Money Burning Improves Mediated Communication*

Yi Liu[†] Yang Yu[‡]

July 29, 2025

Abstract

We propose a novel mediated communication protocol where the Sender can both transmit messages and burn money. We assume that the Sender has state-independent preferences. Our main result proves that increasing the budget must *strictly* improve the Sender's payoff, unless the payoff collapses to the cheap talk value when the budget exceeds the range of the Sender's value function. By this result, we can further show that the money-burning tactic must *strictly* improve the Sender's payoff in almost all scenarios unless the commitment is valueless. We also characterize the Sender's maximum equilibrium payoff. This characterization uncovers a connection to two types of robust Bayesian persuasion. Furthermore, our communication protocol directly applies to Web 3.0 communities, clarifying the value of commitment in these contexts.

Keywords: mediated communication, money burning, mechanism design, commitment

JEL Classification: D82, D83

*We thank Nemanja Antic, Roberto Corrao, Yifan Dai, Piotr Dworczak, Françoise Forges, Alexander Jakobsen, Zhonghong Kuang, Xiao Lin, Elliot Lipnowski, Harry Pei, Fan Wu, Xingye Wu, Junjie Zhou for their helpful comments and discussions. This paper was circulated with the title "Mediated Information Design with Money Burning for Commitment Power."

[†]Department of Economics, Yale University. `yi.liu.yl2859@yale.edu`

[‡]School of Economics and Management, China University of Petroleum-Beijing. `yangyu@cup.edu.cn`

1 Introduction

What is the best the Sender (he) can do *by himself* in strategically communicating with the uninformed Receiver (she)? Previous studies have explored this question by fixing the level of the Sender's commitment. When the Sender can only commit to the message-generating process rather than his type, Myerson [1982], Forges [1986] propose a communication protocol known as the mediated communication (MD) protocol. Since then, most related studies have focused on characterizing the MD protocol, including exploring the associated optimal Sender's equilibrium payoff (Salamanca [2021]). However, we argue that an important but overlooked question remains: Is the MD protocol optimal when the Sender has such an intermediate level of commitment?

Here, we propose a novel communication protocol called *mediated communication with money-burning mechanism* (or MDMB for short), whereby the Sender employs the money-burning tactic to obtain credibility. The MDMB is pre-determined by the Sender, consisting of the information-transmission rule for the mediator sending the message and the extent of money-burning based on the Sender's report. In this protocol, the Sender can independently implement the money-burning tactic through the mediator, without requiring additional assistance from other parties, given the current level of the Sender's commitment power.

In practice, the MDMB has been applied as the communication paradigm in the emerging Web 3.0 economy as proposed by Drakopoulos et al. [2023]. For instance Shaker et al. [2021], some Web 3.0 financial companies sell their products to consumers through smart contracts. Those companies input the risk information into the smart contracts. The smart contracts generate the risk assessment results with randomness to the consumers according to the pre-decided and transparent algorithms. Here, those companies are the Sender while consumers are Receivers. The smart contracts are the mediator. The gas fee or token transferred from the financial companies to the consumers can be regarded as the money-burning tactic. In all Web 3.0 business practices, the transparent and auto-processed algorithms simultaneously play the role of a trustworthy mediator for enforcing the pre-committed message-design protocol and enabling the money burning by gas fee. Therefore, MDMB is a general communication protocol in the Web 3.0 economy.

In this paper, we characterize the MBMD under a substantive assumption: the Sender has state-independent preferences over the Receiver's actions. This transparent-motives assumption simplifies the analysis while retaining substantial real-world economic applications (Chakraborty and Harbaugh [2010], Lipnowski and Ravid [2020], Lipnowski et al. [2022]). Furthermore, in order to capture the limited liability or bounded credit nature for money burning, we set an upper bound on the Sender's budget for money burning. We primarily address two questions. First, what is the Sender's optimal equilibrium payoff for the MDMB, and what is the corresponding optimal design? Second, and more importantly, under what conditions can the money-burning tactic *strictly* improve the Sender's payoff during the MD, or what is the comparative statics of the Sender's payoff with respect to the budget?[1] The second question is crucial for understanding the effect of money burning in MD and justifying the use of the money-burning tactic.

The combination of mechanism design and information transmission significantly increases the difficulty of analyzing the optimal MDMB. The burned money is able to contain the piece of information about the Sender's type, allowing the Receiver to select her action accordingly. Consequently, the outcomes of MDMBs can influence the Receiver's actions in ways that the Sender cannot commit to, which violates the full-commitment condition necessary for applying the revelation principle (Bester and Strausz [2001]). Thus, the revelation principle cannot be directly applied to simplify the optimal MDMB mechanism problem restricted by the equilibrium constraints.

Here, we develop a new revelation principle that decomposes the MBMD into a separable sequential process: the Sender first designs the message and then determines the amount of money to be burned based on that message. This design ensures that the burned money contains no additional information about the Sender's type beyond what is conveyed in the message. It also allows the Receiver to directly check the burned amount based on the received messages, which in turn reduces the commitment requirement for burning money. This decomposition enables us to employ a belief-based approach, transforming the optimal MDMB problem constrained by equilibrium conditions into a simpler optimization problem subject to incentive-compatible, obedience, Bayes-plausible, and

[1] It is clear that when the budget of money burning increases, the Sender's payoff will also weakly increase.

budgetary constraints. The revelation principle for MDMB introduces a novel framework for mechanism design under limited commitment, complementing the revelation principle for the dynamic mechanism selection game developed in Doval and Skreta [2022].

The reduced programming problem enables us to characterize the Sender's maximum equilibrium payoff. Note that under the transparent motives assumption, the incentive-compatible constraints require that all Sender's interim payoffs under an MDMB are the same. The Lagrange multipliers of the incentive-compatible constraints are then the affine combination of the types, where each Lagrange multiplier distorts the corresponding Sender's interim payoff, and the Lagrangian function is then the summation over all the distorted interim payoffs that we denoted as the Sender's generalized subjective payoff function. Theorem 1 states that the Sender's maximum equilibrium payoff is the concavification value of the Sender's generalized subjective payoff function where the corresponding Lagrange multipliers are the affine combination that minimizes such concavification value.

Further analysis on the Sender's generalized subjective payoff function reveals how the signaling scheme and money-burning tactics function. Theorem 1 indicates that if the sum of the distorted probabilities (adjusted by the Lagrange multiplier) at a posterior is positive, the Sender at this posterior will not burn money and will induce the Receiver to take the best response that favors him; in some other scenarios when the sum is negative, the Sender will burn as much money as possible and will induce the Receiver to take the best response that disfavors him. We refer to posteriors of the former as belonging to *the persuasion group*, and those of the latter as belonging to *the credibility-gaining group*. Intuitively, in the persuasion group, the Sender aims to achieve a favorable outcome through persuasion alone, without incurring any cost; whereas in the credibility-gaining group, the Sender uses his entire budget to maximize credibility. Proposition 2 demonstrates that to identify the optimal MDMB, one must simultaneously determine the persuasion group, the credibility-gaining group, and verify the optimality of the signaling scheme.

We then characterize the existence of the credibility-gaining group, which is closely tied to the effect of money burning, as formalized in Proposition 2. Our main result (Theorem 2) establishes that, under a generic condition, when the budget exceeds the range of the Sender's value function, increasing the budget expands the equilibrium payoff set and strictly improves the Sender's payoffs—unless, at that budget level, the Sender's payoff collapses to the cheap talk benchmark. This implies that if money burning does not enhance MD as the budget increases, then the combination of money burning and MD offers no improvement in communication efficiency over cheap talk. Furthermore, under the same condition, we show that for almost all prior beliefs, the Sender's maximum payoff must strictly exceed the cheap talk payoff. Accordingly, Proposition 3 states that when the budget surpasses the threshold, money burning must strictly improve MD and enlarge the equilibrium payoff set for almost all priors, unless commitment has no value.

We further characterize the value of MDMB, which is the supreme payoff of the Sender when the budget is no longer a constraint. We find that as the budget increases, the role of the credibility-gaining group diminishes, as the negative Lagrange multipliers converge to zero. Consequently, the Lagrange multipliers reflect the Sender's subjective beliefs. Therefore, the value of MDMB equals the minimum among all concavification values of the Sender's subjective payoff functions (Proposition 4).[2] This conclusion also reveals a connection between MDMB and robust Bayesian persuasion problems. We further show that the value of MDMB equals the maximum payoff of a cautious Sender with full commitment power.[3] We also demonstrate that the value of MDMB equals the Sender's payoff under the worst-case subjective prior in Bayesian persuasion with heterogeneous beliefs (Alonso and Câmara [2016]). These findings suggest that the analysis of robust Bayesian persuasion can be used to facilitate the MDMB analysis and also justifies the assumptions on min-max utility in the robust Bayesian persuasion problem. This characterization also captures the persuasion group when the budget is sufficiently large. The persuasion group comprises posteriors supported by the optimal scheme in Cautious Bayesian persuasion or the optimal signaling strategy under the worst subjective prior.

As an important application, since MDMB serves as the general model capturing the communication process in the Web 3.0 economy when commitment power is absent, the gap between the values of MDMB and Bayesian Persuasion highlights the necessity of commitment and is referred to as the *refined value of commitment* in the Web 3.0

[2]This refers to the convex combination—i.e., the expectation under the Sender's subjective beliefs—of the truth-adjusted welfare functions defined in Doval and Smolin [2024].

[3]A cautious Sender is one who maximizes his minimum payoff across types (Doval and Smolin [2021]).

economy. Proposition 7 demonstrates that commitment in Web 3.0 is valuable if and only if commitment in the conventional economy is valuable. In spite of this, Corollary 2 still indicates that the refined value of commitment in Web 3.0 is less than the value of commitment in the conventional economy in almost all scenarios. Therefore, the algorithms used in Web 3.0 partially mitigate the losses associated with the absence of commitment power.

1.1 Related Literature

This paper proposes a novel communication protocol where adopting the money-burning tactic enhances the credibility in a limited commitment environment. Our work contributes to the literature that studies communication protocols under various levels of commitment power. When the Sender has no commitment, Crawford and Sobel [1982] introduce the cheap talk (CT) model. In contrast, when the Sender has full commitment, Kamenica and Gentzkow [2011] present the Bayesian persuasion (BP) model. However, obtaining such commitment in real-world contexts is challenging, and the unverifiable nature of cheap talk renders it inefficient in many scenarios. In addition to those literature, Min [2021], Lipnowski et al. [2022] concentrate on the case where Sender's commitment power has a Bernoulli distribution on full commitment and no commitment; and Lin and Liu [2024] examine the situation where the Sender cannot commit to the message-generating process but he can commit to the marginal distribution of types and messages. Bergemann and Morris [2019] summarize information design problems involving persuasion and mediation. Furthermore, our paper is closely related to studies that analyze the effect of commitment on information design under various communication protocols. Fréchette et al. [2022] investigate the effect of communication with different levels of commitment power through experimental methods. Additionally, Corrao and Dai [2023] comparatively analyze different communication protocols at various levels of commitment power, but they do not take money burning into account. Therefore, the main difference is that their work compares protocols across BP, MD, and CT, whereas our analysis can be regarded as a comparison among different MD protocols (different budgets).

The main contribution of this paper is to extend the domain of mediated communication problems. Previous research, such as Salamanca [2021], illustrates the optimal equilibrium payoff of the Sender through MD without burning money. Drakopoulos et al. [2023] establish a blockchain system as a mediator, demonstrating that designing costly messages can improve MD under transparent motives, but they do not identify the optimal Sender's communication efficiency in general as we do and they also do not identify the condition where costly messages improve MD. Additionally, several studies, including Goltsman et al. [2009], Ivanov [2014], have identified the optimal mediation plan for the Receiver. Furthermore, Ivanov [2014] compares the outcomes of mediated communication and cheap talk.

Our paper is also related to the literature on communication with transfers. Some studies discuss cheap talk or signaling games involving money burning, Austen-Smith and Banks [2000], Kartik [2007], Karamychev and Visser [2017], Gersbach [2004], noting that the Sender nearly cannot improve the credibility of cheap talk by the money burning tactic, and chooses not to burn money even in state-independent preferences environments. In their works on cheap talk and signaling games, they find that money burning is a powerful tool to refine the equilibrium and get a more precise equilibrium set. However, in our work, money-burning mechanism plays a key role in enhancing Sender's credibility, expanding the equilibrium payoff set, and even obtaining better communication efficiency. Kolotilin and Li [2021] investigate the application of monetary transfers in repeated cheap talk settings, while Sadakane [2023] examines a model featuring repeated cheap talk games with monetary transfers from the Receiver to the Sender. This latter study observes that the equilibrium set in such settings is larger than that of the original long-term cheap talk setting. Corrao [2023] analyzes the mediation market and characterizes the information and market outcomes of the revenue-maximizing mediator and the Sender-optimal mediator. Additionally, several studies focus on Bayesian persuasion involving transferable utility and the cost of information, such as Li and Shi [2017], Bergemann et al. [2018]. Perez-Richet and Skreta [2022] investigate the Receiver-optimal experiment under the condition that the Sender can costly falsify his private type.

Another important category of literature related to us is about mechanism design with limited commitment. Liu and Wu [2024] examine the implementation problem in general outcome-contingent settings, which is a more generalized context than ours. Bester and Strausz [2001] show that the revelation principle fails to hold in a limited commitment environment, where the principal cannot fully commit to the outcome induced by the mechanism. Doval and Skreta [2022] provide the general revelation principle for limited commitment mechanism design. However, they require the randomization over the mechanism design part given the information outcome. In our revelation principle, we further prove that the money-burning part can be deterministic given the information outcome.

2 Model

In Section 2.1, we develop the basic model of the Sender-Receiver game, and in Section 2.2, we introduce the methodology for simplifying the Sender's programming problem.

2.1 Basic Setup

Primitives. The basic game consists of two players: the Sender (he) and the Receiver (she). The Sender has the private information θ, which denotes his type and belongs to a finite set Θ. Type θ is drawn according to a prior distribution $\mu_0 \in \Delta(\Theta)$, which is common knowledge. The Receiver can choose an action a from a finite set A. The Sender's value function only depends on the Receiver's action, which is denoted as $v(\cdot): A \to \mathbb{R}$; and the Receiver's value function is $u(\cdot,\cdot): A \times \Theta \to \mathbb{R}$.

