players. The main insight of this literature (Myerson (1991), Blume et al. (2007), Goltsman et al. (2009)) is that the mediator can improve welfare by introducing noise in the communication channel such that

[1]Other variations that support the withholding of information includes seller's lack of information and information acquisition by Dye (1986), Matthews and Postlewaite (1985), Farrel (1986), and Shavell (1994), alternative market structure including one-sided market by Fishman and Hagerty (1995), and oligopoly by Board (2009), and Hotz and Xiao (2013), and alternative category of information by Li and Madarasz (2008), and Board (2012).

the incentives for misrepresentation is reduced. In this literature, it is usually assumed that information is unverifiable, while in our model, a pool of objective evidence is available but to access that pool players have to bear the burden of proof. Building upon their framework, we investigate how hard evidence changes the optimal mediation procedure and who should bear the burden of proof. This problem shares the same feature as the information design problems inspired by Kamenica and Gentzkow (2011) where the designer has commitment power over communication. As Kamenica and Gentzkow (2011) have noted, this problem is the same as directly choosing the distribution of posterior beliefs τ.[2] The important divergences from standard Bayesian persuasion are two-folds: i) the mechanism has a screening role such that the designer has to elicit private information from the sender. That is, τ is conditional on the report ω_i, by which we denote $\tau_i(\mu^1)$. ii) the information is costly *verifiable*. If the information is nonverifiable, then given monetary transfer is ruled out, there is no way to incentivize truth-telling. In this problem, truth-telling is plausible precisely because the designer can induce evaluation of the receiver.

2 Model

There is a plaintiff, a defendant, and a mediator. Let $\Omega := [0, 1]$ be the state space, where ω denotes its typical member. Nature randomly selects a state $\omega \in \Omega$ according to some non-degenerate commonly known distribution $\mu^0(\omega)$, and secretly reveals it to the plaintiff only. Upon observing ω, the plaintiff sends a message to the mediator.

Mediator commits whether to pay a cost $C \geqslant 0$ to acquire an evidence $e \in E$, based on m_1, verifying event $\{\omega' | e \in \mathcal{E}(\omega')\} \subset \Omega$. Evidence conclusive with prob. $\tilde{\eta}(e)$, otherwise inconclusive.

Mediator converts the message m_1 and the evidence to another message $m_2 \in \mathcal{M}_2$ according to a committed random mapping $\tilde{\pi}_e(m_2 | m_1, \rho_e)$, and announces m_2 to both parties.

[2] Alternatively, this is the same as appealing to the revelation principle such that we can restrict our attention to $\tilde{D} = \Delta(\Omega)$ w.l.o.g.

Based on $\{m_2, t_1, t_2\}$, the two parties decide upon whether to accept the recommended allocation $(x - t_1, -x - t_2)$, and the mediator collects $\{t_1(m_1), t_2(m_1)\}$. Rejection leads to default allocation $(\omega - L_1, -\omega - L_2)$ (e.g. trial, strike, war).

Plaintiff's payoff:

$$u_1 = \begin{cases} x - t_1 & \text{if an agreement is reached,} \\ \omega - L_1 & \text{otherwise.} \end{cases}$$

Defendant's payoff:

$$u_2 = \begin{cases} -x - t_2 & \text{if an agreement is reached,} \\ -\omega - L_2 & \text{otherwise.} \end{cases}$$

The mediator can commit to any random mapping $\tilde{\pi}(m_r|m_s)$ as the mediation plan. In this game, denote the sender's mixed strategy as $\tilde{\sigma}_s : \Omega \to \Delta(\mathcal{M}_s)$, and the receiver's mixed strategy as $\tilde{\sigma}_r : \mathcal{M}_r \to \Delta(2^{\Omega} \times \mathbb{R}^2)$. To interpret, a pure strategy of the sender is to pick a message m_s given the state ω. A pure strategy of the receiver is a contingent plan that picks a set $D \subset \Omega$ along with two real numbers based on the message m_r, indicating what action to choose if intended evidence is acquired or not acquired, respectively. We use perfect Bayesian equilibrium as the solution concept.

We list several relevant examples below.

Example 1. *Debt and Bankruptcy: creditor and debtor, uncertainty on non-exempt financial assets, evidence is a financial statement, decision is the obligation to repay (White (2007)).*

Example 2. *Tort case: injured and injury party, uncertainty on negligence, decision is the compensation for damage.*

Example 3. *Online transaction: seller and buyer, uncertainty on manufacturing default.*

Example 4. *International conflicts: two countries, uncertainty on the military strength. (Hörner et al. (2015)).*

In reality, mediation is often desired or required even though it is common that the informed party prefers the relevant decision to be as high as possible (sales, compensation, and investment in the areas listed above). In these settings, a pool of objective evidence is usually available, and frequently exploited to one's own advantages. With the inclusion of evidence, one would expect fundamental change to the aforementioned theoretical predictions. Therefore, studying how the communication outcomes change with the usage of evidence is of great policy importance.

We now describe the availability of the evidence. It is common knowledge that the pool of evidence is an exogenous mapping $\mathcal{E}: \Omega \rightarrow 2^{2^{\Omega}}$ that satisfies the following two property:

1. Authenticity property: $E \in \mathcal{E}(\omega)$ implies $\omega \in E$.

2. Consistency property: $\omega' \in E \in \mathcal{E}(\omega)$ implies $E \in \mathcal{E}(\omega')$.

Interpretation of $\mathcal{E}(\omega)$ is the set of events that can be proved conclusively by some documents or other forms of tangible evidence when the true state is ω. The authenticity property says that any evidence must contain truth. The consistency property states that if an evidence available for ω does not rule out ω', that evidence is also available when the true state is ω'.

Because we have the evaluation process, it's important to distinguish two posterior belief profiles. Let $\mu^1(\omega'|m) \equiv \Pr[\omega = \omega'|m]$ be the post-disclosure belief after receiving message m, and $\mu^2(\omega'|\omega, D) \equiv \Pr[\omega = \omega'|D, \rho]$ be the post-evaluation belief after seeing result ρ for the intended evidence D.

We assume the evidence strucuture is complete such that $\mathcal{E}(\omega) = \{E \in 2^{\Omega}|\omega \in E\}$. If the true state is ω_2, the pool of evidence is $\{\omega_2, \{\omega_0, \omega_2\}, \{\omega_1, \omega_2\}, \Omega\}$. Likewise for ω_1 and ω_0. Given $\mu^1(\omega'|m)$, the true state ω, and an intended evidence D, the evaluation technology updates belief as follows:

If $\rho = T$ indicating $\omega \in D$,

$$\mu^2(\omega'|D, T) = \begin{cases} 0 & \text{if } \omega' \notin D, \\ \frac{\mu^1(\omega'|m)}{\sum_{\tilde{\omega} \in D} \mu^1(\tilde{\omega}|m)} & \text{if } \omega' \in D. \end{cases} \tag{1}$$

If $\rho = F$ indicating $\omega \notin D$,

$$\mu^2(\omega'|D,F) = \begin{cases} 0 & \text{if } \omega' \in D, \\ \frac{\mu^1(\omega'|m)}{\sum_{\tilde{\omega}\notin D}\mu^1(\tilde{\omega}|m)} & \text{if } \omega' \notin D. \end{cases} \tag{2}$$

The evaluation technology is formulated such that it disqualifies some states of the world instead of directly representing the truth (which distinguish this model from costly state verification). If the intended evidence is acquired, evaluation informs the receiver that any type outside the evidence cannot be truth. If the intended evidence is not acquired, evaluation is not uninformative but instead tells the receiver that any type within the evidence cannot be truth. From receiver's perspective of information acquisition, the information content of the evidence (and thus the benefit of evaluation) is not fixed and depends crucially on her prior belief, thereby suggesting a welfare improving role of mediation procedure.

3 Mediation

We will focus on direct mechanism where the report is a type and the message is an allocation. We are going to proceed to solve this problem in two steps. First, we solve a closely related problem of costly auditing, where the mediator can directly learn the truth if he pays the cost. And then we show the mediation problem with costly evidence will face stronger constraints, but nevertheless achieves the same value. Let's think intuitively on how to approach this problem. Individual rationality constraints says that the motivator wants both players to get at least their outside option, such that settlement is always weakly better off. Incentive compatibility has two components, depending on whether the mediator audits. Because the mediator can directly learn the truth, he knew whoever is lying. So he can punish lying by refusing to mediate. He also has to ensure when he does not audit, the plaintiff is also telling the truth that requires the p function to be monotone. So that a higher type always have a weekly higher payoff.

Now, let's say we have a candidate pool of all the feasible mediation plans. How should we choose among them? Think about the mediators' objective. The mediator wants the two parties to settle as much as possible. In the meantime, he wants to save auditing cost and facilitative mediation does not cost anything. He will push that to the limit. That's where you find omega star on page nine. Below omega star, all cases are settled, they have exactly the same allocation, and no transfer is ever being paid. But above omega star, without evaluation all cases go to trial. If the cost of auditing is less or equal to the total loss of going to court, then it is worthwhile to audit. Once the mediator audits, omega becomes common knowledge. The plaintiff and the defendant will realize they have a common interest to avoid the loss for both of them. So they are willing to settle, and they are willing to pay for the auditing cost. How do I know it's a threshold but not any other partitions? Weaker cases are always easier to settle. You can tell that from the monotonicity of the p function.

