Persuasion with Strategic Reporting*

Run Li† Wooyoung Lim‡

November, 2024

Abstract

We introduce strategic reporting in Bayesian persuasion. A sender conducts an experiment to acquire information to influence a receiver's action. After committing to an experiment, the sender privately observes its realized result and strategically reports a message. This reporting incurs a cost that depends on the realized result and the message reported and exhibits strictly decreasing differences. We develop a methodology to characterize the optimal experiment choice for the sender and provide a sufficient condition for the sender to choose an experiment whose results cannot be fully revealed to the receiver through reporting. Finally, we find that the sender may ex ante strictly prefer strategic reporting over fully committing to truthful reporting if truthful reporting, which incurs the minimum reporting cost for any realized result, is costly.

Keywords: Bayesian Persuasion; Strategic Manipulation; Reporting Cost; Information Transmission
JEL classification: D81; D82; D83; M37

*This paper supersedes the paper with the same title by the first author. We would like to thank Pak Hung Au, In-Koo Cho, Yuk-fai Fong, Navin Kartik, Jiangtao Li, Yang Lu, Eran Shmaya, Teck Yong Tan, Qinggong Wu, and the participants at the 2019 Asian Meeting of the Econometric Society, 2019 North American Summer Meeting of the Econometric Society, Congress 2019 of the European Economic Association and the Econometric Society, and 2019 Summer School of the Econometric Society for their valuable comments and suggestions. Run Li acknowledges the financial support from the Hong Kong University of Science and Technology (RTG19/20.BM004) at early stages of this research and from the Shenzhen Science and Technology Program (Grant No. RCBS20221008093309020) at later stages.

†First and corresponding author, Southern University of Science and Technology. Email: rliaj@connect.ust.hk
‡Hong Kong University of Science and Technology. Email: wooyoung@ust.hk

1 Introduction

Since Kamenica and Gentzkow (2011), the Bayesian persuasion literature has studied how a sender (she) persuades a receiver (he) by designing an experiment, which is a rule for acquiring information, under the assumption that the sender is fully committed to faithfully executing the experiment and to truthfully reporting the result of the experiment. This setting is useful in situations where information transmission is mandatory, such as prosecutors providing evidence, central banks providing economic data, and so on. On this basis, we want to further explore how the sender designs an experiment when it is not possible to guarantee that the obtained result will be fully revealed at zero cost. Specifically, we consider a scenario in which the sender commits to implementing an experiment and can strategically report its result by sending a message at a certain cost.

The information design problem that precedes strategic and costly reporting has significant economic implications. In reality, numerous organizations gather information under supervision or through public platforms, but how they transmit the information they obtain is associated with the reporting cost that depends on the market environment. For instance, before an open investigation, research institutions often publicly announce their research protocols, including sample sizes and methodologies. After obtaining a result, they may exaggerate it at the expense of integrity. Similarly, when producers build platforms to collect consumer opinions, they may fabricate positive comments if negative reviews emerge and promote consumers' unanimous praise through advertising otherwise.

To study this problem, we build a model that incorporates Bayesian persuasion and costly signaling. We investigate how the sender's incentive to reveal a result through strategic reporting affects information design. How does the sender design an optimal experiment ex ante to induce the most beneficial signaling game? Will the sender design an experiment whose results cannot be fully revealed through strategic reporting? We address these central questions in the following model.

We consider a model with two states, low and high. The realization of the state is unknown to both players. The sender always prefers a higher action while the receiver wants to take an action contingent on the state. To acquire information to persuade the receiver, the sender publicly chooses an experiment that consists of finite results, each of which is a posterior belief or the probability of the high state, and a Bayes-plausible probability distribution over them. The sender commits to conducting the chosen experiment. After privately observing a realized result, the sender reports a message to the receiver with a reporting cost. The exogenously assumed cost structure depends on both the realized result and the message reported, and exhibits strictly decreasing differences. Given any result, sending the sender's cost-minimizing message is considered truth-telling, while sending other costly messages is considered lying, manipulation, or signaling. The receiver takes an action in a continuous action space after observing the experiment and message chosen by the sender.

We first obtain a unique prediction about the sender's reporting strategy and expected payoff for any given experiment. After choosing an experiment, the two players play a signaling game where the realized result becomes the sender's private type. We focus on the unique sequential equilibrium selected by the D1 criterion (Cho and Kreps, 1987; Cho and Sobel, 1990) in each signaling subgame. Given any experiment, there is a cutoff type; any types below it (if any) separate themselves, and all higher types (if any) pool at the highest message (Lemma 1).

Our first contribution is to provide a methodology to characterize the sender's optimal experiment when she cannot commit to truth-telling (Proposition 1). In our framework, for any choice of experiment, the sender's payoffs after observing different results are interdependent and determined by the induced equilibrium characteristics. The typical approaches useful in the Bayesian persuasion literature do not allow us to solve the optimal experiment. By contrast, after establishing the existence of the optimal experiment (Appendix B), we establish the *Expected Pooling Cost Minimization Principle*: for an optimal experiment, pooling types (if any) have to be chosen such that the expected reporting cost of all pooling types reaches the lower bound of the convex hull of the cost function for the highest message. Then, we find an optimal experiment induces at most two types to separate with positive probability, as with more than two separating types, the sender can design a better experiment by deleting unfavorable ones, which reduces the signaling cost for separation.

Applying our methodology, we provide sufficient conditions for an optimal experiment to lead to the incomplete separation of the experiment results (Proposition 2). We show the sender may acquire information that cannot be transmitted in the reporting stage if the cost function is concave in types. Intuitively, for a type that separates at a certain cost, if this type can be substituted with one lower type and one higher type that pool, the lower type induces a higher cost while the higher type can induce a lower cost. Then, as long as the concavity lowers the weighted average of their pooling cost, the substitution is beneficial. By contrast, if the cost function is convex in types, the sender will not acquire any information that cannot be fully transmitted and thus induce full separation because having multiple pooling types leads to higher expected reporting cost relative to having a corresponding separating type.

Our second contribution is that we find the sender may ex ante strictly prefer strategic reporting over commitment to truthful reporting, if we assume truthful reporting, which occurs by sending the cost-minimizing message for any realized result, is costly. A strictly positive minimum reporting cost for any result captures the situations in which telling the truth requires some preparations of sound arguments. Under this assumption, we compare our strategic reporting case with the situation in which the sender commits to truthful reporting. We find that to obtain the same action, strategic reporting may incur a lower expected reporting cost via pooling, which is unfeasible in the corresponding commitment case and makes strategic reporting better (Proposition 3). This complements the well-established conclusion in the Bayesian persuasion literature: commitment to costless truth-telling is always optimal for the sender.

Lastly, we apply our methodology to explicitly characterize the optimal experiment under different restrictions. When separation is always costly and the cost structure is concave in types, it is better to distance the experiment results. Based on this logic, we derive the respective conditions for the fully informative and uninformative experiment to be optimal. We also obtain the condition under which the sender can strictly benefit from information design if she cannot commit to truthful reporting. Finally, we apply our main conclusions to the case in which the sender has linear utility. For different cost function forms, we fully characterize the optimal experiment and conduct a comparative statics analysis, finding that the sender chooses a weakly more informative experiment if the cost intensity is higher.

The remainder of the paper is organized as follows. We discuss the related literature in Section 1.1. Section 2 describes the model. We investigate the subgame after choosing any experiment in Section 3 and characterize the equilibrium of the whole game in Section 4. In Section 5, we consider the setting of costly truth-telling and compare strategic reporting with commitment to truthful reporting. In Section 6, we explicitly characterize the optimal experiment under different conditions. Section 7 concludes. All proofs are relegated to the Appendix.

1.1 Related Literature

This paper contributes to the literature that studies the information design problem faced by the sender who is unable to commit to truthfully reporting the information obtained. The growing literature builds on Bayesian persuasion[1] by the seminal paper Kamenica and Gentzkow (2011) with the key departure of relaxing the truth-telling assumption, and provides different ways to characterize strategic reporting. This paper makes a novel contribution by describing the strategic reporting behavior using signaling games and bridging Bayesian persuasion and signaling games. The classic signaling game analysis provides the theoretical foundation, that enables us to characterize the optimal experiment for the sender and make comparative statics analysis on the cost intensity. This paper is the first to identify the situation where it is optimal for the sender to design an experiment to obtain information that cannot be fully transmitted.

We review several closely related papers. Nguyen and Tan (2021) examine a model in which the sender sends a message, which is potentially costly, after committing to an experiment and privately observing its result. They focus on a cost structure that depends on the label of the realized results and the message. In contrast, our focus is on setting the cost to be dependent on the posterior beliefs of realized results and the message sent, such that more beneficial results lead to a lower marginal cost. They explore full separation of results after committing to an experiment, and our emphasis lies in characterizing the optimal experiment for the sender based on signaling subgames. Guo and Shmaya (2021) introduce a miscalibration cost and assume no commitment power in both the information design stage and the reporting stage. This means that the information structure chosen by the sender is also unobservable for the receiver. Min (2021) and Lipnowski et al. (2022) consider a situation where the sender commits to a signaling rule, after which she reports the true signal realization with a given probability and privately chooses a signal to send to the receiver with a complementary probability. Ederer and Min (2024) also consider partial commitment to a signaling rule when lying can be detected with certain probability. Lyu and Suen (2022) study the information design problem that precedes the cheap talk communication of the acquired information.[2] Lastly, Pei (2023), Best and Quigley (2024) and

[1]See Kamenica (2019) and Bergemann and Morris (2019) that provide comprehensive surveys of the Bayesian persuasion literature.

[2]Ivanov (2010) also studies information control before cheap talk but the sender's information structure is chosen by the receiver.

Mathevet et al. (2024) relax the commitment assumption in dynamic settings and provide justifications for the sender's commitment power in Bayesian persuasion.

Further, Hedlund (2017) and Perez-Richet and Skreta (2022) explore the situation in which the sender has private information before information design. Hedlund (2017) studies a Bayesian persuasion model in which the sender has private payoff-relevant information and then her choice of experiment signals her private information. Perez-Richet and Skreta (2022) introduce a designer to design an experiment after receiving the sender's state-related information. The designer aims to maximize the receiver's welfare while the informed sender can incur costs to falsify the true state.

Our notion of the reporting cost is conceptually related to money burning in Austen-Smith and Banks (2000) and the lying cost in Kartik et al. (2007) and Kartik (2009). These studies examine the costly communication between a perfectly informed sender and an uninformed receiver without the additional layer of an information design problem.

As noted above, our model bridges information design and signaling. Compared with the standard signaling game since Spence (1973), in our model, the information structure, considered as the set of sender types, is endogenously chosen by the sender. The signaling equilibria in Cho and Kreps (1987) and Cho and Sobel (1990) provide the theoretical foundation for the analysis of our model. In and Wright (2018) also consider an extended signaling game, in which a sender chooses her private type, rather than a publicly observable information structure, before a signaling game is played.

Our model is also related to the endogenized (covertly or overtly) information acquisition problem. Che and Kartik (2009), Pei (2015), Argenziano et al. (2016), and Kreutzkamp (2022) consider models in which the sender can obtain information with a cost before information transmission. Our model can be considered a costless information acquisition problem that precedes strategic reporting.

2 Model

2.1 Setup

There are two players, a sender (she) and a receiver (he). The state of the world $\theta \in \{L = 0, H = 1\}$. The realization of the state is unknown to both players; while they have the same prior belief μ, the probability that H is realized. The sender's utility only depends on the receiver's action while the receiver's utility also depends on the state. The sender's utility denoted by $U : A \to \mathbb{R}$ is continuous and strictly increasing in the receiver's action $a \in A \equiv [\underline{a}, +\infty)$, where $\underline{a}$ is the lower bound of the action space. When the state is $\theta = L, H$, the receiver's utility is denoted by $V^\theta : A \to \mathbb{R}$. We assume that V^θ is twice differentiable in a with $\frac{\partial^2 V^\theta}{\partial a^2} < 0$ and $\frac{\partial V^L}{\partial a} < \frac{\partial V^H}{\partial a}$, to let the receiver's ideal action, denoted by $a_\theta = \arg\max_{a \in A} V^\theta$, be higher for the high state, i.e., $a_L \leq a_H$. To guarantee the existence of the two ideal actions, we further assume there exists $a_H > \underline{a}$ such that $\frac{\partial V^H}{\partial a}\big|_{a=a_H} = 0$. In this model, the sender always prefers a higher action and the receiver would like to choose an action contingent on the state.

Information Design The sender publicly chooses an *information structure* that includes a finite set $\mathbb{S}$ of signal realizations and a signaling structure $\pi : \{L, H\} \to \Delta(\mathbb{S})$, a family of distributions over $\mathbb{S}$ conditional on each state. Then, a signal in $\mathbb{S}$ is realized according to π. The sender privately observes the realized signal and forms her posterior belief about the state.

According to Proposition 1 of Kamenica and Gentzkow (2011), choosing an information structure is equivalent to choosing a Bayes-plausible distribution of posterior beliefs $(1-t, t) \in \Delta(\{L, H\})$, where t denotes the probability of H. The sender's choice of information structure can then be transformed to choosing an *experiment* $\boldsymbol{\tau} = (T \equiv \{t_1, \ldots, t_n\}; \tau(\cdot) \equiv (\tau_1, \ldots, \tau_n))$,[3] where

- $0 \le t_1 \le \cdots \le t_n \le 1$;
- $\tau(\cdot) \in \text{int}(\Delta(T))$ is a probability distribution over T that satisfies

$$\sum_{i=1}^{n} t_i \tau_i = \mu.$$

T contains finite elements and $||T|| = n$ is chosen by the sender. We call $t \in T$ a *result* and T a result set. By definition, $\tau(\cdot)$ is a Bayes-plausible probability distribution over the result set.[4] After an experiment $\boldsymbol{\tau}$ is chosen, it becomes common knowledge between the players. Then, one result $t \in T$ is drawn according to $\tau(\cdot)$, which is *privately* observed by the sender.

Strategic (Mis-)Reporting The sender then reports a message $m \in M$ to the receiver at a reporting cost $c(t, m)$. The message space $M = [0, \overline{m}]$ with the highest message $\overline{m} > 0$ and cost function $c : [0, 1] \times M \to [0, +\infty)$ are exogenously given and independent of the sender's choice of experiment. The sender's reporting cost depends on both the realized result and the message sent. c is continuous in both variables and strictly quasi-convex in m for any given t, implying that, for any $t \in [0, 1]$, there exists a unique cost-minimizing message $m_c(t) \equiv \arg\min_{m \in M} c(t, m)$. The cost function also satisfies a single-crossing condition: $\frac{\partial^2 c}{\partial t \partial m} < 0$, that is, the sender's marginal cost of sending a message is higher after observing a lower result. We assume that the sender sending $m_c(t)$ after observing t, which is considered *truth-telling*, is always costless, i.e., the minimum cost $c(t, m_c(t)) = 0$ for any result t. To eliminate trivial cases, assume $m_c < \overline{m}$.

The receiver observes the experiment $\boldsymbol{\tau}$ and the message m chosen by the sender, after which he takes an action $a \in A$, determining both players' payoffs. All the information in the game except the realized result is common knowledge between the players. The timing of the game is summarized as follows.

- *Stage 1: Information Design*
  - The sender designs and faithfully implements an experiment $\boldsymbol{\tau} = (T; \tau(\cdot))$.

[3] In our model, the sender's reporting cost depends on the message and posterior. Therefore, it is without loss of generality to define experiments as distributions of posteriors.

[4] For an experiment with two results $t_1 < t_2$, $\tau(\cdot)$ is determined by these two results and the experiment should be $(t_1, t_2; \frac{t_2-\mu}{t_2-t_1}, \frac{\mu-t_1}{t_2-t_1})$.

- The experiment chosen by the sender becomes common knowledge.

- *Stage 2: Strategic Reporting*
  - Nature draws one experiment result $t \in T$ according to $\tau(\cdot)$.
  - The sender privately observes the result t and sends a message $m \in M$ to the receiver that incurs a cost $c(t, m)$.
  - The receiver observes $\boldsymbol{\tau}$ and m and takes action $a \in A$.

2.2 Discussion of the Assumptions

The reporting cost mainly captures the cost of transmitting or manipulating information and the message can be natural language, advertisement, persuasive argument, and so on. The cost structure is exogenously assumed. The interpretation is that the cost of sending messages depends on the market environment. In a market filled with advertising and promotion, the sender needs to transmit information through advertising media. The cost is associated with the signaling cost incurred by promotion, and the higher the cost, the more credible the report. In a market in which there are ways to manipulate information, the cost incurred is associated with the fabrication or lying cost.

The single-crossing condition imposed on the cost function (i.e., $\frac{\partial^2 c}{\partial t \partial m} < 0$) is crucial but natural. This condition is standard in signaling games and requires the cost exhibits strictly decreasing differences: for any two messages, if the sender strictly prefers the higher message after obtaining a result t, she must strictly prefer the higher message after obtaining any result higher than t. Roughly speaking, it is less costly for the sender to send a high message when she obtains a higher result. Since the sender's preference for the receiver's action is state-independent, i.e., U strictly increases in a, the condition $\frac{\partial^2 c}{\partial t \partial m} < 0$ guarantees the indifference curves $U - c$ of different sender types through a fixed action-message pair intersect only once. This single-crossing condition is satisfied by a wide range of functional forms which have rich economic implications. For instance, the cost of advertising can be represented by $c = (1 - bt)m$, where $b \in (0, 1)$ is constant, meaning that, the higher the result, the lower the cost of advertising. The function captures one kind of signaling cost and the sender can utilize costly messages to signal her obtained result. Another example is that the sender's cost from lying or manipulation of experiment results can be represented by $c = (m - t)^2$, that captures the cost depending on the "size" of a lie, similar to the lying cost in Kartik (2009).[5]

2.3 Solution Concept

The solution concept is *the D1 subgame perfect equilibrium*. For any experiment that may be chosen in Stage 1, a corresponding subgame is played in Stage 2. Since no information asymmetry exists in Stage 1, we mainly employ the notion of subgame perfection to obtain the solution of the entire game. In any subgame, we focus on the sequential equilibrium selected by the D1 criterion (Cho and Kreps, 1987),

[5] The lying cost function form, e.g., $c = (m - t)^2$, satisfies the single-crossing condition, and the sender's payoff function $U(a) - c$ has the single-crossing property. This is different from the setting of Kartik (2009), in which the sender's preference for the action is state-dependent and the sender's payoff function does not has the single-crossing property.

called the *D1 equilibrium*, the existence and uniqueness of which are proved by Cho and Sobel (1990). The next section offers a detailed illustration.

Given any experiment $\boldsymbol{\tau} = (T; \tau(\cdot))$, let $\sigma(\cdot|t) : T \to \Delta(M)$ denote the sender's reporting strategy, $a^R(\cdot) : M \to A$ denote the receiver's action strategy, and $\rho(\cdot|m) : M \to \Delta(T)$ denote the receiver's posterior belief.[6] We restrict attention to equilibria in which σ has finite support. Then, $\sigma(m|t)$ is the probability that the sender sends m after observing t. The D1 subgame perfect equilibrium is represented by $(\boldsymbol{\tau}^*, (\sigma^*, a^{R*}, \rho)_{\boldsymbol{\tau}})$, where

- $(\sigma^*, a^{R*}, \rho)_{\boldsymbol{\tau}}$ is the D1 equilibrium of the subgame given any $\boldsymbol{\tau}$; and
- $\boldsymbol{\tau}^*$ is the experiment chosen by the sender in Stage 1 by taking $(\sigma^*, a^{R*}, \rho)_{\boldsymbol{\tau}}$ as given.