Communication with money-burning mechanisms. Before the game, the Sender can commit to a mediated communication with money-burning mechanism (or MDMB), which consists of an input set M, an output message set S, and a corresponding mechanism $\phi: M \to \Delta(S \times T)$, where $T \triangleq [0,C]$ and $C \in \mathbb{R}_{\geq 0} \cup +\infty$ denotes the Sender's exogenously given total budget. The MDMB prescribes how the Sender designs the message and determines the money-burning amount based on his private input. We restrict attention to the case where M, S, and the support of ϕ are all finite.

The Sender-Receiver game. At the beginning of the game, the Sender commits to the MDMB (M,S,ϕ) with a mediator. Then, the Sender's type θ is realized, and the Sender sends an input message $m \in M$ to the mediator. With probability $\phi(s,t|m)$, the mediator sends an output message $s \in S$ to the Receiver and burns $t \in T$ money from the Sender's account. After observing the money burned by the Sender and the message m, the Receiver updates her belief and chooses an action $a \in A$. The ex-post payoffs of the Sender and the Receiver are $v(a)-t$ and $u(a,\theta)$, respectively.

Remark 1. In our model, the Receiver observes the money burning. Compared to the model where only the Sender knows the money burning, our model requires less commitment from the Sender and offers a broader range of applications, especially when the mediator is replaced by transparent algorithms or when the money is provided as a subsidy to the Receiver.

After the commitment of MDMB (M,S,ϕ), the sub-game is denoted as $\mathcal{G}_{(M,S,\phi)}$.

Beliefs and strategies. The Sender's strategy in $\mathcal{G}_{(M,S,\phi)}$ prescribes a transition probability $\sigma:\Theta \to \Delta(M)$. As for the Receiver, the output message s and the money burning amount t together form the information set. For each information set (s,t), the Receiver's strategy prescribes a transition probability $\alpha:S\times T \to \Delta(A)$. In each information set (s,t), the Receiver must form a belief $\mu:S\times T\to\Delta(\Theta)$. We call the triple (σ,α,μ) an assessment.

Equilibrium. In this paper, we use Perfect Bayesian equilibrium (henceforth, PBE) as the solution concept of game $\mathcal{G}_{(M,S,\phi)}$. We denote the set of PBE of game $\mathcal{G}_{(M,S,\phi)}$ as $\mathcal{E}[\mathcal{G}_{(M,S,\phi)}]$. Formally, $(\sigma^*,\alpha^*,\mu^*)\in\mathcal{E}[\mathcal{G}_{(M,S,\phi)}]$ if the following three conditions hold:

Sender's optimality: for any $\theta\in\Theta$,

$$\sigma^*(\theta)\in\arg\max_{\sigma(\theta)\in\Delta(M)}\sum_{m\in M,s\in S,t\in T,a\in A}\sigma(m|\theta)\phi(s,t|m)\alpha^*(a|s,t)(v(a)-t). \tag{1}$$

Receiver's optimality: for any $s\in S,t\in T$,

$$\alpha^*(s,t)\in\arg\max_{\alpha(s,t)\in\Delta(A)}\sum_{\theta\in\Theta,a\in A}\mu^*(\theta|s,t)\alpha(a|s,t)u(a,\theta). \tag{2}$$

Bayesian updating: for any $s\in S,t\in T$ and $\theta\in\Theta$,

$$\mu^*(\theta|s,t)\sum_{\theta'\in\Theta,m\in M}\mu_0(\theta')\sigma^*(m|\theta')\phi(s,t|m)=\mu_0(\theta)\sum_{m\in M}\sigma^*(m|\theta)\phi(s,t|m). \tag{3}$$

Communication efficiency. The Sender seeks to maximize his ex ante expected payoff across all possible PBEs. We can formulate the Sender's optimization problem as follows:

$$\begin{aligned}&\sup_{M,S,\phi}\sum_{\theta\in\Theta}\mu_0(\theta)\sum_{m\in M,s\in S,t\in T,a\in A}\sigma^*(m|\theta)\phi(s,t|m)\alpha^*(a|s,t)(v(a)-t)\\&s.t.(\sigma^*,\alpha^*,\mu^*)\in\mathcal{E}[\mathcal{G}_{(M,S,\phi)}].\end{aligned} \tag{4}$$

We focus on the effects of the prior μ_0 and the budget C on the Sender's maximum equilibrium payoff. Thus, we denote the value of (4) by $\mathcal{V}_C^*(\mu_0)$. When $C=0$, $\mathcal{V}_0^*(\mu_0)$ represents *the value of MD*; while when $C=+\infty$, we simply write $\mathcal{V}^*(\mu_0)$, representing *the value of MDMB*.

2.2 Simplifying the Problem

The optimization problem (4) is complex due to its non-convex and equilibrium-selection complexity. In addition, the revelation principle developed by Myerson [1982], Forges [1986] cannot be applied to simplify the problem since in MDMB, the Sender is not contractible with the Receiver and the money burning may contain additional information that influences the Receiver's action. In this section, we develop a new technique to simplify the problem. Inspired by Doval and Skreta [2022], we apply the method of canonical mechanisms and canonical assessments to develop a new revelation principle for MDMB.

Here, we first formally define the canonical MDMBs and canonical assessments.

Definition 1 (Canonical MDMBs)**.** *An MDMB is canonical if $M=\Theta$, $S=\Delta(\Theta)$, and there exists a signaling scheme $\pi:\Theta\to\Delta(\Delta(\Theta))$ and a deterministic function $x:\Delta(\Theta)\to T$ such that π satisfies the Bayesian updating condition.*[4] *and $\phi(u,x(u)|\theta)=\pi(u|\theta)$ for all $\theta\in\Theta$ and $u\in\boldsymbol{supp}\{\pi(\theta)\}$.*

In canonical MDMBs, the input sets are the type sets while the output sets are the sets of distributions of types. The output message of the canonical MDMB contains *all* information transmitted to the Receiver, and the amount of money burning does not provide any additional information. Hence, ϕ in a canonical MDMB can be decomposed into two parts. The first part is a *signaling scheme* π, and the second part is a *money-burning scheme* x which is contingent on the output message. This decomposition has a similar structure to the revelation principle in Doval and Skreta [2022]. Henceforth, we use (π,x) to refer to a canonical MDMB.

In a canonical MDMB, the canonical assessment ensures that the Sender's strategy is truthful-telling and the Receiver's posterior belief coincides with the output message.

[4] $u(\theta)\sum_{\theta'\in\Theta}\mu_0(\theta')\pi(u|\theta')=\mu_0(\theta)\pi(u|\theta)$ for all $u\in\Delta(\Theta)$.

Definition 2 (Canonical assessments)**.** *For a canonical MDMB, an assessment* (σ,α,μ) *is canonical if* $\sigma(\theta|\theta)=1$ *and* $\mu(u,x(u))=u$ *for any* $u\in \boldsymbol{supp}\{\pi(\theta)\}$.

The following proposition shows that every MDMB has a corresponding canonical MDMB that maintains the same expected payoffs of the Sender and the Receiver. This proposition allows us to focus exclusively on the canonical MDMBs and the associated canonical assessment without loss of generality, as depicted in Figure 1.

Proposition 1. *For any MDMB* (M,S,ϕ) *and* $(\sigma,\alpha,\mu)\in\mathcal{E}[\mathcal{G}_{(M,S,\phi)}]$, *there exists a canonical MDMB* (π,x) *and a canonical assessment* $(\sigma^*,\alpha^*,\mu^*)\in\mathcal{E}[\mathcal{G}_{(\pi,x)}]$ *such that the two assessments are payoff-equivalent.*[5]

$$\underset{\theta}{\text{Sender}}\quad M\xrightarrow{\phi(\cdot|m)}\quad S\times T\quad\text{Receiver}\quad\Leftrightarrow\quad\underset{\theta}{\text{Sender}}\quad\Theta\xrightarrow{\pi(\cdot|\theta)}\Delta(\Theta)\xrightarrow{x(\cdot)}T\quad\text{Receiver}$$

Figure 1: Revelation Principle

The core idea of Proposition 1 is that for any MDMB (M,S,ϕ) with its corresponding PBE assessment (σ,α,μ), each pair $(s,t)\in S\times T$ can be viewed as a new signal $s'\in S'$, where $S'=S\times T$ and $s'=(s,t)$, with the associated money burning defined as $x(s')=t$. Thus, we have $\mu(s,t)=\mu(s')$, leading to a decomposition structure. Furthermore, since the Receiver's actions depend on beliefs, each signal s' can be replaced with the posterior belief $\mu(s')$, allowing us to substitute the signal space S' with the belief space $\Delta(\Theta)$. The remaining challenge is handling cases where different signals s' correspond to the same posterior belief but require a consistent money burning amount. As shown in the appendix, by defining the money burning at each posterior $\mu(s')$ as the expected total payment conditional on $\mu(s')$, the payoff-equivalence property is maintained. A central feature of Proposition 1 is its decomposition structure, which facilitates the calculation of the value in (4). Moreover, interpreting the signal set as a set of posterior beliefs not only simplifies the notation but also lays the foundation for a belief-based framework.

In the rest of this section, we explain how to apply Proposition 1 to simplify the optimization problem via the belief-based approach. This approach considers the ex ante

[5] Two assessments are payoff-equivalent if they induce the same ex ante payoffs of the Sender and the Receiver.

distribution over Receiver's posterior beliefs, $p\in\Delta(\Delta(\Theta))$. By Kamenica and Gentzkow [2011], the Bayesian updating constraint is equivalent to requiring $p\in BP(\mu_0)$ where $BP(\mu_0)\triangleq\left\{p\in\Delta(\Delta(\Theta))\middle|\int_\mu \mu dp(\mu)=\mu_0\right\}$. In other words, there is a mapping between signaling scheme π and $p\in BP(\mu_0)$ through Bayesian updating. Therefore, we also call $p\in BP(\mu_0)$ the signaling scheme.

For any posterior belief $\mu\in\Delta(\Theta)$, the Receiver's best response can be summarized in $RO(\mu)\triangleq\left\{\alpha\in\Delta(A)\middle|\mathbf{supp}\{\alpha\}\subset\mathrm{argmax}_{a'\in A}\sum_{\theta\in\Theta}\mu(\theta)u(a',\theta)\right\}$. Therefore, in PBEs, the Sender's value of the Receiver's action can be summarized in the Sender's *belief-value correspondence*, $\mathbb{V}:\Delta(\Theta)\rightrightarrows\mathbb{R}$, where $\mathbb{V}(\mu)\triangleq\left\{q\middle|\exists\alpha\in RO(\mu),q=\sum_{a\in A}\alpha(a)v(a)\right\}$. Hence, the Receiver's optimality constraints can be converted to obedience constraints that $V(\mu)\in\mathbb{V}(\mu)$ where $V(\mu)$ denotes the Sender's ex-post payoff at posterior μ.

Regarding the Sender's optimality constraints, suppose the canonical MDMB (π,x) induces a distribution over posteriors denoted by p. The Sender's expected payoff when reporting type θ is given by $\int_\mu \frac{\mu(\theta)}{\mu_0(\theta)}(V(\mu)-x(\mu))dp(\mu)$.[6] Therefore, the Sender's optimality condition requires type-indifference: that is, there exists a constant $k\in\mathbb{R}$ such that $\int_\mu \frac{\mu(\theta)}{\mu_0(\theta)}(V(\mu)-x(\mu))dp(\mu)=k$ for all $\theta\in\Theta$. We conclude with the following corollary.

Corollary 1. *$\mathcal{V}_C^*(\mu_0)$ can be calculated using the following optimization problem.*

$$\begin{aligned}
\max\quad & k && && (5)\\
s.t\quad & k=\int_\mu \frac{\mu(\theta)}{\mu_0(\theta)}(V(\mu)-x(\mu))dp(\mu) && \forall\theta\in\Theta && \text{(IC)}\\
& p\in BP(\mu_0) && && \text{(BP)}\\
& V(\mu)\in\mathbb{V}(\mu) && \forall\mu\in\Delta(\Theta) && \text{(O)}\\
& 0\leq x(\mu)\leq C && \forall\mu\in\Delta(\Theta) && \text{(Bgt)}
\end{aligned}$$

In this paper, we will use the following salesman's example to illustrate our main characterizations and intuitions. In this example, the Sender cannot benefit from cheap talk or mediated communication.

Example 1. *Consider a salesman problem between a consumer and a salesman. The consumer decides whether to purchase a product, whose quality is either high (θ^H) or low*

[6] The Bayes plausibility condition requires that $\pi(\mu|\theta)=\frac{\mu(\theta)}{\mu_0(\theta)}dp(\mu)$.

(θ^L). The salesman has private information about the true quality of the product, while the consumer has a prior belief that the product is high-quality with probability $0<\mu_0<\frac{1}{2}$. The market price of the product is fixed at 5. The consumer's payoff depends on the quality of the product: if she purchases a high-quality product, she gets a payoff of 10; if she purchases a low-quality product, she gets a payoff of 0. The salesman's payoff is determined by the consumer's decision: he receives a reward of 1 from the producer as his commission for selling the product and receives nothing otherwise.

3 Sender-optimal Equilibrium Payoffs

In this section, we characterize the Sender's maximum equilibrium payoff $\mathcal{V}_C^*(\mu_0)$ and the economics of optimal MDMB. Using the characterization, we can find the value of MDMB $\mathcal{V}^*(\mu_0)$ and discuss its implications.

We define the Sender's *generalized subjective payoff function* at posterior $\mu\in\Delta(\Theta)$, denoted by $\hat{V}_{\lambda,C}(\mu)$, as follows

$$\hat{V}_{\lambda,C}(\mu)\triangleq\max\{\sum_{\theta\in\Theta}\lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}\max\mathbb{V}(\mu),\sum_{\theta\in\Theta}\lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}(\min\mathbb{V}(\mu)-C)\}. \tag{6}$$

Here, $\lambda:\Theta\to\mathbb{R}$ is the Lagrange multiplier of constraints (IC) that satisfies $\sum_\theta\lambda(\theta)=1$, or in other words $\lambda\in\text{aff}(\Theta)$. In addition, let $cav(f)$ denote the concave envelope of function f. The characterization of the Sender's maximum equilibrium payoff is based on the generalized subjective payoff function and the concave envelope.