Given a direct mediation plan $\{\pi_0(x|\cdot), \pi_1(y|\cdot), I(\cdot), t_i(\cdot)\}$, we define some quantities that prove to be useful later. Define

$$p_\pi(\omega) = \int_{\underline{\omega}-L_1}^{\overline{\omega}+L_2} \pi_0(x|\omega)dx, \quad p_\pi(\omega)x_\pi(\omega) = \int_{\underline{\omega}-L_1}^{\overline{\omega}+L_2} x\pi_0(x|\omega)dx$$
$$q_\pi(\omega) = \int_{\underline{\omega}-L_1}^{\overline{\omega}+L_2} \pi_1(y|\omega)dy, \quad q_\pi(\omega)y_\pi(\omega) = \int_{\underline{\omega}-L_1}^{\overline{\omega}+L_2} y\pi_1(y|\omega)dy$$

as the probability of reaching an agreement and the expected settlement if the mediator does not request evidence, and the probability of reaching an agreement and the expected settlement if the mediator finds an evidence.

The expected payoffs for the plaintiff:

$$X_1(\omega, \hat{\omega}) = p_\pi(\hat{\omega})[x_\pi(\hat{\omega}) - t_1(\hat{\omega})] + (1 - p_\pi(\hat{\omega}))(\omega - L_1)$$
$$Y_1(\omega, \hat{\omega}) = \begin{cases} q_\pi(\omega)[y_\pi(\omega) - t_1(\omega)] + (1 - q_\pi(\omega))(\omega - L_1) & \text{if } \hat{\omega} = \omega, \\ \omega - L_1 & \text{if } \hat{\omega} \neq \omega. \end{cases}$$

Similarly, The expected payoffs for the defendant:

$$X_2(\omega,\hat{\omega}) = -p_\pi(\hat{\omega})[x_\pi(\hat{\omega}) + t_2(\hat{\omega})] - (1 - p_\pi(\hat{\omega}))(\omega + L_2)$$

$$Y_2(\omega,\hat{\omega}) = \begin{cases} -q_\pi(\omega)[y_\pi(\omega) + t_2(\omega)] - (1 - q_\pi(\omega))(\omega + L_2) & \text{if } \hat{\omega} = \omega, \\ -\omega + L_2 & \text{if } \hat{\omega} \neq \omega. \end{cases}$$

Note that because mediation is self-enforcing, the harshest punishment in the off equilibrium Y_1 is to ask the two parties to proceed to outside options.

$$\max_{\substack{\pi_0(x|\cdot),\pi_1(y|\cdot),\\ I(\cdot),t_i(\cdot)}} [L_1 + L_2] \int_0^1 \left[I(\omega)\left(1 - q_\pi(\omega)\right) + \left(1 - I(\omega)\right)\left(1 - p_\pi(\omega)\right)\right] \mu^0(\omega) d\omega$$

$$\begin{aligned} \text{s.t. } & C \int_0^1 I(\omega)\mu^0(\omega) d\omega \leqslant T \\ & X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \omega - L_1, \ \ \forall\omega \\ & X_2(\omega,\omega) + I(\omega)[Y_2(\omega,\omega) - X_2(\omega,\omega)] \geqslant -\mathbb{E}_{\mu^1}[\omega] - L_2, \ \ \forall\omega \\ & X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \\ & \quad \max\{X_1(\omega,\hat{\omega}) + I(\hat{\omega})[Y_1(\omega,\hat{\omega}) - X_1(\omega,\hat{\omega})],\ \omega - L_1\}, \ \ \forall\omega,\hat{\omega} \end{aligned}$$

The first constraint is the budget constraint for evidence acquisition, the second set of constraints is individual rationality for player 1, the third constraint is individual rationality for player 2, the last set of constraints is incentive compatibility for double deviation of truth-telling and opting out.

Note that $x(\omega), p(\omega)$ are defined on $\{\hat{\omega}|I(\hat{\omega}) = 0\}$, and are free to choose on the complement $\{\hat{\omega}|I(\hat{\omega}) = 1\}$. Likewise, $y(\omega), q(\omega)$ are defined on $\{\hat{\omega}|I(\hat{\omega}) = 1\}$, and are free to choose on $\{\hat{\omega}|I(\hat{\omega}) = 0\}$.

We can break down IC for truth-telling into four cases:

$$Y_1(\omega,\omega) \geqslant Y_1(\omega,\hat{\omega}) \quad \text{if } I(\omega) = I(\hat{\omega}) = 1. \tag{3}$$

This is equivalent to $q(\omega)y(\omega) \geqslant q(\omega)(\omega - L)$, which is implied by IR for Player 1.

$$X_1(\omega,\omega) \geqslant Y_1(\omega,\hat{\omega}) \quad \text{if } I(\omega) = 0, I(\hat{\omega}) = 1. \tag{4}$$

This is equivalent to $p(\omega)x(\omega) \geqslant p(\omega)(\omega - L)$, which is implied by IR for Player 1.

$$X_1(\omega,\omega) \geqslant X_1(\omega,\hat{\omega}) \quad \text{if } I(\omega) = I(\hat{\omega}) = 0. \tag{5}$$

This is similar to a screening problem, and we can simplifies them as Myerson (1981).

$$Y_1(\omega,\omega) \geqslant X_1(\omega,\hat{\omega}) \quad \text{if } I(\omega) = 1, I(\hat{\omega}) = 0. \tag{6}$$

This is similar to an auditing problem, and we can simplifies them as Townsend (1979).

We start with 5.

Lemma 1. *$\{x(\omega), p(\omega)\}$ is incentive compatible, if and only if (i) $p(\omega)$ is non-increasing in ω, and (ii) for any $\omega \in [\underline{\omega}, \overline{\omega}]$, $X_1 = \int_0^{\omega}[1 - p(\tilde{\omega})]d\tilde{\omega} + X_1(0)$.*

Proof. Consider two types ω and $\hat{\omega}$ where $\omega > \hat{\omega}$. 5 requires

$$\begin{aligned} p(\omega)x(\omega) + (1 - p(\omega))(\omega - L) &\geqslant p(\hat{\omega})x(\hat{\omega}) + (1 - p(\hat{\omega}))(\omega - L) \\ p(\hat{\omega})x(\hat{\omega}) + (1 - p(\hat{\omega}))(\hat{\omega} - L) &\geqslant p(\omega)x(\omega) + (1 - p(\omega))(\hat{\omega} - L) \end{aligned} \tag{7}$$

Adding the two inequalities and rearranging, we get

$$(p(\hat{\omega}) - p(\omega))(\omega - \hat{\omega}) \geqslant 0. \tag{8}$$

Since $\omega > \hat{\omega}$, we have $p(\omega) \leqslant p(\hat{\omega})$.

5 means that for all ω, we have

$$X_1(\omega) = \max_{\hat{\omega} \in \{\hat{\omega} | I(\hat{\omega}) = 0\}} p(\hat{\omega})x(\hat{\omega}) - (1 - p(\hat{\omega}))L + (1 - p(\hat{\omega}))\omega \tag{9}$$

By envelop theorem, we have $X_1'(\omega) = 1 - p(\omega)$ whenever differentiable.

By the second fundamental theorem of calculus (Royden and Fitzpatrick (2010) shows it extends to piece-wise differentiability), we obtain

$$X_1(\omega) = \int_0^{\omega} [1 - p(\tilde{\omega})] d\tilde{\omega} + X_1(0) \quad \forall \omega. \tag{10}$$

□

It is easy to see that $Y_1(\omega) \geqslant X_1(\omega)$ plus 5 imply 6; thus given 5 we only have to deal with $Y_1(\omega) \geqslant X_1(\omega)$. Combining it with IR for player 1, we have

$$q(\omega)y(\omega) \geqslant \max\{X_1(\omega) - (1 - q(\omega)(\omega - L)), q(\omega)(\omega - L)\}. \tag{11}$$

The next lemma tells us that this constraint is binding.

Lemma 2. *For any* $\omega \in [0,1]$, $q(\omega)y(\omega) = \max\{X_1(\omega) - (1 - q(\omega))(\omega - L), q(\omega)(\omega - L)\}$ *if* $q(\omega) < 1$, $y(\omega) \geqslant \max\{X_1(\omega), (\omega - L)\}$ *if* $q(\omega) = 1$.

Proof. If $q(\omega) = 0$, since $X_1(\omega) - (1 - q(\omega))(\omega - L) \leqslant 0$, the equation holds.