3 Preliminaries: Signaling Subgames

In this section, we reformulate each subgame as a signaling game and characterize its D1 equilibrium.

Given any experiment $\boldsymbol{\tau}$, the subsequent subgame in Stage 2 is a signaling game. The result $t_i \in T$ drawn according to $\tau(\cdot)$ is considered the sender's private *type*[7], and the receiver holds the prior belief $(\tau_1, \ldots, \tau_n)$ over all possible types, where τ_i is the probability assigned to t_i. Denote the receiver's expected utility conditional on any result as

$$V(t_i, a) \equiv t_i V^H + (1 - t_i) V^L.$$

Then, $V : [0,1] \times A \to \mathbb{R}$ satisfies $V_{aa} < 0$ and $V_{at} > 0$, so that the receiver has a unique *optimal action for any type*, denoted by $\alpha^R(t_i) \equiv \arg\max_{a \in A} V(t_i, a)$, which weakly increases in $t_i \in [0, 1]$.[8] After observing t_i, the sender reports a message with a cost, and the receiver takes an action. A sequential equilibrium of the subgame induced by experiment $\boldsymbol{\tau}$ is a triple $(\sigma, a^R, \rho)_{\boldsymbol{\tau}}$ that satisfies

1. for all $t_i \in T$, if $\sigma(m'|t_i) > 0$, then $m' \in \arg\max_{m \in M}\ U(a^R(m)) - c(t_i, m)$; and
2. for all $m \in M$, $a^R(m) = \arg\max_{a \in A}\ \sum_{i=1}^n V(t_i, a)\rho(t_i|m)$, where
3. $\rho(t_i|m) = \frac{\sigma(m|t_i)\tau_i}{\sum_{j=1}^n \sigma(m|t_j)\tau_j}$ if $\sum_{j=1}^n \sigma(m|t_j)\tau_j > 0$.

In any subgame, however, there could be multiple sequential equilibria, which brings about difficulties in comparing different experiments as the sender's expected payoff from choosing any experiment cannot be uniquely determined. To make equilibrium selection, we apply the D1 criterion, which restricts the off-equilibrium-path beliefs and has been widely applied in the signaling game literature.[9] Especially for this subgame, it neither constrains too much to make the equilibrium non-existent nor constrains

[6] The triple (σ, a^R, ρ) varies according to different $\boldsymbol{\tau}$. To be rigorous, we can represent the sender's strategy after choosing $\boldsymbol{\tau}$ as $\sigma_{\boldsymbol{\tau}} : T \to \Delta(M)$, where T is the result set of $\boldsymbol{\tau}$. We remove the subscript for the sake of brevity. Moreover, owing to the strict concavity of V in a, we consider pure strategies of the receiver action without loss of generality.

[7] We use the two terms "result" and "type" interchangeably.

[8] See Lemma A.1 in the Appendix for the proof.

[9] For example, Hedlund (2017) and Heese and Liu (2023) also apply the D1 criterion to consider signaling in information design problems.

too little to bring about multiple equilibria, as the D1 equilibrium always exists and is unique. Roughly speaking, it requires the receiver to believe that any off-path message is sent by the type that is most likely to benefit from such deviation. The formal definition is as follows.

In the subgame after choosing $\boldsymbol{\tau}$, given a sequential equilibrium (σ, a^R, ρ), for message $m, \widetilde{m}$ s.t. $\sigma(m|t_i) > 0$ and $\sum_{j=1}^{n} \sigma(\widetilde{m}|t_j)\tau_j = 0$, define the set $D(\widetilde{m}, t_i) = \{a \in [a_L, a_H] \big| \ U(a^R(m)) - c(t_i, m) < U(a) - c(t_i, \widetilde{m})\}$ and $D^0(\widetilde{m}, t_i) = \{a \in [a_L, a_H] \big| \ U(a^R(m)) - c(t_i, m) = U(a) - c(t_i, \widetilde{m})\}$. Given the receiver responds optimally, $D(\widetilde{m}, t_i) \cup D^0(\widetilde{m}, t_i)$ is the set of the actions in response to $\widetilde{m}$ that makes type t_i weakly prefer to deviate from her equilibrium message and send $\widetilde{m}$.

Definition 1. *In the signaling subgame induced by experiment $\boldsymbol{\tau}$, a sequential equilibrium $(\sigma, a^R, \rho)_{\boldsymbol{\tau}}$ satisfies the D1 criterion if for any off-equilibrium-path message $\widetilde{m}$, $\rho(t_j|\widetilde{m}) = 0$ whenever there exists $t_k \in T$ such that $D(\widetilde{m}, t_j) \cup D^0(\widetilde{m}, t_j) \subset D(\widetilde{m}, t_k)$ and $D(\widetilde{m}, t_k) \neq \varnothing$.*

Next, based on Cho and Sobel (1990, Proposition 4.1-4.4),[10] we provide the full characterization of the D1 equilibrium induced by $\boldsymbol{\tau}$. To simplify the notation, denote

$$\hat{U}(t) \equiv U(\alpha^R(t))$$

as the sender's utility given that the receiver believes the realized result is t. We then calculate the *separating message* $m_i(\boldsymbol{\tau})$, which is the least costly message utilized by type t_i for full separation of $\{t_1, \dots, t_i\}$, recursively. First, $m_1(\boldsymbol{\tau}) := m_c(t_1)$. Next, for $i \geq 2$,

$$\begin{aligned} m_i(\boldsymbol{\tau}) &:= \underset{m \in M}{\arg\max} \ \hat{U}(t_i) - c(t_i, m) = \underset{m \in M}{\arg\min} \ c(t_i, m) \\ \text{s.t. } & \hat{U}(t_{i-1}) - c(t_{i-1}, m_{i-1}(\boldsymbol{\tau})) \geq \hat{U}(t_i) - c(t_{i-1}, m) \\ & \text{and } m \geq m_{i-1}(\boldsymbol{\tau}). \end{aligned} \tag{IC}$$

The inequality is the incentive compatibility (IC) condition for separation, which guarantees types t_{i-1} and t_i have no incentive to mimic each other. Define the expected value of the types, conditional on the type being higher than t_i or exactly at t_i with probability $x \in (0, 1]$, as

$$\phi_i(\boldsymbol{\tau}, x) = \mathbb{E}_\tau\Big(t \in T \Big| \ t = t_i|_{\text{prob } x} \text{ , or } t > t_i\Big) = \frac{t_i \tau_i x + \sum_{k=i+1}^{n} t_k \tau_k}{\tau_i x + \sum_{k=i+1}^{n} \tau_k}.$$

Lemma 1 (Cho and Sobel (1990))**.** *In the signaling subgame induced by experiment $\boldsymbol{\tau} = (t_1, \dots, t_n; \tau_1, \dots, \tau_n)$, there exists a unique D1 equilibrium outcome derived by the following induction procedure starting from type t_1. The recursive induction of the strategy of type t_i is as follows.*

1. *If $\hat{U}(t_i) - c(t_i, m_i(\boldsymbol{\tau})) \leq \hat{U}(\phi_i(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m})$, then t_j, $j \geq i$, sends $\overline{m}$.*
2. *If $\hat{U}(\phi_i(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m}) < \hat{U}(t_i) - c(t_i, m_i(\boldsymbol{\tau})) < \hat{U}(\phi_{i+1}(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m})$, t_i sends $m_i(\boldsymbol{\tau})$ with*

[10] The setting of our model satisfies their assumptions (A0-A4 in Section 4 in Cho and Sobel (1990)), and their conclusions about the characterization and uniqueness of the D1 equilibrium apply.

probability $1-q$ *and* $\overline{m}$ *with probability* q*, where* q *satisfies* $\hat{U}(t_i) - c(t_i, m_i(\boldsymbol{\tau})) = \hat{U}(\phi_i(\boldsymbol{\tau}, q)) - c(t_i, \overline{m})$*. Then,* t_j*,* $j \geq i+1$*, sends* $\overline{m}$*.*

3. *If* $\hat{U}(t_i) - c(t_i, m_i(\boldsymbol{\tau})) \geq \hat{U}(\phi_{i+1}(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m})$*,* t_i *separates by sending* $m_i(\boldsymbol{\tau})$*. We then continue the induction procedure to analyze the strategy for type* t_{i+1}*.*

Lemma 1 guarantees each experiment choice leads to a unique D1 equilibrium in the subsequent subgame. Then, in Stage 1, the sender chooses an experiment that can induce the highest expected payoff for her, denoted by $\boldsymbol{\tau}^*$ and called an *optimal experiment*. The recursive induction of the sender's equilibrium strategy provides a foundation for the subsequent analysis of different experiment choices. The D1 equilibrium must be one of two kinds:

1. A *separating equilibrium* in which every type of the sender separates herself by sending her separating message so that the acquired information is fully transmitted.[11]

2. A *pooling equilibrium* in which there exists a threshold type t_p such that all lower types (if any) separate and all higher types pool at $\overline{m}$. To determine t_p, we need to sequentially check types t_1 to t_{n-1}, to find out the lowest type that has an incentive to pool with all higher types. This equilibrium can be partial-pooling or total-pooling, in which partial or no information is transmitted.

Full separation is not a mere consequence of the expansion of the message space. It depends on both $\overline{m}$ and the characteristic of the cost function. All types need to utilize messages *costly enough* to separate themselves from all lower types. If the cost function is bounded above, even with $\overline{m} \to +\infty$, there may exist an experiment that induces a pooling equilibrium.

4 Optimal Experiment

Note that in Kamenica and Gentzkow (2011), as no strategic incentive is allowed in the reporting stage, the equilibrium payoff of each type can be considered a function merely depending on the posterior or the type itself, in which case the concavification approach is useful. In our setting, however, the equilibrium payoff of each type is also determined by the choice of experiment. Since given any experiment, all types' equilibrium payoffs are interdependent, the standard approach cannot be applied directly and we need to compare the different experiments to narrow the range of the optimal experiment stepwise. We investigate how to design an optimal experiment (Section 4.1) and when the sender will design an experiment to obtain information that cannot be fully transmitted through reporting (Section 4.2).

While a unique D1 equilibrium exists in each signaling subgame, the existence of a D1 subgame perfect equilibrium or an optimal experiment is not automatically guaranteed, as there are infinite subgames or experiment choices.[12] We establish the existence of the optimal experiment by proving the

[11] The D1 equilibrium induced by the experiment that contains only one result, $(\mu; 1)$, is still called a separating equilibrium.

[12] In our game, the existence of an optimal experiment is not trivial because of two challenges. First, the set of all possible experiments is not closed. Second, the sender's expected payoff is discontinuous in the choice of experiment. This discontinuity is caused by the discontinuous changes in the equilibrium strategy as the number of results decreases. In Appendix B, we show an optimal experiment needs finite results and transform the set of all experiments with at most n

sender's expected payoff is upper semi-continuous across all experiment choices. The formal proof of the existence is provided in Appendix B.

4.1 Characterization

Let us call an experiment that induces a signaling subgame with a unique separating (pooling) D1 equilibrium a *separating (pooling) experiment.* A pooling experiment can be either a total-pooling or partial-pooling experiment.

We introduce several notations. In the signaling subgame induced by an experiment $\boldsymbol{\tau}$, denote the equilibrium payoff function of each type $t_i \in T$ as $f_{\boldsymbol{\tau}} : T \to \mathbb{R}$. Note that for any given t_i, the function $f_{\boldsymbol{\tau}}(t_i)$ varies depending on $\boldsymbol{\tau}$. Then, the sender's expected payoff from experiment $\boldsymbol{\tau}$, denoted by $E(\boldsymbol{\tau}) = \sum_{i=1}^{n} \tau_i f_{\boldsymbol{\tau}}(t_i)$, is a convex combination of them. In Stage 1, the sender chooses an *optimal experiment* $\boldsymbol{\tau}^*$ that maximizes her expected payoff.

If the sender intends to produce any pooling in the strategic reporting stage, we find her expected cost from pooling would always approach a certain minimum, stated as the *Expected Pooling Cost Minimization Principle*. To describe this property, for any pooling experiment $\boldsymbol{\tau}$, let t_p denote the lowest type with $\sigma^*(\overline{m}|t) > 0$. Denote the expected value of all pooling types as $\phi_p^*(\boldsymbol{\tau}) = \frac{\sum_{i=p}^{n} t_i \tau_i \sigma^*(\overline{m}|t_i)}{\sum_{i=p}^{n} \tau_i \sigma^*(\overline{m}|t_i)}$ and the *expected reporting cost from pooling* (at $\overline{m}$) as

$$ECP_{\boldsymbol{\tau}} \equiv \frac{\sum_{i=p}^{n} c(t_i, \overline{m}) \tau_i \sigma^*(\overline{m}|t_i)}{\sum_{i=p}^{n} \tau_i \sigma^*(\overline{m}|t_i)}.$$

Denote the reporting cost at $\overline{m}$ for result $t \in B$, where $B \subseteq [0,1]$ is a closed convex set, as function $c(t, \overline{m})|_B : B \to [0, +\infty)$. Define the *convex lower closure* of the function $c(t, \overline{m})|_B$ as

$$\mathcal{C}(\hat{t}, \overline{m})|_B \equiv \inf \left\{ z | (\hat{t}, z) \in co(c(t, \overline{m})|_B) \right\},$$

where $co(c(t, \overline{m})|_B)$ denotes the convex hull of the graph of $c(t, \overline{m})|_B$. $\mathcal{C}(\hat{t}, \overline{m})|_B$ is the largest convex function that is weakly smaller than $c(\cdot, \overline{m})$ everywhere on B. Figure 1 illustrates the convex lower closure of $c(t, \overline{m})|_{[0,1]}$ and $c(t, \overline{m})|_{[\mu,1]}$, respectively.

Clearly, $ECP_{\boldsymbol{\tau}} \geq \mathcal{C}(\phi_p^*(\boldsymbol{\tau}), \overline{m})\big|_{[t_p,1]} \geq \mathcal{C}(\phi_p^*(\boldsymbol{\tau}), \overline{m})\big|_{[\underline{t},1]}$, $\underline{t} \leq t_p$. The next lemma shows when a pooling experiment is optimal, the expected reporting cost from pooling must be minimized to a certain convex lower closure.

Lemma 2 (Expected Pooling Cost Minimization Principle)**.** *If an optimal experiment* $\boldsymbol{\tau}^*$ *is a pooling experiment, then*

$$ECP_{\boldsymbol{\tau}^*} = \mathcal{C}(\phi_p^*(\boldsymbol{\tau}^*), \overline{m})\big|_{[t_p,1]}.$$

results to a closed set, for any finite $n \in \mathbb{N}$. Then, we prove the expected payoff satisfies the upper semi-continuity.

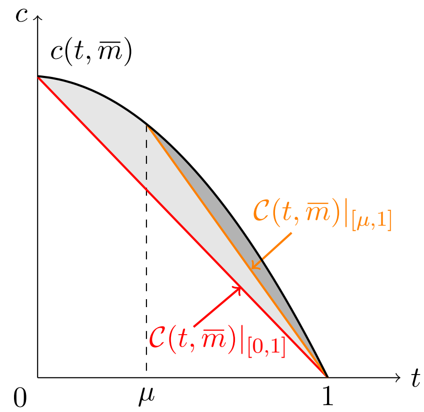

Figure 1: Convex Lower Closure

Moreover, when $\sigma^*(\overline{m}|t_p) = 1$,

$$ECP_{\boldsymbol{\tau}^*} = \begin{cases} \mathcal{C}(\phi_p^*(\boldsymbol{\tau}^*), \overline{m})\big|_{[0,1]} & \textit{if } p = 1, \\ \mathcal{C}(\phi_p^*(\boldsymbol{\tau}^*), \overline{m})\big|_{[t_{p-1},1]} & \textit{otherwise.} \end{cases}$$

To understand why this principle must hold, we consider a case of total-pooling.[13] Suppose $\boldsymbol{\tau} = (t_1, t_2; \tau_1, \tau_2)$ induces t_1, t_2 to pool at $\overline{m}$ and is optimal. Then, $t_p = t_1$, $\phi_p^*(\boldsymbol{\tau}) = \mu$, and

$$E(\boldsymbol{\tau}) = \tau_1[\hat{U}(\mu) - c(t_1, \overline{m})] + \tau_2[\hat{U}(\mu) - c(t_2, \overline{m})] = \hat{U}(\mu) - ECP_{\boldsymbol{\tau}}.$$

Suppose $ECP_{\boldsymbol{\tau}} > \mathcal{C}(\mu, \overline{m})\big|_{[0,1]}$ and there are two points $(t_1', c(t_1', \overline{m}))$ and $(t_2', c(t_2', \overline{m}))$, $t_1' < t_2'$, such that

$$\begin{aligned} &\tau_1' c(t_1', \overline{m}) + \tau_2' c(t_2', \overline{m}) = \mathcal{C}(\mu, \overline{m})\big|_{[0,1]}, \\ &\tau_1' t_1' + \tau_2' t_2' = \mu. \end{aligned}$$

Then, we can show experiment $\boldsymbol{\tau'} = (t_1', t_2'; \tau_1', \tau_2')$ is strictly better than $\boldsymbol{\tau}$, which leads to a contradiction, for the following reason.

If $\boldsymbol{\tau'}$ induces a total-pooling equilibrium in which t_1' and t_2' pool at $\overline{m}$, we have $E(\boldsymbol{\tau'}) = \hat{U}(\mu) - \mathcal{C}(\mu, \overline{m})\big|_{[0,1]} > E(\boldsymbol{\tau})$. Otherwise, $\boldsymbol{\tau'}$ induces a separating equilibrium, or a partial-pooling equilibrium in which t_1' pools at $\overline{m}$ with probability $q \in (0, 1)$. By the IC condition for any separation, we have

$$f_{\boldsymbol{\tau'}}(t_1') = \hat{U}(t_1') > \hat{U}(\mu) - c(t_1', \overline{m}).$$

Since

$$f_{\boldsymbol{\tau'}}(t_2') \geq \hat{U}(t_2') - c(t_2', \overline{m}) \quad \text{or} \quad f_{\boldsymbol{\tau'}}(t_2') = \hat{U}(\phi_1(\boldsymbol{\tau'}, q)) - c(t_2', \overline{m}),$$

[13] Although a total-pooling experiment is never optimal because it is always worse than the experiment $(\mu; 1)$, for the sake of clarity, we use it as an example to illustrate the logic of the proof.

then

$$f_{\boldsymbol{\tau'}}(t_2') > \hat{U}(\mu) - c(t_2', \overline{m}).$$

Therefore, $E(\boldsymbol{\tau'}) > E(\boldsymbol{\tau})$.

The newly constructed experiment $\boldsymbol{\tau'}$ is better regardless of which kind of equilibrium it induces. If it still induces total-pooling, the expected reporting cost is reduced. If it induces any separation, by the IC condition, the lower type's payoff from separation must be higher than that from total-pooling and the remaining pooling types can obtain a higher receiver action as there is less pooling. By the same logic, for any partial-pooling experiment that does not satisfy the Expected Pooling Cost Minimization Principle, we can always substitute its pooling types similarly to above, without changing its separating part, to construct a strictly better experiment.