Theorem 1.

$$\mathcal{V}_C^*(\mu_0)=\min_{\lambda\in\mathit{aff}(\Theta)}cav(\hat{V}_{\lambda,C})(\mu_0).$$

To compute the Sender's maximum equilibrium payoff, we conduct the two-step optimization approach. First, given any signaling scheme $p\in BP(\mu_0)$, we calculate the optimal money-burning scheme and the Receiver's response $V(\mu)$. Note that at this stage, the variables $x(\mu)$, $V(\mu)$, and k are linear in the objective and constraints. By introducing the Lagrange multiplier of (IC) λ, the Lagrangian function is

$$k+\sum_{\theta\in\Theta}\lambda(\theta)\left[\int_\mu\frac{\mu(\theta)}{\mu_0(\theta)}(V(\mu)-x(\mu))dp(\mu)-k\right].$$

Thus, in the optimal solution, $(V(\mu),x(\mu))$ can be either $(\max\mathbb{V}(\mu),0)$ or $(\min\mathbb{V}(\mu),C)$. We know that $x(\mu)=0$ and $V(\mu)=\max\mathbb{V}(\mu)$ when $\sum_\theta\lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}>0$; $x(\mu)=C$ and

$V(\mu)=\max\mathbb{V}(\mu)$ when $\sum_\theta \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}<0$. In addition, by the first-order condition of variable k, the Lagrange multiplier must satisfy that $\lambda\in \text{aff}(\Theta)$. Therefore, after optimizing over the money burning scheme and the Receiver's response, the objective becomes

$$\min_{\lambda\in\text{aff}(\Theta)}\int_\mu \max\{\sum_{\theta\in\Theta}\lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}\max\mathbb{V}(\mu),\sum_{\theta\in\Theta}\lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}(\min\mathbb{V}(\mu)-C)\}dp(\mu)$$

and the Sender aims to maximize the objective over $p\in BP(\mu_0)$. The second step applies Sion's minimax theorem, which permits the interchange of $\min_{\lambda\in\text{aff}(\Theta)}$ and $\max_{p\in BP(\mu_0)}$.

The generalized subjective payoff function $\hat{V}_{\lambda,C}(\mu)$ also reveals the economic logic behind the optimal MDMB, denoted by (p^*,x^*). The structure of $\hat{V}_{\lambda,C}$ allows us to classify the output messages ($\mathbf{supp}\{p^*\}$) of the optimal MDMB into three distinct groups. Suppose λ^* solves $\min_{\lambda\in\text{aff}(\Theta)} cav(\hat{V}_{\lambda,C})(\mu_0)$. We can then identify the money-burning amount and the Receiver's best response across two core categories: the *persuasion group* and the *credibility-gaining group*. Given a posterior belief μ, we say μ belongs to the *persuasion group* if $\sum_\theta \lambda^*(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}>0$. At such beliefs, the Sender seeks to maximize persuasive payoff, and the associated messages are costless. Conversely, μ falls into the *credibility-gaining group* if $\sum_\theta \lambda^*(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}<0$. Here, the Sender aims to get credibility by sacrificing his payoff. Formally, when a message lies in the persuasion group, the Sender prefers the Receiver to take the action most favorable to him; in contrast, for the credibility-gaining group, the Sender maximizes money burning and prefers the Receiver to take the action least favorable to him. Beyond these two, when $\sum_\theta \lambda^*(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}=0$, the message belongs to a third group, which balances interim payoffs to satisfy the incentive compatibility condition (IC). For these beliefs, the money-burning amount and Receiver's best response cannot be simply determined, but this indeterminacy does not affect the Sender's equilibrium payoff since $\hat{V}_{\lambda,C}(\mu)=0$ for such μ. Based on the above analysis, we present the following proposition that characterizes the structure of the optimal MDMB.

Proposition 2. *The solution consisting of an MDMB (p,x) and V that is feasible to the program (5) is optimal if and only if there exists $\lambda\in$ aff(Θ) such that:*

*(i) for any $\mu\in$ **supp**$\{p\}$ and $\sum_{\theta\in\Theta}\lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}>0$, $V(\mu)=\max\mathbb{V}(\mu)$ and $x(\mu)=0$;*

*(ii) for any $\mu\in$ **supp**$\{p\}$ and $\sum_{\theta\in\Theta}\lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}<0$, $V(\mu)=\min\mathbb{V}(\mu)$ and $x(\mu)=C$;*

(iii) $p\in$ $\text{argmax}_{\tau\in BP(\mu_0)}\int_\mu \hat{V}_{\lambda,C}(\mu)d\tau(\mu)$.

The necessity of this proposition has been discussed in the last paragraph. The sufficiency of this proposition is based on the minimax theory. Let

$$\mathcal{L}_C[\lambda,(p,x,V)] = \int_\mu \sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}(V(\mu)-x(\mu))dp(\mu).$$

Following the discussions below Theorem 1, we have shown that

$$\min_{\lambda\in\text{aff}(\Theta)} \max_{V\in\mathbb{V},x(\mu)\in T,p\in BP(\mu_0)} \mathcal{L}_C[\lambda,(p,x,V)] = \max_{V\in\mathbb{V},x(\mu)\in T,p\in BP(\mu_0)} \min_{\lambda\in\text{aff}(\Theta)} \mathcal{L}_C[\lambda,(p,x,V)].$$

Therefore, the optimal solution of program (5), denoted by (p^*,x^*,V), along with the corresponding Lagrange multiplier, can be interpreted as a Nash equilibrium of a zero-sum game. In this game, one player selects the MDMB (p,x) and the Receiver's best responses V to maximize $\mathcal{L}_C[\lambda,(p,x,V)]$, while the other player chooses $\lambda\in\text{aff}(\Theta)$ to minimize it. Hence, the condition for optimality is equivalent to that of a Nash equilibrium. Consequently, for (p^*,x^*,V), it is not possible to further reduce $\mathcal{L}_C[\lambda,(p^*,x^*,V)]$ by selecting a different $\lambda\in\text{aff}(\Theta)$, because (IC) implies that $\mathcal{L}_C[\lambda,(p^*,x^*,V)]$ remains constant over $\text{aff}(\Theta)$. Conversely, given a $\lambda\in\text{aff}(\Theta)$ and (p^*,x^*,V) satisfying conditions (i), (ii), and (iii), no other feasible solution can yield a higher value than $\mathcal{L}_C[\lambda,(p^*,x^*,V)]$. Therefore, Proposition 2 holds, as the feasible solution and the associated multiplier together form a Nash equilibrium.

4 Money Burning Improves Mediated Communication

To have a better understanding of the effect of money burning and the budget C, we have to characterize the property of the credibility-gaining group. In this section, we focus on the existence of the credibility-gaining group that is closely related to the comparative statics of the budget C.

Intuitively, the credibility-gaining group is non-empty for almost all $C\in(C_1,C_2)$ if and only if $\mathcal{V}^*_{C_2}(\mu_0) > \mathcal{V}^*_{C_1}(\mu_0)$, as suggested by an approximation inspired by the "envelope theorem":

$$\frac{\partial \mathcal{V}^*_C(\mu_0)}{\partial C} = \int_{\mu:\sum_\theta \lambda^*(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}<0} -\sum_\theta \lambda^*(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}dp^*(\mu)$$

where λ^* denotes the Lagrange multiplier and p^* represents the optimal signaling scheme under budget C in the MDMB problem. Therefore, $\frac{\partial \mathcal{V}^*_C(\mu_0)}{\partial C} > 0$ if and only if the credibility-gaining group is non-empty, at least in an approximate sense.[7] Hence, the

[7]This statement may not strictly hold, as it does not directly apply the envelope theorem. However,

existence of a credibility-gaining group is equivalent to the strict monotonicity of $\mathcal{V}_C^*(\mu_0)$ with respect to C.

To introduce our approach to characterizing the credibility-gaining group, we impose the following assumption on the Receiver's payoff functions. This assumption holds generically.[8] Whenever this assumption holds, we can show that, generically, the credibility-gaining group is non-empty, implying that money burning improves mediated communication.

Assumption 1. *For any belief μ and any $a \in RO(\mu)$, there exists μ' such that $\boldsymbol{supp}\{\mu'\} = \boldsymbol{supp}\{\mu\}$ and $RO(\mu') = \{a\}$.*[9]

Let $\mathcal{V}_{CT}^*(\mu_0)$ denote the Sender's maximum payoff under cheap talk equilibrium given the prior μ_0. According to Lipnowski and Ravid [2020], $\mathcal{V}_{CT}^*(\mu_0) = qcav(\mathbb{V})(\mu_0)$ where $qcav(f)$ is the quasi-concave envelope of function f. Then, we state the following theorem, which asserts that when the budget C exceeds the range of ex-post payoffs, one of two outcomes must occur: either the credibility-gaining group is non-empty, or it is empty for all weakly smaller budgets and $\mathcal{V}_C^*(\mu_0)$ collapses to $\mathcal{V}_{CT}^*(\mu_0)$.

Theorem 2. *Under Assumption 1, for any $C_2 > C_1 > \max_{\mu, V(\mu)\in\mathbb{V}(\mu)} V(\mu) - \min_{\mu, V(\mu)\in\mathbb{V}(\mu)} V(\mu)$, either $\mathcal{V}_{C_2}^*(\mu_0) > \mathcal{V}_{C_1}^*(\mu_0)$ or $\mathcal{V}_{C_2}^*(\mu_0) = \mathcal{V}_0^*(\mu_0) = \mathcal{V}_{CT}^*(\mu_0)$.*

Proof Overview. The proof of Theorem 2 has the following sketch. We first show that if the credibility-gaining group is empty under budget C_2, then the Sender's maximum equilibrium payoff under C_2 must collapse to the value of cheap talk. To establish this, we partition the set of posterior beliefs in the optimal MDMB into two subsets: the persuasion group and the balancing group, where the latter consists of posteriors not included in the persuasion group. Now, we require two main claims. Claim 2 shows that the Sender receives the same payoff at all posteriors in the persuasion group. Claim 3 further establishes that if the budget exceeds the range of ex-post payoffs, the aggregate

a related argument is rigorously established in the proof of Theorem 2 in the appendix.

[8] The Lebesgue measure of the set of Receiver payoff functions that do not satisfy this condition is zero in the entire function space.

[9] This condition is also present in Lipnowski et al. [2024], where they employ it as a generic sufficient criterion for the uniqueness of the Sender's payoff under perfect Bayesian equilibrium in the Bayesian persuasion game.

belief of the balancing group must lie on the boundary of $\Delta(\Theta)$. To recover a cheap talk equilibrium, we focus on the aggregate belief of the balancing group, which has a lower dimension than the prior. If the credibility-gaining group remains empty in the optimal MDMB corresponding to this aggregate belief, the same argument can be recursively applied. The dimension-reduction property guarantees that this recursive procedure terminates. Finally, if at some step the credibility-gaining group becomes non-empty under the optimal MDMB for the aggregate belief, we invoke Assumption 1 to derive a contradiction: the balancing group, being located on the boundary, cannot be part of an optimal MDMB because of the property of concavification. □

Theorem 2 characterizes the existence of the credibility-gaining group when the budget exceeds a certain threshold. In particular, it shows that money burning must expand the equilibrium payoff set and strictly improve the Sender's payoff unless communication via a trusted mediator collapses to that of cheap talk. Building on this characterization, we provide a mild condition under which the credibility-gaining group is guaranteed to be non-empty, thereby ensuring that money burning strictly improves mediated communication.

Proposition 3. *Under Assumption 1, if* $qcav(\mathbb{V})(\mu_0) \neq cav(\mathbb{V})(\mu_0)$ *and there is a sufficiently small* $\varepsilon > 0$ *such that* $qcav(\mathbb{V})(\mu_0 + \varepsilon(\mu - \mu_0)) = qcav(\mathbb{V})(\mu_0)$ *for all* $\mu \in \Delta(\Theta)$*, it follows that* $\mathcal{V}_0^*(\mu_0) < \mathcal{V}_C^*(\mu_0)$ *for any* $C > \max_{\mu, V(\mu) \in \mathbb{V}(\mu)} V(\mu) - \min_{\mu, V(\mu) \in \mathbb{V}(\mu)} V(\mu)$.

By Proposition 3, and following a similar argument to Corollary 2 in Lipnowski and Ravid [2020], we immediately obtain that the value of MDMB serves as a strict upper bound on the maximum Sender's equilibrium payoff achievable under MDMB with finite budgets, for almost all prior beliefs. We summarize this finding in Corollary 2, which shows that, generically, money burning (without budgetary constraints) strictly improves communication efficiency relative to smaller budgets, including the case of classical mediated communication.

Corollary 2. *Under Assumption 1 and for any* $C \in [0, +\infty)$*, for almost all prior beliefs* $\mu_0 \in \Delta(\Theta)$*, either* $\mathcal{V}_{CT}^*(\mu_0) = \max_{\mu \in \Delta(\Theta)} \max \mathbb{V}(\mu)$ *or* $\mathcal{V}_C^*(\mu_0) < \mathcal{V}^*(\mu_0)$.

5 The Value of MDMB

In this section, we apply Theorem 1 to characterize the value of MDMB and give the construction that approaches the value of MDMB based on the philosophy of Proposition 2 when there are no budget constraints. Our results for the value of MDMB also give approximations of the persuasion group when the budget is sufficiently large.

Since $\mathcal{V}^*_C(\mu_0)$ is bounded, $\hat{V}_{\lambda,C}(\mu)$ must also be bounded. Intuitively, when C is sufficiently large, the Lagrange multiplier λ^* associated with the optimal MDMB must satisfy $-\lambda^*(\theta) \leq \mathcal{O}(1/C)$ for all $\theta \in \Theta$. Therefore, although some multipliers may satisfy $\lambda^*(\theta) < 0$, they necessarily converge to zero as the budget increases. Hence, when $C = +\infty$, the value of MDMB can be obtained directly as follows by Theorem 1.

We define the type θ's share of ex-post payoff $\max \mathbb{V}(\mu)$ given the posterior μ as $\hat{V}_\theta(\mu) = \frac{\mu(\theta)}{\mu_0(\theta)} \max \mathbb{V}(\mu)$.[10] Based on this adjusted ex-post payoff, we introduce the Sender's *subjective payoff function* under the posterior μ and the subjective prior $\lambda \in \Delta(\Theta)$ as

$$\hat{V}_\lambda(\mu) \triangleq \mathbb{E}_{\theta \sim \lambda}\{\hat{V}_\theta(\mu)\} = \sum_{\theta \in \Theta} \lambda(\theta) \frac{\mu(\theta)}{\mu_0(\theta)} \max \mathbb{V}(\mu).^{11} \tag{7}$$

Given a canonical MDMB (π, x), we define the *interim signaling payoff* of type θ under signaling scheme π as follows,

$$V_\pi(\theta) \triangleq \sum_{\mu \in \mathbf{supp}\{\pi(\theta)\}} \pi(\mu|\theta) \max \mathbb{V}(\mu). \tag{8}$$

The following proposition characterizes the value of MDMB.