Assume $q(\omega) \in (0,1)$, suppose $y(\omega) > \max\{\frac{X_1(\omega)}{q(\omega)} - \frac{1-q(\omega)}{q(\omega)}(\omega - L), (\omega - L)\}$. There exists $\tilde{y}(\omega) \geqslant \omega - L$ and $\tilde{y}(\omega) > y(\omega)$.

Observe that $\tilde{y}(\omega)$ is feasible but would relax the IR constraint for player 2 and thus make it feasible to raise the agreement probability $\tilde{q}(\omega) > q(\omega)$. This contradicts the fact that $q(\omega)$ is a minimizer.

Therefore, $q(\omega)y(\omega) = \max\{X_1(\omega) - (1 - q(\omega))(\omega - L), q(\omega)(\omega - L)\}$ for all ω where $q(\omega) < 1$. If $q(\omega) = 1$, then IC is not binding and $y(\omega) \geqslant \max\{X_1(\omega), (\omega - L)\}$. □

We know $X_1(\omega)$ is determined by $p(\omega)$ and $X_1(0)$. Now consider the following program

$$\begin{aligned} \min_{p(\omega)} \quad & (L_1 + L_2) \int_0^1 [1 - I(\omega)][1 - p(\omega)] \mu^0(\omega) d\omega \\ s.t. \quad & p(\omega) \text{is non-increasing} \end{aligned} \tag{12}$$

Lemma 3. *$p(\omega)$ solves the above program if there exists a ω^* such that*

$$p(\omega) = \begin{cases} 1 & \textit{if } \omega \leqslant \omega^*, \\ 0 & \textit{if } \omega > \omega^*. \end{cases} \tag{13}$$

Proof. Let $\mathcal{P}$ be the set of all bounded non-increasing functions such that $p(\omega) \in [0,1] \forall \omega \in [0,1]$. This is the set the designer can choose. We endow $\mathcal{P}$ with the linear structure and the metrics induced by L^1-norm.

Notice that $\mathcal{P}$ is compact and convex. Denote L_0 as the objective function and notice further that L_0 is continuous and linear in $p(\omega)$. By the Extreme Point Theorem (Ok, 2007, p.658), the set of extreme points of $\mathcal{P}$ is nonempty which includes a $p(\omega)$ such that

$$L_0(p(\omega)) \geqslant L_0(\tilde{p}(\omega)) \quad \forall p \in \mathcal{P}. \tag{14}$$

Therefore a function $p(\omega)$ that is an extreme point of $\mathcal{P}$ and that minimizes L_0 among all extreme points of $\mathcal{P}$ also minimizes L_0 among all functions in $\mathcal{P}$. A function $p(\omega)$ is an extreme point of $\mathcal{P}$ if $p(\omega) \in \{0,1\}$ for almost all $\omega \in [0,1]$. The designer can thus restrict attention to non-stochastic $p(\omega)$.

An extreme point $p(\omega)$ is non-increasing if and only if there exists a ω^* such that

$$p(\omega) = \begin{cases} 1 & \text{if } \omega \leqslant \omega^*, \\ 0 & \text{if } \omega > \omega^*. \end{cases} \tag{15}$$

□

From lemma 1, we have

$$X_1(\omega) = \begin{cases} X_1(0) & \text{if } \omega \leqslant \omega^*, \\ \omega - \omega^* + X_1(0) & \text{if } \omega > \omega^*. \end{cases} \tag{16}$$

By definition, this implies $x(\omega) = x(0) = \omega^* - L$ for $\omega \in [0, \omega^*]$. Therefore,

$$x(\omega) = \begin{cases} \omega^* - L & \text{if } \omega \leqslant \omega^*, \\ \omega - L & \text{if } \omega > \omega^*. \end{cases} \tag{17}$$

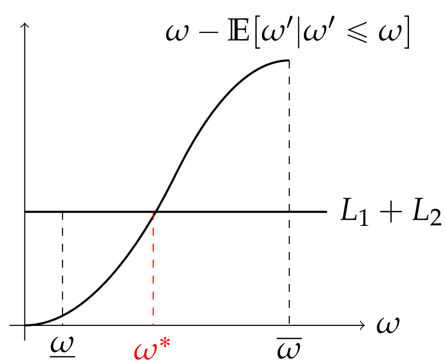
It is straightforward to check that this satisfies IR for player 1, i.e. $x(0) \geqslant \omega - L$ for $\omega \in [0, \omega^*]$ and $\omega - L \geqslant \omega - L$ for $\omega \in (\omega^*, 1]$. IR for player 2 requires whenever $\omega \in [0, \omega^*]$, $x(0) \leqslant \mathbb{E}[\omega | \omega \leqslant \omega^*] + L$. That is, for a ω^* to be feasible, it has to satisfy

$$\omega^* - \mathbb{E}[\omega | \omega \leqslant \omega^*] \leqslant L_1 + L_2. \tag{18}$$

This is a critical condition that we have to explore more.

Corollary 1. $\omega^* \in \{\omega' | \omega' - \mathbb{E}[\omega | \omega \leqslant \omega'] \leqslant L_1 + L_2\}$.

Notice that $L_1 + L_2 \geqslant 0$ so such a ω^* always exist. As either L_1 or L_2 becomes larger, this set possibly grows larger.



Lemma 4. *$\{\omega | I(\omega) = 1\}$ is a nonempty connected set if (i) the shadow price of T is strictly larger than $(L_1 + L_2)/C$, (ii) $q(\omega)$ is non-decreasing.*

Proof. Suppose the function $I^*(\omega)$ is a solution to the arbitration problem. Define $h(\omega)$ to be the deviation between $I^*(\omega)$ and some other feasible function $I(\omega)$. For any constant a, the function $I(\omega) = I^*(\omega) + ah(\omega)$ is also feasible. With both $I^*(\omega)$ and $h(\omega)$ held fixed, consider the Lagrangian as a function of a,

$$\begin{aligned}\mathcal{L}(a) = &-(L_1+L_2)\int_0^1 \left[1-p(\omega)+\left[I^*(\omega)+ah(\omega)\right]\left[p(\omega)-q(\omega)\right]\right]\mu^0(\omega)d\omega \\ &+\lambda_T\left[T - C\int_0^1 (I^*(\omega)+ah(\omega))\mu^0(\omega)d\omega\right]\end{aligned} \tag{19}$$

By assumption, the function $\mathcal{L}(a)$ obtains its optimum at $a=0$. This implies that for all $h(\omega)$, the first order derivative

$$\mathcal{L}'(a) = \int_0^1 \{[L_1+L_2][q(\omega)-p(\omega)] - \lambda_T C\}h(\omega)\mu^0(\omega)d\omega \leqslant 0 \quad \forall a > 0 \tag{20}$$

Therefore, if $h(\omega) > 0$, $[L_1+L_2][q(\omega)-p(\omega)] \leqslant \lambda_T c$; if $h(\omega) < 0$, $[L_1+L_2][q(\omega)-p(\omega)] \geqslant \lambda_T c$. Since $I^*(\omega)+ah(\omega)$ is feasible, it follows that

$$I^*(\omega) = \begin{cases} 1 & \text{if } q(\omega)-p(\omega) \geqslant \frac{\lambda_T c}{[L_1+L_2]}, \\ 0 & \text{if } q(\omega)-p(\omega) < \frac{\lambda_T c}{[L_1+L_2]}. \end{cases} \tag{21}$$

Since $p(\omega)$ is non-increasing, $\{\omega | I(\omega)=1\}$ is a nonempty connected set if (i) there exists $\underline{\omega}_I$ such that $q(\underline{\omega}_I) - p(\underline{\omega}_I) \geqslant \frac{\lambda_T c}{[L_1+L_2]}$, (ii) there exists $\overline{\omega}$ such that $q(\omega)$ is non-decreasing in $[0,\overline{\omega}]$, (iii) $q(\omega)-p(\omega) \leqslant \frac{\lambda_T c}{[L_1+L_2]}$ for $\omega \in (\overline{\omega}, 1]$.

(i) cannot be satisfied if $\lambda_T c > [L_1+L_2]$. If (ii) holds for $\overline{\omega}=1$, then (iii) is irrelevant and for any $\omega \in [\underline{\omega}_I, \overline{\omega}]$, $q(\omega)-p(\omega) \geqslant \frac{\lambda_T c}{[L_1+L_2]}$. □

Assuming the conditions on $q(\omega)$ holds (which we will check back later), by lemma 4 $\{\omega | I(\omega)=1\} = [\underline{\omega}_I, 1]$. The budget constraint becomes

$$1 - M^0(\underline{\omega}_I) = \frac{T}{C} \tag{22}$$

where $M(\omega)$ is the CDF of prior belief.

Lemma 5. *If* $T < \min\{\frac{\int_0^{\omega_I} M^0(\tilde{\omega})d\tilde{\omega} - [L_1+L_2]}{\underline{\omega}_I - [L_1+L_2]}C, [1-M^0([L_1+L_2])]C\}$, *then* $\omega^* = \sup\{\omega' | \omega' - \mathbb{E}[\omega | \omega \leqslant \omega'] = [L_1+L_2]\}$. *If* $T \geqslant \min\{\frac{\int_0^{\omega_I} M^0(\tilde{\omega})d\tilde{\omega} - [L_1+L_2]}{\underline{\omega}_I - [L_1+L_2]}C, [1-M^0([L_1+L_2])]C\}$, *then*

$\omega^* = \underline{\omega}_I$.