Proposition 1. *The optimal experiment $\boldsymbol{\tau}^*$ needs at most three types. Specifically,*

1. *$\boldsymbol{\tau}^*$ needs at most two types to induce a separating equilibrium; or*
2. *$\boldsymbol{\tau}^*$ needs at most three types to induce a pooling equilibrium, in which at most two types pool with positive probability and at most one type separates with positive probability, and satisfies the Expected Pooling Cost Minimization Principle.*

To achieve this conclusion, we compare payoffs of each type across experiments. Given any $\boldsymbol{\tau}$, define the *concave closure* of $f_{\boldsymbol{\tau}}(t_i)$, $t_i \in T$ as

$$\mathcal{F}_{\boldsymbol{\tau}}(\hat{t}) \equiv \sup\left\{z|(\hat{t}, z) \in co(f_{\boldsymbol{\tau}})\right\},$$

where $co(f_{\boldsymbol{\tau}})$ denotes the convex hull of the graph of $f_{\boldsymbol{\tau}}$. $f_{\boldsymbol{\tau}}$ is defined on T, while $\mathcal{F}_{\boldsymbol{\tau}}$ is defined over $[t_1, t_n]$, the smallest convex set that contains T. Then, $E(\boldsymbol{\tau}) \leq \mathcal{F}_{\boldsymbol{\tau}}(\mu)$.

First, Proposition 1 implies it is without loss of generality to limit our attention to the set of experiments with only two types or one type to characterize the optimal experiment if it is a separating experiment. The reason is as follows. Consider any separating experiment $\boldsymbol{\tau}$ with $\|T\| > 2$. There must exist two points[14] $(t_j, f_{\boldsymbol{\tau}}(t_j))$ and $(t_k, f_{\boldsymbol{\tau}}(t_k))$, $t_j, t_k \in T$, $t_j < \mu < t_k$, such that

$$\begin{aligned} r f_{\boldsymbol{\tau}}(t_j) + (1-r) f_{\boldsymbol{\tau}}(t_k) &= \mathcal{F}_{\boldsymbol{\tau}}(\mu), \\ r t_j + (1-r) t_k &= \mu, \end{aligned}$$

where $r = \frac{t_k - \mu}{t_k - t_j}$. We can utilize the two types to construct a new experiment $\boldsymbol{\tau'} = (t_j, t_k; r, 1-r)$, which would induce a separating equilibrium with $f_{\boldsymbol{\tau'}}(t_j) \geq f_{\boldsymbol{\tau}}(t_j)$ and $f_{\boldsymbol{\tau'}}(t_k) \geq f_{\boldsymbol{\tau}}(t_k)$ because fewer types can spend weakly lower reporting cost to signal their types due to fewer IC conditions for separation. Thus, its expected payoff $E(\boldsymbol{\tau'}) \geq \mathcal{F}_{\boldsymbol{\tau}}(\mu) \geq E(\boldsymbol{\tau})$. For any separating experiment with more than two types, we can delete all "unbeneficial" types to let the remaining type(s) construct a weakly better separating experiment that consists of at most two types, as illustrated in Figure 2.

[14] Or one point $(\mu, f_{\boldsymbol{\tau}}(\mu))$ s.t. $f_{\boldsymbol{\tau}}(\mu) = \mathcal{F}_{\boldsymbol{\tau}}(\mu)$. Then, experiment $(\mu; 1)$ induces a weakly higher expected payoff than $\boldsymbol{\tau}$.

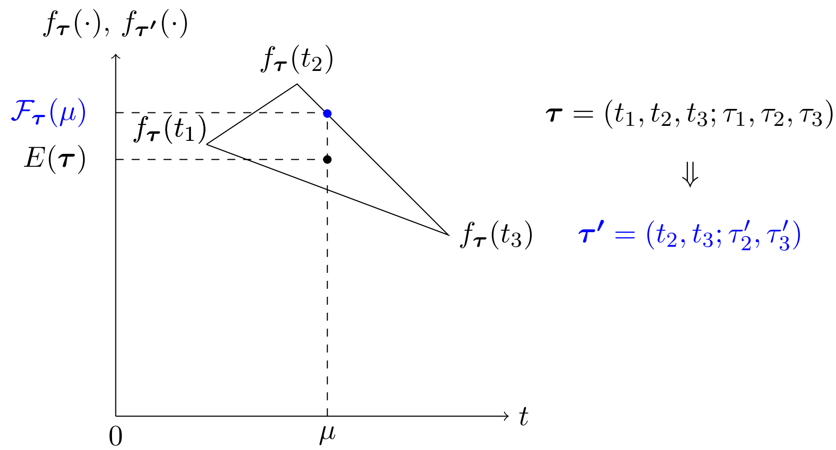

Figure 2: A Separating Experiment with Three Types

Next, we investigate the case in which $\boldsymbol{\tau}^*$ is a pooling experiment. Proposition 1 indicates that neither too much separation nor too much pooling is beneficial. On the one hand, if the number of separating types exceeds the number of states, positive probabilities are allocated to unbeneficial types and some signaling costs are wasted. Moreover, having fewer separating types weakens the IC constraints and enables some pooling types to choose separation whenever it is more beneficial than pooling, which also lets the remaining pooling types gain a weakly higher action. On the other hand, based on the Expected Pooling Cost Minimization Principle, at most two types are needed to approach any convex lower closure. To summarize, there are three cases: the optimal experiment needs two types that pool; it needs two types and the lower type mixes between separation and pooling with the higher type; it needs three types, with the lowest type separating and the other two pooling.

4.2 Sufficient Conditions for Full/Partial Information Transmission

In this subsection, we focus on providing sufficient conditions for $\boldsymbol{\tau}^*$ to be a separating experiment and to be a partial-pooling experiment. We denote $t_L = 0$ and $t_H = 1$.

Corollary 1 (Sufficient Condition for Separation)**.** *The sender must choose a separating experiment if either of the following conditions holds:*

1. $c(t, \overline{m})$ *is strictly convex in* $t \in [\mu, 1]$, *and* $\mathcal{C}(t, \overline{m})\big|_{[\mu,1]} = \mathcal{C}(t, \overline{m})\big|_{[0,1]}$, $\forall\, t \in [\mu, 1]$;
2. $c(t, \overline{m}) \geq \hat{U}(t_H) - \hat{U}(t_L)$, $\forall\, t$.

The first condition restricts the shape of the cost structure. To obtain the condition, let us start with a stronger version of it: $c(t, \overline{m})$ is strictly convex in t everywhere. Under this condition, any pooling experiment is dominated by a corresponding separating experiment, derived by substituting all the pooling types with one type that equals their expectation. To show this, consider a partial-pooling experiment $\boldsymbol{\tau} = (t_1, \ldots, t_n; \tau_1, \ldots, \tau_n)$ such that $t_1, \ldots, t_{p-1}$ separate and $t_p, \ldots, t_n$ pool. Recall that

$\phi_p^*(\boldsymbol{\tau})$ is the expected value of all pooling types. Since

$$\hat{U}(t_{p-1}) - c(t_{p-1}, m_{p-1}(\boldsymbol{\tau})) \geq \hat{U}(\phi_p^*(\boldsymbol{\tau})) - c(t_{p-1}, \overline{m}),$$

the new experiment $\hat{\boldsymbol{\tau}} = (t_1, \dots, t_{p-1}, \hat{t}_p; \tau_1, \dots, \tau_{p-1}, \hat{\tau}_p)$, where $\hat{t}_p = \phi_p^*(\boldsymbol{\tau})$, $\hat{\tau}_p = \sum_{j=p}^n \tau_j$, must induce a separating equilibrium. Since $m_i(\boldsymbol{\tau}) = m_i(\hat{\boldsymbol{\tau}})$, $i \leq p-1$, we have $f_{\boldsymbol{\tau}}(t_i) = f_{\hat{\boldsymbol{\tau}}}(t_i)$. By the convexity of the cost function, $\hat{\tau}_p f_{\hat{\boldsymbol{\tau}}}(\hat{t}_p) > \sum_{j=p}^n \tau_j f_{\boldsymbol{\tau}}(t_j)$, i.e., the new experiment is strictly better. Furthermore, because the expectation of pooling types is always weakly higher than μ, the convexity of $c(t, \overline{m})$ can be relaxed, as the first condition in Corollary 1 states. This sufficient condition can also be considered a direct consequence of Lemma 2: when the cost function is strictly convex, any pooling cannot minimize the expected reporting cost, so it is never optimal.

The second condition restricts the scale of the reporting cost to ensure full separation. In any experiment $(t_1, \dots, t_n; \tau_1, \dots, \tau_n)$, for any type t_i, if it sends $\overline{m}$, its payoff is weakly lower than $\hat{U}(t_n) - c(t_i, \overline{m})$ regardless of whether it separates or pools; if it sends its costless message, which may be an off-path message, its payoff is weakly higher than $\hat{U}(t_1)$. Hence, the condition $\hat{U}(t_L) > \hat{U}(t_H) - c(t, \overline{m})$ guarantees any type has no incentive to send $\overline{m}$ and any experiment induces a separating equilibrium. Under either condition in Corollary 1, the sender never acquires information that cannot be transmitted.

We now move our attention to characterizing the sufficient conditions for the sender choosing a partial-pooling experiment. We define the *best separating experiment* $\boldsymbol{\tau^s}$ as the separating experiment that achieves the highest expected payoff for the sender among all separating experiments. $\boldsymbol{\tau^s}$ always exists (Lemma B.1, Appendix B). The following proposition guarantees there exists a partial-pooling experiment that is better than $\boldsymbol{\tau^s}$, which means an optimal experiment must be partial-pooling.

Proposition 2 (Sufficient Condition for Pooling)**.** *The sender must choose a partial-pooling experiment if the best separating experiment* $\boldsymbol{\tau^s}$ *is informative and* $\boldsymbol{\tau^s} = (t_1, t_2; \tau_1, \tau_2)$ *satisfies* $c(t_2, m_2(\boldsymbol{\tau^s})) > \mathcal{C}(t_2, \overline{m})\big|_{[t_1,1]}$.

Let us sketch the proof. If the condition in Proposition 2 is satisfied, we can substitute type t_2 of $\boldsymbol{\tau^s}$ with two types, t' and t'', to construct a new experiment $\boldsymbol{\tau'} = (t_1, t', t''; \tau_1, r\tau_2, (1-r)\tau_2)$[15], where $t' < t_2 < t''$ satisfy

$$\begin{aligned} &rt' + (1-r)t'' = t_2, \\ &rc(t', \overline{m}) + (1-r)c(t'', \overline{m}) = \mathcal{C}(t_2, \overline{m})\big|_{[t_1,1]}. \end{aligned}$$

We then show experiment $\boldsymbol{\tau'}$ induces a partial-pooling equilibrium with $E(\boldsymbol{\tau'}) > E(\boldsymbol{\tau^s})$, that means the optimal experiment must be partial-pooling and the sender acquires information that cannot be transmitted.

Let us consider the equilibrium induced by $\boldsymbol{\tau'}$. t_1 will still separate and obtain the same payoff as in experiment $\boldsymbol{\tau^s}$. (1) If $\boldsymbol{\tau'}$ induces t' and t'' to pool, the sender's expected utility from the two types equals

[15] We elaborate on the situation in which $t' > t_1$. If $t' = t_1$, our analysis also applies.

$\hat{U}(t_2)$. Compared with $c(t_2, m_2(\boldsymbol{\tau^s}))$, the reporting cost of type t', $c(t', \overline{m})$, is higher, while that of type t'', $c(t'', \overline{m})$, may be lower. As long as the weighted average of the cost from the two types, $\mathcal{C}(t_2, \overline{m})\big|_{[t_1,1]}$, is lower than $c(t_2, m_2(\boldsymbol{\tau^s}))$, we have $E(\boldsymbol{\tau'}) > E(\boldsymbol{\tau^s})$. As depicted in Figure 3, when the cost is concave in types, $c(t, \overline{m})$ decreases significantly with t such that $c(t'', \overline{m})$ is much lower than $c(t_2, m_2(\boldsymbol{\tau^s}))$. Then, after inducing pooling, the sender can save the reporting cost by introducing the higher type t'' though she incurs a higher cost by introducing the lower type t'. (2) If $\boldsymbol{\tau'}$ induces t' to separate, t' obtains a higher payoff from separation than from pooling. Then, $\boldsymbol{\tau'}$ induces a separating equilibrium with $E(\boldsymbol{\tau'}) > E(\boldsymbol{\tau^s})$, which contradicts the assumption that $\boldsymbol{\tau^s}$ is the best separating experiment.

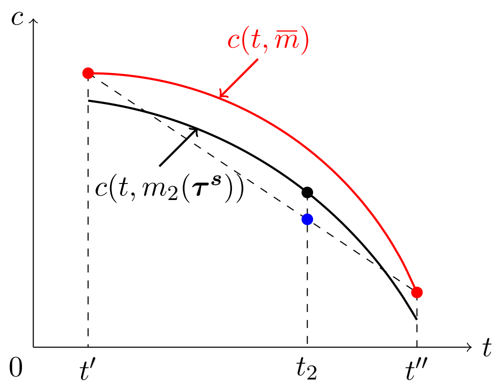

Figure 3: Beneficial Pooling

In the sufficient conditions provided by Corollary 1 and Proposition 2, the property of the cost c in $t \in [0, 1]$ plays an important role, which is ignored in standard signaling games. The reason is that the sender's private information is often exogenously assumed in signaling games.[16] By contrast, in this model, though the message space and cost function are exogenously given, the sender's experiment choice or the distribution of sender types is endogenized.

Note that the convexity in types is sufficient for endogenized full information transmission, while the concavity in types is necessary but not sufficient for partial information transmission. The condition that $c(t, \overline{m})$ is concave in t only ensures $c(t_2, \overline{m}) \geq \mathcal{C}(t_2, \overline{m})\big|_{[t_1,1]}$, which does not suffice when $m_2(\boldsymbol{\tau^s}) < \overline{m}$. In Corollary 2 (in Section 6.1), we identify a condition under which there exists a partial-pooling experiment better than $\boldsymbol{\tau^s}$ when the cost is concave in types.

5 Strategic Reporting vs. Commitment to Truth-telling

So far, the model we analyzed has maintained the assumption that the minimized reporting cost for any given result is zero. In this section, we relax this assumption. We allow the possibility that the lowest reporting cost is strictly positive for some result and only assume $c(t, m_c(t)) \geq 0$, for $t \in [0, 1]$.

[16]One exception is In and Wright (2018) that endogenizes the sender's private type in the signaling game.

To elaborate on its economic implications, we decompose the reporting cost function as follows:

$$c(t,m) = \underbrace{c(t,m_c(t))}_{\text{communication cost}} + \underbrace{\left(c(t,m) - c(t,m_c(t))\right)}_{\text{manipulation cost}}.$$

The first component $c(t,m_c(t))$ depends on the result and a positive value reflects the situation where a non-negligible *communication cost* is incurred (e.g., Oniki, 1974; Hutter, 1986) regardless of whether communication involves only truth-telling or not.[17] The communication cost can have rich function forms. It can be increasing in t if a more positive result needs to be delivered with more careful argument and detailed evidence, which reflects underlying skepticism the receiver holds. It can be concave in t, e.g., communication cost $t(1-t)$ reflects that a more extreme result can be delivered at a lower cost. The second component is considered *manipulation cost*. An example of the cost function is $c(t,m) = c_0(t) + (m-t)^2$, where communication cost $c_0(t) \geq 0$ is a function of t and $(m-t)^2$ depends on the degree of manipulation.

Recall that sending the cost-minimizing message $m_c(t)$ after observing t is considered truth-telling. We further assume $m_c(t) > 0$ for any $t > 0$. Under this assumption, $m_c(t)$ strictly increases with t (see Lemma A.2), in which case there is one-to-one mapping between $m_c(t)$ and t. Thus, if the sender commits to truth-telling, her type will always be fully revealed to the receiver. Based on the assumption that $c(t,m_c(t)) \geq 0$, truth-telling can be costly capturing the situation where telling the truth requires some preparations of sound arguments. When truth-telling is always costless, for the sender ex ante, strategic reporting is weakly worse than commitment to truth-telling in the reporting stage. Based on this, we wonder given the same cost structure, when truth-telling (that always incurs the minimum cost) can be costly, whether it is possible for the sender to ex ante strictly prefer strategic reporting over the corresponding commitment to truth-telling.

We begin the following analysis. Define $g(t) := \hat{U}(t) - c(t,m_c(t))$, for any $t \in [0,1]$. Let $\mathcal{G}(t)$ denote the concave closure of $g(t)$ on $[0,1]$. With commitment to truth-telling, the type t sender obtains payoff $g(t)$ and the optimal experiment induces the expected payoff $\mathcal{G}(\mu)$. In strategic reporting, after choosing any separating experiment, any type t obtains a payoff weakly lower than $g(t)$, and therefore, the sender's expected payoff from any separating experiment cannot exceed $\mathcal{G}(\mu)$.

What is the necessary condition for strategic reporting to be better than commitment to truth-telling? We mainly consider that from the cost function aspect. First, in strategic reporting, based on Corollary 1, if $c(t,\overline{m})$ is convex in types, a separating experiment is optimal, so the sender can never achieve an expected payoff higher than $\mathcal{G}(\mu)$. Second, if truth-telling is costless for all types, commitment is always weakly better than strategic reporting. We summarize as follows.

Observation: *Given the same cost structure, the sender ex ante strictly prefers strategic reporting over commitment to truth-telling, which incurs the minimum cost for any realized result, only if neither of the*

[17] Communication costs can be regarded as a kind of transaction or institutional cost á la Coase (1960) and Demsetz (1964). For example, when an experiment is designed by a pharmaceutical company to persuade investors and stockholders, writing a report summarizing the scientific findings from the experiment to ensure that the audience can understand is costly regardless of whether the report contains truthful information only.

following conditions is satisfied: (1) $c(t,\overline{m})$ is convex in $t \in [0,1]$; (2) $c(t,m_c(t)) = 0$ for any $t \in [0,1]$.

In strategic reporting, any separating experiment cannot induce an expected payoff higher than $\mathcal{G}(\mu)$. However, the sender's payoff from a pooling experiment cannot be directly compared with $\mathcal{G}(\mu)$. We find if a pooling experiment is optimal, it can induce an expected payoff higher than $\mathcal{G}(\mu)$. Proposition 3 provides two sufficient conditions under which there exists a pooling experiment in our strategic reporting that can achieve strictly higher expected payoff than $\mathcal{G}(\mu)$.

Proposition 3. *Given the same cost structure, for the sender ex ante, strategic reporting is strictly better than commitment to truth-telling, which incurs the minimum cost for any realized result, if either of the following conditions holds.*

1. $\hat{U}(\mu) - \mathcal{C}(\mu,\overline{m})\big|_{[0,1]} > \mathcal{G}(\mu)$.
2. *There exists experiment $\tilde{\boldsymbol{\tau}} = (\tilde{t}_1,\tilde{t}_2;\tilde{\tau}_1,\tilde{\tau}_2)$, where $\tilde{\tau}_1 g(\tilde{t}_1) + \tilde{\tau}_2 g(\tilde{t}_2) = \mathcal{G}(\mu)$, that satisfies $\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1)) \geq \hat{U}(\tilde{t}_2) - c(\tilde{t}_1,\overline{m})$ and $c(\tilde{t}_2, m_c(\tilde{t}_2)) > \mathcal{C}(\tilde{t}_2,\overline{m})\big|_{[\tilde{t}_1,1]}$.*

Note that one of the main conclusions from the Bayesian persuasion literature (e.g., Lipnowski et al., 2022; Nguyen and Tan, 2021; Guo and Shmaya, 2021) is that fully committing to truthful reporting is optimal for the sender ex ante when truthful reporting is costless. Our finding based on costly truthful reporting complements that conclusion. Next, we provide an example to show how strategic reporting benefits the sender, compared with truth-telling.