Proposition 4. *$\mathcal{V}^*(\mu_0) = \min_{\lambda \in \Delta(\Theta)} cav(\hat{V}_\lambda)(\mu_0) = \max_\pi \min_{\theta \in \Theta} V_\pi(\theta)$.*

We have already discussed the intuition behind the first equation. For the second equation, we observe that $V_\pi(\theta) \leq \max_\pi \min_{\theta \in \Theta} V_\pi(\theta)$ because negative money burning is not permitted. Then, we can derive the second equation by applying the reasoning in Proposition 2, which shows that we can credibilize any signaling scheme by slightly perturbing it and incorporating money burning.

Claim 1. *Given any signaling scheme $\pi : \Theta \to \Delta(\Delta(\Theta))$, we construct an MDMB $(\bar{\pi}, \bar{x})$,*

$$\bar{\pi}(\mu|\theta) = \begin{cases} (1-\delta)\pi(\mu|\theta) & \mu(\theta) \neq 1 \\ \delta + (1-\delta)\pi(\mu|\theta) & \mu(\theta) = 1 \end{cases}, x(\cdot) = \begin{cases} 0 & \mu(\theta) \neq 1 \\ \frac{1}{\bar{\pi}(\mu|\theta)}(V_{\bar{\pi}}(\theta) - \min_{\theta' \in \Theta} V_{\bar{\pi}}(\theta')) & \mu(\theta) = 1 \end{cases}.$$

[10] This follows from the fact that $\mathbb{E}_{\theta \sim \mu_0}\{\hat{V}_\theta(\mu)\} = \max \mathbb{V}(\mu)$. The notation $\hat{V}_\theta$ is also known as truth-adjust welfare function introduced by Doval and Smolin [2024].

[11] Note that when $\lambda = \mu_0$ the subjective payoff function under the posterior μ becomes $\max \mathbb{V}(\mu)$.

The mechanism $(\bar{\pi}, \bar{x})$ *is incentive-compatible for* $\delta \in (0,1)$ *and the Sender's payoff under this mechanism converges to* $\min_{\theta\in\Theta} V_\pi(\theta)$ *as* $\delta \to 0^+$*.*

Given any signaling scheme π, we retain the original message in the persuasion group and add new messages to the credibility-gaining group that perfectly reveal the types. Although this construction may not be optimal for any finite C (as shown in Proposition 2, since the money-burning amounts differ across messages), as $C \to \infty$, the messages outside the persuasion group play an increasingly negligible role in the optimal design. Therefore, the construction in Claim 1 provides a good approximation, and the Sender's payoff under this construction converges to the value of MDMB as $\delta \to 0^+$ and $C \to +\infty$. Up to this point, we have formally established that $\mathcal{V}^*(\mu_0) = \max_\pi \min_{\theta\in\Theta} V_\pi(\theta)$. The formal proof of the first equation also follows from Sion's minimax theorem.

This section also highlights that when C is sufficiently large, the persuasion group consists of posteriors supported by the signaling scheme that maximizes the Sender's minimum interim payoff. Moreover, the Sender's payoff approaches $\mathcal{V}^*(\mu_0)$, as characterized in Proposition 4, as $C \to \infty$.

5.1 Money Burning Secures Commitment

We demonstrate that the value of MDMB can be interpreted as the values of two types of robust Bayesian persuasion. In both scenarios, the Sender achieves full commitment power while prioritizing his worst-case utility guarantee. Therefore, compared to classical MD, MDMB represents an implementation of robust Bayesian persuasion, which secures greater commitment power for the Sender through the money-burning tactic. Additionally, it offers a theoretical micro-foundation for the max-min utility.

The first equation of Proposition 4 indicates that the value of MDMB is the same as the value of the Sender who has full commitment power and opts for a signaling scheme that maximizes his minimum interim payoff. The model of the Sender opting for such a signaling scheme is named cautious Bayesian persuasion in which the Sender only focuses on his lowest possible welfare, Doval and Smolin [2021, 2024]. In the cautious Bayesian persuasion setting, the Sender has full commitment power but acts robustly to the type realization. The following corollary constitutes our first implication of Proposition 4.

Corollary 3. *The value of MDMB equals the payoff of the Sender with full commitment power but who is cautious.*

The second equation of Proposition 4 relates the value of MDMB to the Sender's payoff in robust Bayesian persuasion with heterogeneous beliefs. To elucidate this, we call a subjective distribution $\lambda \in \Delta(\Theta)$ as the *worst Sender's subjective prior* if it minimizes $cav(\hat{V}_\lambda)(\mu_0)$ which we refer to as *the worst Sender's subjective expected payoff.* Based on the model introduced by Alonso and Câmara [2016] in which the Sender's and the Receiver's subjective priors are heterogeneous, the second equation of Proposition 4 shows that the value of the MDMB coincides with the worst Sender's subjective expected payoff in heterogeneous belief Bayesian persuasion. Hence, the following corollary is our second implication.

Corollary 4. *The value of MDMB $\mathcal{V}^*(\mu_0)$ equals the payoff of the Sender under Bayesian persuasion with heterogeneous priors, in which the Sender holds the worst subjective prior and the Receiver has prior μ_0.*

5.2 More on the Persuasion Group

In Section 5, we have examined the properties of the persuasion group under a sufficiently large budget. Note that Proposition 4 also provides a minimax characterization, where the Lagrange multiplier is non-negative. Accordingly, in this section, we further characterize the persuasion group under a large budget. According to Proposition 4, the key characteristic of the persuasion group is identifying a signaling scheme that maximizes the minimum interim payoff—referred to as *the optimal signaling scheme*. We can compute the optimal signaling scheme based on the equivalence between the min-max and max-min formulations.

The min-max and max-min equality can be modeled as a Nash equilibrium in a zero-sum game between the Sender and Nature. In this game, the Sender chooses a signaling scheme $p \in BP(\mu_0)$, while Nature selects the Sender's subjective prior $\lambda \in \Delta(\Theta)$. The Sender seeks to maximize the following payoff function, whereas Nature aims to minimize it.

$$\mathcal{L}(\lambda,p) \triangleq \int_\mu \hat{V}_\lambda(\mu)dp(\mu). \tag{9}$$

This function represents the Sender's subjective expected payoff under the signaling scheme p and the prior belief λ. The following proposition states the indifference condition that characterizes Nature's optimal strategy at the Nash equilibrium of the zero-sum game.

Proposition 5. *A subjective prior $\lambda^* \in \Delta(\Theta)$ is the worst Sender's subjective prior if and only if there exists $p^* \in BP(\mu_0)$ such that $\mathcal{L}(\lambda^*,p^*) = cav(\hat{V}_{\lambda^*})(\mu_0)$, and for any $\theta \in \boldsymbol{supp}(\lambda^*)$, $\mathcal{L}(\lambda^*,p^*) = \mathcal{L}(\mu_\theta,p^*)$, and for any $\theta \notin \boldsymbol{supp}(\lambda^*)$, $\mathcal{L}(\lambda^*,p^*) \leq \mathcal{L}(\mu_\theta,p^*)$.*[12]

We also get the condition for Sender's optimal strategy at the Nash equilibrium and explain it in the following proposition.

Proposition 6. *A signaling scheme $p^* \in BP(\mu_0)$ is optimal if and only if there exists $\lambda^* \in \Delta(\Theta)$ such that $\mathcal{L}(\lambda^*, p^*) = cav(\hat{V}_{\lambda^*})(\mu_0)$, and for any $\theta \in \boldsymbol{supp}(\lambda^*)$, $\mathcal{L}(\lambda^*,p^*) = \mathcal{L}(\mu_\theta,p^*)$, and for any $\theta \notin \boldsymbol{supp}(\lambda^*)$, $\mathcal{L}(\lambda^*,p^*) \leq \mathcal{L}(\mu_\theta,p^*)$.*

These two propositions jointly characterize the worst subjective prior and the optimal signaling scheme. The condition, "$\mathcal{L}(\lambda^*,p^*) = cav(\hat{V}_{\lambda^*})(\mu_0)$, and for any $\theta \in \mathbf{supp}(\lambda^*)$, $\mathcal{L}(\lambda^*,p^*) = \mathcal{L}(\mu_\theta,p^*)$, and for any $\theta \notin \mathbf{supp}(\lambda^*)$, $\mathcal{L}(\lambda^*,p^*) \leq \mathcal{L}(\mu_\theta,p^*)$", implies that λ^* and p^* form a Nash equilibrium in the "zero-sum game", where λ^* is the best response to p^* and vice versa.

We can apply the characterization of the worst subjective prior to narrow the possible range of the set $\operatorname{argmin}_{\lambda \in \Delta(\Theta)} cav(\hat{V}_\lambda)(\mu_0)$ when there are only two possible types of the Sender in the appendix.

5.3 The Salesman Example

At the end of this section, we use our previous results to compare the MDMBs with different budgets, Bayesian persuasion, mediated communication, and cheap talk using Example 1.

We begin by characterizing $\mathcal{V}_C^*(\mu_0)$. By Theorem 1, we need to consider four lines for a given parameter $\lambda \in \mathbb{R}$, namely $l_1(\mu) = 0, l_2(\mu) = -C(\frac{\lambda\mu}{\mu_0} + \frac{(1-\lambda)(1-\mu)}{1-\mu_0}), l_3(\mu) = \frac{\lambda\mu}{\mu_0} + \frac{(1-\lambda)(1-\mu)}{1-\mu_0}$ and $l_4(\mu) = (1-C)(\frac{\lambda\mu}{\mu_0} + \frac{(1-\lambda)(1-\mu)}{1-\mu_0})$. Correspondingly, we have that for $\mu \in [0,1]$,

$$\hat{V}_{\lambda,C}(\mu) = \begin{cases} \max\{l_1(\mu), l_2(\mu)\} & \mu < \frac{1}{2} \\ \max\{l_2(\mu), l_3(\mu)\} & \mu = \frac{1}{2} \\ \max\{l_3(\mu), l_4(\mu)\} & \mu > \frac{1}{2} \end{cases}.$$

Since $\hat{V}_{\lambda,C}$ is convex and upper semi-continuous on $\mu \in [0, \frac{1}{2})$ and $\mu \in (\frac{1}{2}, 1]$, to compute $cav(\hat{V}_{\lambda,C})(\mu_0)$ we only need to evaluate $\hat{V}_{\lambda,C}(0) = \max\{0, -C\frac{1-\lambda}{1-\mu_0}\}$, $\hat{V}_{\lambda,C}(\frac{1}{2}) =$

[12] μ_θ is the distribution in $\Delta(\Theta)$ with a singleton support $\{\theta\}$.

$\max\{\frac{-C}{2}(\frac{\lambda}{\mu_0}+\frac{1-\lambda}{1-\mu_0}),\frac{1}{2}(\frac{\lambda}{\mu_0}+\frac{1-\lambda}{1-\mu_0})\}$ and $\hat{V}_{\lambda,C}(1)=\max\{\frac{\lambda}{\mu_0},(1-C)\frac{\lambda}{\mu_0}\}$. Assuming $\mu_0<\frac{1}{2}$, we can only partition μ_0 into $0,\frac{1}{2}$ or $0,1$. Since $\hat{V}_{\lambda,C}(0),\hat{V}_{\lambda,C}(\frac{1}{2}),\hat{V}_{\lambda,C}(1)$ are all decreasing in λ for $\lambda\geq 0$, to find the minimum concavification value, we only need to consider the case of $\lambda\leq 0$. We then divide this case into two subcases: $\lambda\in[-\frac{\mu_0}{1-2\mu_0},0]$ and $\lambda\in(-\infty,-\frac{\mu_0}{1-2\mu_0})$. By considering two subcases, we solve the Lagrange multiplier $\lambda^*=\frac{-\mu_0}{(1-\mu_0)C-\mu_0}$ for $C>1$. Thus, we can solve for the result and obtain that for $\mu_0<\frac{1}{2}$,

$$\mathcal{V}_C^*(\mu_0)=\begin{cases}0 & C\leq 1\\ \frac{(C-1)\mu_0}{C(1-\mu_0)-\mu_0} & C>1\end{cases}.$$

We thus show the optimal Sender's payoffs of different communication protocols in Figure 2. The red line is the concave envelope of $\max\mathbb{V}(\mu)$, which is the result of Bayesian persuasion. The black line is the result of the MDMB with bound $C=+\infty$, i.e. $\mathcal{V}^*(\mu_0)$. The blue line is the result of $\mathcal{V}_2^*(\mu_0)$. Finally, we can see that regardless of what we use among the MDMB with bounded budget $C\leq 1$, cheap talk, or classical mediated communication, we can only get the results as the green line, which cannot benefit from those protocols.

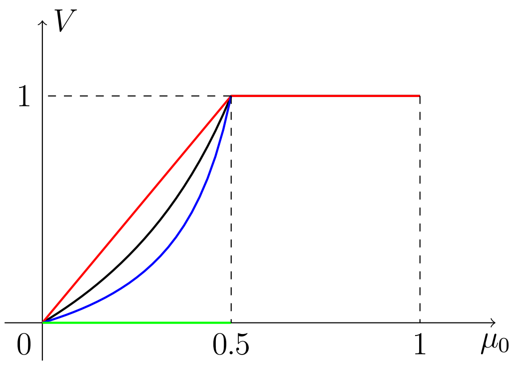

Figure 2: Comparison of Different Protocols

6 Discussions

6.1 The Value of Commitment in Web 3 Communities

Let $\mathcal{V}_{BP}^*(\mu_0)$ represent the value of Bayesian persuasion, that is $\mathcal{V}_{BP}^*(\mu_0)=cav(\mathbb{V})(\mu_0)$. This section directly compares $\mathcal{V}_{BP}^*(\mu_0)$ with $\mathcal{V}^*(\mu_0)$, which encapsulates the value of commitment within Web 3 communities.