Proof.

$$\underline{\omega}_I - \mathbb{E}[\omega|\omega \leqslant \underline{\omega}_I] = \underline{\omega}_I - \frac{\int_0^{\underline{\omega}_I}[1 - M^0(\tilde{\omega})]d\tilde{\omega}}{M(\underline{\omega}_I)} = -\frac{T}{C}\underline{\omega}_I + \frac{\int_0^{\underline{\omega}_I} M^0(\tilde{\omega})d\tilde{\omega}}{1 - \frac{T}{C}}. \tag{23}$$

If $T < \min\{\frac{\int_0^{\underline{\omega}_I} M^0(\tilde{\omega})d\tilde{\omega} - [L_1+L_2]}{\underline{\omega}_I - [L_1+L_2]}C, [1 - M^0([L_1 + L_2])]C\}$, we have

$$\begin{aligned}
&(\underline{\omega}_I - [L_1 + L_2])T < [\int_0^{\underline{\omega}_I} M^0(\tilde{\omega})d\tilde{\omega} - [L_1 + L_2]]C, \text{ and } \underline{\omega}_I > [L_1 + L_2] \\
\Longleftrightarrow &(1 - \frac{T}{C})([L_1 + L_2]) + \frac{T}{C}\underline{\omega}_I < \int_0^{\underline{\omega}_I} M^0(\tilde{\omega})d\tilde{\omega}, \text{ and } \underline{\omega}_I > L_1 + L_2 \\
\Longleftrightarrow &\underline{\omega}_I - \mathbb{E}[\omega|\omega \leqslant \underline{\omega}_I] > L_1 + L_2
\end{aligned} \tag{24}$$

The feasibility constraint in Corollary 1 is binding, and $\omega^* = \sup\{\omega'|\omega' - \mathbb{E}[\omega|\omega \leqslant \omega'] = L_1 + L_2\}$ such that $1 - p(\omega)$ is minimized for $I(\omega) = 0$ whenever possible.

If $T \geqslant \min\{\frac{\int_0^{\underline{\omega}_I} M^0(\tilde{\omega})d\tilde{\omega} - (L_1+L_2)}{\underline{\omega}_I - (L_1+L_2)}C, [1 - M^0(L_1 + L_2)]C\}$, we have $\underline{\omega}_I - \mathbb{E}[\omega|\omega \leqslant \underline{\omega}_I] \leqslant L_1 + L_2$. The feasibility constraint is non-binding, and we choose $p(\omega) = 1 \forall \omega \in [0, \underline{\omega}_I]$. It follows that $\omega^* = \underline{\omega}_I$. □

It only remains to determine $q(\omega), y(\omega)$. Consider the following program

$$\begin{aligned}
\min_{q(\omega),y(\omega)} \quad & L_1 + L_2 \int_0^1 I(\omega)[1 - q(\omega)]\mu^0(\omega)d\omega \\
s.t. \quad & q(\omega)y(\omega) + (1 - q(\omega))(\omega + L) \leqslant \mathbb{E}_{\mu^1}[\omega] + L \quad \forall \omega \in [\underline{\omega}_I, 1], \\
& q(\omega)y(\omega) = \max\{X_1(\omega) - (1 - q(\omega))(\omega - L)), q(\omega)(\omega - L)\} \quad \forall \omega.
\end{aligned} \tag{25}$$

Lemma 6. *$q(\omega)$ is non-decreasing and $q(\omega) = 1$ for $\omega \in [\underline{\omega}_I, 1]$.*

Proof. From lemma 5, we know that $\omega^* \leqslant \underline{\omega}_I$. Thus for $\omega \in [\underline{\omega}_I, 1]$, $X_1(\omega) - (1 - q(\omega))(\omega - L) = q(\omega)(\omega - L)$, and $q(\omega)y(\omega) = q(\omega)(\omega - L)$.

$q(\omega) = 1$ for $\omega \in [\underline{\omega}_I, 1]$ obviously solves the unconstrained program. We show it is also feasible in the constrained program, in particular it satisfies the IR constraint for player 2. Since $q(\omega) = 1, y(\omega) = \omega - L$ for $\omega \in [\underline{\omega}_I, 1]$. By Bayes rule, $\mathbb{E}[\omega | I(\omega) = 1] = \omega$. IR is non-binding since $\omega - L < \omega + L$.

Set $q(\omega) = 0$ for $\omega \in [0, \underline{\omega}_I)$. $q(\omega)$ is non-decreasing. □

Proposition 1. *In any optimal arbitration plan with given budget T,*
(i) an agreement is reached for all $\omega \in [0,1]$, if $T < \min\{\frac{\int_0^{\underline{\omega}_I} M^0(\tilde{\omega})d\tilde{\omega} - (L_1+L_2)}{\underline{\omega}_I - (L_1+L_2)} C, [1 - M^0(L_1 + L_2)]C\}$,
(ii) an agreement is reached except $\omega \in [\omega^*, \underline{\omega}_I]$ where $\omega^* = \sup\{\omega' | \omega' - \mathbb{E}[\omega | \omega \leqslant \omega'] = L_1 + L_2\}$ and $1 - M^0(\underline{\omega}_I) = \frac{T}{C}$, if $T \geqslant \min\{\frac{\int_0^{\underline{\omega}_I} M^0(\tilde{\omega})d\tilde{\omega} - (L_1+L_2)}{\underline{\omega}_I - (L_1+L_2)} C, [1 - M^0(L_1 + L_2)]C\}$.

3.1 Budget Balance

The budget constraint now changes to

$$C \int_0^1 I(\omega)\mu^0(\omega)d\omega \leqslant \int_0^1 \left[I(\omega)\left[t_1 + t_2\right] + (1 - I(\omega))\left[t_1 + t_2\right]\right]\mu^0(\omega)d\omega \tag{26}$$

$$\begin{aligned}
X_1(\omega, \hat{\omega}) &= p(\hat{\omega})[x(\hat{\omega}) - t_1(\hat{\omega})] + (1 - p(\hat{\omega}))(\omega - L) \\
X_2(\omega, \hat{\omega}) &= -p(\hat{\omega})[x(\hat{\omega}) + t_2(\hat{\omega})] - (1 - p(\hat{\omega}))(\omega + L) \\
Y_1(\omega, \hat{\omega}) &= \begin{cases} q(\omega)[y(\omega) - \tau_1(\omega)] + (1 - q(\omega))(\omega - L) & \text{if } \hat{\omega} = \omega, \\ \omega - L & \text{if } \hat{\omega} \neq \omega. \end{cases} \\
Y_2(\omega, \hat{\omega}) &= \begin{cases} -q(\omega)[y(\omega) + \tau_2(\omega)] - (1 - q(\omega))(\omega + L) & \text{if } \hat{\omega} = \omega, \\ -\omega + L & \text{if } \hat{\omega} \neq \omega. \end{cases}
\end{aligned} \tag{27}$$

Individual rationality for player 1 is then

$$\begin{aligned}
p(\omega)[x(\omega) - t_1] &\geqslant p(\omega)(\omega - L) \\
q(\omega)[y(\omega) - t_1] &\geqslant q(\omega)(\omega - L)
\end{aligned} \tag{28}$$

Individual rationality for player 2 is then

$$\begin{aligned} p(\omega)[x(\omega)+t_2]+(1-p(\omega))(\omega-L) &\leqslant \mathbb{E}_{\mu^1}[\omega]+L, \qquad \text{or } p(\omega)=0 \\ q(\omega)[y(\omega)+t_2]+(1-q(\omega))(\omega-L) &\leqslant \mathbb{E}_{\mu^1}[\omega]+L, \qquad \text{or } q(\omega)=0 \end{aligned} \tag{29}$$

We proceed as last section. The IC can be divided into four cases.

Lemma 1 holds.

Lemma 2 becomes for any $\omega \in [0,1]$, $q(\omega)[y(\omega)-t_1] = \max\{X_1(\omega)-(1-q(\omega))(\omega-L), q(\omega)(\omega-L)\}$.

Lemma 3 holds, and it follows that for $\omega \in [0,\omega^*]$, $x(\omega)-t_1 = x(0)-t_1(0) = \omega^*-L$. Combined with IR for player 2, we have for $\omega \in [0,\omega^*]$,

$$t_1+t_2 \leqslant L_1+L_2-(\omega^*-\mathbb{E}[\omega|\omega \leqslant \omega^*]). \tag{30}$$

We prove the updated version of Lemma 6 first, and bring it to the proof of lemma 4.