Example 1. *For $a \geq 0$, $U(a) = \sqrt{a+1}$, $V(t,a) = -(a-t)^2$. $M = [0,1]$, $\mu = \frac{1}{2}$, and $c(t,m) = \frac{(t-m)^2}{40} + t(1-t)$. Then $\alpha^R(t) = t$ and the cost function is strictly* concave *in $t \in [0,1]$ for any $m \in M$. We have $g(t) = \hat{U}(t) - c(t,m_c(t)) = \sqrt{t+1} - t(1-t)$, strictly convex in t. Then, if committed to truth-telling, the sender chooses the fully informative experiment $\bar{\boldsymbol{\tau}} = (t_L = 0, t_H = 1; \frac{1}{2}, \frac{1}{2})$ and obtains*

$$\mathcal{G}(\mu) = \frac{1}{2}g(t_L) + \frac{1}{2}g(t_H) = \frac{1+\sqrt{2}}{2} \approx 1.207.$$

As in Figure 4, the distance between the blue dot and the red dot is $\mathcal{G}(\mu)$.

In strategic reporting, let us consider the equilibrium induced by $\bar{\boldsymbol{\tau}}$. Since

$$\hat{U}(t_L) - c(t_L, m_c(t_L)) < \hat{U}(\mu) - c(t_L,\overline{m}),$$

t_L and t_H pool at $\overline{m}$. Then, the sender's expected payoff is

$$E(\bar{\boldsymbol{\tau}}) = \frac{1}{2}[\hat{U}(\mu) - c(t_L,\overline{m})] + \frac{1}{2}[\hat{U}(\mu) - c(t_H,\overline{m})] \approx 1.212 > \mathcal{G}(\mu).$$

Thus, the sender strictly prefers strategic reporting.

Figure 4 shows in strategic reporting, how experiment $\bar{\boldsymbol{\tau}}$ achieves an expected payoff higher than $\mathcal{G}(\mu)$. The expected utility and expected reporting cost from $\bar{\boldsymbol{\tau}}$ are depicted by the two black dots,

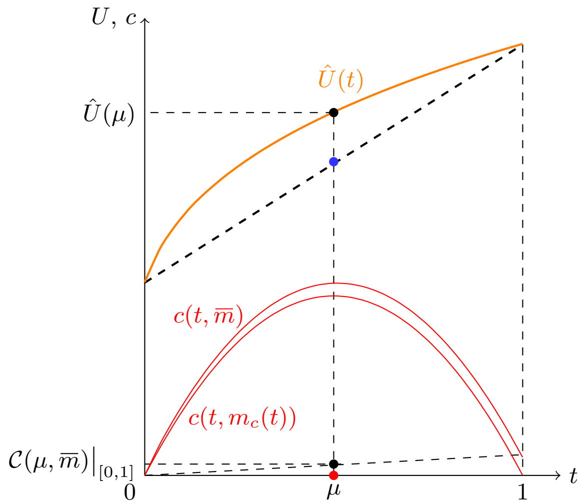

Figure 4: Benefits of Strategic Reporting

respectively. Compared with $\mathcal{G}(\mu)$, $\bar{\boldsymbol{\tau}}$ induces a total-pooling equilibrium such that the sender obtains higher expected utility and a higher expected reporting cost. As long as the increase of the expected utility exceeds the increase of the expected reporting cost, $E(\bar{\boldsymbol{\tau}}) > \mathcal{G}(\mu)$. In strategic reporting, by choosing $\bar{\boldsymbol{\tau}}$, the sender can obtain the highest expected utility at μ, while she pays zero communication cost and a relatively low manipulation cost. In contrast, with commitment to truth-telling, if the sender chooses the uninformative experiment to induce utility $\hat{U}(\mu)$, she needs to pay the communication cost $c(\mu, m_c(\mu))$; if she chooses $\bar{\boldsymbol{\tau}}$ to get rid of communication cost, her expected utility (depicted by the blue dot) will be lower than $\hat{U}(\mu)$. Thus, strategic reporting can be better than the corresponding commitment to truth-telling for the sender ex ante due to the above-mentioned cost-saving benefit.

6 Informativeness of the Optimal Experiment

In this section, we maintain the assumption that the minimized reporting cost for any realized result is zero and we apply Proposition 1 to explicitly characterize the optimal experiment, which is dependent on both the utility and cost structure. We divide all experiments into two categories: the *uninformative experiment*, denoted by $\boldsymbol{\tau_0} = (\mu; 1)$, and *informative experiments*, which are all experiments except $\boldsymbol{\tau_0}$. Specially, the *fully informative experiment*, denoted by $\bar{\boldsymbol{\tau}} = (t_L = 0, t_H = 1; 1-\mu, \mu)$, enables the sender to accurately know the state.

We study the following questions. When does the sender select the fully informative experiment? When does she want to select an informative experiment, that is, when does she benefit from persuasion? How does the sender's experiment choice vary with the intensity of her cost?

6.1 Choosing the Fully Informative Experiment

Compared with the sender's utility U, if the reporting cost is either too high or too low, it plays no role in the construction of the optimal experiment. i) When strategic reporting is too costly for the sender so that truth-telling is always the best choice after observing any realized result, the optimal experiment is determined by the shape of the utility, the same as the situation in which the sender commits to truthful reporting. ii) When the cost is too low so that the sender cannot credibly transmit information from any informative experiment, the uninformative experiment is always optimal. Besides the two extreme cases, our interest is concentrated on how the characteristics of the cost affect the optimal experiment.

To ascertain the determinant of the sender's expected payoff, let us take a separating experiment $\boldsymbol{\tau} = (t_1, t_2; \tau_1, \tau_2)$ as an example. Suppose $m_2(\boldsymbol{\tau}) > m_c(t_2)$, that implies the binding IC condition

$$\hat{U}(t_1) = \hat{U}(t_2) - c(t_1, m_2(\boldsymbol{\tau})).$$

Then, the sender's expected payoff

$$\begin{aligned} E(\boldsymbol{\tau}) &= \frac{t_2 - \mu}{t_2 - t_1}\hat{U}(t_1) + \frac{\mu - t_1}{t_2 - t_1}[\hat{U}(t_2) - c(t_2, m_2(\boldsymbol{\tau}))] \\ &= \hat{U}(t_1) + \frac{\hat{U}(t_2) - c(t_2, m_2(\boldsymbol{\tau})) - \hat{U}(t_1)}{t_2 - t_1}(\mu - t_1) \end{aligned}$$

depends on two factors:

1. $\hat{U}(t_1)$, the lower bound of $E(\boldsymbol{\tau})$, the lowest payoff either type is able to obtain. Type t_1 obtains $\hat{U}(t_1)$ if it separates. Type t_2 can take less cost to send the separating message than type t_1, and thus, by the IC condition for type t_1 to separate, t_2 achieves a payoff higher than $\hat{U}(t_1)$.

2. The effect of t_2, which is reflected in the slope

$$\frac{\hat{U}(t_2) - c(t_2, m_2(\boldsymbol{\tau})) - \hat{U}(t_1)}{t_2 - t_1} = \frac{c(t_1, m_2(\boldsymbol{\tau})) - c(t_2, m_2(\boldsymbol{\tau}))}{t_2 - t_1}.$$ [18]

   The numerator is the higher payoff that t_2 can obtain, compared to t_1, and it equals the amount of the cost t_2 can save for sending the separating message, compared with t_1.

Given any low type t_1, $E(\boldsymbol{\tau})$ increases with the slope. The slope represents the ability of the high type to distinguish itself from the low type and depends on the property of the cost structure. When the cost function is concave in types, the sender's cost decreases faster with types for any given message. Then, the slope increases with t_2 if $m_2(\boldsymbol{\tau})$ remains constant. Since higher t_2 lets type t_1 benefit more from mimicking t_2, $m_2(\boldsymbol{\tau})$ increases with t_2, which further raises the slope because $\frac{\partial^2 c}{\partial t \partial m} < 0$.

Therefore, for any given low type t_1, the sender wants the high type t_2 to be as high as possible if her ability to distinguish herself increases vastly with t_2. This conclusion also applies when $\boldsymbol{\tau}$ is a pooling experiment. Formally, if the sender chooses from the experiments with two types, we summarize the following condition under which the sender wants the high type to distance from the low type (HD).

HD Condition:

[18] It can be considered the slope of the line connecting the two payoff points $(t_1, \hat{U}(t_1))$ and $(t_2, \hat{U}(t_2) - c(t_2, m_2(\boldsymbol{\tau})))$. The equality is derived from the IC condition for separation.

1. Given any m, $c(t,m)$ is strictly concave in t; and

2. $\hat{U}$ is strictly convex, or
   $\forall\ t_1 < \mu < t_2$, $\hat{U}(t_1) \leq \hat{U}(t_2) - c(t_1, m_c(t_2))$.

Lemma 3. *If the **HD** condition is satisfied, the sender's expected payoff from any two-result experiment* $\boldsymbol{\tau} = (t_1, t_2; \tau_1, \tau_2)$ *increases with* t_2 *for any given* t_1.

The second part of the HD condition either restricts the shape of the utility or limits the scope and sensitivity of the cost. When the cost is relatively low such that any separation is costly, i.e., $\hat{U}(t_1) \leq \hat{U}(t_2) - c(t_1, m_c(t_2))$, the shape of the cost function is the determinant of the optimal choice of t_2. The binding IC condition and the concavity of c imply that the sender's expected payoff increases with t_2, for any given t_1. When the cost is high such that some separation is costless, both the utility and cost are the determinants. If t_2 separates with no cost, convex $\hat{U}$ guarantees the sender prefers a higher t_2; otherwise, concavity of c can guarantee that.

The HD condition guarantees that for all experiments containing two types, $(t_1, t_2 = t_H; \tau_1, \tau_2)$ is optimal. After further imposing some restrictions on the sender's utility and cost structure, we achieve the full characterization of $\boldsymbol{\tau}^*$ as follows.

Proposition 4. *Assume the **HD** condition is satisfied. The sender must choose the fully informative experiment if* $\hat{U}(t) + c(t,m)$ *is decreasing in* t*, for* $t \leq \mu$, $m \geq m_c(\mu)$.

The HD condition requires c to be strictly concave in t, which implies the optimal experiment needs at most two types regardless of whether it leads to a separating equilibrium or pooling equilibrium. This is a direct application of Proposition 1. From Lemma 3, the HD condition indicates the sender chooses either the uninformative experiment $\boldsymbol{\tau_0}$ or an experiment containing t_H, denoted by $(t_1, t_H; \tau_1, \tau_2)$.

We then solve the optimal experiment under the condition in Proposition 4 that $\hat{U}' \leq -\frac{\partial c}{\partial t}$, for $t \leq \mu$. Similarly to the demonstration for Lemma 3, let us take a separating experiment $\boldsymbol{\tau} = (t_1, t_H; \tau_1, \tau_2)$ as an example. Suppose t_H sends $m_2(\boldsymbol{\tau}) > m_c(t_H)$. The sender's expected payoff

$$E(\boldsymbol{\tau}) = \hat{U}(t_H) - c(t_H, m_2(\boldsymbol{\tau})) - \frac{\hat{U}(t_H) - c(t_H, m_2(\boldsymbol{\tau})) - \hat{U}(t_1)}{t_H - t_1}(t_H - \mu)$$

depends on 1) the payoff of type t_H, $\hat{U}(t_H) - c(t_H, m_2(\boldsymbol{\tau}))$, considered the possible upper bound, and 2) the slope

$$\frac{\hat{U}(t_H) - c(t_H, m_2(\boldsymbol{\tau})) - \hat{U}(t_1)}{t_H - t_1} = \frac{c(t_1, m_2(\boldsymbol{\tau})) - c(t_H, m_2(\boldsymbol{\tau}))}{t_H - t_1}.$$

We then consider how the choice of the low type t_1 affects $E(\boldsymbol{\tau})$. Given any upper bound, a lower slope means higher expected payoff from the experiment, so the sender wants t_1 to induce the lowest slope. If the separating message remains unchanged, the payoff of type t_H is fixed. Since c is concave in t, the slope is minimized at $t_1 = t_L$ for any given message m_2. Further, the condition $\hat{U}' \leq -\frac{\partial c}{\partial t}$ guarantees that the payoff of type t_H decreases with t_1 by making the separating message increase with

t_1, i.e.,

$$\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1} = -\frac{\hat{U}'(t_1) + \frac{\partial c(t_1, m_2(\boldsymbol{\tau}))}{\partial t_1}}{\frac{\partial c(t_1, m_2(\boldsymbol{\tau}))}{\partial m}} > 0,$$

which is derived from the IC condition by the implicit function theorem. In conclusion, $t_1 = t_L$ makes type t_H obtain the highest payoff, conditional on which, the slope is minimized, and thus generates the highest expected payoff. Since the uninformative experiment can be considered as $(t_1 = \mu, t_H; 1, 0)$, then it is strictly worse than the fully informative experiment.

Intuitively, this condition $\hat{U}' \leq -\frac{\partial c}{\partial t}$ ensures that the benefit from lowering t_1 is greater than the loss. The loss is from the reduction of $\hat{U}(t_1)$, while the benefit is the raise of the payoff of type t_H as well as the higher probability allocated to t_H. The analysis can also be applied to the case where $\boldsymbol{\tau}$ is a pooling experiment. In summary, the sender chooses the fully informative experiment regardless of whether the acquired information can be fully transmitted or not.

Applying Lemma 3, under the HD condition, we also identify a situation in which it is optimal for the sender to design a partial-pooling experiment.

Corollary 2. *Assume the* ***HD*** *condition is satisfied. The optimal experiment is partial-pooling if* $\boldsymbol{\tau^s} = (t_1', t_2'; \tau_1', \tau_2')$ *is informative and* $t_2' < t_H$.

According to Lemma 3, $m_2(\boldsymbol{\tau^s}) = \overline{m}$, and then, the experiment $(t_1', t_H; \tau_1, \tau_2)$ induces a partial-pooling equilibrium with an expected payoff higher than $\boldsymbol{\tau^s}$. Hence, the optimal experiment is partial-pooling. This corollary can also be derived from Proposition 2.

6.2 Value of Persuasion: Choosing an Informative Experiment

Does the sender have an incentive to participate in the persuasion process when she cannot commit to full transmission of the obtained information? As she can always abandon persuasion through choosing the uninformative experiment without incurring any cost, we need to explore when the sender can strictly benefit from persuasion followed by strategic reporting.

We define the *value* of the persuasion process as the difference between the sender's expected payoff from her optimal experiment choice and that from choosing the uninformative experiment $\boldsymbol{\tau_0}$. The sender *benefits from persuasion* if the value of persuasion is strictly positive. We assume the sender always chooses $\boldsymbol{\tau_0}$ if the value of persuasion is zero. Then, she benefits from the persuasion process if and only if she selects an informative experiment in the information design stage. Obviously, if $\hat{U}$ is concave, the sender does not benefit from the persuasion process for any prior belief. We then have the following finding.

Proposition 5. *The sender benefits from the persuasion process if and only if she transmits information in the reporting stage. For any prior* μ *that satisfies* $\alpha^R(\mu) = \underline{a}$*, the sender benefits from the persuasion process.*

The first statement in Proposition 5 comes from the fact that a total-pooling experiment is always worse than $\boldsymbol{\tau_0}$, because a total-pooling experiment does not transmit any information to the receiver but incurs positive reporting cost while $\boldsymbol{\tau_0}$ transmits no information with no cost. Hence, the sender never chooses a total-pooling experiment.[19] This indicates the sender benefits from persuasion if and only if she transmits information in the signaling subgame.

The second statement provides one condition that guarantees positive value of persuasion. The condition is that the receiver takes the lowest action after obtaining no information, under which, for all $t \leq \mu$, $\hat{U}(t) = U(\underline{a})$ because the receiver's optimal action α^R weakly increases in types. Compared with choosing $\boldsymbol{\tau_0}$, the sender can be better off through transmitting information to the receiver. For example, choosing the fully informative experiment $\bar{\boldsymbol{\tau}} = (t_L, t_H; 1-\mu, \mu)$ can transmit information and achieve $E(\bar{\boldsymbol{\tau}}) > E(\boldsymbol{\tau_0})$. In the equilibrium induced by $\bar{\boldsymbol{\tau}}$, type t_L separates or partially pools, because separation is more beneficial than total-pooling, i.e., $\hat{U}(t_L) > \hat{U}(\mu) - c(t_L, \overline{m})$. Then, t_L obtains a payoff $\hat{U}(t_L) = U(\underline{a})$. If type t_H sends $m_c(t_H)$, it separates and its payoff is $\hat{U}(t_H) > U(\underline{a})$; otherwise, t_L must be indifferent between sending $m_c(t_L)$ and t_H's equilibrium message, and thus, the single-crossing condition guarantees t_H obtains a payoff strictly higher than $\hat{U}(t_L)$. Therefore, $E(\bar{\boldsymbol{\tau}}) > U(\underline{a})$.

Further, we derive one category of the sufficient condition for the sender selecting an informative experiment. Proposition 4 provides the conditions under which it is always more advantageous for the sender to distance the two types. Using this, our exercise is making the conditions held at the prior belief μ, that is, we ensure that there exists $\tilde{t}_2$ such that among all experiments $(t_1, \tilde{t}_2; \tau_1, \tau_2)$, the sender's expected payoff decreases with t_1 around $t_1 = \mu$.

Corollary 3. *The sender benefits from the persuasion process if the following three conditions are satisfied.*

1. *Given any m, $c(t, m)$ is concave in t.*
2. *There exists $t_2 > \mu$ such that $\hat{U}(\mu) \leq \hat{U}(t_2) - c(\mu, m_c(t_2))$.*
3. *$\frac{\partial}{\partial t}[\hat{U}(t) + c(t, m)] < 0$ at $t = \mu$, for $m \geq m_c(\mu)$.*

By analogy with Proposition 4, we also derive the "symmetric" conditions for $\boldsymbol{\tau_0}$ being optimal. Obviously, when $\hat{U}$ is concave in t, $\boldsymbol{\tau_0}$ is optimal. The following conditions sustain the property that for any experiment with types t_1 and t_2, given any t_2, the expected payoff increases with t_1, and thus, any informative experiment is worse than $\boldsymbol{\tau_0}$.

Corollary 4. *The sender must choose the uninformative experiment if the following three conditions are satisfied.*

1. *Given any m, $c(t, m)$ is strictly convex in t.*
2. *$\forall\ t_1 < \mu < t_2$, $\hat{U}(t_1) \leq \hat{U}(t_2) - c(t_1, m_c(t_2))$.*

[19] This conclusion relies on the costless truth-telling assumption. If $c(t, m_c(t)) \geq 0$ (as the costless truth-telling assumption relaxed in Section 5), a total-pooling experiment can be optimal.

3. *$\hat{U}(t) + c(t,m)$ is increasing in t, for $t \le \mu$, $m \ge m_c(\mu)$.*

6.3 Optimal Experiment Varying with Cost Intensity: the Case of Linear Utility

This subsection studies a specification of our framework that permits explicit characterization of an optimal experiment and comparative statics with respect to the intensity of the reporting cost.