In conventional societies, the paradigm of communication without commitment is epitomized by the cheap talk model. Consequently, $\mathcal{V}^*_{BP}(\mu_0) - \mathcal{V}^*_{CT}(\mu_0)$ quantifies the value of commitment inherent in communication within conventional societies. However, as delineated in Drakopoulos et al. [2023], the paradigm of communication in Web 3.0 communities, facilitated by Blockchain systems and smart contracts, presents a radically altered landscape. In Web 3.0 communities, users are characterized by full decentralization and a potential for high unreliability. Consequently, a viable approach to facilitating communication among these users is through the deployment of smart contracts, which serve as transparent algorithms. The Sender leverages smart contracts to integrate money-burning mechanisms via subsidies and gas fees. Therefore, the communication milieu of Web 3.0 communities is not amenable to modeling as cheap talk but rather as MDMB. Hence, we denote $\mathcal{V}_{BP}(\mu_0) - \mathcal{V}^*(\mu_0)$ as the *refined value of commitment* in Web 3.0 communities. It is important to note that $\mathcal{V}_{BP}(\mu_0) - \mathcal{V}^*(\mu_0)$ also represents the cost of obtaining credibility through money-burning tactics.

Our first result establishes the condition under which the refined value of commitment does not exist.

Proposition 7. *If $\mathcal{V}^*(\mu_0) = \mathcal{V}^*_{BP}(\mu_0)$ then $\mathcal{V}^*_{CT}(\mu_0) = \mathcal{V}^*_0(\mu_0) = \mathcal{V}^*(\mu_0) = \mathcal{V}^*_{BP}(\mu_0)$.*

Intuitively, we can derive this proposition from the result of Corrao and Dai [2023]. When $\mathcal{V}^*(\mu_0) = \mathcal{V}^*_{BP}(\mu_0)$, there must be no money burning. Hence, $\mathcal{V}^*_0(\mu_0) = \mathcal{V}^*_{BP}(\mu_0)$, which leads to $\mathcal{V}^*_{CT}(\mu_0) = \mathcal{V}^*_{BP}(\mu_0)$ by Corrao and Dai [2023]. By Proposition 7, we deduce that if $\mathcal{V}^*_{CT}(\mu_0) < \mathcal{V}^*_{BP}(\mu_0)$, then $\mathcal{V}^*(\mu_0) < \mathcal{V}^*_{BP}(\mu_0)$. Consequently, we promptly arrive at the subsequent corollary, which juxtaposes the value of commitment with the refined value of commitment.

Corollary 5. *There is a positive value of commitment, i.e. $\mathcal{V}^*_{CT}(\mu_0) < \mathcal{V}^*_{BP}(\mu_0)$, if and only if there is a positive refined value of commitment in the Web 3.0 community, i.e. $\mathcal{V}^*(\mu_0) < \mathcal{V}^*_{BP}(\mu_0)$.*

Technically, we provide Example 4 showing that the relation $\mathcal{V}^*_{CT}(\mu_0) = \mathcal{V}^*_0(\mu_0) = \mathcal{V}^*(\mu_0) < \mathcal{V}^*_{BP}(\mu_0)$ can hold even under generic settings, which refers to the case that the refined value of commitment is positive and the same as the value of commitment.

Corollary 5 has an important implication that if there is the value of commitment in conventional societies, commitment is still valuable in Web 3.0 communities. In addition, according to Corollary 2 in Lipnowski and Ravid [2020], the refined value of commitment is strictly positive under almost all prior beliefs as long as the value of commitment is strictly positive.

6.2 Relations to Money Burning in Cheap Talk

Numerous studies have explored the role of money burning in cheap talk equilibria Austen-Smith and Banks [2000], Kartik [2007], Karamychev and Visser [2017], where it can be viewed as a mechanism for designing a costly signaling game. These studies reveal that while money burning can enhance the precision of a cheap talk equilibrium, it only makes cheap talk influential[13] if some cheap talk equilibrium themselves, without money burning, are already influential.

Under the assumption of transparent motives, these results apply directly, demonstrating that the Sender cannot leverage money burning to improve communication efficiency in a cheap talk equilibrium. In the binary-type example, Example 1, we observe that, under a prior of $\mu_0 < 0.5$, the Sender cannot achieve a higher overall payoff than the no-communication payoff by using protocols such as cheap talk, cheap talk with money burning, or mediated communication. More specifically, these communication protocols fail to produce an influential equilibrium. In other words, their works treat money burning as a tool for equilibrium refinement. However, in our construction, money burning does expand the credibility boundary of mediated communication, resulting in an influential equilibrium. This distinction highlights the differing roles of money burning in cheap talk versus mediated communication. As noted by Theorem 1 in Austen-Smith and Banks [2000], money burning can enhance the precision of cheap talk communication but does not improve its credibility. In contrast, Theorem 2 demonstrates that money burning can indeed bolster the credibility of mediated communication and expand the equilibrium payoff set rather than refinement .

Beyond the assumption of transparent motives, a study on the implementation problem Liu and Wu [2024] demonstrates a significant difference in the implementability conditions between implementing a cheap talk equilibrium with money burning and a mediated

[13] An equilibrium is considered influential if there are at least two of the Receiver's actions taken along the equilibrium path with the same amount of money burning.

communication equilibrium with money burning. In addition, in Example 3, we show that money burning cannot make mediated communication credible in any case, implying that money burning cannot enhance mediated communication and the necessity of the transparent motives assumption.

7 Concluding Remarks

In this paper, we introduce and examine a novel communication protocol called mediated communication with money-burning mechanism (MDMB). In this protocol, the Sender not only utilizes a trusted mediator to transmit messages but also voluntarily forgoes part of his utility to enhance credibility. By generalizing the revelation principle from mechanism design with limited commitment, we characterize the communication efficiency of this protocol under the transparent motives assumption.

The main finding of this paper is that, under the transparent motives assumption, money burning can strictly enhance the Sender's communication efficiency in nearly all environments where commitment has value. Our results suggest that when balancing the trade-off between incurring costs and gaining credibility, the Sender generically benefits from engaging in money burning—provided the budget exceeds a certain threshold. This highlights a distinct role for money burning in expanding the credibility of mediated communication and enlarging the equilibrium payoff set, in contrast to its role in cheap talk, where money burning serves as a method of equilibrium refinement rather than expansion. Additionally, this insight sets a new benchmark for the communication efficiency of an unreliable Sender, since money burning constitutes an action the Sender can undertake independently to enhance credibility.

Furthermore, the MDMB framework is directly applicable to Web 3.0 communities, where our results reveal a refined notion of commitment value. We demonstrate that this refined commitment is valuable if and only if commitment remains valuable in the absence of cryptographic infrastructure. Communication efficiency in Web 3.0 environments is, in general, higher than in conventional settings, underscoring the critical role of algorithmic and cryptographic technologies.

It is important to emphasize, however, that these conclusions hinge on the strong assumption of state-independent Sender's preferences. In broader contexts, the function of money burning merits further exploration. Moreover, while we provide a quantifiable threshold for money burning, deriving a quantifiable result of the improvement remains an open problem.

References

Ricardo Alonso and Odilon Câmara. Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory*, 165:672–706, 2016.

David Austen-Smith and Jeffrey S Banks. Cheap talk and burned money. *Journal of Economic Theory*, 91(1):1–16, 2000.

Dirk Bergemann and Stephen Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019.

Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. The design and price of information. *American Economic Review*, 108(1):1–48, 2018.

Helmut Bester and Roland Strausz. Contracting with imperfect commitment and the revelation principle: the single agent case. *Econometrica*, 69(4):1077–1098, 2001.

Archishman Chakraborty and Rick Harbaugh. Persuasion by cheap talk. *American Economic Review*, 100(5):2361–2382, 2010.

Roberto Corrao. Mediation markets: The case of soft information. *Working Paper*, 2023.

Roberto Corrao and Yifan Dai. The bounds of mediated communication. *Working Paper*, 2023.

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

Laura Doval and Vasiliki Skreta. Mechanism design with limited commitment. *Econometrica*, 90(4):1463–1500, 2022.

Laura Doval and Alex Smolin. Information payoffs: An interim perspective. *arXiv preprint arXiv:2109.03061*, 2021.

Laura Doval and Alex Smolin. Persuasion and welfare. *Journal of Political Economy*, 132(7):2451–2487, 2024.

Kimon Drakopoulos, Irene Lo, and Justin Mulvany. Blockchain mediated persuasion. In *Proceedings of the 24th ACM Conference on Economics and Computation*, 2023.

Francoise Forges. An approach to communication equilibria. *Econometrica: Journal of the Econometric Society*, pages 1375–1385, 1986.

Guillaume R Fréchette, Alessandro Lizzeri, and Jacopo Perego. Rules and commitment in communication: An experimental analysis. *Econometrica*, 90(5):2283–2318, 2022.

Hans Gersbach. The money-burning refinement: with an application to a political signalling game. *International Journal of Game Theory*, 33:67–87, 2004.

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

Maxim Ivanov. Beneficial mediated communication in cheap talk. *Journal of Mathematical Economics*, 55:129–135, 2014.

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

Vladimir Karamychev and Bauke Visser. Optimal signaling with cheap talk and money burning. *International Journal of Game Theory*, 46(3):813–850, 2017.

Navin Kartik. A note on cheap talk and burned money. *Journal of Economic Theory*, 136(1):749–758, 2007.

Anton Kolotilin and Hongyi Li. Relational communication. *Theoretical Economics*, 16(4):1391–1430, 2021.

Hao Li and Xianwen Shi. Discriminatory information disclosure. *American Economic Review*, 107(11):3363–3385, 2017.

Xiao Lin and Ce Liu. Credible persuasion. *Journal of Political Economy*, 132(7): 2228–2273, 2024.

Elliot Lipnowski and Doron Ravid. Cheap talk with transparent motives. *Econometrica*, 88(4):1631–1660, 2020.

Elliot Lipnowski, Doron Ravid, and Denis Shishkin. Persuasion via weak institutions. *Journal of Political Economy*, 130(10):2705–2730, 2022.

Elliot Lipnowski, Doron Ravid, and Denis Shishkin. Perfect bayesian persuasion. *arXiv preprint arXiv:2402.06765*, 2024.

Yi Liu and Fan Wu. Implementing randomized allocation rules with outcome-contingent transfers. *Journal of Economic Theory*, 220:105878, 2024.

Daehong Min. Bayesian persuasion under partial commitment. *Economic Theory*, 72 (3):743–764, 2021.

Roger B Myerson. Optimal coordination mechanisms in generalized principal–agent problems. *Journal of Mathematical Economics*, 10(1):67–81, 1982.

Eduardo Perez-Richet and Vasiliki Skreta. Test design under falsification. *Econometrica*, 90(3):1109–1142, 2022.

Hitoshi Sadakane. Multistage information transmission with voluntary monetary transfers. *Theoretical Economics*, 18(1):267–301, 2023.

Andrés Salamanca. The value of mediated communication. *Journal of Economic Theory*, 192:105191, 2021.

Monir Shaker, Fereidoon Shams Aliee, and Reza Fotohi. Online rating system development using blockchain-based distributed ledger technology. *Wireless Networks*, 27(3):1715–1737, 2021.

Appendix

A Omitted Examples

A.1 The Necessity of Generic Assumption

In this section, we provide an example to show how the results in Section 4 fail if Assumption 1 doesn't hold.

Example 2. *We present an abstract setting in this example, where we only specify the belief-value function and ensure the existence of the basic settings of A,u,v,Θ by imposing the upper-semi continuity of the belief-value function.*

We assume that there are three distinct types θ_1,θ_2 and θ_3. The maximum of belief-value correspondence is

$$V(\mu)=\begin{cases} \frac{7}{3} & \mu(\theta_1)=1 \\ 2 & \mu(\theta_1)=0,\mu(\theta_2)\in[0,\frac{1}{2}) \\ 3 & \mu(\theta_1)=0,\mu(\theta_2)\in[\frac{1}{2},\frac{3}{4}] \\ 1 & \mu(\theta_1)=0,\mu(\theta_2)\in(\frac{3}{4},1] \\ 1 & \text{otherwise} \end{cases}.$$

By restricting the support to $\{\theta_2,\theta_3\}$, the value function of Example 2 coincides with Example 3 or Figure 7 in Salamanca [2021]. For $\mu_0=(\frac{1}{2},\frac{1}{6},\frac{1}{3})$, since $\mathcal{V}_0^*((0,\frac{1}{3},\frac{2}{3}))=\frac{7}{3}$ as shown by Salamanca [2021], splitting μ_0 into $(1,0,0)$ and $(0,\frac{1}{3},\frac{2}{3})$ yields $\mathcal{V}_{MD}^*(\mu_0)=\frac{7}{3}$. Furthermore, the interim payoff of θ_1 cannot exceed $\frac{7}{3}$, implying that $\mathcal{V}^*(\mu_0)\le\frac{7}{3}$. Hence, we obtain $\mathcal{V}^*(\mu_0)=\mathcal{V}_{MD}^*(\mu_0)=\frac{7}{3}$. However, to find a cheap talk equilibrium with $\frac{7}{3}$ as the Sender's payoff, we need to split μ_0 into $(1,0,0)$ and $(0,\frac{1}{3},\frac{2}{3})$ and keep $(1,0,0)$ unchanged. Since $\mathcal{V}_{CT}^*((0,\frac{1}{3},\frac{2}{3}))=2<\frac{7}{3}=V((1,0,0))$, no cheap talk equilibrium achieves $\frac{7}{3}$ for the Sender, and thus $\mathcal{V}_{CT}^*(\mu_0)<\mathcal{V}_{MD}^*(\mu_0)=\mathcal{V}^*(\mu_0)$ for $\mu_0=(\frac{1}{2},\frac{1}{6},\frac{1}{3})$.

A.2 The Necessity of Transparent-motives Assumption

In this section, we provide an example to show that without the transparent-motives assumption, money burning may not improve mediated communication at all.

Example 3. *Consider a buyer (Sender) and a seller (Receiver) in a market. The buyer's valuation of the seller's product, which is the buyer's private information, is drawn from a set* $\Theta = \{1,2\}$. *The probability that the valuation* v *is 2 is given by* μ_0.

The seller can set a price p *chosen from the set* $A = \{1,2\}$. *When the seller sets a price* p *and the buyer's valuation is* v, *the payoffs are defined as follows:*

- ***Buyer's payoff****:* $v(p,v) = \max\{0, v-p\}$, *representing the buyer's surplus.*
- ***Seller's payoff****:* $u(p,v) = \mathbb{I}(v \geq p) \cdot p$, *where* $\mathbb{I}(v \geq p)$ *is an indicator function equal to 1 if* $v \geq p$ *and 0 otherwise, representing the seller's revenue.*

In this example, we show that for any $\mu_0 > 0.5$, there is no MDMB that can induce an outcome where the seller sets a price of 1. Therefore, MDMB cannot improve the buyer's total payoff; however, BP can. Further, we show that even if the money burning is private, which is a stronger setting, the money burning still cannot improve mediated communication. In this setting, we can merge the posteriors that induce the same action of the seller. It is without loss of generality to assume that there are two possible posteriors $x_1 < x_2$ of the seller induced by an MDMB. If the MDMB is influential, then we must have $x_1 \leq 0.5 \leq x_2$.