Lemma 7. *$q(\omega)=1$ for any $\omega \in \{\omega|I(\omega)=1\}$ if $t_1+t_2 \leqslant L_1+L_2-(\omega^*-\omega)$ for $\omega \leqslant \omega^*$, and $t_1+t_2 \leqslant L_1+L_2$ for $\omega > \omega^*$.*

Proof. $q(\omega)=1$ for $\omega \in \{\omega|I(\omega)=1\}$ obviously solves the unconstrained program. We show it is also feasible in the constrained program if $t_1+t_2 \leqslant L_1+L_2-(\omega^*-\omega)$ for $\omega \leqslant \omega^*$, and $t_1+t_2 \leqslant L_1+L_2$ for $\omega > \omega^*$.

By lemma 2,

$$y(\omega)-t_1 = \begin{cases} \max\{\omega^*-L, \omega-L\} = \omega^*-L & \text{if } I(\omega)=1 \& \omega \leqslant \omega^*, \\ \omega-L & \text{if } I(\omega)=1 \& \omega > \omega^*. \end{cases} \tag{31}$$

It remains to check whether IR constraint for player 2 is satisfied. By Bayes rule, $\mathbb{E}[\omega|I(\omega) =$

$1] = \omega$. If $I(\omega) = 1$ and $\omega \leqslant \omega^*$, we have

$$\begin{aligned} y(\omega) + t_2 &= \omega^* - L + t_1 + t_2 \\ &\leqslant \omega^* + L - (\omega^* - \omega) \\ &= \omega + L \end{aligned} \tag{32}$$

If $I(\omega) = 1$ and $\omega > \omega^*$, we have

$$\begin{aligned} y(\omega) + t_2 &= \omega - L + t_1 + t_2 \\ &\leqslant \omega + L \end{aligned} \tag{33}$$

□

Lemma 8. $\{\omega | I(\omega) = 1\}$ *is a nonempty connected set if (i) the shadow price of T is strictly larger than* $L_1 + L_2/C$*, (ii)* $q(\omega)$ *is non-decreasing, (iii)* $t_1 + t_2 - t_1 - t_2$ *is non-increasing.*

Proof. Suppose the function $I^*(\omega)$ is a solution to the arbitration problem. Define $h(\omega)$ to be the deviation between $I^*(\omega)$ and some other feasible function $I(\omega)$. For any constant a, the function $I(\omega) = I^*(\omega) + ah(\omega)$ is also feasible. With both $I^*(\omega)$ and $h(\omega)$ held fixed, consider the Lagrangian as a function of a,

$$\begin{aligned} \mathcal{L}(a) = &- (L_1 + L_2) \int_0^1 \left[1 - p(\omega) + [I^*(\omega) + ah(\omega)]\,[p(\omega) - q(\omega)]\right] \mu^0(\omega) d\omega \\ &+ \lambda_T \int_0^1 \left\{t_1 + t_2 + [I^*(\omega) + ah(\omega)]\,[t_1 + t_2 - t_1 - t_2 - C]\right\} \mu^0(\omega) d\omega \end{aligned} \tag{34}$$

By assumption, the function $\mathcal{L}(a)$ obtains its optimum at $a = 0$. This implies that for all $h(\omega)$, the first order derivative

$$\mathcal{L}'(a) = \int_0^1 \{L_1 + L_2[q(\omega) - p(\omega)] + \lambda_T\,[t_1 + t_2 - t_1 - t_2 - C]\} h(\omega) \mu^0(\omega) d\omega \leqslant 0 \quad \forall a > 0 \tag{35}$$

Therefore, if $h(\omega) > 0$, $2Lq(\omega) + \lambda_T[t_1 + t_2] \leqslant 2Lp(\omega) + \lambda_T[t_1 + t_2 + C]$; if $h(\omega) < 0$, $2Lq(\omega) + \lambda_T[t_1 + t_2] \geqslant 2Lp(\omega) + \lambda_T[t_1 + t_2 + C]$. Since $I^*(\omega) + ah(\omega)$ is feasible, it follows that

$$I^*(\omega) = \begin{cases} 1 & \text{if } q(\omega) - p(\omega) \geqslant \frac{\lambda_T[t_1+t_2-t_1-t_2+C]}{L_1+L_2}, \\ 0 & \text{if } q(\omega) - p(\omega) < \frac{\lambda_T[t_1+t_2-t_1-t_2+C]}{L_1+L_2}. \end{cases} \tag{36}$$

Since $p(\omega)$ is non-increasing, $\{\omega | I(\omega) = 1\}$ is a nonempty connected set if (i) there exists $\underline{\omega}_I$ such that $q(\underline{\omega}_I) - p(\underline{\omega}_I) \geqslant \frac{\lambda_T[t_1(\underline{\omega}_I)+t_2(\underline{\omega}_I)-\tau_1(\underline{\omega}_I)-\tau_2(\underline{\omega}_I)+C]}{L_1+L_2}$, (ii) $q(\omega)$ is non-decreasing in $[0, 1]$, (iii) $t_1 + t_2 - t_1 - t_2$ is non-increasing. □

Set $q(\omega) = 0$ for $\omega \in [0, \underline{\omega}_I)$. $q(\omega)$ is non-decreasing. The three constraints on transfers are

$$\begin{aligned} t_1 + t_2 &\leqslant L_1 + L_2 - (\omega^* - \mathbb{E}[\omega | \omega \leqslant \omega^*]) && \text{if } \omega \leqslant \omega^*, \\ t_1 + t_2 &\leqslant L_1 + L_2 - (\omega^* - \omega) && \text{if } \omega \leqslant \omega^*, \\ t_1 + t_2 &\leqslant L_1 + L_2 && \text{if } \omega > \omega^*. \end{aligned} \tag{37}$$

It is straightforward to see that if all the constraints are binding and set $t_1 + t_2 = 0$ for $\omega > \omega^*$, $t_1 + t_2 - t_1 - t_2$ is non-increasing. By lemma 4, $\{\omega | I(\omega) = 1\} = [\underline{\omega}_I, 1]$. The budget constraint implies

$$1 - M^0(\underline{\omega}_I) = \frac{\int_0^{\underline{\omega}_I} [t_1 + t_2] \, \mu^0(\omega) d\omega + \int_{\underline{\omega}_I}^1 [t_1 + t_2] \, \mu^0(\omega) d\omega}{C} \tag{38}$$

where $M^0(\omega)$ is the CDF of prior belief.

Proposition 2. *If $C \leqslant L_1 + L_2$, in any budget balanced optimal arbitration plan with costly state verification, an agreement is reached for all $\omega \in [0, 1]$.*

Proof. It only requires to show that $C \leqslant L_1 + L_2$ implies $T \geqslant \min\{\frac{\int_0^{\underline{\omega}_I} M^0(\tilde{\omega}) d\tilde{\omega} - (L_1+L_2)}{\underline{\omega}_I - (L_1+L_2)} C, [1 - M^0(L_1 + L_2)] C\}$ where $T = \int_0^{\underline{\omega}_I} [t_1 + t_2] \, \mu^0(\omega) d\omega + \int_{\underline{\omega}_I}^1 [t_1 + t_2] \, \mu^0(\omega) d\omega$.

Since $\underline{\omega}_I \geqslant \omega^*$, we have

$$\begin{aligned}
T &= (1 - M^0(\underline{\omega}_I))(L_1 + L_2) + M^0(\omega^*)[L_1 + L_2 - (\omega^* - \mathbb{E}[\omega|\omega \leqslant \omega^*])] \\
&> (1 - M^0(\underline{\omega}_I))(L_1 + L_2) \\
&\geqslant (1 - M^0(\underline{\omega}_I))C \\
&\geqslant \begin{cases} \frac{\int_0^{\underline{\omega}_I} M^0(\tilde{\omega})d\tilde{\omega} - (L_1+L_2)}{\underline{\omega}_I - (L_1+L_2)} C & \text{if } \underline{\omega}_I > L_1 + L_2, \\ [1 - M^0(L_1 + L_2)]C & \text{if } \underline{\omega}_I \leqslant L_1 + L_2. \end{cases}
\end{aligned} \tag{39}$$

where the third inequality follows from the assumption, and the last inequality stems from the fact that $\int_0^{\underline{\omega}_I} M^0(\tilde{\omega})d\tilde{\omega} \leqslant M^0(\underline{\omega}_I)\underline{\omega}_I$. □

3.2 Cost Minimization

The objective function now changes to maximization of the total payoffs, which is equivalent to minimizing the total costs. The goal of the mediation is to help the parties make the most informed decision with the least cost of evidence. As such, the optimality consists of two intuitive requirements: provide more information and optimize the use of evaluation.