We consider the following specification. The sender's utility $U = a$, linear in the receiver's action.[20] The receiver's utility is $V^H = -(a-H)^2$ and $V^L = -(a-L)^2$, for the respective two states. The two players' prior belief is $P(H) = \mu$. The highest message $\overline{m} = 1$, and then $m \in [0,1]$. The sender's reporting cost is $K \cdot c(t,m)$, where $K > 0$ is a scale parameter. We maintain the assumption that the minimum cost $Kc(t, m_c(t)) = 0$ for any result t. The receiver's action space is $A = [\underline{a}, 1]$, where $0 < \underline{a} < \mu$. The receiver's optimal action given any t is depicted in Figure 5:

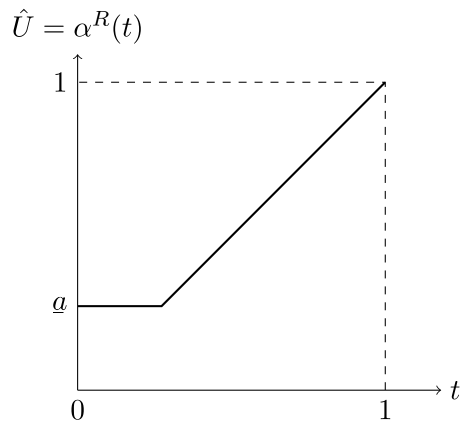

Figure 5: Bounded action space

Concave Cost Structure First, we illustrate the situation in which the cost structure $Kc(t,m)$ is concave in types. Suppose $\boldsymbol{\tau^*}$ is informative; then, $\boldsymbol{\tau^*}$ consists of two types denoted by t_1^* and t_2^*. Since $\hat{U}$ is convex and c is concave in t, by Lemma 3, $t_2^* = t_H$ is optimal for any given t_1^*. Then, let us prove $t_1^* = t_L$. If $\underline{a} \le t_1^* < \mu$, $E(\boldsymbol{\tau^*}) \le E(\boldsymbol{\tau_0})$. If $t_L \le t_1^* < \underline{a}$ and $\alpha^R(t_1^*) = \underline{a}$, as t_1^* decreases, t_2^* can utilize less costly message for separation or obtain a higher pooling action, that increases both types' equilibrium payoffs. Moreover, t_2^* is allocated with the highest probability when $t_1^* = t_L$.

Corollary 5. *In this linear specification, if given any m, the cost function $Kc(t,m)$ is strictly concave in types, the optimal experiment is either the fully informative experiment or the uninformative experiment. The fully informative experiment induces a separating equilibrium if $K \ge \frac{1-\underline{a}}{c(t_L,\overline{m})}$, and induces a pooling equilibrium otherwise.*

Based on this corollary, we only need to compare the fully informative experiment $\bar{\boldsymbol{\tau}} = (t_L, t_H; 1-\mu, \mu)$ and $\boldsymbol{\tau_0}$ to determine the optimal experiment. $E(\bar{\boldsymbol{\tau}})$ changes continuously with the scale of cost

[20] The sender's utility can be any linear and increasing function of a, which does not affect the conclusions.

K, while $E(\boldsymbol{\tau_0}) = \hat{U}(\mu)$ is fixed. Therefore, as K increases, the sender's optimal experiment choice will change from $\boldsymbol{\tau_0}$ to $\bar{\boldsymbol{\tau}}$ at a certain threshold value of K. Consequently, the receiver's expected payoff is discontinuous in K such that as the scale of reporting cost expands, the receiver's payoff would jump at a certain threshold because the sender will suddenly choose the fully informative experiment and transmit information.

Convex Cost Structure Then, we study the situation in which the cost is strictly convex in types.

Corollary 6. *In this linear specification, if given any m, the cost function $Kc(t, m)$ is strictly convex in types, the optimal experiment is either the uninformative experiment or a separating experiment containing t_L, denoted by $(t_L, t_2; \tau_1, \tau_2)$. Moreover:*

1. *When $K \geq \frac{1-\underline{a}}{c(t_L,\overline{m})}$, the sender chooses the fully informative experiment.*
2. *When $K \leq \frac{\mu-\underline{a}}{c(t_L,\overline{m})}$, the sender chooses the uninformative experiment.*

By Proposition 1 and Corollary 1, the sender chooses a separating experiment with at most two types. Intuitively, when the cost is high, the optimal experiment is $\bar{\boldsymbol{\tau}}$, the same as that in Bayesian persuasion. When the cost is low, any experiment $(t_L, t_2; \tau_1, \tau_2)$ induces a pooling equilibrium, so $\boldsymbol{\tau_0}$ is optimal. In Corollary 7, we explicitly characterize the optimal experiment when $c = (m - t)^2$.

Corollary 7. *In this linear specification, when the cost function is $K(m - t)^2$, the sender's expected payoff from the optimal experiment weakly increases with K.*

1. *When $K \geq 1 - \underline{a}$, the sender chooses the fully informative experiment.*
2. *When $1 - \frac{1}{2}\underline{a} - \sqrt{\frac{1}{4}\underline{a}^2 - \underline{a} + \frac{\underline{a}}{\mu}} < K \leq 1 - \underline{a}$, the sender chooses experiment $(t_L, \tilde{t}_2; \tau_1, \tau_2)$, where $\tilde{t}_2 = \underline{a} + K$.*
3. *When $K \leq 1 - \frac{1}{2}\underline{a} - \sqrt{\frac{1}{4}\underline{a}^2 - \underline{a} + \frac{\underline{a}}{\mu}}$, the sender chooses the uninformative experiment.*

By Corollary 6, the optimal experiment is either $\boldsymbol{\tau_0}$ or $(t_L, t_2; \tau_1, \tau_2)$. For experiment $(t_L, t_2; \tau_1, \tau_2)$, higher t_2 has stronger ability to distinguish itself and improve the sender's expected payoff from this experiment. Therefore, t_2 is the highest type that can separate from t_L. Hence, $t_2 = t_H$ or satisfies

$$\hat{U}(t_L) = \hat{U}(t_2) - c(t_L, \overline{m}),$$

that is, $t_2 = \min\{t_H, \underline{a} + K\}$. That t_2 weakly increases with K means when the sender has to spend higher cost manipulating, she would select a more informative experiment. The reason is that she is able to credibly transmit a higher result through reporting if she has stronger commitment power.

We further compare these two potentially optimal experiments to ultimately obtain the optimal one. The expected payoff from $(t_L, t_2 = \underline{a} + K; \tau_1, \tau_2)$, $K \leq 1 - \underline{a}$, is increasing in K because given the lower bound $\hat{U}(t_L) = \underline{a}$, higher reporting cost lets t_2 distinguish itself more easily. When K is so low that t_2 needs to waste much cost to get rid of the low type's manipulation, the sender chooses $\boldsymbol{\tau_0}$ and both players gain no information. Only when K is higher than a threshold value, the sender chooses the informative experiment $(t_L, \underline{a} + K; \tau_1, \tau_2)$. As the reporting cost becomes high enough, the sender always chooses the fully informative experiment.

Note that the increase in fabrication costs may not change the optimal experiment choice, as the sender will always choose the uninformative experiment until the fabrication cost reaches a critical value. Moreover, the sender's choice does not change continuously with the scale of cost, as her choice will jump to an informative experiment at a critical value and the receiver will suddenly receive useful information.

7 Conclusion

We study the optimal information design when a sender has commitment to conducting an experiment but cannot commit to reporting the obtained result truthfully. In our model, to persuade a receiver, the sender can report a message to reveal information and has to bear a cost that depends on both the realized result and the message reported. The cost has strictly decreasing differences, that implies the sender's marginal cost with respect to messages is higher if she obtains a worse result. The cost function we set has many economic implications that can represent the sender's manipulation cost, signaling cost, and so on. This model bridges Bayesian persuasion and costly lying (or signaling).

In this framework, our methodology (Proposition 1) to characterize the optimal experiment allows us to greatly simplify the analysis. We further find whether the sender will choose an experiment whose results can be fully revealed is determined by the properties of the cost structure. Built on this, we show that it is possible for the sender ex ante to prefer strategic reporting over commitment to truthful reporting if truth-telling, although it incurs the minimum cost, is costly.

In Section 6.3, we conduct comparative statics analysis with respect to cost intensity when the sender has linear utility. The cost intensity can influence the sender's strategy through two channels. First, increasing the cost intensity can make the sender design a more informative experiment. Second, higher cost intensity can enable the sender to transmit more information without changing the choice of experiment. Our findings rely on the assumption that the sender's utility is linear. If we consider a more general utility function form, we may obtain the opposite conclusion: higher cost intensity leads to the sender choosing a less informative experiment and cannot facilitate information transmission. In Lipnowski et al. (2022), a crucial conclusion is that lower credibility in the result reporting stage may make the receiver better off. This intuition may also be derived in a strategic and costly reporting scenario, which we leave for future research.

References

Argenziano, Rossella, Sergei Severinov, and Francesco Squintani (2016), "Strategic information acquisition and transmission." *American Economic Journal: Microeconomics*, 8, 119–155.

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

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

Best, James and Daniel Quigley (2024), "Persuasion for the long run." *Journal of Political Economy*, 132, 1740–1791.

Che, Yeon-Koo and Navin Kartik (2009), "Opinions as incentives." *Journal of Political Economy*, 117, 815–860.

Cho, In-Koo and David M Kreps (1987), "Signaling games and stable equilibria." *The Quarterly Journal of Economics*, 102, 179–221.

Cho, In-Koo and Joel Sobel (1990), "Strategic stability and uniqueness in signaling games." *Journal of Economic Theory*, 50, 381–413.

Coase, Ronald H (1960), "The problem of social cost." In *Classic papers in natural resource economics*, 87–137, Springer.

Demsetz, Harold (1964), "The exchange and enforcement of property rights." *The Journal of Law and Economics*, 7, 11–26.

Ederer, Florian and Weicheng Min (2024), "Bayesian persuasion with lie detection." Technical report, National Bureau of Economic Research.

Guo, Yingni and Eran Shmaya (2021), "Costly miscalibration." *Theoretical Economics*, 16, 477–506.

Hedlund, Jonas (2017), "Bayesian persuasion by a privately informed sender." *Journal of Economic Theory*, 167, 229–268.

Heese, Carl and Shuo Liu (2023), "Persuasion versus presentation." *Available at SSRN 4680477.*

Hutter, Michael (1986), "Transaction cost and communication: A theory of institutional change, applied to the case of patent law." In *Law and economics and the economics of legal regulation*, 113–129, Springer.

In, Younghwan and Julian Wright (2018), "Signaling private choices." *The Review of Economic Studies*, 85, 558–580.

Ivanov, Maxim (2010), "Informational control and organizational design." *Journal of Economic Theory*, 145, 721–751.

Kamenica, Emir (2019), "Bayesian persuasion and information design." *Annual Review of Economics*, 11, 249–272.

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

Kartik, Navin (2009), "Strategic communication with lying costs." *The Review of Economic Studies*, 76, 1359–1395.

Kartik, Navin, Marco Ottaviani, and Francesco Squintani (2007), "Credulity, lies, and costly talk." *Journal of Economic theory*, 134, 93–116.

Kreutzkamp, Sophie (2022), "Endogenous information acquisition in cheap-talk games."

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

Lyu, Qianjun and Wing Suen (2022), "Information design in cheap talk." *arXiv preprint arXiv:2207.04929*.

Mathevet, Laurent, David Pearce, and Ennio Stacchetti (2024), "Reputation and information design."

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

Nguyen, Anh and Teck Yong Tan (2021), "Bayesian persuasion with costly messages." *Journal of Economic Theory*, 193, 105212.

Oniki, Hajime (1974), "The cost of communication in economic organization." *The Quarterly Journal of Economics*, 88, 529–550.

Pei, Harry (2023), "Repeated communication with private lying costs." *Journal of Economic Theory*, 210, 105668.

Pei, Harry Di (2015), "Communication with endogenous information acquisition." *Journal of Economic Theory*, 160, 132–149.

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

Spence, Michael (1973), "Job market signaling." *The Quarterly Journal of Economics*, 87, 355–374.

Appendix A. Proofs

First, we provide three preliminary lemmas useful for follow-up proofs. Then, we provide proofs of all the conclusions.

Lemma A.1. *For any $t \in [0,1]$, there exists a unique $\alpha^R(t_i) \equiv \arg\max_{a\in[\underline{a},+\infty)} V(t_i,a)$. Moreover,*

- *if $\alpha^R(t_L) > \underline{a}$, α^R is continuously differentiable and strictly increasing in $t \in [0,1]$;*
- *if $\alpha^R(t_L) = \underline{a}$, there exists a unique $\tilde{t}$ such that for $t \le \tilde{t}$, $\alpha^R = \underline{a}$, and for $\tilde{t} \le t \le t_H$, α^R is continuously differentiable and strictly increasing in t.*

Proof. From $V(t,a) \equiv tV^H + (1-t)V^L$, we have

$$\begin{aligned}\frac{\partial V}{\partial a} &= t\frac{\partial V^H}{\partial a} + (1-t)\frac{\partial V^L}{\partial a},\\ \frac{\partial^2 V}{\partial a^2} &= t\frac{\partial^2 V^H}{\partial a^2} + (1-t)\frac{\partial^2 V^L}{\partial a^2} < 0,\\ \frac{\partial^2 V}{\partial a \partial t} &= \frac{\partial V^H}{\partial a} - \frac{\partial V^L}{\partial a} > 0.\end{aligned}$$

Since $\frac{\partial V(t_H,a)}{\partial a}\big|_{a=a_H} = \frac{\partial V^H}{\partial a}\big|_{a=a_H} = 0$, $\alpha^R(t_H) = a_H > \underline{a}$. Then, $\frac{\partial V(t_L,a)}{\partial a}\big|_{a=a_H} < \frac{\partial V(t_H,a)}{\partial a}\big|_{a=a_H} = 0$ and $\frac{\partial^2 V}{\partial a^2} < 0$ implies that a unique $\alpha^R(t_L) < a_H$ exists.

If $\alpha^R(t_L) > \underline{a}$, $\frac{\partial V(t_L,a)}{\partial a}\big|_{a=\alpha^R(t_L)} = \frac{\partial V^L}{\partial a}\big|_{a=\alpha^R(t_L)} = 0$. Then,

$$\begin{aligned}\frac{\partial V(t,a)}{\partial a}\Big|_{a=\alpha^R(t_L)} &= t\frac{\partial V^H}{\partial a}\Big|_{a=\alpha^R(t_L)} + (1-t)\frac{\partial V^L}{\partial a}\Big|_{a=\alpha^R(t_L)} \ge 0,\\ \frac{\partial V(t,a)}{\partial a}\Big|_{a=\alpha^R(t_H)} &= t\frac{\partial V^H}{\partial a}\Big|_{a=\alpha^R(t_H)} + (1-t)\frac{\partial V^L}{\partial a}\Big|_{a=\alpha^R(t_H)} \le 0.\end{aligned}$$

Therefore, for any t, there exists a unique $\alpha^R(t)$ that satisfies $\frac{\partial V(t,a)}{\partial a}\big|_{a=\alpha^R(t)} = 0$. According to the implicit function theorem, $\alpha^R(t)$ is continuously differentiable with

$$\frac{d\alpha^R(t)}{dt} = -\frac{V_{at}(t,\alpha^R(t))}{V_{aa}(t,\alpha^R(t))} > 0.$$

If $\alpha^R(t_L) = \underline{a}$, $\frac{\partial V(t_L,a)}{\partial a}\big|_{a=\underline{a}} \le 0$. Because $\frac{\partial V(t_H,a)}{\partial a}\big|_{a=\underline{a}} > 0$, there exists a unique $\tilde{t}$ such that $\frac{\partial V(\tilde{t},a)}{\partial a}\big|_{a=\underline{a}} = 0$. For $t \le \tilde{t}$, $\frac{\partial V(t,a)}{\partial a}\big|_{a=\underline{a}} \le \frac{\partial V(\tilde{t},a)}{\partial a}\big|_{a=\underline{a}} \le 0$, so $\alpha^R(t) = \underline{a}$. For $t \ge \tilde{t}$, $\frac{\partial V(\tilde{t},a)}{\partial a}\big|_{a=\alpha^R(\tilde{t})} = 0$, and then, similarly to the proof above, α^R is continuously differentiable and strictly increasing in t. □

Lemma A.2. *The message strategies have the following properties.*

1. *$m_c(t)$ weakly increases in t. For any two types $\tilde{t} < \tilde{t}'$, if $m_c(\tilde{t}) = m_c(\tilde{t}')$, then $m_c(\tilde{t}) = m_c(\tilde{t}') = 0$.*
2. (Message Monotonicity) *In a sequential equilibrium of any signaling subgame, if $\sigma(m_j^*|t_j) > 0$ and $\sigma(m_k^*|t_k) > 0$, then the two equilibrium messages satisfy $m_j^* \le m_k^*$ for any $t_j < t_k$, $t_j, t_k \in T$.*

Proof. 1. Suppose $\exists\, t_1 < t_2$ s.t. $m_c(t_1) > m_c(t_2)$. Since $\frac{\partial^2 c}{\partial t \partial m} < 0$,

$$\frac{\partial}{\partial t}[c(t, m_c(t_1)) - c(t, m_c(t_2))] < 0,$$

that implies

$$c(t_2, m_c(t_1)) - c(t_2, m_c(t_2)) < c(t_1, m_c(t_1)) - c(t_1, m_c(t_2)).$$

Because $c(t_1, m_c(t_1)) - c(t_1, m_c(t_2)) < 0$, we have $c(t_2, m_c(t_1)) < c(t_2, m_c(t_2))$, which leads to a contradiction.

Suppose there exist two types $\tilde{t} < \tilde{t}'$ such that $m_c(\tilde{t}) = m_c(\tilde{t}') > 0$. Since $m_c < \overline{m}$, we have $c_m(\tilde{t}, m_c(\tilde{t})) = c_m(\tilde{t}', m_c(\tilde{t}')) = 0$, which contradicts the assumption that $\frac{\partial^2 c}{\partial t \partial m} < 0$. Therefore, $m_c(\tilde{t}) = m_c(\tilde{t}')$ must be the lowest message 0.

2. Let $a^*(m_j^*)$ and $a^*(m_k^*)$ be the receiver's actions in response to m_j^* and m_k^* in a sequential equilibrium, respectively. Suppose $m_j^* > m_k^*$, $t_j < t_k$. Since $U(a^*(m_j^*)) - c(t_j, m_j^*) \ge U(a^*(m_k^*)) - c(t_j, m_k^*)$, then $U(a^*(m_j^*)) - c(t_k, m_j^*) > U(a^*(m_k^*)) - c(t_k, m_k^*)$. Type t_k will deviate from reporting m_k^*, leading to a contradiction.

□

Lemma A.3. *If $m_i(\boldsymbol{\tau})$ exists, we have $m_i(\boldsymbol{\tau}) \ge m_c(t_i)$.*

Proof. (Remind that the separating message $m_i(\boldsymbol{\tau})$ is defined in Section 3.) $m_1(\boldsymbol{\tau}) = m_c(t_1)$ always exists. If $m_2(\boldsymbol{\tau})$ exists, $m_1(\boldsymbol{\tau}) = m_c(t_1) < \overline{m}$. Suppose $m_2(\boldsymbol{\tau}) < m_c(t_2)$. By the message monotonicity in Lemma A.2, $m_2(\boldsymbol{\tau}) > m_1(\boldsymbol{\tau}) = m_c(t_1)$. Because c is strictly quasi-convex in m, $c(t_1, m_2(\boldsymbol{\tau})) < \max\{c(t_1, m_c(t_1)), c(t_1, m_c(t_2))\} = c(t_1, m_c(t_2))$. Then

$$\hat{U}(t_1) - c(t_1, m_1(\boldsymbol{\tau})) \ge \hat{U}(t_2) - c(t_1, m_2(\boldsymbol{\tau})) > \hat{U}(t_2) - c(t_1, m_c(t_2)).$$

That is, t_2 will report $m_c(t_2)$ instead of $m_2(\boldsymbol{\tau})$, which leads to a contradiction. Thus, $m_2(\boldsymbol{\tau}) \ge m_c(t_2)$.