If the unconditional probability of the posterior x_1 is p, then we must have $px_1 + (1-p)x_2 = \mu_0$. Suppose the expected amount of money burned by a valuation-1 buyer is $T(1)$, and the expected amount of money burned by a valuation-2 buyer is $T(2)$. Then, the incentive-compatible constraints can be written as

$$T(2) \geq T(1)$$

and

$$\frac{px_1}{px_1 + (1-p)x_2} - T(2) \geq \frac{p(1-x_1)}{p(1-x_1) + (1-p)(1-x_2)} - T(1).$$

Thus, we can deduce that

$$\frac{px_1}{px_1 + (1-p)x_2} \geq \frac{p(1-x_1)}{p(1-x_1) + (1-p)(1-x_2)},$$

which is equivalent to $x_1 \geq x_2$. A contradiction!

A.3 No Value of MDMB

In this section, we provide an example that the MDMB may not improve the communication efficiency even under Assumption 1. Thus, the phrases, "almost all" or "generically", in Section 4 are necessary.

Example 4. *The Receiver has three possible actions a_1,a_2,a_3 and the Sender has two possible types H,L. The prior belief assigns probability μ_0 to the Sender's type being H. The Sender's values for the actions are $v(a_1) = 0$,$v(a_2) = \frac{1}{4}$,$v(a_3) = 1$. We summarize the Receiver's payoffs in Table 1.*

$u(a,\theta)$	H	L
a_1	-4	1
a_2	0	0
a_3	1	-2

Table 1: Receiver's payoff matrix.

In this example, the belief-value correspondence is

$$\mathbb{V}(\mu) = \begin{cases} 1 & \mu \in (\frac{2}{3},1] \\ [\frac{1}{4},1] & \mu = \frac{2}{3} \\ \frac{1}{4} & \mu \in (\frac{1}{5},\frac{2}{3}) \\ [0,\frac{1}{4}] & \mu = 1/5 \\ 0 & \mu < \frac{1}{5} \end{cases}.$$

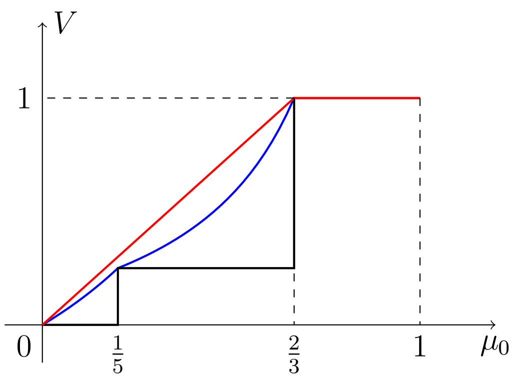

Figure 3: Results of Example 4.

We depict $\mathbb{V}(\mu)$ on Figure 3 as the black line, which corresponds to the outcome under a mediator without money burning and cheap talk. Based on Proposition 8, we display the result of $\mathcal{V}^*$ on Figure 3 as the blue line and $\mathcal{V}^*_{BP}$ as the red line, with the procedure omitted. We observe that, when $\mu_0 = \frac{1}{5}$, the case of $\mathcal{V}^*_{CT}(\mu_0) = \mathcal{V}^*_0(\mu_0) = \mathcal{V}^*(\mu_0) < \mathcal{V}^*_{BP}(\mu_0)$ arises under the generic setting.

B Omitted Proofs

B.1 Omitted Proofs in Section 2.2

Proof of Proposition 1. For any MDMB (M,S,ϕ) and a corresponding PBE assessment $(\sigma,\alpha,\mu)\in\mathcal{E}[\mathcal{G}_{M,S,\phi}(\mu_0)]$, we will directly construct a canonical MDMB (π,x) and a corresponding PBE canonical assessment $(\sigma^*,\alpha^*,\mu^*)\in\mathcal{E}[\mathcal{G}_{(\pi,x)}(\mu_0)]$ such that the constructed canonical assessment is payoff-equivalent to the original assessment.

The canonical MDMB we constructed is as follows: for any $\mu\in\Delta(\Theta),\theta\in\Theta$,

$$\pi(\mu|\theta)=\sum_{s\in S,t\in T,\mu(s,t)=\mu,m\in M}\phi(s,t|m)\sigma(m|\theta), \tag{10}$$

and for any $\mu\in\Delta(\Theta)$

$$x(\mu)=\begin{cases}\frac{\sum_{s\in S,t\in T,\mu(s,t)=\mu,m\in M,\theta\in\Theta}\phi(s,t|m)\sigma(m|\theta)\mu_0(\theta)t}{\sum_{s\in S,t\in T,\mu(s,t)=\mu,m\in M,\theta\in\Theta}\phi(s,t|m)\sigma(m|\theta)\mu_0(\theta)} & \sum_{s,t,\mu(s,t)=\mu,m,\theta}\phi(s,t|m)\sigma(m|\theta)\mu_0(\theta)\neq 0\\ 0 & \text{otherwise}\end{cases}. \tag{11}$$

Note that the above canonical MDMB is well-defined since the support of $\phi(m)$ is finite and $x(\mu)\in T$.

The canonical assessment $(\sigma^*,\alpha^*,\mu^*)$ we constructed is as follows: for all θ, $\sigma^*(\theta|\theta)=1$; for all $\mu\in\Delta(\Theta)$, $\mu^*(\mu)=\mu$; for all $\mu\in\textbf{supp}\{\pi(\theta)\}$ for some $\theta\in\Theta$,

$$\alpha^*(\mu)=\sum_{s\in S,t\in T,\mu(s,t)=\mu}\frac{\sum_{\theta\in\Theta,m\in M}\mu_0(\theta)\sigma(m|\theta)\phi(s,t|m)}{\sum_{s'\in S,t'\geq 0,\mu(s',t')=\mu}\sum_{\theta\in\Theta,m\in M}\mu_0(\theta)\sigma(m|\theta)\phi(s',t'|m)}\alpha(s,t).$$

For $\mu\notin\textbf{supp}\{\pi(\theta)\}$ for any $\theta\in\Theta$, $\alpha^*(\mu)$ is any best response given posterior belief μ.

Subsequently, we have to verify two things. The first is that the canonical assessments $(\sigma^*,\alpha^*,\mu^*)\in\mathcal{E}[\mathcal{G}_{\pi,x}(\mu_0)]$. The second is that $(\sigma^*,\alpha^*,\mu^*)$ and (σ,α,μ) are payoff-equivalent.

Before the verification, we prove the following lemma.

Lemma 1. *Suppose $\mu=\mu(s,t)\in\textbf{supp}\{\pi(\hat\theta)\}$ for some $\hat\theta$, then*

$$\frac{\sum_{\theta\in\Theta,m\in M}\mu_0(\theta)\sigma(m|\theta)\phi(s,t|m)}{\sum_{\mu(s',t')=\mu}\sum_{\theta\in\Theta,m\in M}\mu_0(\theta)\sigma(m|\theta)\phi(s',t'|m)}=\frac{\sum_{m\in M}\sigma(m|\hat\theta)\phi(s,t|m)}{\sum_{\mu(s',t')=\mu}\sum_{m\in M}\sigma(m|\hat\theta)\phi(s',t'|m)}.$$

Proof of Lemma 1. By the theorem on equal ratios, it is sufficed to show that for any $\bar\theta\in\textbf{supp}\{\mu\}$, we have that

$$\frac{\sum_{m\in M}\sigma(m|\hat\theta)\phi(s,t|m)}{\sum_{\mu(s',t')=\mu}\sum_{m\in M}\sigma(m|\hat\theta)\phi(s',t'|m)}=\frac{\sum_{m\in M}\sigma(m|\bar\theta)\phi(s,t|m)}{\sum_{\mu(s',t')=\mu}\sum_{m\in M}\sigma(m|\bar\theta)\phi(s',t'|m)}. \tag{12}$$

According to Bayesian updating, for any $s''\in S,t''\geq 0$ such that $\mu(s,t)=\mu(s'',t'')$, we

have that
$$\frac{\mu_0(\hat{\theta})\sum_{m\in M}\sigma(m|\hat{\theta})\phi(s,t|m)}{\mu_0(\bar{\theta})\sum_{m\in M}\sigma(m|\bar{\theta})\phi(s,t|m)}=\frac{\mu(\hat{\theta}|s,t)}{\mu(\bar{\theta}|s,t)}=\frac{\mu(\hat{\theta}|s'',t'')}{\mu(\bar{\theta}|s'',t'')}=\frac{\mu_0(\hat{\theta})\sum_{m\in M}\sigma(m|\hat{\theta})\phi(s'',t''|m)}{\mu_0(\bar{\theta})\sum_{m\in M}\sigma(m|\bar{\theta})\phi(s'',t''|m)}.$$
Thus,
$$\frac{\sum_{m\in M}\sigma(m|\hat{\theta})\phi(s,t|m)}{\sum_{m\in M}\sigma(m|\hat{\theta})\phi(s'',t''|m)}=\frac{\sum_{m\in M}\sigma(m|\bar{\theta})\phi(s,t|m)}{\sum_{m\in M}\sigma(m|\bar{\theta})\phi(s'',t''|m)}.$$
Since s'',t'' can be any one satisfying that $\mu(s'',t'')=\mu$, by the theorem on equal ratios, Equation 12 holds. □

Sender's optimality and payoff-equivalence for the Sender. To show Sender's optimality and Sender's payoff-equivalence, it is sufficient to show that the expected payoffs of type θ Sender under both assessments are the same. The expected payoff of type θ Sender under the assessment (σ,α,μ) is
$$\sum_{m\in M,s\in S,t\in T,a\in A}\sigma(m|\theta)\phi(s,t|m)\alpha(a|s,t)(v(a)-t).$$
The expected payoff of type θ Sender under the assessment $(\sigma^*,\alpha^*,\mu^*)$ is
$$\sum_{\mu\in\mathbf{supp}\{\pi(\theta)\},a\in A}\pi(\mu|\theta)\alpha^*(a|\mu)(v(a)-x(\mu)).$$
These two expected payoffs are the same, since by Lemma 1, for any $a\in A,s\in S,t\in T$ the coefficient of $\alpha(a|s,t)$ where $\mu(s,t)=\mu\in\mathbf{supp}\{\pi(\theta)\}$ in the expression
$$\sum_{\mu\in\mathbf{supp}\{\pi(\theta)\},a\in A}\pi(\mu|\theta)\alpha^*(a|\mu)v(a)$$
is
$$\sum_{s',t',m,\mu(s',t')=\mu}\sigma(m|\theta)\phi(s',t'|m)\frac{\sum_{\theta'\in\Theta,m\in M}\mu_0(\theta')\sigma(m|\theta')\phi(s,t|m)}{\sum_{s'\in S,t'\geq 0,\mu(s',t')=\mu}\sum_{\theta'\in\Theta,m\in M}\mu_0(\theta')\sigma(m|\theta')\phi(s',t'|m)}v(a)$$
$$=\sum_{s',t',m,\mu(s',t')=\mu}\sigma(m|\theta)\phi(s',t'|m)\frac{\sum_{m}\sigma(m|\theta)\phi(s,t|m)}{\sum_{s',t',m,\mu(s',t')=\mu}\sigma(m|\theta)\phi(s',t'|m)}v(a)$$
$$=\sum_{m}\sigma(m|\theta)\phi(s,t|m)v(a)$$
where the second equation holds by Lemma 1, and expected money burning of type θ Sender of $(\sigma^*,\alpha^*,\mu^*)$ is
$$\begin{aligned}\sum_{\mu\in\mathbf{supp}\{\pi(\theta)\},a\in A}\pi(\mu|\theta)\alpha^*(a|\mu)x(\mu)&=\sum_{\mu\in\mathbf{supp}\{\pi(\theta)\}}\pi(\mu|\theta)x(\mu)\\&=\sum_{\mu\in\mathbf{supp}\{\pi(\theta)\}}\pi(\mu|\theta)\frac{\sum_{s,t,\mu(s,t)=\mu,m\in M}\sigma(m|\theta)\phi(s,t|m)t}{\sum_{\mu(s,t)=\mu}\sum_{m\in M}\sigma(m|\theta)\phi(s,t|m)}\\&=\sum_{m\in M,s\in S,t\in T}\sigma(m|\theta)\phi(s,t|m)t,\end{aligned}$$
where the second equation holds by Lemma 1.

Receiver's optimality. Since $\alpha(s,t)$ is the best response under the belief $\mu(s,t)$ and $\alpha^*(\mu)$ is a convex combination of some $\alpha(s',t')$ where $\mu(s',t')=\mu$, by the convexity of the best response set, $\alpha^*(\mu)$ must satisfy the Receiver's optimality condition.