The mediation problem is

$$\begin{aligned}
\max_{\substack{\pi_0(x|\cdot),\pi_1(y|\cdot), \\ I(\cdot),t_i(\cdot)}} & \int_0^1 \Big(I(\omega) \sum_i X_i(\omega) + (1 - I(\omega)) \sum_i Y_i(\omega) \Big) \mu^0(\omega) \mathrm{d}\omega \\
\text{s.t.} & \int_0^1 \left(I(\omega) q_\pi(\omega) \sum_i t_i(\omega) + (1 - I(\omega))\, p_\pi(\omega) \sum_i t_i(\omega) \right) \mu^0(\omega) \mathrm{d}\omega \geqslant \\
& \quad C \int_0^1 I(\omega) \mu^0(\omega) \mathrm{d}\omega \\
& X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \omega - L_1, \;\; \forall \omega \\
& X_2(\omega,\omega) + I(\omega)[Y_2(\omega,\omega) - X_2(\omega,\omega)] \geqslant -\mathbb{E}_{\mu^1}[\omega] - L_2, \;\; \forall \omega \\
& X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \\
& \quad \max\{X_1(\omega,\hat{\omega}) + I(\hat{\omega})[Y_1(\omega,\hat{\omega}) - X_1(\omega,\hat{\omega})],\; \omega - L_1\}, \;\; \forall \omega, \hat{\omega}
\end{aligned}$$

The first constraint is the budget constraint for evidence acquisition, the second set of constraints is individual rationality for player 1, the third constraint is individual rationality for player 2, the last set of constraints is incentive compatibility for double deviation of truth-telling and opting out.

Applying Lemma 9, we have

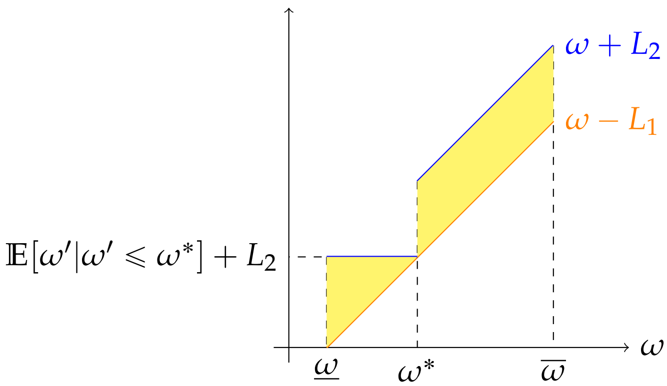
$$I^*(\omega) = \begin{cases} 1 & \text{if } q(\omega) - p(\omega) \geqslant \frac{(1-\lambda_T)[t_1+t_2-t_1-t_2]+\lambda_T C}{L_1+L_2}, \\ 0 & \text{if } q(\omega) - p(\omega) < \frac{(1-\lambda_T)[t_1+t_2-t_1-t_2]+\lambda_T C}{L_1+L_2}. \end{cases} \tag{40}$$

$$I(\omega) = \begin{cases} 1 & \text{if } q_\pi(\omega)[\sum_i L_i - \sum_i \tau_i(\omega)] - p_\pi(\omega)[\sum_i L_i - \sum_i t_i(\omega)] \\ & \quad \geqslant C + q_\pi(\omega)\sum_i t_i(\omega) - p_\pi(\omega)\sum_i \tau_i(\omega), \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 9. *In any efficient mediation plan, (i)* $t_1 = t_2 = 0$ *for any* $\omega \in [\underline{\omega}, \overline{\omega}]$*, (ii)* $t_1 + t_2 = C$ *whenever* $I(\omega) = 1$*, (iii)* $\omega^* = \sup\{\omega' | \omega' - \mathbb{E}[\omega | \omega \leqslant \omega'] = L_1 + L_2\}$.

Proof. unconstrained optimization gives $t_1 = t_2 = 0$ for any $\omega \in [\underline{\omega}, \overline{\omega}]$ and it is feasible. $t_1 + t_2$ is bounded below by the budget requirement. Thus, we want the domain of $I(\omega) = 0$ as large as possible. check back the lemma 9 conditions. □

Proposition 3. *An efficient mediation plan with costly auditing is characterized by a threshold* ω^* *such that*
$\omega^* = \sup\{\omega | \omega - \mathbb{E}[\omega' | \omega' \leqslant \omega] = L_1 + L_2\}$,
for any $\omega \leqslant \omega^*$, $I(\omega) = 0$, $t_1 = t_2 = 0$, $p_\pi(\omega) = 1$, *and* $x_\pi(\omega) = \omega^* - L_1$,
for any $\omega > \omega^*$, *if* $C \leqslant L_1 + L_2$, *then* $I(\omega) = 1$, $t_1 + t_2 = C$, $q_\pi(\omega) = 1$, $y_\pi(\omega) \in [\omega - L_1, \omega + L_2 - C]$; *if* $C > L_1 + L_2$, *then* $I(\omega) = 0$, $t_1 = t_2 = 0$, $p_\pi(\omega) = 0$.



Proposition 4. *$\{I(\omega), \pi_0(x|\omega), \pi_1(y|\omega)\}$ solves the mediation problem with costly state verification if $p_\pi(\omega) = p(\omega)$, $x_\pi(\omega) = p(\omega)x(\omega)$, $q_\pi(\omega) = q(\omega)$, $y_\pi(\omega) = q(\omega)y(\omega)$.*

Proof. We start by observing that the objective functions and constraints of the two problem are the same except the stronger incentive compatibility in the mediation program, therefore the value of the mediation program is bounded above by the value of the arbitration program.

Next we show that $\{p_\pi(\omega), x_\pi(\omega), q_\pi(\omega), y_\pi(\omega)\}$ would satisfy the incentive compatibility of double deviation. By assumption, we have

$$X_1(\omega) = \begin{cases} \omega^* - L & \text{if } \omega \leqslant \omega^*, \\ \omega - L & \text{if } \omega > \omega^*. \end{cases} \geqslant \omega - L. \tag{41}$$

and

$$Y_1(\omega) = \begin{cases} \omega - L & \text{if } \omega < \underline{\omega}_I, \\ \omega - L & \text{if } \omega \geqslant \underline{\omega}_I. \end{cases} \geqslant \omega - L. \tag{42}$$

Since $\{p(\omega), x(\omega), q(\omega), y(\omega)\}$ solves the arbitration problem, we know the IC of arbitration are satisfied. Therefore, $X_1(\omega) \geqslant \max\{X_1(\omega, \hat{\omega}), \omega - L\}$ and $Y_1(\omega) \geqslant \max\{Y_1(\omega, \hat{\omega}), \omega - L\}$.

Now the value of the objective function under $\{I(\omega), \pi_0(x|\omega), \pi_1(y|\omega)\}$ is equal to the value of the arbitration program, which is weakly larger than any mediation solution, thus $\{I(\omega), \pi_0(x|\omega), \pi_1(y|\omega)\}$ is optimal. □

3.3 Evidence Acquisition

We now consider the evidence acquisition technology we are primarily interested in. Once we move from costly auditing to costly evidence, two complications arise. The first one is the evidence acquired will be a set instead of the true state. But that itself doesn't stop us from learning the truth, if D is chosen properly. In particular, when the plaintiff reports omega, the mediator will check whether the true state is at least omega. If that's true, he will treat the type of the plaintiff as the smallest element in that set. So a higher type will be assigned a finer set. No one wants to misreport. Otherwise, he will either be found lying or he will receive a lower payoff that's actually consistent with the classic unraveling result in voluntary disclosure. The other complication is that evaluation may be inconclusive. And that gives rise to the second threshold. That problem is similar to facilitative mediation with one critical difference: the defendant, in his posterior belief, thinks being inconclusive higher types are more likely, because the evidence assigned to them is harder to find. In fact, we prove in the paper that if η satisfies a monotonicity condition, then the second threshold can even reach the highest type, such that all cases can be settled if and only if evaluative mediation is used.

Let $\underline{\omega}_D = \inf D(\omega)$ be the smallest element of the acquired evidence $D(\omega)$.

$$Y_1(\omega,\hat{\omega},D(\hat{\omega})) = \begin{cases} [y_\pi(\underline{\omega}_D) - q_\pi(\underline{\omega}_D)\tau_1(\underline{\omega}_D)] + (1-q_\pi(\underline{\omega}_D))(\omega - L) & \text{if } \hat{\omega} \in D(\hat{\omega}), \\ \omega - L & \text{if } \hat{\omega} \notin D(\hat{\omega}). \end{cases}$$
$$Y_2(\omega,\hat{\omega},D(\hat{\omega})) = \begin{cases} -[y_\pi(\underline{\omega}_D) + q_\pi(\underline{\omega}_D)\tau_2(\underline{\omega}_D)] - (1-q_\pi(\underline{\omega}_D))(\omega + L) & \text{if } \hat{\omega} \in D(\hat{\omega}), \\ -\omega + L & \text{if } \hat{\omega} \notin D(\hat{\omega}). \end{cases} \tag{43}$$

Proposition 5. $\{D(\omega), \pi_D(y|\omega)\}$ *solves the mediation problem with costly evidence if* $D(\omega) = [\omega, 1]$ *for any* ω *such that* $I(\omega) = 1$, $\pi_D(y|\omega) = \pi_0(x|\omega)$ *for* $D = \emptyset$, $\pi_D(y|\omega) = \pi_1(y|\omega)$ *for* $D \neq \emptyset$.