If $m_3(\boldsymbol{\tau})$ exists, $m_2(\boldsymbol{\tau}) < \overline{m}$. Then we can prove $m_3(\boldsymbol{\tau}) \geq m_c(t_3)$ by the same logic as above. The same argument applies to any $m_i(\boldsymbol{\tau})$, $i > 2$. ☐

Proof of Lemma 1

Proof. See Proposition 4.1-4.4 in Cho and Sobel (1990). Since the setting in our model satisfies their assumptions (A0-A4 in Section 4 in Cho and Sobel (1990)), their conclusions about the characterization and uniqueness of the D1 equilibrium apply. ☐

Proof of Lemma 2

Proof. Consider any pooling experiment $\boldsymbol{\tau} = (t_1, \dots, t_n; \tau_1, \dots, \tau_n)$, $n \geq 2$, that induces the D1 equilibrium in which types $t_i \geq t_p$ pool and type(s) $t_i < t_p$ (if any) separate. That is, type $t_i < t_p$ sends $m_i(\boldsymbol{\tau})$, and type $t_i \geq t_{p+1}$ sends $\overline{m}$, and type t_p sends $\overline{m}$ with probability $q \in (0, 1]$ and $m_p(\boldsymbol{\tau})$ with probability $1 - q$. In the following three cases, we show $\boldsymbol{\tau}$ cannot be optimal if the condition in Lemma 2 is violated.

1. **When $q = 1$ and $p > 1$**, type $t_i \geq t_p$ sends $\overline{m}$ and the receiver takes the action $\alpha^R(\phi_p(\boldsymbol{\tau}, 1))$ after receiving $\overline{m}$. Suppose $ECP_{\boldsymbol{\tau}} > \mathcal{C}(\phi_p(\boldsymbol{\tau}, 1), \overline{m})\big|_{[t_{p-1},1]}$.

   There must exist two points $(t'_p, c(t'_p, \overline{m}))$ and $(t'_{p+1}, c(t'_{p+1}, \overline{m}))$, where $t'_p, t'_{p+1} \in [t_{p-1}, 1]$ and $t'_p \leq \phi_p(\boldsymbol{\tau}, 1) < t'_{p+1}$, such that $rc(t'_p, \overline{m}) + (1-r)c(t'_{p+1}, \overline{m}) = \mathcal{C}(\phi_p(\boldsymbol{\tau}, 1), \overline{m})\big|_{[t_{p-1},1]}$ and $rt'_p + (1-r)t'_{p+1} = \phi_p(\boldsymbol{\tau}, 1)$, where $r = \frac{t'_{p+1} - \phi_p(\boldsymbol{\tau}, 1)}{t'_{p+1} - t'_p} \in (0, 1]$. Then, we have the following three cases:

   (a) $t'_p > t_{p-1}$ and $r < 1$

   We construct a new experiment $\boldsymbol{\tau'} = (t_1, \dots, t_{p-1}, t'_p, t'_{p+1}; \tau_1, \dots, \tau_{p-1}, r\sum_{i=p}^n \tau_i, (1-r)\sum_{i=p}^n \tau_i)$ and prove $E(\boldsymbol{\tau}) < E(\boldsymbol{\tau'})$ as follows. Since $\forall\, i < p$, $m_i(\boldsymbol{\tau'}) = m_i(\boldsymbol{\tau})$, and $\phi_p(\boldsymbol{\tau'}, 1) = \phi_p(\boldsymbol{\tau}, 1)$, then in the D1 equilibrium $\boldsymbol{\tau'}$ induces, $t_1, \dots, t_{p-1}$ still separate and $m_p(\boldsymbol{\tau'})$ exists.

   When $\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'})) \leq \hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - c(t'_p, \overline{m})$, $\boldsymbol{\tau'}$ induces a pooling equilibrium in which type t_i, $\forall\, i < p$, sends $m_i(\boldsymbol{\tau'})$ while types t_p and t_{p+1} pool at $\overline{m}$. Then, $f_{\boldsymbol{\tau'}}(t_i) = f_{\boldsymbol{\tau}}(t_i)$, $i < p$. Thus,

$$\begin{aligned}
E(\boldsymbol{\tau}) =& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau}}(t_i) + \sum_{i=p}^{n} \tau_i [\hat{U}(\phi_p(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m})] \\
=& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + r\Big(\sum_{i=p}^{n} \tau_i\Big)\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) + (1-r)\Big(\sum_{i=p}^{n} \tau_i\Big)\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - \sum_{i=p}^{n} \tau_i c(t_i, \overline{m}) \\
<& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \Big(\sum_{i=p}^{n} \tau_i\Big)\Big[r\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) + (1-r)\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - \mathcal{C}(\phi_p(\boldsymbol{\tau'}, 1), \overline{m})\big|_{[t_{p-1},1]}\Big] \\
=& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \Big[r\sum_{i=p}^{n} \tau_i\Big]\Big[\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - c(t'_p, \overline{m})\Big] + \Big[(1-r)\sum_{i=p}^{n} \tau_i\Big]\Big[\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - c(t'_{p+1}, \overline{m})\Big] \\
=& E(\boldsymbol{\tau'}).
\end{aligned} \tag{A.1}$$

   When $\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'})) \geq \hat{U}(t'_{p+1}) - c(t'_p, \overline{m})$, both t'_p and t'_{p+1} separate. Because $t'_{p+1} > \phi_p(\boldsymbol{\tau'}, 1)$

and $m_c(t'_{p+1}) \leq m_{p+1}(\boldsymbol{\tau'}) \leq \overline{m}$, based on equation (A.1), we have

$$\begin{aligned}
E(\boldsymbol{\tau}) <& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \Big[r \sum_{i=p}^{n} \tau_i\Big] \Big[\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - c(t'_p, \overline{m})\Big] \\
&+ \Big[(1-r) \sum_{i=p}^{n} \tau_i\Big] \Big[\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - c(t'_{p+1}, \overline{m})\Big] \\
<& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \Big[r \sum_{i=p}^{n} \tau_i\Big] \Big[\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'}))\Big] \\
&+ \Big[(1-r) \sum_{i=p}^{n} \tau_i\Big] \Big[\hat{U}(t'_{p+1}) - c(t'_{p+1}, m_{p+1}(\boldsymbol{\tau'}))\Big] \\
=& E(\boldsymbol{\tau'}).
\end{aligned}$$

When $\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - c(t'_p, \overline{m}) < \hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'})) < \hat{U}(t'_{p+1}) - c(t'_p, \overline{m})$, t'_{p+1} would report $\overline{m}$, and t'_p would report $m_p(\boldsymbol{\tau'})$ and $\overline{m}$ with probability $1-q'$ and q', respectively, where $q' \in (0,1)$ satisfies $\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'})) = \hat{U}(\phi_p(\boldsymbol{\tau'}, q')) - c(t'_p, \overline{m})$. Since $\phi_p(\boldsymbol{\tau'}, q') > \phi_p(\boldsymbol{\tau'}, 1)$,

$$\begin{aligned}
E(\boldsymbol{\tau}) <& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \Big[r \sum_{i=p}^{n} \tau_i\Big] \Big[\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - c(t'_p, \overline{m})\Big] \\
&+ \Big[(1-r) \sum_{i=p}^{n} \tau_i\Big] \Big[\hat{U}(\phi_p(\boldsymbol{\tau'}, 1)) - c(t'_{p+1}, \overline{m})\Big] \\
<& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \Big[r \sum_{i=p}^{n} \tau_i\Big] \Big[\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'}))\Big] \\
&+ \Big[(1-r) \sum_{i=p}^{n} \tau_i\Big] \Big[\hat{U}(\phi_p(\boldsymbol{\tau'}, q')) - c(t'_{p+1}, \overline{m})\Big] \\
=& E(\boldsymbol{\tau'}).
\end{aligned}$$

(b) $r = 1$ or $t'_p = \phi_p(\boldsymbol{\tau}, 1) > t_{p-1}$
We construct a new experiment $\boldsymbol{\tau'} = (t_1, \ldots, t_{p-1}, t'_p; \tau_1, \ldots, \tau_{p-1}, \sum_{i=p}^{n} \tau_i)$. It would induce a separating equilibrium in which $m_i(\boldsymbol{\tau'}) = m_i(\boldsymbol{\tau})$, $i < p$, and $m_p(\boldsymbol{\tau'}) \leq \overline{m}$. Thus,

$$\begin{aligned}
E(\boldsymbol{\tau}) =& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau}}(t_i) + \sum_{i=p}^{n} \tau_i [\hat{U}(\phi_p(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m})] \\
=& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \Big(\sum_{i=p}^{n} \tau_i\Big) \hat{U}(t'_p) - \sum_{i=p}^{n} \tau_i c(t_i, \overline{m}) \\
<& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \Big(\sum_{i=p}^{n} \tau_i\Big) \Big[\hat{U}(t'_p) - \mathcal{C}(\phi_p(\boldsymbol{\tau'}, 1), \overline{m})\big|_{[t_{p-1}, 1]}\Big] \\
\leq& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \Big(\sum_{i=p}^{n} \tau_i\Big) \Big[\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'}))\Big] \\
=& E(\boldsymbol{\tau'}).
\end{aligned}$$

(c) $t'_p = t_{p-1}$, $t'_{p+1} > \phi_p(\boldsymbol{\tau}, 1)$

We construct a new experiment $\boldsymbol{\tau'} = (t_1, \ldots, t_{p-1}, t'_{p+1}; \tau_1, \ldots, \tau_{p-2}, \tau_{p-1} + r\sum_{i=p}^n \tau_i, (1-r)\sum_{i=p}^n \tau_i)$, then $m_i(\boldsymbol{\tau'}) = m_i(\boldsymbol{\tau})$, $i \leq p-1$. Since $\hat{U}(t_{p-1}) - c(t_{p-1}, m_{p-1}(\boldsymbol{\tau})) \geq \hat{U}(\phi_p(\boldsymbol{\tau}, 1)) - c(t_{p-1}, \overline{m})$ and $\phi_p(\boldsymbol{\tau}, 1) > \phi_{p-1}(\boldsymbol{\tau'}, 1)$, we have $\hat{U}(t_{p-1}) - c(t_{p-1}, m_{p-1}(\boldsymbol{\tau'})) > \hat{U}(\phi_{p-1}(\boldsymbol{\tau'}, 1)) - c(t_{p-1}, \overline{m})$.

If $\hat{U}(t_{p-1}) - c(t_{p-1}, m_{p-1}(\boldsymbol{\tau'})) \geq \hat{U}(t'_{p+1}) - c(t_{p-1}, \overline{m})$, $\boldsymbol{\tau'}$ induces a separating equilibrium and

$$\begin{aligned}
E(\boldsymbol{\tau}) <& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau}}(t_i) + \left[r\sum_{i=p}^{n}\tau_i\right]\left[\hat{U}(\phi_p(\boldsymbol{\tau}, 1)) - c(t_{p-1}, \overline{m})\right] \\
&+ \left[(1-r)\sum_{i=p}^{n}\tau_i\right]\left[\hat{U}(\phi_p(\boldsymbol{\tau}, 1)) - c(t'_{p+1}, \overline{m})\right] \\
<& \sum_{i=1}^{p-1} \tau_i f_{\boldsymbol{\tau'}}(t_i) + \left[r\sum_{i=p}^{n}\tau_i\right]\left[\hat{U}(t_{p-1}) - c(t_{p-1}, m_{p-1}(\boldsymbol{\tau'}))\right] \\
&+ \left[(1-r)\sum_{i=p}^{n}\tau_i\right]\left[\hat{U}(t'_{p+1}) - c(t'_{p+1}, m_p(\boldsymbol{\tau'}))\right] \\
=& E(\boldsymbol{\tau'}).
\end{aligned}$$

If $\hat{U}(t_{p-1}) - c(t_{p-1}, m_{p-1}(\boldsymbol{\tau'})) < \hat{U}(t'_{p+1}) - c(t_{p-1}, \overline{m})$, $\boldsymbol{\tau'}$ induces a partial-pooling equilibrium in which t_{p-1} reports $\overline{m}$ with probability $q' \in (0,1)$ that satisfies $\hat{U}(t_{p-1}) - c(t_{p-1}, m_{p-1}(\boldsymbol{\tau'})) = \hat{U}(\phi_{p-1}(\boldsymbol{\tau'}, q')) - c(t_{p-1}, \overline{m})$. Because $\phi_{p-1}(\boldsymbol{\tau'}, q') \geq \phi_p(\boldsymbol{\tau}, 1)$, we must have $E(\boldsymbol{\tau}) < E(\boldsymbol{\tau'})$.

2. **When $q = 1$ and $p = 1$**, all types pool at $\overline{m}$. Suppose $ECP_{\boldsymbol{\tau}} > \mathcal{C}(\mu, \overline{m})\big|_{[0,1]}$.

   There must exist two points $(t'_1, c(t'_1, \overline{m}))$ and $(t'_2, c(t'_2, \overline{m}))$, $0 \leq t'_1 \leq \mu < t'_2 \leq 1$, such that $rc(t'_1, \overline{m}) + (1-r)c(t'_2, \overline{m}) = \mathcal{C}(\mu, \overline{m})\big|_{[0,1]}$ and $rt'_1 + (1-r)t'_2 = \mu$, where $r = \frac{t'_2 - \mu}{t'_2 - t'_1} \in (0,1]$. If $r < 1$, as the above case (a) and (b), we can similarly construct a new experiment $\boldsymbol{\tau'} = (t'_1, t'_2; r, 1-r)$ and prove $E(\boldsymbol{\tau}) < E(\boldsymbol{\tau'})$, no matter which kind of D1 equilibrium $\boldsymbol{\tau'}$ would induce. If $r = 1$, the uninformative experiment $(\mu; 1)$ is strictly better than $\boldsymbol{\tau}$.

3. **When $0 < q < 1$**, we have $\hat{U}(t_p) - c(t_p, m_p(\boldsymbol{\tau})) = \hat{U}(\phi_p(\boldsymbol{\tau}, q)) - c(t_p, \overline{m})$. If $ECP_{\boldsymbol{\tau}} > \mathcal{C}(\phi_p(\boldsymbol{\tau}, q), \overline{m})\big|_{[t_p,1]}$, we can similarly prove $\boldsymbol{\tau}$ is not optimal based on the above analysis.

□

Proof of Proposition 1

Proof. ***Part 1.*** *We prove for any separating experiment with more than two types, we can construct a separating experiment with two (or one) types that induces weakly higher expected payoff.*

Consider any separating experiment $\boldsymbol{\tau} = (t_1, \ldots, t_n; \tau_1, \ldots, \tau_n)$, $n \geq 3$. The sender's expected payoff is $E(\boldsymbol{\tau}) = \sum_{i=1}^n \tau_i f_{\boldsymbol{\tau}}(t_i)$, where $f_{\boldsymbol{\tau}}(t_i) = \hat{U}(t_i) - c(t_i, m_i(\boldsymbol{\tau}))$, and $\mathcal{F}_{\boldsymbol{\tau}}(\mu) \geq E(\boldsymbol{\tau})$. There must exist $t_j, t_k \in T$ with $t_j \leq \mu < t_k$ such that $rf_{\boldsymbol{\tau}}(t_j) + (1-r)f_{\boldsymbol{\tau}}(t_k) = \mathcal{F}_{\boldsymbol{\tau}}(\mu)$ and $rt_j + (1-r)t_k = \mu$, where $r = \frac{t_k - \mu}{t_k - t_j} \in (0,1]$.

If $t_j = \mu$, the uninformative experiment $(\mu; 1)$ is weakly better than $\boldsymbol{\tau}$. Next, we show if $t_j < \mu$, the experiment $\boldsymbol{\tau'} = (t_j, t_k; r, 1-r)$ is weakly better than $\boldsymbol{\tau}$. Since $m_1(\boldsymbol{\tau'}) = m_c(t_j) \leq m_j(\boldsymbol{\tau})$, by the IC conditions and single-crossing condition, $m_2(\boldsymbol{\tau'})$ must exist and $m_2(\boldsymbol{\tau'}) \leq m_k(\boldsymbol{\tau})$. Thus, $\boldsymbol{\tau'}$ is a separating experiment and

$$\begin{aligned} E(\boldsymbol{\tau'}) &= r\big[\hat{U}(t_j) - c(t_j, m_1(\boldsymbol{\tau'}))\big] + (1-r)\big[\hat{U}(t_k) - c(t_k, m_2(\boldsymbol{\tau'}))\big] \\ &\geq r f_{\boldsymbol{\tau}}(t_j) + (1-r) f_{\boldsymbol{\tau}}(t_k) \\ &= \mathcal{F}_{\boldsymbol{\tau}}(\mu) \geq E(\boldsymbol{\tau}). \end{aligned}$$

Part 2. *We prove that if a pooling experiment $\boldsymbol{\tau}$ is optimal, it needs at most one type that separates with positive probability.*

Based on Lemma 2, we only need to consider the experiment that includes two pooling types because at most two types are needed to approach any convex lower closure. Then, we consider any pooling experiment $\boldsymbol{\tau} = (t_1, \ldots, t_n; \tau_1, \ldots, \tau_n)$, $n \geq 2$, that induces t_{n-1} and t_n to pool. Specifically, type t_{n-1} reports $\overline{m}$ with probability $q \in (0,1]$ and $m_{n-1}(\boldsymbol{\tau})$ with probability $1-q$. Next, we show $\boldsymbol{\tau}$ is weakly dominated by other experiment if $\boldsymbol{\tau}$ induces more than one type to separate with positive probability.

1. **When $\boldsymbol{q = 1}$**, $\boldsymbol{\tau}$ induces $t_i \leq t_{n-2}$ to separate with $m_i(\boldsymbol{\tau})$ and types t_{n-1} and t_n to pool at $\overline{m}$. Suppose $n \geq 4$.

   There must exist two points $(t_j, f_{\boldsymbol{\tau}}(t_j))$ and $(t_k, f_{\boldsymbol{\tau}}(t_k))$, $t_j, t_k \in T$, $t_j \leq \mu < t_k$, such that $r f_{\boldsymbol{\tau}}(t_j) + (1-r) f_{\boldsymbol{\tau}}(t_k) = \mathcal{F}_{\boldsymbol{\tau}}(\mu)$ and $r t_j + (1-r) t_k = \mu$, where $r = \frac{t_k - \mu}{t_k - t_j} \in (0,1]$. If $t_j = \mu$, the uninformative experiment is weakly better than $\boldsymbol{\tau}$. If $t_j < \mu$, we consider the following four cases.

   (a) $k \leq n-2$

   Based on the proof in Part 1, experiment $\boldsymbol{\tau'} = (t_j, t_k; r, 1-r)$ is weakly better than $\boldsymbol{\tau}$.