Bayesian updating. Given $\mu\in\Delta(\Theta)$, for any s,t such that $\mu(s,t)=\mu=\mu^*(\mu)$, by Bayesian updating, we have that for any $\theta\in\Theta$

$$\mu(\theta|s,t)\sum_{\theta'\in\Theta,m\in M}\mu_0(\theta')\sigma(m|\theta')\phi(s,t|m)=\mu_0(\theta)\sum_{m\in M}\sigma(m|\theta)\phi(s,t|m).$$

Hence,

$$\mu(\theta)\sum_{s,t,\mu(s,t)=\mu}\sum_{\theta'\in\Theta,m\in M}\mu_0(\theta')\sigma(m|\theta')\phi(s,t|m)=\sum_{s,t,\mu(s,t)=\mu}\mu_0(\theta)\sum_{m\in M}\sigma(m|\theta)\phi(s,t|m).$$

That is

$$\mu^*(\theta|\mu)\sum_{\theta'}\mu_0(\theta')\pi(\mu|\theta')=\mu_0(\theta)\pi(\mu|\theta).$$

Receiver's payoff-equivalence. The receiver's payoff under assessment (σ,α,μ) is

$$\sum_{\theta\in\Theta}\mu_0(\theta)\sum_{m,s,t,a}\sigma(m|\theta)\phi(s,t|m)\alpha(a|s,t)u(a,\theta).$$

The receiver's payoff under assessment $(\sigma^*,\alpha^*,\mu^*)$ is

$$\sum_{\theta\in\Theta}\mu_0(\theta)\sum_{\mu,a}\pi(\mu|\theta)\alpha(a|\mu)u(a,\theta).$$

These two payoffs are the same since, by definition,

$$\begin{aligned}\sum_{\theta\in\Theta}\mu_0(\theta)\sum_{\mu,a}\pi(\mu|\theta)\alpha(a|\mu)u(a,\theta)&=\sum_{\theta\in\Theta}\sum_{\mu}\sum_{a,s,t:\mu(s,t)=\mu}\mu_0(\theta)\phi(s,t|m)\sigma(m|\theta)\alpha(a|\mu)u(a,\theta)\\&=\sum_{\theta\in\Theta}\sum_{\mu}\sum_{a,s,t:\mu(s,t)=\mu}\mu_0(\theta)\sigma(m|\theta)\phi(s,t|m)\alpha(a|s,t)u(a,\theta)\\&=\sum_{\theta\in\Theta}\mu_0(\theta)\sum_{m,s,t,a}\sigma(m|\theta)\phi(s,t|m)\alpha(a|s,t)u(a,\theta).\end{aligned}$$

□

B.2 Omitted Proofs in Section 3

Proof of Theorem 1. According to Corollary 1, we need to solve the following optimization problem.

$$\max \quad k \tag{13}$$

$$\text{s.t} \quad k = \int_\mu \frac{\mu(\theta)}{\mu_0(\theta)}(V(\mu)-x(\mu))dp(\mu) \quad \forall \theta \in \Theta \tag{IC}$$

$$p \in BP(\mu_0) \tag{BP}$$

$$V(\mu) \in \mathbb{V}(\mu) \quad \forall \mu \in \Delta(\Theta) \tag{O}$$

$$0 \le x(\mu) \le C \quad \forall \mu \in \Delta(\Theta) \tag{Bgt}$$

We adopt a two-step optimization approach. First, fixing the signaling scheme $p \in BP(\mu_0)$, we try to find the optimal burning scheme $x(\mu)$ where $0 \le x(\mu) \le C$ and the Receiver's best response $V(\mu) \in \mathbb{V}(\mu)$. Now, it is a linear programming problem. By the fundamental theorem of linear programming, we can also obtain $\mathcal{V}_C^*(\mu_0)$ from the following max-min problem.

$$\max_{p \in BP(\mu_0)} \min_{\lambda \in \text{aff}(\Theta)} \int_\mu \max\{\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)} \max \mathbb{V}(\mu), \sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}(\min \mathbb{V}(\mu) - C)\} dp(\mu).$$

This implies that we choose $V(\mu) = \max \mathbb{V}(\mu), x(\mu) = 0$ if $\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)} > 0$ and $V(\mu) = \min \mathbb{V}(\mu), x(\mu) = C$ if $\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)} < 0$, which determines the choice of the receiver's best responses.

The rest of the proof relies on Sion's minimax theorem as well. It is easy to verify that $BP(\mu_0)$ is a compact and convex set, and $\{\lambda | \lambda \in \text{aff}(\Theta)\}$ is a convex set. Moreover, $\int_\mu \max\{\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)} \max \mathbb{V}(\mu), \sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}(\min \mathbb{V}(\mu) - C)\} dp(\mu)$ is linear in p and continuous convex in λ since it is the maximum of two linear functions. Since $\max \mathbb{V}(\mu)$ and $\min \mathbb{V}(\mu) - C$ are upper and lower semi-continuous, respectively, we have that $\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)} \max \mathbb{V}(\mu)$ is upper semi-continuous when $\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)} > 0$ and $\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}(\min \mathbb{V}(\mu) - C)$ is upper semi-continuous when $\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)} < 0$. Therefore,

$$\hat{V}_{\lambda,C}(\mu) = \max\{\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)} \max \mathbb{V}(\mu), \sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}(\min \mathbb{V}(\mu) - C)\}$$

is upper semi-continuous and so is $\int_\mu \max\{\sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)} \max \mathbb{V}(\mu), \sum_{\theta\in\Theta} \lambda(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}(\min \mathbb{V}(\mu) - C)\} dp(\mu)$ in p. Hence, we can apply Sion's minimax theorem directly and complete the proof. $\square$

B.3 Omitted Proofs in Section 4

Proof of Theorem 2. Without loss of generality, we suppose $v(a)>0$ for all $a\in A$ and $C_2>C_1>\max_a v(a)$. In the following proof, we prove that either $\mathcal{V}^*_{C_2}(\mu_0)>\mathcal{V}^*_{C_1}(\mu_0)$ or $\mathcal{V}^*_{C_2}(\mu_0)=\mathcal{V}^*_{C_1}(\mu_0)=\mathcal{V}^*_{0}(\mu_0)=\mathcal{V}^*_{CT}(\mu_0)$.

We adopt a proof by contradiction. Let μ_0 be a belief with the *smallest* support such that $\mathcal{V}^*_{CT}(\mu_0)<\mathcal{V}^*_{C_1}(\mu_0)=\mathcal{V}^*_{C_2}(\mu_0)$ where $C_2>C_1>\max_a v(a)$. Let $(p_{C_1},x_{C_1},V_{C_1})$ denote the optimal solution to the program (5) given the budget C_1 and let λ_{C_1} denote the corresponding Lagrange multiplier. Since $\mathcal{V}^*_{C_2}(\mu_0)=\mathcal{V}^*_{C_1}(\mu_0)$, $(p_{C_1},x_{C_1},V_{C_1})$ is also optimal to the program (5) given the budget C_2 and let λ_{C_2} denote the corresponding Lagrange multiplier. Because $x_{C_1}(\mu)\le C_1<C_2$, by Proposition 2 (ii), we know that for any $\mu\in\mathbf{supp}\{p_{C_1}\}$, $\sum_{\theta\in\Theta}\lambda_{C_2}(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}\ge 0$. Now, let $U_1=\{\mu\in\mathbf{supp}\{p_{C_1}\}\big|\sum_{\theta\in\Theta}\lambda_{C_2}(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}>0\}$ and $U_2=\{\mu\in\mathbf{supp}\{p_{C_1}\}\big|\sum_{\theta\in\Theta}\lambda_{C_2}(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}=0\}$. We know that U_1 and U_2 are a partition of $\mathbf{supp}\{p_{C_1}\}$. Next, we prove two claims.

Claim 2. *For any $\mu,\mu'\in U_1$, $\max\mathbb{V}(\mu)=\max\mathbb{V}(\mu')$.*

Proof of Claim 2. By the optimality of $(p_{C_1},x_{C_1},V_{C_1})$ and λ_{C_2}, we have that
$$\int_\mu\frac{\mu(\theta)}{\mu_0(\theta)}(V_{C_1}(\mu)-x_{C_1}(\mu))dp_{C_1}=\int_\mu\sum_\theta\frac{\lambda_{C_2}(\theta)\mu(\theta)}{\mu_0(\theta)}(V_{C_1}(\mu)-x_{C_1}(\mu))dp_{C_1}(\mu),$$
and p_{C_1} is the concavification of $\hat{V}_{\lambda_{C_2},C_2}$ at μ_0. Then, by Proposition 9 of the working paper version of Kamenica and Gentzkow [2011], we have that there exist parameters A_θ for $\theta\in\Theta$ and for any $\mu\in\mathbf{supp}\{p_{C_1}\}$,
$$\sum_\theta\frac{\lambda_{C_2}(\theta)\mu(\theta)}{\mu_0(\theta)}(V_{C_1}(\mu)-x_{C_1}(\mu))=\sum_\theta A_\theta\mu(\theta).$$
Hence, we obtain that for any $\theta\in\Theta$,
$$A_\theta\mu_0(\theta)\int_\mu\frac{\mu(\theta)}{\mu_0(\theta)}(V_{C_1}(\mu)-x_{C_1}(\mu))dp_{C_1}(\mu)=A_\theta\mu_0(\theta)\sum_{\theta'}A_{\theta'}\mu_0(\theta').$$
Summing them up and since $0=\sum_{\theta\in\Theta}\lambda_{C_2}(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}=\sum_{\theta\in\Theta}A_\theta\mu(\theta)$ for any $\mu\in U_2$, we have that
$$\int_{\mu\in U_1}\sum_\theta A_\theta\mu(\theta)\frac{\sum_\theta A_\theta\mu(\theta)}{\sum_\theta\frac{\lambda_{C_2}(\theta)\mu(\theta)}{\mu_0(\theta)}}dp_{C_1}(\mu)=(\sum_\theta A_\theta\mu_0(\theta))^2.$$
Since $\int_{\mu\in U_1}\sum_\theta\frac{\lambda_{C_2}(\theta)\mu(\theta)}{\mu_0(\theta)}dp_{C_1}(\mu)=\int_\mu\sum_\theta\frac{\lambda_{C_2}(\theta)\mu(\theta)}{\mu_0(\theta)}dp_{C_1}(\mu)=1$, by the equality condition of Cauchy's inequality and the above equation, we have that for any $\mu,\mu'\in U_1$,
$$\frac{\sum_\theta A_\theta\mu(\theta)}{\sum_\theta\frac{\lambda_{C_2}(\theta)\mu(\theta)}{\mu_0(\theta)}}=\frac{\sum_\theta A_\theta\mu'(\theta)}{\sum_\theta\frac{\lambda_{C_2}(\theta)\mu'(\theta)}{\mu_0(\theta)}}.$$

By Proposition 2, this implies that $\max\mathbb{V}(\mu)=\max\mathbb{V}(\mu')$ for $\mu,\mu'\in U_1$. To simplify the notation, we set this value to be R, which coincides with $\mathcal{V}^*_{C_1}(\mu_0)$ and $\mathcal{V}^*_{C_2}(\mu_0)$. □

Claim 3. *For any $\theta\in\textbf{supp}\{\mu\},\mu\in U_2$, $\lambda_{C_2}(\theta)=0$.*

Proof of Claim 3. Let $q_1=\int_{\mu\in U_1}dp_{C_1}(\mu),q_2=\int_{\mu\in U_2}dp_{C_1}(\mu)$ and $\mu_1=\int_{\mu\in U_1}\mu dp_{C_1}(\mu)/q_1,\mu_2=\int_{\mu\in U_2}\mu dp_{C_1}(\mu)/q_2$. We only have to show that for any $\theta\in\textbf{supp}\{\mu_2\}$, $\lambda_{C_2}(\theta)=0$.

We proof by contradiction. Suppose there is $\theta\in\textbf{supp}\{\mu_2\}$, $\lambda_{C_2}(\theta)\neq 0$, thus there is $\hat{\theta}\in\textbf{supp}\{\mu_2\}$ such that $\lambda_{C_2}(\hat{\theta})<0$. Therefore,

$$\sum_{\theta\in\textbf{supp}\{\mu_2\}}\hat{V}_{\lambda_{C_2},C_2}(\mu_\theta)\mu_2(\theta)\geq\mu_2(\hat{\theta})\lambda_{C_2}(\hat{\theta})(\min\mathbb{V}(\mu_{\hat{\theta}})-C_2)>0=\int_{\mu\in U_2}\hat{V}_{\lambda_{C_2},C_2}(\mu)dp_{C_1}(\mu)$$

where the last inequality is because $C_2>\max_a v(a)$. This contradicts to p_{C_1} is the concavification of $\hat{V}_{\lambda_{C_2},C_2}$. □

Now, back to our proof. For any $\theta\in\textbf{supp}\{\mu_2\}$, we have that

$$R=\int_\mu\frac{\mu(\theta)}{\mu_0(\theta)}(V_{C_1}(\mu)-x_{C_1}(\mu))dp_{C_1}(\mu)=q_1\frac{\mu_1(\theta)}{\mu_0(\theta)}R+q_2\frac{\mu_2(\theta)}{\mu_0(\theta)}\int_{\mu\in U_2}\frac{\mu(\theta)}{\mu_2(\theta)}(V_{C_1}(\mu)-x_{C_1}(\mu))d\frac{p_{C_1}(\mu)}{q_2}.$$

Hence, for any $\theta\in\textbf{supp}\{\mu_2\}$,

$$\int_{\mu\in U_2}\frac{\mu(\theta)}{\mu_2(\theta)}(V_{C_1}(\mu)-x_{C_1}(\mu))d\frac{p_{C_1}(\mu)}{q_2}=R.$$

This implies that $\mathcal{V}^*_{C_1}(\mu_2)\geq R$. Next, we divide the rest of the proof into two cases.

Case 1: If $\mathcal{V}^*_{C_2}(\mu_2)=R$, we have $\mathcal{V}^*_{C_2}(\mu_2)=\mathcal{V}^*_{C_1}(\mu_2)$. Since $|\textbf{supp}\{\mu_2\}|<|\textbf{supp}\{\mu_0\}|$ (otherwise $\mathcal{V}^*_{C_2}(\mu_0)=0$ which is clearly impossible), we have that $\mathcal{V}_{CT}(\mu_2)=R$ by the smallest support property of μ_0. Since $\max\mathbb{V}(\mu)=R$ for all $\mu\in U_1$, we can deduce that there exists a cheap talk equilibrium for μ_0 where the Sender achieves payoff R. This means that we have shown that $\mathcal{V}_{CT}(\mu_0)=R$, which contradicts our assumption.