Proof. We first show that $\underline{\omega}_D = \omega$. Equation 5 is now changed to

$$Y_1(\omega, \omega, D(\omega)) \geqslant Y_1(\omega, \hat{\omega}, D(\hat{\omega})) \quad \text{if } I(\omega) = 1, I(\hat{\omega}) = 1. \tag{44}$$

This is similar to a disclosure problem, and we can simplifies them as Grossman (1981).

Next we show it satisfies all other constraints □

Theorem 1. *If $C \leqslant L_1 + L_2$, in any budget balanced optimal mediation plan with costly evidence, an agreement is reached for all $\omega \in [0, 1]$.*

Theorem 2. *An efficient mediation plan with costly evidence is characterized by two thresholds $\{\omega^*, \omega_I^*\}$ such that:*

$$\begin{aligned} \omega^* &= \sup\{\omega | \omega - \mathbb{E}[\omega' | \omega' \leqslant \omega] = L_1 + L_2\}, \\ \omega_I^* &= \min\left\{\overline{\omega}, \sup\{\omega | \omega - \mathbb{E}_\eta[\omega' | \omega' \in [\omega^*, \omega]] = L_1 + L_2 - C\}\right\}. \end{aligned}$$

(i) For any $\omega \leqslant \omega^*$, $D(\omega) = \varnothing$, $t_1 = t_2 = 0$, $p_\pi(\omega) = 1$, and $x_\pi(\omega) = \omega^* - L_1$;
(ii) For any $\omega > \omega^*$:

- *if $C > L_1 + L_2$, then $D(\omega) = \varnothing$, $t_1 = t_2 = 0$, $p_\pi(\omega) = 0$;*
- *if $C \leqslant L_1 + L_2$, then $D(\omega) = [\omega, \overline{\omega}]$, $t_1 + t_2 = C$, $q_\pi(\omega) = 1$, $y_\pi(\omega) \in [\omega - L_1 + t_1, \omega + L_2 - t_2]$.*

(iii) For any $\omega \in [\omega^*, \omega_I^*]$, $r_\pi(\omega) = 1$, $z_\pi(\omega) = \omega_I^* - L_1$,
(iv) For any $\omega > \omega_I^*$, $r_\pi(\omega) = 0$.

4 Other Procedures of ADR

4.1 Arbitration

$\omega^* = \min\{\overline{\omega}, L_1 + L_2 - \mathbb{E}_{\mu^0}[\omega]\}$

$$
\begin{aligned}
X_1(\omega, \hat{\omega}) &= p(\hat{\omega})x(\hat{\omega}) + (1 - p(\hat{\omega}))(\omega - L) \\
X_2(\omega, \hat{\omega}) &= -p(\hat{\omega})x(\hat{\omega}) - (1 - p(\hat{\omega}))(\omega + L) \\
Y_1(\omega, \hat{\omega}) &= \begin{cases} q(\omega)y(\omega) + (1 - q(\omega))(\omega - L) & \text{if } \hat{\omega} = \omega, \\ \omega - L & \text{if } \hat{\omega} \neq \omega. \end{cases} \\
Y_2(\omega, \hat{\omega}) &= \begin{cases} -q(\omega)y(\omega) - (1 - q(\omega))(\omega + L) & \text{if } \hat{\omega} = \omega, \\ -\omega + L & \text{if } \hat{\omega} \neq \omega. \end{cases}
\end{aligned} \tag{45}
$$

The arbitration problem with ex-ante IR is

$$
\begin{aligned}
\min_{I(\omega),x(\omega),p(\omega),y(\omega),q(\omega)} \quad & [L_1 + L_2] \int_0^1 [I(\omega)\,(1 - q(\omega)) + (1 - I(\omega))\,(1 - p(\omega))]\, \mu^0(\omega) d\omega \\
s.t. \quad & C \int_0^1 I(\omega)\mu^0(\omega) d\omega \leqslant T \\
& X_1(\omega, \omega) + I(\omega)[Y_1(\omega, \omega) - X_1(\omega, \omega)] \geqslant \omega - L_1 \quad \forall \omega \\
& X_2(\omega, \omega) + I(\omega)[Y_2(\omega, \omega) - X_2(\omega, \omega)] \geqslant -\mathbb{E}_{\mu^0}[\omega] - L_2 \quad \forall \omega \\
& X_1(\omega, \omega) + I(\omega)[Y_1(\omega, \omega) - X_1(\omega, \omega)] \geqslant \\
& X_1(\omega, \hat{\omega}) + I(\hat{\omega})[Y_1(\omega, \hat{\omega}) - X_1(\omega, \hat{\omega})] \quad \forall \omega, \hat{\omega}
\end{aligned}
$$

The first constraint is the budget constraint for evidence acquisition, the second set of constraints is individual rationality for player 1, the third constraint is individual rationality for player 2, the last set of constraints is incentive compatibility for truth-telling.

The only difference with mediation is ex-ante IR and ex-interim IR.

4.2 Unmediated Negotiation

Shavell (1989) considers a screening game where the informed plaintiff can costless disclose verifiable private information and the uninformed defendant makes the offer. He finds an equilibrium under which there are no trials: plaintiffs with strong cases reveal their type, while plaintiffs with weak cases remain silent and receive a pooling offer that all accept.

Sobel (1989), however, shows that costly voluntary disclosure will not take place if the opposing party makes the final offer. If the plaintiff reveals her type but the defendant makes the final offer, the defendant will gain all the benefits from settlement through this final offer. Thus, there is no benefit to the plaintiff from revealing her type, and with positive cost of disclosure, she strictly prefers to remain silent. As a consequence, no trial equilibrium in Shavell (1989) will take place if and only if disclosure is costless.

5 Concluding Remarks

This paper studies the efficient mediation procedure where disputants are asymmetrically informed and hard evidence can be acquired probabilistically at a cost. A mediator commits ex-ante to a mediation plan that generates stochastic messages for the uninformed party, based on the informed party's reports, and acquired costly evidence. The model encompasses both facilitative mediation where mediator only transmits information, and evaluative mediation where mediator bases recommendation on evidence, thus has to acquire information.

The efficient mediation plan features two threshold where the lower threshold determines whether an evidence should be acquired, and the higher threshold determines whether the two parties can settle if evaluation turns out inconclusive. We show that (i) facilitation is always involved in efficient mediation, (ii) evaluation is required by efficient mediation if and only if efficiency demands all cases to be settled, and (iii) in an efficient mediation plan, weak cases are settled by facilitative mediation, strong cases are settled by evaluative mediation if required, and settlement for stronger case hinges on more precise yet risky evidence. Furthermore, who bears the burden of proof is irrelevant for efficiency.

While facilitation is the exclusive focus of previous literature, our findings suggest that evaluation is equally important for efficiency. Our results speak directly to the evaluative-facilitative debate central in mediation. Our findings highlight mediation default, which bears implications for the design of online platform, dispute resolution, ratings, and international relations.

Our efficient mediation appears consistent with several empirical findings (McDermott & Obar, 2004; Klerman & Klerman, 2015). Substantial percent of mediators use both evaluative and facilitative techniques, and/or "hybrid" techniques. A mediators proposal, when used, leads to very high settlement rate (over 99 percent). Pure facilitation has a comparatively narrow range of settlement. Evaluation results in higher amount of settlement.

We conclude this paper by relating it to the literature that no unmediated negotiation procedures can achieve the same mediated result. Mediation has a strict benefit. We also advocate the policy of mediation default to resolve costly disputes, especially in developing countries where the legal costs are high, and with the advancement in information technology, the cost of evaluation is getting lower and lower.

A Direct Mechanism

The following revelation principle is in the spirit of Myerson (1991). The following lemma tells us that we can focus only on direct, truthful, and obedient mediation plans.

Lemma 10. *Given an arbitrary mediation plan $\tilde{\pi}(m_r|m_s)$ which implements a random mapping from states to joint distributions of (D, a) as an outcome of a perfect Bayesian equilibrium, there exists a direct, truthful, and obedient mediation plan $\pi(D, a_T, a_F|\omega)$ that implements the same random mapping as an outcome of a perfect Bayesian equilibrium.*

Proof. Construct a direct mechanism π using type report ω as input and a recommended set D and recommended action a_z contingent on $\rho = \{T, F\}$ as the output:

$$\pi(D, a_T, a_F|\omega) \equiv \sum_{m_r}\sum_{m_s}\left(\prod_{z}\tilde{\sigma}_a(a_z|D, m_r, \rho)\right)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega).$$

In the following, we verify that this mechanism implements the same distribution of outcomes by means of a truthful and obedient PBE, where $\sigma_s(\omega|\omega) = 1$, $\sigma_r(D|D, a_T, a_F) = 1$, and $\sigma_a(a_z|D, a_T, a_F, \rho) = 1$.