   (b) $j = n-1$

   Denote

   $$t'_{n-1} = \phi_{n-1}(\boldsymbol{\tau}, 1) = \frac{\tau_{n-1} t_{n-1} + \tau_n t_n}{\tau_{n-1} + \tau_n}$$

   and $T' = \{t_1, \ldots, t_{n-2}, t'_{n-1}\}$. Let $f : T' \to \mathbb{R}$ be a function such that $f(t_i) = f_{\boldsymbol{\tau}}(t_i)$, $i \leq n-2$, and $f(t'_{n-1}) = \frac{\tau_{n-1} f_{\boldsymbol{\tau}}(t_{n-1}) + \tau_n f_{\boldsymbol{\tau}}(t_n)}{\tau_{n-1} + \tau_n}$. Denote $\mathcal{F}(\hat{t}, f)$ as the concave closure of f, where $\hat{t} \in co(T')$ and f is defined on T'. Since $t_1 < \cdots < t_{n-2} < \mu < t'_{n-1}$, there exists $t_h \in T'$, $h \leq n-2$ that satisfies $r' f(t_h) + (1-r') f(t'_{n-1}) = \mathcal{F}(\mu, f)$ and $r' t_h + (1-r') t'_{n-1} = \mu$, where $r' = \frac{t'_{n-1} - \mu}{t'_{n-1} - t_h} \in (0,1)$. Then, we can construct a new experiment $\boldsymbol{\tau'} = (t_h, t_{n-1}, t_n; r', \frac{\tau_{n-1}}{\tau_{n-1} + \tau_n}(1-r'), \frac{\tau_n}{\tau_{n-1} + \tau_n}(1-r'))$ and show its expected payoff is weakly higher than $\boldsymbol{\tau}$ as follows.

   Since t_h separates in the equilibrium induced by $\boldsymbol{\tau}$,

   $$\hat{U}(t_h) - c(t_h, m_h(\boldsymbol{\tau})) \geq \hat{U}(\phi_{n-1}(\boldsymbol{\tau}, 1)) - c(t_h, \overline{m}).$$

   Because $\phi_2(\boldsymbol{\tau'}, 1) = \phi_{n-1}(\boldsymbol{\tau}, 1)$, we have

   $$\hat{U}(t_h) - c(t_h, m_c(t_h)) \geq \hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(t_h, \overline{m}),$$

   which means $\boldsymbol{\tau'}$ induces t_h to separate with her cost-minimizing message. If $\boldsymbol{\tau'}$ induces t_{n-1} and t_n to pool at $\overline{m}$, we have $f_{\boldsymbol{\tau'}}(t_{n-1}) = f_{\boldsymbol{\tau}}(t_{n-1})$ and $f_{\boldsymbol{\tau'}}(t_n) = f_{\boldsymbol{\tau}}(t_n)$. Thus, $E(\boldsymbol{\tau'}) \geq \mathcal{F}(\mu, f) \geq E(\boldsymbol{\tau})$. Otherwise, t_{n-1} separates with positive probability, implying $\hat{U}(t_{n-1}) - c(t_{n-1}, m_2(\boldsymbol{\tau'})) > \hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(t_{n-1}, \overline{m})$. Hence, $f_{\boldsymbol{\tau'}}(t_{n-1}) > f_{\boldsymbol{\tau}}(t_{n-1})$ and t_n gets a higher receiver action such that $f_{\boldsymbol{\tau'}}(t_n) > f_{\boldsymbol{\tau}}(t_n)$, which makes $E(\boldsymbol{\tau'}) > \mathcal{F}(\mu, f) \geq E(\boldsymbol{\tau})$.

(c) $k = n-1$

The proof is similar to that for the above case (b) $j = n-1$.

(d) $j \leq n-2$ and $k = n$

We can construct an experiment $\boldsymbol{\tau'} = (t_j, t_n; r, 1-r)$. Since in the equilibrium induced by experiment $\boldsymbol{\tau}$, t_j separates, we have

$$\hat{U}(t_j) - c(t_j, m_c(t_j)) \geq \hat{U}(t_j) - c(t_j, m_j(\boldsymbol{\tau})) \geq \hat{U}(\phi_{n-1}(\boldsymbol{\tau}, 1)) - c(t_j, \overline{m}).$$

Thus, whether $\boldsymbol{\tau'}$ induces a separating or pooling equilibrium, $E(\boldsymbol{\tau'}) \geq \mathcal{F}_{\boldsymbol{\tau}}(\mu) \geq E(\boldsymbol{\tau})$.

2. **When $0 < q < 1$**, $\boldsymbol{\tau}$ induces t_{n-1} to separate with probability $1-q$. Similarly, we can prove $\boldsymbol{\tau}$ is weakly dominated by other experiment if it contains any other type(s) besides types t_{n-1} and t_n.

□

Proof of Corollary 1

Proof. **The first condition:** Consider any pooling experiment $\boldsymbol{\tau} = (t_1, \ldots, t_n; \tau_1, \ldots, \tau_n)$. If $\boldsymbol{\tau}$ induces a pure strategy pooling equilibrium in which type t_i, $\forall\, i \geq p$, pool at $\overline{m}$ and all other types (if any) separate, according to the proof of Lemma 2, we can construct experiment $\boldsymbol{\tau'} = (t_1, \ldots, t_{p-1}, \phi_p(\boldsymbol{\tau}, 1); \tau_1, \ldots, \tau_{p-1}, \sum_{i=p}^{n} \tau_i)$, where $\phi_p(\boldsymbol{\tau}, 1) \geq \mu$ must hold. Then, $\boldsymbol{\tau'}$ would induce a separating equilibrium with an expected payoff higher than $\boldsymbol{\tau}$. If $\boldsymbol{\tau}$ induces a mixed strategy partial-pooling equilibrium in which type t_i, $\forall\, i > p$, reports $\overline{m}$ and type t_p reports $\overline{m}$ with probability $q \in (0,1)$, we can construct experiment $\widetilde{\boldsymbol{\tau}} = (t_1, \ldots, t_{p-1}, t_p, \phi_p(\boldsymbol{\tau}, q); \tau_1, \ldots, \tau_{p-1}, (1-q)\tau_p, q\tau_p + \sum_{i=p+1}^{n} \tau_i)$, where $\phi_p(\boldsymbol{\tau}, q) \geq \mu$ must hold. Then, $\widetilde{\boldsymbol{\tau}}$ would induce a separating equilibrium with an expected payoff higher than $\boldsymbol{\tau}$. Therefore, any pooling experiment cannot be optimal, that is, the optimal experiment must be separating.

The second condition: By Proposition 1, the optimal experiment contains at most three results. Since $\hat{U}(t) \geq \hat{U}(t_L) > \hat{U}(t_H) - c(t, \overline{m})$, any experiment containing two results induces a separating equilibrium. Consider any experiment with three results, denoted by $\boldsymbol{\tau} = (t_1, t_2, t_3; \tau_1, \tau_2, \tau_3)$. Since $\hat{U}(t_1) > \hat{U}(t_H) - c(t_1, \overline{m})$, t_1 separates with $m_c(t_1)$. If $\hat{U}(t_1) \geq \hat{U}(t_2) - c(t_1, m_c(t_2))$, t_2 separates with $m_c(t_2)$ because $\hat{U}(t_2) > \hat{U}(t_H) - c(t_2, \overline{m})$. If $\hat{U}(t_1) < \hat{U}(t_2) - c(t_1, m_c(t_2))$, we have $\hat{U}(t_1) = \hat{U}(t_2) - c(t_1, m_2(\boldsymbol{\tau}))$. Because

$$\hat{U}(t_2) - c(t_2, m_2(\boldsymbol{\tau})) > \hat{U}(t_2) - c(t_1, m_2(\boldsymbol{\tau})) = \hat{U}(t_1) > \hat{U}(t_H) - c(t_2, \overline{m}),$$

$\boldsymbol{\tau}$ induces a separating equilibrium. □

Proof of Proposition 2

Proof. If the optimal separating experiment $\boldsymbol{\tau^s} = (t_1, t_2; \tau_1, \tau_2)$ satisfies $c(t_2, m_2(\boldsymbol{\tau^s})) > \mathcal{C}(t_2, \overline{m})\big|_{[t_1,1]}$, we can find a partial-pooling experiment that is strictly better than $\boldsymbol{\tau^s}$, which means the optimal experiment must induce a partial-pooling equilibrium.

Since $c(t_2, \overline{m}) \geq c(t_2, m_2(\boldsymbol{\tau^s})) > \mathcal{C}(t_2, \overline{m})\big|_{[t_1,1]}$, there must exist two points $(t_2', c(t_2', \overline{m}))$ and $(t_2'', c(t_2'', \overline{m}))$, $t_1 \leq t_2' < t_2 < t_2'' \leq 1$, such that $rc(t_2', \overline{m}) + (1-r)c(t_2'', \overline{m}) = \mathcal{C}(t_2, \overline{m})\big|_{[t_1,1]}$ and $rt_2' + (1-r)t_2'' = t_2$, where $r = \frac{t_2'' - t_2}{t_2'' - t_2'}$.

If $t_2' > t_1$, we then show there exists a partial-pooling experiment $\boldsymbol{\tau'} = (t_1, t_2', t_2''; \tau_1, \tau_2 r, \tau_2(1-r))$ that induces $E(\boldsymbol{\tau'}) > E(\boldsymbol{\tau^s})$. Since $\phi_2(\boldsymbol{\tau'}, 1) = t_2$, then $\hat{U}(t_1) \geq \hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(t_1, \overline{m})$, implying in the D1 equilibrium induced by $\boldsymbol{\tau'}$, t_1 separates and $m_2(\boldsymbol{\tau'})$ exists.

- If $\hat{U}(t_2') - c(t_2', m_2(\boldsymbol{\tau'})) \geq \hat{U}(t_2'') - c(t_2', \overline{m})$, t_2' would separate. Then, $\boldsymbol{\tau'}$ induces a separating equilibrium with an expected payoff strictly higher than $E(\boldsymbol{\tau^s})$, which is impossible.
- If $\hat{U}(t_2') - c(t_2', m_2(\boldsymbol{\tau'})) < \hat{U}(t_2'') - c(t_2', \overline{m})$, $\boldsymbol{\tau'}$ would induce a partial-pooling equilibrium. When $\hat{U}(t_2') - c(t_2', m_2(\boldsymbol{\tau'})) \leq \hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(t_2', \overline{m})$, t_2' and t_2'' would pool at $\overline{m}$, so $E(\boldsymbol{\tau'}) > E(\boldsymbol{\tau^s})$. When $\hat{U}(t_2') - c(t_2', m_2(\boldsymbol{\tau'})) > \hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(t_2', \overline{m})$, $\exists\ q \in (0,1)$ s.t. $\hat{U}(t_2') - c(t_2', m_2(\boldsymbol{\tau'})) = \hat{U}(\phi_2(\boldsymbol{\tau'}, q)) - c(t_2', \overline{m})$, that means t_2' would report $\overline{m}$ with probability q. Thus,

$$\begin{aligned} E(\boldsymbol{\tau'}) &= \tau_1 \hat{U}(t_1) + \tau_2 r[\hat{U}(\phi_2(\boldsymbol{\tau'}, q)) - c(t_2', \overline{m})] + \tau_2(1-r)[\hat{U}(\phi_2(\boldsymbol{\tau'}, q)) - c(t_2'', \overline{m})] \\ &> \tau_1 \hat{U}(t_1) + \tau_2 r[\hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(t_2', \overline{m})] + \tau_2(1-r)[\hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(t_2'', \overline{m})] \\ &> E(\boldsymbol{\tau^s}). \end{aligned}$$

<u>If $t_2' = t_1$</u>, we then show experiment $\boldsymbol{\tau''} = (t_1, t_2''; \tau_1 + \tau_2 r, \tau_2(1-r))$ induces a partial-pooling equilibrium with $E(\boldsymbol{\tau''}) > E(\boldsymbol{\tau^s})$. Since

$$\hat{U}(t_1) \geq \hat{U}(t_2) - c(t_1, \overline{m}) > \hat{U}(\mu) - c(t_1, \overline{m}),$$

t_1 would not report $\overline{m}$ with probability 1.

- If $\hat{U}(t_1) \geq \hat{U}(t_2'') - c(t_1, \overline{m})$, $\boldsymbol{\tau''}$ induces a separating equilibrium. Then,

$$\begin{aligned} E(\boldsymbol{\tau''}) &= (\tau_1 + \tau_2 r)\hat{U}(t_1) + \tau_2(1-r)[\hat{U}(t_2'') - c(t_2'', m_2(\boldsymbol{\tau''}))] \\ &\geq \tau_1 \hat{U}(t_1) + \tau_2 r[\hat{U}(t_2'') - c(t_1, \overline{m})] + \tau_2(1-r)[\hat{U}(t_2'') - c(t_2'', \overline{m})] \\ &> E(\boldsymbol{\tau^s}), \end{aligned}$$

  which is impossible.
- If $\hat{U}(t_1) < \hat{U}(t_2'') - c(t_1, \overline{m})$, $\boldsymbol{\tau''}$ induces a partial-pooling equilibrium. There exists $q' \in (0,1)$ s.t. $\hat{U}(t_1) = \hat{U}(\phi_1(\boldsymbol{\tau''}, q')) - c(t_1, \overline{m})$, where $\phi_1(\boldsymbol{\tau''}, q') \geq t_2$. Then, t_1 sends $\overline{m}$ with probability q' and

$$\begin{aligned} E(\boldsymbol{\tau''}) &= (\tau_1 + \tau_2 r)\hat{U}(t_1) + \tau_2(1-r)[\hat{U}(\phi_1(\boldsymbol{\tau''}, q')) - c(t_2'', \overline{m})] \\ &\geq \tau_1 \hat{U}(t_1) + \tau_2 r[\hat{U}(t_2) - c(t_1, \overline{m})] + \tau_2(1-r)[\hat{U}(t_2) - c(t_2'', \overline{m})] \\ &> E(\boldsymbol{\tau^s}). \end{aligned}$$

□

Proof of Proposition 3

Proof. In this proof, we show that under either Condition 1 or Condition 2, there exists a pooling experiment that induces an expected payoff strictly higher than $\mathcal{G}(\mu)$.

<u>**Under Condition 1**</u> that $\hat{U}(\mu) - \mathcal{C}(\mu, \overline{m})\big|_{[0,1]} > \mathcal{G}(\mu)$.

If $\mathcal{C}(\mu, \overline{m})\big|_{[0,1]} = c(\mu, \overline{m})$, $\hat{U}(\mu) - \mathcal{C}(\mu, \overline{m})\big|_{[0,1]} = \hat{U}(\mu) - c(\mu, \overline{m}) < g(\mu) \leq \mathcal{G}(\mu)$, which contradicts Condition 1. If $\mathcal{C}(\mu, \overline{m})\big|_{[0,1]} < c(\mu, \overline{m})$, there exist two points $(\tilde{t}, c(\tilde{t}, \overline{m}))$ and $(\tilde{t}', c(\tilde{t}', \overline{m}))$, $0 \leq \tilde{t} < \mu < \tilde{t}' \leq 1$, such that $rc(\tilde{t}, \overline{m}) + (1-r)c(\tilde{t}', \overline{m}) = \mathcal{C}(\mu, \overline{m})\big|_{[0,1]}$ and $r\tilde{t} + (1-r)\tilde{t}' = \mu$, where $r = \frac{\tilde{t}' - \mu}{\tilde{t}' - \tilde{t}} \in (0,1)$. Then, we consider the

D1 equilibrium induced by experiment $\boldsymbol{\tau'} = (\tilde{t}, \tilde{t}'; r, 1-r)$. Since

$$\begin{aligned} & r[\hat{U}(\mu) - c(\tilde{t}, \overline{m})] + (1-r)[\hat{U}(\mu) - c(\tilde{t}', \overline{m})] \\ = & \hat{U}(\mu) - \mathcal{C}(\mu, \overline{m})\big|_{[0,1]} \\ > & \mathcal{G}(\mu) \geq rg(\tilde{t}) + (1-r)g(\tilde{t}') \end{aligned}$$

and $g(\tilde{t}') = \hat{U}(\tilde{t}') - c(\tilde{t}', m_c(\tilde{t}')) > \hat{U}(\mu) - c(\tilde{t}', \overline{m})$, we have $\hat{U}(\tilde{t}) - c(\tilde{t}, m_c(\tilde{t})) = g(\tilde{t}) < \hat{U}(\mu) - c(\tilde{t}, \overline{m})$. Therefore, we find the experiment $\boldsymbol{\tau'}$ induces a total-pooling equilibrium and

$$E(\boldsymbol{\tau'}) = r[\hat{U}(\mu) - c(\tilde{t}, \overline{m})] + (1-r)[\hat{U}(\mu) - c(\tilde{t}', \overline{m})] > \mathcal{G}(\mu).$$

Under Condition 2 that there exists experiment $\tilde{\boldsymbol{\tau}} = (\tilde{t}_1, \tilde{t}_2; \tilde{\tau}_1, \tilde{\tau}_2)$, where $\tilde{\tau}_1 g(\tilde{t}_1) + \tilde{\tau}_2 g(\tilde{t}_2) = \mathcal{G}(\mu)$, that satisfies $\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1)) \geq \hat{U}(\tilde{t}_2) - c(\tilde{t}_1, \overline{m})$ and $c(\tilde{t}_2, m_c(\tilde{t}_2)) > \mathcal{C}(\tilde{t}_2, \overline{m})\big|_{[\tilde{t}_1,1]}$.

Because $c(\tilde{t}_2, \overline{m})) \geq c(\tilde{t}_2, m_c(\tilde{t}_2)) > \mathcal{C}(\tilde{t}_2, \overline{m})\big|_{[\tilde{t}_1,1]}$, there exist two points $(\tilde{t}, c(\tilde{t}, \overline{m}))$ and $(\tilde{t}', c(\tilde{t}', \overline{m}))$, $\tilde{t}_1 \leq \tilde{t} < \tilde{t}_2 < \tilde{t}' \leq 1$, such that $\gamma c(\tilde{t}, \overline{m}) + (1-\gamma) c(\tilde{t}', \overline{m}) = \mathcal{C}(\tilde{t}_2, \overline{m})\big|_{[\tilde{t}_1,1]}$ and $\gamma \tilde{t} + (1-\gamma)\tilde{t}' = \tilde{t}_2$, where $\gamma = \frac{\tilde{t}' - \tilde{t}_2}{\tilde{t}' - \tilde{t}}$.

<u>If $\tilde{t} > \tilde{t}_1$</u>, we then show experiment $\boldsymbol{\tau'} = (\tilde{t}_1, \tilde{t}, \tilde{t}'; \tilde{\tau}_1, \tilde{\tau}_2\gamma, \tilde{\tau}_2(1-\gamma))$ induces a partial-pooling equilibrium with $E(\boldsymbol{\tau'}) > \mathcal{G}(\mu)$. Since $\phi_2(\boldsymbol{\tau'}, 1) = \tilde{t}_2$, then $\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1)) \geq \hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(\tilde{t}_1, \overline{m})$, implying in the D1 equilibrium induced by $\boldsymbol{\tau'}$, $\tilde{t}_1$ separates and $m_2(\boldsymbol{\tau'})$ exists.