Case 2: Suppose that $\mathcal{V}^*_{C_2}(\mu_2)>R$. Then, there must exist μ_3 where $\textbf{supp}\{\mu_3\}\subseteq\textbf{supp}\{\mu_2\}$ such that $\max\mathbb{V}(\mu_3)>R$. By Assumption 1, we choose μ_3 such that $RO(\mu_3)$ is a singleton. So, we can select small enough $\varepsilon>0$ such that $\max\mathbb{V}(\hat{\mu}_3)=\max\mathbb{V}(\mu_3)$ where $\hat{\mu}_3=(1-\varepsilon)\mu_3+\varepsilon\mu_1$. On the one hand, we have

$$\begin{aligned}\hat{V}_{\lambda_{C_2},C_2}(\hat{\mu}_3)&=\sum_{\theta\in\Theta}\lambda_{C_2}(\theta)\frac{(1-\varepsilon)\mu_3(\theta)+\varepsilon\int_{\mu\in U_1}\mu d\frac{p_{C_1}(\mu)}{q_1}}{\mu_0(\theta)}\max\mathbb{V}(\mu_3)\\&=\varepsilon\int_{\mu\in U_1}\sum_{\theta\in\Theta}\lambda_{C_2}(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}\max\mathbb{V}(\mu_3)d\frac{p_{C_1}(\mu)}{q_1}\\&>\varepsilon\sum_\theta\lambda_{C_2}(\theta)\frac{\mu_1(\theta)}{\mu_0(\theta)}R.\end{aligned}$$

On the other hand, by p_{C_1} is the concavification of $\hat{V}_{\lambda_{C_2},C_2}$, we have

$$\hat{V}_{\lambda_{C_2},C_2}(\hat{\mu}_3) \leq \sum_\theta A_\theta \hat{\mu}_3(\theta) = \varepsilon \sum_\theta A_\theta \mu_1(\theta) = \varepsilon \int_{\mu \in U_1} \sum_{\theta \in \Theta} \lambda_{C_2}(\theta) \frac{\mu(\theta)}{\mu_0(\theta)} \max \mathbb{V}(\mu) d\frac{p_{C_1}(\mu)}{q_1}$$
$$= \varepsilon \sum_\theta \lambda_{C_2}(\theta) \frac{\mu_1(\theta)}{\mu_0(\theta)} R$$

where the last equation is because of Claim 2. So far, we have obtained a contradiction, and we complete the proof. □

Proof of Proposition 3. By Assumption 1, there must be a full-support $\hat{\mu}$ such that $V(\hat{\mu}) = \max_{\mu \in \Delta(\Theta)} \max \mathbb{V}(\mu)$. Using the terminology from Corrao and Dai [2023], since there is a sufficiently small $\varepsilon > 0$ such that $qcav(\mathbb{V})(\mu_0 + \varepsilon(\mu - \mu_0)) = qcav(\mathbb{V})(\mu_0)$ for all $\mu \in \Delta(\Theta)$, then μ_0 satisfies the full-dimensionality condition; and since $C > \max_{\mu, V(\mu) \in \mathbb{V}(\mu)} V(\mu) - \min_{\mu, V(\mu) \in \mathbb{V}(\mu)} V(\mu)$, for $\hat{\mu}' = (1-t)\mu_0 + t\hat{\mu}$ where $t > 1$, the Sender can induce a payoff $\min \mathbb{V}(\hat{\mu}') - C < qcav(\mathbb{V})(\mu_0)$ at $\hat{\mu}'$. Therefore, the prior μ_0 is also locally improvable under the budget C. Thus, according to Theorem 3 in Corrao and Dai [2023], we have that $\mathcal{V}_C^*(\mu_0) > \mathcal{V}_{CT}^*(\mu_0)$. Hence, $\mathcal{V}_C^*(\mu_0) > \mathcal{V}_0^*(\mu_0)$ by Theorem 2. □

B.4 Omitted Proofs in Section 5

Proof of Claim 1. Given that $V_{\bar{\pi}}(\theta) - \min_{\theta' \in \Theta} V_{\bar{\pi}}(\theta') \geq 0$, it follows that $x(\mu) \geq 0$ for all $\mu \in \Delta(\Theta)$. Therefore, for any $\theta \in \Theta$, the Sender's payoff of type θ under the mechanism $(\bar{\pi}, x)$ is equal to

$$V_{\bar{\pi}}(\theta) - \sum_{\mu \in \mathbf{supp}\{\bar{\pi}(\theta)\}} \bar{\pi}(\mu|\theta) x(\mu) = \min_{\theta' \in \Theta} V_{\bar{\pi}}(\theta').$$

This implies that the mechanism is incentive-compatible. Moreover, the Sender's expected payoff is $\min_{\theta \in \Theta} V_{\bar{\pi}}(\theta)$.

Next, we show that the corresponding canonical assessments also satisfy the Bayesian updating condition under this MDMB mechanism. Let $\mu_\theta \in \Delta(\Theta)$ be such that $\mu_\theta(\theta) = 1$ and $\mu_\theta(\theta') = 0$ for any $\theta' \neq \theta$. Let $\mu_\pi^*(\mu)$ and $\mu_{\bar{\pi}}^*(\mu)$ denote the posterior beliefs under the signaling schemes π and $\bar{\pi}$ for any $\mu \in \Delta(\Theta)$, respectively. For any $\mu \in \Delta(\Theta)$, we have that

$$\mu_{\bar{\pi}}^*(\theta|\mu) = \frac{\bar{\pi}(\mu|\theta)\mu_0(\theta)}{\sum_{\theta' \in \Theta} \bar{\pi}(\mu|\theta')\mu_0(\theta')} = \frac{\pi(\mu|\theta)\mu_0(\theta)}{\sum_{\theta' \in \Theta} \pi(\mu|\theta')\mu_0(\theta')} = \mu_\pi^*(\theta|\mu).$$

Thus, $\mu_\pi^*(\mu) = \mu$ if and only if $\mu_{\bar{\pi}}^*(\mu) = \mu$. Hence, $V_{\bar{\pi}}(\theta) = \sum_{\mu \in \mathbf{supp}\{\pi(\theta)\}} \bar{\pi}(\mu|\theta) V(\mu) = (1-\delta)V_\pi(\theta) + \delta V(\mu_\theta)$. It follows that

$$\lim_{\delta \to 0^+} \min_{\theta \in \Theta} \{V_{\bar{\pi}}(\theta)\} = \lim_{\delta \to 0^+} \min_{\theta \in \Theta} \{(1-\delta)V_\pi(\theta) + \delta V(\mu_\theta)\} = \min_{\theta \in \Theta} V_\pi(\theta).$$

$\square$

C Binary Types Cases

In this section, we assume that the Sender has only two possible types, $\Theta = \{\theta_1, \theta_2\}$. In this case, we can give a clearer characterization of the persuasion group and the credibility-gaining group.

First, the following proposition states that the worst subjective prior must be one of the extreme points of $\Delta(\Theta)$. Thus, when we search for the persuasion group if the budget C is sufficiently large, we only need to consider the concavification over $\hat{V}_{\theta_1}$ or $\hat{V}_{\theta_2}$.

Proposition 8. *Suppose the Sender has a binary type set $\Theta = \{\theta_1, \theta_2\}$. Then, for any prior distribution $\mu_0 \in \Delta(\Theta)$, we have $\mathcal{V}^*(\mu_0) = \min\{cav(\hat{V}_{\theta_1})(\mu_0), cav(\hat{V}_{\theta_2})(\mu_0)\}$.*

Proof of Proposition 8. By Proposition 5, it suffices to show that there exists $\theta_i \in \{\theta_1, \theta_2\}$ such that for any $p \in BP(\mu_0)$ with $\mathcal{L}(\theta_i, p) = cav(\hat{V}_{\theta_i})(\mu_0)$, we have $\mathcal{L}(\theta_i, p) \leq \mathcal{L}(\theta_{3-i}, p)$.

Let $U = \{\mu | V(\mu) = \max_{x \in [0,1]} V(x)\}$ denote the range of posteriors that yield the maximum value for Sender, for $\mu \in [0,1]$. Since U is convex and $V(\cdot)$ is upper semi-continuous, U can be expressed as the union of closed intervals. We assume that $l = \min U$ and $r = \max U$. If $l \leq \mu_0 \leq r$, then it is clear that $V^*(\mu_0) = \max_{x \in [0,1]} V(x)$ and for any θ_i, $cav(\hat{V}_{\theta_i})(\mu_0) = V^*(\mu_0)$, which is a constant. Hence, our statement holds trivially. In the following proof, we consider $\mu_0 > r$ or $\mu_0 < l$.

Without loss of generality, we assume that $\mu_0 > r \geq 0$ by symmetry. We focus on θ_1. We prove by contradiction that if there exists $p \in BP(\mu_0)$ with $\mathcal{L}(\theta_1, p) = cav(\hat{V}_{\theta_1})(\mu_0)$ and $\mathcal{L}(\theta_1, p) > \mathcal{L}(\theta_2, p)$, then we reach a contradiction. Since $p \in BP(\mu_0)$ performs the concavification of the function $\hat{V}_{\theta_1}$ at point μ_0, by Proposition 9 of the working paper version of Kamenica and Gentzkow [2011], we have that the points $(\mu, \hat{V}_{\theta_1}(\mu))$ for $\mu \in \mathbf{supp}\{p\}$ are collinear. This means that there exist parameters k, b such that for any $\mu \in \mathbf{supp}\{p\}$,

$$\frac{\mu}{\mu_0} V(\mu) = k\mu + b.$$

Since $\mathcal{L}(\theta_1, p) > \mathcal{L}(\theta_2, p)$, we have

$$\int_{\mu} (\mu - \mu_0) V(\mu) dp(\mu) > 0.$$

Substituting $V(\mu)$ with $\mu_0(k+\frac{b}{\mu})$ and using $\int_\mu(\mu-\mu_0)dp(\mu)=0$, we obtain
$$b(1-\int_\mu \frac{\mu_0}{\mu}dp(\mu))=\int_\mu(\mu-\mu_0)\frac{b}{\mu}dp(\mu)>0.$$
By Cauchy's inequality,
$$\int_\mu \frac{\mu_0}{\mu}dp(\mu)=\int_\mu \frac{\mu}{\mu_0}dp(\mu)\int_\mu \frac{\mu_0}{\mu}dp(\mu)\geq(\int_\mu dp(\mu))^2=1.$$
Therefore, we must have $b<0$. However, $k\mu+b$ is the concavification line of $\hat{V}_{\theta_1}(\cdot)$ at μ_0. Thus, it must satisfy that for any $\mu\in[0,1]$, $\hat{V}_{\theta_1}(\mu)\leq k\mu+b$. Choosing $\mu=0$, we get $b\geq 0$. This is a contradiction. □

In Example 5, however, we show that the worst subjective prior is not necessarily an extreme point of $\Delta(\Theta)$ in general.

Example 5. *We consider an example with three parties: a seller, a buyer, and an influencer. The seller wants to sell a zero-cost product to the buyer. The buyer's valuation of the product is v. The seller only knows that v is distributed uniformly in $\{1,2,3\}$. The buyer is a fan of the influencer, who wants to help the buyer reduce the price of the product by disclosing information about the buyer's type and subsidizing the seller. The influencer acts as the Sender who uses our MDMB to influence the seller's action as the Receiver. To fit our model, we let the type set be $\Theta=\{v_1=1,v_2=2,v_3=3\}$, the prior distribution be $\mu_0=(\frac{1}{3},\frac{1}{3},\frac{1}{3})$, and the seller's action set be $A=\{p_1=1,p_2=2,p_3=3\}$.*

We assume that the influencer's objective is to minimize the price of the product. If the seller charges a price p_i to the buyer, the influencer's valuation function is $v(p_i)=4-p_i$.

We use Proposition 4 to examine the extreme point subjective priors of $\Delta(\Theta)$ at first. We then use Proposition 5 to find the worst Sender's subjective prior and the corresponding maximum payoff of the influencer achieved by the MDMB.

We consider three extreme point subjective priors $\lambda_i\in\Delta(\Theta),i=1,2,3$, where $\lambda_1=(1,0,0),\lambda_2=(0,1,0),\lambda_3=(0,0,1)$. For any $\lambda\in\Delta(\Theta)$, to find the concavification value of $\hat{V}_\lambda(\mu)$ at μ_0, we can assume without loss of generality that we only need to find the distribution of posterior $\tau\in BP(\mu_0)$ that induces different actions of the Receiver.[14] Then finding $cav(\hat{V}_\lambda)(\mu_0)$ becomes a linear programming problem. We obtain that $cav(\hat{V}_{\lambda_1})(\mu_0)=3$, where $\{(1,0,0),(0,1,0),(0,0,1)\}$ forms the support of Receiver's posterior distribution and

[14]This is true because if two posteriors μ_1,μ_2 in the support of τ lead to the same action of the Receiver, we can merge them as posterior $\frac{\tau(\mu_1)}{\tau(\mu_1)+\tau(\mu_2)}\mu_1+\frac{\tau(\mu_2)}{\tau(\mu_1)+\tau(\mu_2)}\mu_2$ with probability $\tau(\mu_1)+\tau(\mu_2)$.

they are realized with equal probability; $cav(\hat{V}_{\lambda_2})(\mu_0)=3$, where $(\frac{1}{2},\frac{1}{2},0),(0,0,1)$ forms the support of Receiver's posterior distribution and they are realized with probability $\frac{2}{3},\frac{1}{3}$ respectively; and $cav(\hat{V}_{\lambda_3})(\mu_0)=\frac{8}{3}$. So we can conclude that in this example $\mathcal{V}^*(\mu_0)\leq\frac{8}{3}$.

However, λ_3 is not the worst subjective prior in this case, even though it minimizes $cav(\hat{V}_{\lambda})(\mu_0)$ among the extreme points of $\Delta(\Theta)$. Next we show that $\lambda^*=(0,\frac{1}{2},\frac{1}{2})$ is the worst subjective prior by Proposition 5. We first calculate that $cav(\hat{V}_{\lambda^*})=\frac{5}{2}$ and the process of concavification splits μ_0 into $(\frac{1}{2},\frac{1}{4},\frac{1}{4}),(0,\frac{1}{2},\frac{1}{2})$ with probability $\frac{2}{3},\frac{1}{3}$ respectively. We denote this distribution over posterior as τ^*. We verify that $\mathcal{L}(\lambda_2,\tau^*)=\mathcal{L}(\lambda_3,\tau^*)=\frac{5}{2}<3=\mathcal{L}(\lambda_1,\tau^*)$. So by Proposition 5, λ^* is the worst subjective prior.

Moreover, we can show that under the binary-type cases, the credibility-gaining group exists for any C unless MDMB collapses to CT. This is more general than Theorem 2 as we don't impose the threshold on the budget.

Theorem 3. *Suppose the Sender has a binary type set. Under Assumption 1, for any $C_2>C_1$, either $\mathcal{V}^*_{C_2}(\mu_0)>\mathcal{V}^*_{C_1}(\mu_0)$ or $\mathcal{V}^*_{C_2}(\mu_0)=\mathcal{V}^*_0(\mu_0)=\mathcal{V}^*_{CT}(\mu_0)$.*

Proof of Theorem 3. Following the proof of Theorem 2, it suffices to observe that $U_2=\left\{\mu\in\mathbf{supp}\{p_{C_1}\}\middle|\sum_{\theta\in\Theta}\lambda_{C_2}(\theta)\frac{\mu(\theta)}{\mu_0(\theta)}=0\right\}$ contains at most one element. Therefore, by Claim 2, any posterior that does not belong to the persuasion group must yield the same payoff as those within the persuasion group. This directly constructs a cheap talk equilibrium.

□