First, the sender finds truth-telling optimal:

$$
\begin{aligned}
&U_s(\omega|\omega) \equiv \sum_{D,\rho,a_z} a_z \Pr(\rho|D,\omega)\pi(D,a_T,a_F|\omega) \\
= &\sum_{m_s,m_r,D,\rho,a_z} a_z\tilde{\sigma}_a(a_T|D,m_r,T)\tilde{\sigma}_a(a_F|D,m_r,F)\Pr(\rho|D,\omega)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega) \\
\geqslant &\sum_{m_s,m_r,D,\rho,a_z} a_z\tilde{\sigma}_a(a_T|D,m_r,T)\tilde{\sigma}_a(a_F|D,m_r,F)\Pr(\rho|D,\omega)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega') \\
= &\sum_{D,\rho,a_z} a_z \Pr(\rho|D,\omega)\pi(D,a_T,a_F|\omega') = U_s(\omega'|\omega),
\end{aligned}
$$

where the inequality follows because in the original PBE, $\tilde{\sigma}_s(\cdot|\omega)$ is a better strategy than $\tilde{\sigma}_s(\cdot|\omega')$ when the true state is ω.

Second, the receiver's beliefs are consistent with the equilibrium. Upon receiving recommendation (D, a_T, a_F),

$$
\begin{aligned}
&\mu^1(\omega|D,a_T,a_F) = \frac{\pi(D,a_T,a_F|\omega)\mu^0(\omega)}{\sum_\omega \pi(D,a_T,a_F|\omega)\mu^0(\omega)} \\
= &\frac{\mu^0(\omega)\sum_{m_r}\sum_{m_s}\left(\prod_z \tilde{\sigma}_a(a_z|D,m_r,\rho)\right)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega)}{\sum_\omega \mu^0(\omega)\sum_{m_r}\sum_{m_s}\left(\prod_z \tilde{\sigma}_a(a_z|D,m_r,\rho)\right)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega)}.
\end{aligned}
$$

Upon receiving recommendation (D, a_T, a_F) and seeing the test outcome ρ,

$$
\begin{aligned}
\mu^2(\omega|D,a_T,a_F,T) &= \begin{cases} 0 & \text{if } \omega \notin D, \\ \frac{\mu^1(\omega|D,a_T,a_F)}{\sum_{\omega'\in D}\mu^1(\omega|D,a_T,a_F)} & \text{if } \omega \in D. \end{cases} \\
\mu^2(\omega|D,a_T,a_F,F) &= \begin{cases} 0 & \text{if } \omega \in D, \\ \frac{\mu^1(\omega|D,a_T,a_F)}{\sum_{\omega'\in D^c}\mu^1(\omega|D,a_T,a_F)} & \text{if } \omega \notin D. \end{cases}
\end{aligned} \tag{46}
$$

Third, the receiver finds obedience optimal. Notice that by Bayes' rule,

$$
\mu^1(\omega|D,a_T,a_F) = \sum_{m_r} \Pr(m_r|D,a_T,a_F)\tilde{\mu}^1(\omega|m_r),
$$

where $\tilde{\mu}^1(\omega|m_r)$ is the receiver's belief upon receiving message m_r in the original PBE. Therefore, for every m_r such that the receiver's strategy selects plan (D, a_T, a_F) with positive probability in the original PBE, we have:

$$\begin{aligned} & \sum_{\rho,\omega} v(a_z, w) \Pr(\rho|D,\omega)\tilde{\mu}^1(\omega|m_r) - c(D) \\ \geqslant \; & \sum_{\rho,\omega} v(a'_z, w) \Pr(\rho|D',\omega)\tilde{\mu}^1(\omega|m_r) - c(D'), \end{aligned}$$

for any (D', a'_T, a'_F). Averaging over all such m_r, we know that in the direct mediation plan:

$$\begin{aligned} & U_r(D, a_T, a_F|D, a_T, a_F) \equiv \sum_{\rho,\omega} v(a_z, w) \Pr(\rho|D,\omega)\mu^1(\omega|D, a_T, a_F) - c(D) \\ = \; & \sum_{m_r} \Pr(m_r|D, a_T, a_F) \sum_{\rho,\omega} v(a_z, w) \Pr(\rho|D,\omega)\tilde{\mu}^1(\omega|m_r) - c(D) \\ \geqslant \; & \sum_{m_r} \Pr(m_r|D, a_T, a_F) \sum_{\rho,\omega} v(a'_z, w) \Pr(\rho|D',\omega)\tilde{\mu}^1(\omega|m_r) - c(D') \\ = \; & \sum_{\rho,\omega} v(a'_z, w) \Pr(\rho|D',\omega)\mu^1(\omega|D, a_T, a_F) - c(D') \\ = \; & U_r(D', a'_T, a'_F|D, a_T, a_F). \end{aligned}$$

□

References

Balzer, Benjamin and Johannes Schneider, "Managing A Conflict: Optimal Alternative Dispute Resolution," 2020.

Ben-Porath, Elchanan, Eddie Dekel, and Barton L Lipman, "Optimal allocation with costly verification," *American Economic Review*, 2014, *104* (12), 3779–3813.

_ , _ , and _ , "Mechanisms with evidence: Commitment and robustness," *Econometrica*, 2019, *87* (2), 529–566.

Bergemann, Dirk and Stephen Morris, "Information design, Bayesian persuasion, and Bayes correlated equilibrium," *American Economic Review*, 2016, *106* (5), 586–91.

Blume, Andreas, Oliver J Board, and Kohei Kawamura, "Noisy talk," *Theoretical Economics*, 2007, 2 (4), 395–440.

Border, Kim C and Joel Sobel, "Samurai accountant: A theory of auditing and plunder," *The Review of economic studies*, 1987, *54* (4), 525–540.

Brown, Jennifer Gerarda and Ian Ayres, "Economic rationales for mediation," *Virginia Law Review*, 1994, pp. 323–402.

Crawford, Vincent P and Joel Sobel, "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1982, pp. 1431–1451.

Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Courts," *The Quarterly Journal of Economics*, 2003, *118* (2), 453–517.

Dye, Ronald A, "Disclosure of nonproprietary information," *Journal of accounting research*, 1985, pp. 123–145.

Gentzkow, Matthew and Emir Kamenica, "Costly persuasion," *American Economic Review*, 2014, *104* (5), 457–62.

Glazer, Jacob and Ariel Rubinstein, "On optimal rules of persuasion," *Econometrica*, 2004, *72* (6), 1715–1736.

Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani, "Mediation, arbitration and negotiation," *Journal of Economic Theory*, 2009, *144* (4), 1397–1420.

Grossman, Sanford J, "The informational role of warranties and private disclosure about product quality," *The Journal of Law and Economics*, 1981, *24* (3), 461–483.

_ **and Oliver D Hart**, "Disclosure laws and takeover bids," *The Journal of Finance*, 1980, *35* (2), 323–334.

Hart, Sergiu, Ilan Kremer, and Motty Perry, "Evidence games: Truth and commitment," *American Economic Review*, 2017, *107* (3), 690–713.

Hörner, Johannes, Massimo Morelli, and Francesco Squintani, "Mediation and peace," *The Review of Economic Studies*, 2015, *82* (4), 1483–1501.

Jovanovic, Boyan, "Truthful disclosure of information," *The Bell Journal of Economics*, 1982, pp. 36–44.

Kamenica, Emir and Matthew Gentzkow, "Bayesian persuasion," *American Economic Review*, 2011, *101* (6), 2590–2615.

Klement, Alon and Zvika Neeman, "Against compromise: A mechanism design approach," *Journal of Law, Economics, and Organization*, 2005, *21* (2), 285–314.

McDermott, E Patrick and Ruth Obar, "What's going on in mediation: An empirical analysis of the influence of a mediator's style on party satisfaction and monetary benefit," *Harv. Negot. L. Rev.*, 2004, *9*, 75.

Milgrom, Paul R, "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 1981, pp. 380–391.

Mookherjee, Dilip and Ivan Png, "Optimal auditing, insurance, and redistribution," *The Quarterly Journal of Economics*, 1989, *104* (2), 399–415.

Myerson, R, "Game Theory: Analysis of Conflict Harvard Univ," *Press, Cambridge*, 1991.

Okuno-Fujiwara, Masahiro, Andrew Postlewaite, and Kotaro Suzumura, "Strategic information revelation," *The Review of Economic Studies*, 1990, *57* (1), 25–47.

Roberts, Kenneth M, "Mediating the evaluative-facilitative debate: Why both parties are wrong and a proposal for settlement," *Loy. U. Chi. LJ*, 2007, *39*, 187.

Shin, Hyun Song, "The burden of proof in a game of persuasion," *Journal of Economic Theory*, 1994, *64* (1), 253–264.

Spier, Kathryn E, "Pretrial bargaining and the design of fee-shifting rules," *The RAND Journal of Economics*, 1994, pp. 197–214.

Townsend, Robert M, "Optimal contracts and competitive markets with costly state verification," *Journal of Economic theory*, 1979, *21* (2), 265–293.

White, Michelle J, "Bankruptcy law," *Handbook of law and economics*, 2007, *2*, 1013–1072.