- If $\hat{U}(\tilde{t}) - c(\tilde{t}, m_2(\boldsymbol{\tau'})) \geq \hat{U}(\tilde{t}') - c(\tilde{t}, \overline{m})$, $\tilde{t}$ would separate. Then, $\boldsymbol{\tau'}$ induces a separating equilibrium with an expected payoff strictly higher than $\mathcal{G}(\mu)$, which is impossible.
- If $\hat{U}(\tilde{t}) - c(\tilde{t}, m_2(\boldsymbol{\tau'})) < \hat{U}(\tilde{t}') - c(\tilde{t}, \overline{m})$, $\boldsymbol{\tau'}$ would induce a partial-pooling equilibrium. When $\hat{U}(\tilde{t}) - c(\tilde{t}, m_2(\boldsymbol{\tau'})) \leq \hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(\tilde{t}, \overline{m})$, $\tilde{t}$ and $\tilde{t}'$ would pool at $\overline{m}$, so $E(\boldsymbol{\tau'}) > \mathcal{G}(\mu)$. When $\hat{U}(\tilde{t}) - c(\tilde{t}, m_2(\boldsymbol{\tau'})) > \hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(\tilde{t}, \overline{m})$, $\exists\, q \in (0,1)$ s.t. $\hat{U}(\tilde{t}) - c(\tilde{t}, m_2(\boldsymbol{\tau'})) = \hat{U}(\phi_2(\boldsymbol{\tau'}, q)) - c(\tilde{t}, \overline{m})$, that means $\tilde{t}$ would report $\overline{m}$ with probability q. Thus,

$$\begin{aligned} E(\boldsymbol{\tau'}) &= \tilde{\tau}_1[\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1))] + \tilde{\tau}_2\gamma[\hat{U}(\phi_2(\boldsymbol{\tau'}, q)) - c(\tilde{t}, \overline{m})] + \tilde{\tau}_2(1-\gamma)[\hat{U}(\phi_2(\boldsymbol{\tau'}, q)) - c(\tilde{t}', \overline{m})] \\ &> \tilde{\tau}_1[\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1))] + \tilde{\tau}_2\gamma[\hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(\tilde{t}, \overline{m})] + \tilde{\tau}_2(1-\gamma)[\hat{U}(\phi_2(\boldsymbol{\tau'}, 1)) - c(\tilde{t}', \overline{m})] \\ &> \mathcal{G}(\mu). \end{aligned}$$

<u>If $\tilde{t} = \tilde{t}_1$</u>, we then show experiment $\boldsymbol{\tau''} = (\tilde{t}_1, \tilde{t}'; \tilde{\tau}_1 + \tilde{\tau}_2\gamma, \tilde{\tau}_2(1-\gamma))$ induces a partial-pooling equilibrium with $E(\boldsymbol{\tau''}) > \mathcal{G}(\mu)$. Since

$$\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1)) \geq \hat{U}(\tilde{t}_2) - c(\tilde{t}_1, \overline{m}) > \hat{U}(\mu) - c(\tilde{t}_1, \overline{m}),$$

$\tilde{t}_1$ would not report $\overline{m}$ with probability 1.

- If $\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1)) \geq \hat{U}(\tilde{t}') - c(\tilde{t}_1, \overline{m})$, $\boldsymbol{\tau''}$ induces a separating equilibrium. Then,

$$\begin{aligned} E(\boldsymbol{\tau''}) &= (\tilde{\tau}_1 + \tilde{\tau}_2\gamma)[\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1))] + \tilde{\tau}_2(1-\gamma)[\hat{U}(\tilde{t}') - c(\tilde{t}', m_2(\boldsymbol{\tau''}))] \\ &\geq \tilde{\tau}_1[\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1))] + \tilde{\tau}_2\gamma[\hat{U}(\tilde{t}') - c(\tilde{t}_1, \overline{m})] + \tilde{\tau}_2(1-\gamma)[\hat{U}(\tilde{t}') - c(\tilde{t}', \overline{m})] \\ &> \mathcal{G}(\mu), \end{aligned}$$

  which is impossible.

- If $\hat{U}(\tilde{t}_1)-c(\tilde{t}_1,m_c(\tilde{t}_1)) < \hat{U}(\tilde{t}')-c(\tilde{t}_1,\overline{m})$, $\boldsymbol{\tau''}$ induces a partial-pooling equilibrium. There exists $q' \in (0,1)$ s.t. $\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1)) = \hat{U}(\phi_1(\boldsymbol{\tau''}, q')) - c(\tilde{t}_1, \overline{m})$, where $\phi_1(\boldsymbol{\tau''}, q') \geq \tilde{t}_2$. Then, $\tilde{t}_1$ sends $\overline{m}$ with probability q' and

$$
\begin{aligned}
E(\boldsymbol{\tau''}) &= (\tilde{\tau}_1 + \tilde{\tau}_2\gamma)[\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1))] + \tilde{\tau}_2(1-\gamma)[\hat{U}(\phi_1(\boldsymbol{\tau''}, q')) - c(\tilde{t}', \overline{m})] \\
&\geq \tilde{\tau}_1[\hat{U}(\tilde{t}_1) - c(\tilde{t}_1, m_c(\tilde{t}_1))] + \tilde{\tau}_2\gamma[\hat{U}(\tilde{t}_2) - c(\tilde{t}_1, \overline{m})] + \tilde{\tau}_2(1-\gamma)[\hat{U}(\tilde{t}_2) - c(\tilde{t}', \overline{m})] \\
&> \mathcal{G}(\mu).
\end{aligned}
$$

□

Proof of Lemma 3

Proof. For any experiment $\boldsymbol{\tau} = (t_1, t_2; \tau_1, \tau_2)$, we consider the following three cases for different given value of t_1.

1. If $\hat{U}(t_1) \leq \hat{U}(\mu) - c(t_1, \overline{m})$, $\boldsymbol{\tau}$ induces a total-pooling equilibrium, in which both types sends $\overline{m}$. Since the expected payoff from $\boldsymbol{\tau}$ is $E(\boldsymbol{\tau}) = \hat{U}(\mu) - [\tau_1 c(t_1, \overline{m}) + \tau_2 c(t_2, \overline{m})]$ and $c(t, \overline{m})$ is concave in t, then $E(\boldsymbol{\tau})$ is increasing with t_2.

2. If $\hat{U}(t_1) \geq \hat{U}(t_H) - c(t_1, \overline{m})$, $\boldsymbol{\tau}$ induces a separating equilibrium.

   (a) When for any $t_1 < \mu < t_2$, $\hat{U}(t_1) \leq \hat{U}(t_2) - c(t_1, m_c(t_2))$ holds, we have

   $$\hat{U}(t_1) = \hat{U}(t_2) - c(t_1, m_2(\boldsymbol{\tau})). \tag{A.2}$$

   Then,

   $$
   \begin{aligned}
   E(\boldsymbol{\tau}) &= \hat{U}(t_1) + \frac{[\hat{U}(t_2) - c(t_2, m_2(\boldsymbol{\tau}))] - \hat{U}(t_1)}{t_2 - t_1} \cdot (\mu - t_1) \\
   &= \hat{U}(t_1) + \frac{c(t_1, m_2(\boldsymbol{\tau})) - c(t_2, m_2(\boldsymbol{\tau}))}{t_2 - t_1} \cdot (\mu - t_1)
   \end{aligned}
   $$

   and

   $$
   \begin{aligned}
   \frac{\partial E(\boldsymbol{\tau})}{\partial t_2} =& \frac{(\mu - t_1)}{(t_2 - t_1)^2}\Big[\Big(\frac{\partial c(t_1, m_2(\boldsymbol{\tau}))}{\partial m}\frac{\partial m}{\partial t_2} - \frac{\partial c(t_2, m_2(\boldsymbol{\tau}))}{\partial t_2} - \frac{\partial c(t_2, m_2(\boldsymbol{\tau}))}{\partial m}\frac{\partial m}{\partial t_2}\Big)(t_2 - t_1) \\
   & - c(t_1, m_2(\boldsymbol{\tau})) + c(t_2, m_2(\boldsymbol{\tau}))\Big] \\
   =& \frac{(\mu - t_1)}{(t_2 - t_1)^2}\Big[\Big(\frac{\partial c(t_1, m_2(\boldsymbol{\tau}))}{\partial m} - \frac{\partial c(t_2, m_2(\boldsymbol{\tau}))}{\partial m}\Big)\frac{\partial m}{\partial t_2}(t_2 - t_1) \\
   & + c(t_2, m_2(\boldsymbol{\tau})) - \frac{\partial c(t_2, m_2(\boldsymbol{\tau}))}{\partial t_2}(t_2 - t_1) - c(t_1, m_2(\boldsymbol{\tau}))\Big].
   \end{aligned}
   $$

   Since $\frac{\partial^2 c}{\partial m \partial t} < 0$, $\frac{\partial c(t_1, m_2(\boldsymbol{\tau}))}{\partial m} > \frac{\partial c(t_2, m_2(\boldsymbol{\tau}))}{\partial m}$. Also, since c is concave in t, then

   $$c(t_2, m_2(\boldsymbol{\tau})) - \frac{\partial c(t_2, m_2(\boldsymbol{\tau}))}{\partial t_2}(t_2 - t_1) > c(t_1, m_2(\boldsymbol{\tau})).$$

   By the implicit function theorem, from equation (A.2), we have

   $$\frac{\partial \hat{U}(t_2)}{\partial t_2} - \frac{\partial c(t_1, m_2(\boldsymbol{\tau}))}{\partial m}\frac{\partial m}{\partial t_2} = 0,$$

so

$$\frac{\partial m}{\partial t_2} = \frac{\hat{U}'(t_2)}{\frac{\partial c(t_1, m_2(\boldsymbol{\tau}))}{\partial m}} > 0.$$

Therefore, $\frac{\partial E(\boldsymbol{\tau})}{\partial t_2} > 0$, which means the expected payoff from $\boldsymbol{\tau}$ increases with t_2, for any given t_1.

(b) When the condition that $\hat{U}$ is convex in t is satisfied, the proof is as follows.
Given any $t_1 < \mu$,
if $\hat{U}(t_1) \geq \hat{U}(t_2) - c(t_1, m_c(t_2))$ for t_2 in any convex set (e.g., $t_2 \in [\underline{t}, \tilde{t}]$), the two types separate with their respective costless messages and thus, $E(\boldsymbol{\tau})$ increases with t_2;
if $\hat{U}(t_1) \leq \hat{U}(t_2) - c(t_1, m_c(t_2))$ for t_2 in any convex set, by the analysis in (a) above, $E(\boldsymbol{\tau})$ increases with t_2.
Therefore, whether t_2 can separate with her cost-minimizing message, given t_1, the experiment $\boldsymbol{\tau}$ with higher $t_2 \in (\mu, t_H]$ has higher expected payoff.

3. If $\hat{U}(\mu) - c(t_1, \overline{m}) < \hat{U}(t_1) < \hat{U}(t_H) - c(t_1, \overline{m})$, the proof is as follows.
If $\boldsymbol{\tau}$ induces a partial-pooling experiment, $a^R(\overline{m})$ must satisfy

$$\hat{U}(t_1) = U(a^R(\overline{m})) - c(t_1, \overline{m}).$$

Then, there exists $\tilde{t} \in (\mu, t_H)$ s.t. $\alpha^R(\tilde{t}) = a^R(\overline{m})$.

(a) When $t_2 \leq \tilde{t}$, $\hat{U}(t_1) \geq \hat{U}(\tilde{t}) - c(t_1, \overline{m})$. $\boldsymbol{\tau}$ induces a separating equilibrium. Based on the analysis in case 2, $E(\boldsymbol{\tau})$ increases with t_2 and when $t_2 = \tilde{t}$,

$$E(\boldsymbol{\tau}) = \hat{U}(t_1) + \frac{[\hat{U}(\tilde{t}) - c(\tilde{t}, \overline{m})] - \hat{U}(t_1)}{\tilde{t} - t_1}(\mu - t_1).$$

(b) When $t_2 > \tilde{t}$, $\hat{U}(t_1) < \hat{U}(t_2) - c(t_1, \overline{m})$. $\boldsymbol{\tau}$ induces a partial-pooling equilibrium in which t_1 sends $\overline{m}$ with probability q. Given t_1,

$$E(\boldsymbol{\tau}) = \hat{U}(t_1) + (\mu - t_1) \cdot \frac{1}{\tilde{t} - t_1} \cdot \Big[\hat{U}(\tilde{t}) - \frac{\tau_1 q c(t_1, \overline{m}) + \tau_2 c(t_2, \overline{m})}{\tau_1 q + \tau_2} - \hat{U}(t_1)\Big],$$

where $\frac{\tau_1 q t_1 + \tau_2 t_2}{\tau_1 q + \tau_2} = \tilde{t}$ for any $\boldsymbol{\tau}$. The expected cost from pooling $\frac{\tau_1 q c(t_1, \overline{m}) + \tau_2 c(t_2, \overline{m})}{\tau_1 q + \tau_2}$ decreases with t_2 and is lower than $c(\tilde{t}, \overline{m})$.

Thus, $E(\boldsymbol{\tau})$ increases with t_2.

□

Proof of Proposition 4

Proof. c is concave in t (HD condition) indicates a pooling experiment has two results if it is optimal. Then, the optimal experiment has two results or one result. In this proof, we first consider the optimal experiment that consists of two results, and then, compare that with the uninformative experiment $(\mu; 1)$.

Consider any experiment containing two results, denoted by $\boldsymbol{\tau} = (t_1, t_2; \tau_1, \tau_2)$. According to the condition in Proposition 4,

$$\hat{U}(t_1) + c(t_1, \overline{m}) > \hat{U}(\mu) + c(\mu, \overline{m}) > \hat{U}(\mu), \ \forall \ t_1 < \mu.$$

Therefore, there is no experiment with two results that induces a total-pooling equilibrium. We then have the following three cases.

1. If $\boldsymbol{\tau}$ induces a separating equilibrium in which the two types sends their respective costless message, according to the second part of the HD condition, $\hat{U}$ is convex. Thus, given any t_1, $E(\boldsymbol{\tau})$ decreases with t_1.

2. If $\boldsymbol{\tau}$ induces a separating equilibrium and

$$\hat{U}(t_1) = \hat{U}(t_2) - c(t_1, m_2(\boldsymbol{\tau})), \tag{A.3}$$

the expected payoff of $\boldsymbol{\tau}$ is

$$\begin{aligned} E(\boldsymbol{\tau}) &= \hat{U}(t_1) + [\hat{U}(t_2) - c(t_2, m_2(\boldsymbol{\tau})) - \hat{U}(t_1)] \cdot \frac{\mu - t_1}{t_2 - t_1} \\ &= \hat{U}(t_1) + [c(t_1, m_2(\boldsymbol{\tau})) - c(t_2, m_2(\boldsymbol{\tau}))] \cdot \frac{\mu - t_1}{t_2 - t_1}. \end{aligned}$$

Then,

$$\begin{aligned} \frac{\partial E(\boldsymbol{\tau})}{\partial t_1} =& \hat{U}'(t_1) + [c_t(t_1, m_2(\boldsymbol{\tau})) + c_m(t_1, m_2(\boldsymbol{\tau}))\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1} - c_m(t_2, m_2(\boldsymbol{\tau}))\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1}]\frac{\mu - t_1}{t_2 - t_1} \\ &+ \frac{-t_2 + t_1 + \mu - t_1}{(t_2 - t_1)^2}[c(t_1, m_2(\boldsymbol{\tau})) - c(t_2, m_2(\boldsymbol{\tau}))]. \end{aligned}$$

By equation (A.3),

$$\hat{U}'(t_1) = -c_t(t_1, m_2(\boldsymbol{\tau})) - c_m(t_1, m_2(\boldsymbol{\tau}))\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1},$$

implying

$$\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1} = -\frac{\hat{U}'(t_1) + c_t(t_1, m_2(\boldsymbol{\tau}))}{c_m(t_1, m_2(\boldsymbol{\tau}))}.$$

Since $m_2(\boldsymbol{\tau}) \geq m_c(t_2)$ and $m_2(\boldsymbol{\tau}) \geq m_1(\boldsymbol{\tau}) = m_c(t_1)$, $c_m(t_1, m_2(\boldsymbol{\tau})) > c_m(t_2, m_2(\boldsymbol{\tau})) \geq 0$. Hence, $\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1} > 0$. Substitute the value of $\hat{U}'(t_1)$, and then,

$$\begin{aligned} \frac{\partial E(\boldsymbol{\tau})}{\partial t_1} &= -\frac{t_2 - \mu}{t_2 - t_1}[c_t(t_1, m_2(\boldsymbol{\tau})) + c_m(t_1, m_2(\boldsymbol{\tau}))\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1}] - c_m(t_2, m_2(\boldsymbol{\tau}))\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1}\frac{\mu - t_1}{t_2 - t_1} \\ &\quad - \frac{t_2 - \mu}{(t_2 - t_1)^2}[c(t_1, m_2(\boldsymbol{\tau})) - c(t_2, m_2(\boldsymbol{\tau}))] \\ &= -\frac{t_2 - \mu}{(t_2 - t_1)^2}[c(t_1, m_2(\boldsymbol{\tau})) + c_t(t_1, m_2(\boldsymbol{\tau})) \cdot (t_2 - t_1) - c(t_2, m_2(\boldsymbol{\tau}))] \\ &\quad - \frac{t_2 - \mu}{t_2 - t_1}c_m(t_1, m_2(\boldsymbol{\tau}))\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1} - \frac{\mu - t_1}{t_2 - t_1}c_m(t_2, m_2(\boldsymbol{\tau}))\frac{\partial m_2(\boldsymbol{\tau})}{\partial t_1}. \end{aligned} \tag{A.4}$$

Since $c(t, \overline{m})$ is concave in t,

$$c(t_1, m_2(\boldsymbol{\tau})) + c_t(t_1, m_2(\boldsymbol{\tau})) \cdot (t_2 - t_1) - c(t_2, m_2(\boldsymbol{\tau})) > 0.$$

Thus, $\frac{\partial E(\boldsymbol{\tau})}{\partial t_1} < 0$.

3. If $\boldsymbol{\tau}$ induces a partial-pooling equilibrium, $\hat{U}(t_1) = U(a^R(\overline{m})) - c(t_1, \overline{m})$. Then,

$$\begin{aligned} E(\boldsymbol{\tau}) =& \hat{U}(t_1) + [U(a^R(\overline{m})) - c(t_2, \overline{m}) - \hat{U}(t_1)]\frac{\mu - t_1}{t_2 - t_1} \\ =& \hat{U}(t_1) + [c(t_1, \overline{m}) - c(t_2, \overline{m})]\frac{\mu - t_1}{t_2 - t_1} \end{aligned}$$

and

$$\begin{aligned}\frac{\partial E(\boldsymbol{\tau})}{\partial t_1} =&\hat{U}'(t_1) + c_t(t_1, \overline{m}) \cdot \frac{\mu - t_1}{t_2 - t_1} + \frac{-t_2 + t_1 + \mu - t_1}{(t_2 - t_1)^2}[c(t_1, \overline{m}) - c(t_2, \overline{m})] \\ =&\hat{U}'(t_1) + c_t(t_1, \overline{m}) - \frac{t_2 - \mu}{(t_2 - t_1)^2}[c(t_1, \overline{m}) - c(t_2, \overline{m}) + c_t(t_1, \overline{m}) \cdot (t_2 - t_1)].\end{aligned} \tag{A.5}$$

Since $c(t, \overline{m})$ is concave in t, $c(t_1, \overline{m}) + c_t(t_1, \overline{m}) \cdot (t_2 - t_1) - c(t_2, \overline{m}) > 0$. Thus, $\frac{\partial E(\boldsymbol{\tau})}{\partial t_1} < 0$.

Therefore, for all experiments with two results $\boldsymbol{\tau} = (t_1, t_2; \tau_1, \tau_2)$, given any t_2, the sender's expected payoff from $\boldsymbol{\tau}$, $E(\boldsymbol{\tau})$, decreases with t_1 so that $E(\boldsymbol{\tau})$ is maximized when $t_1 = t_L$. By Lemma 3, given any t_1, $E(\boldsymbol{\tau})$ increases with t_2 so that $E(\boldsymbol{\tau})$ is maximized when $t_2 = t_H$.

Moreover, the sender's expected payoff from the uninformative experiment $(\mu; 1)$ is $\hat{U}(\mu)$. The uninformative experiment can be considered as $(t_1 = \mu, t_H; 1, 0)$. Since given t_2, $E(\boldsymbol{\tau})$ is decreasing in $t_1 \in [0, \mu]$, no matter if $\boldsymbol{\tau}$ induces a separating equilibrium or partial-pooling equilibrium, then the uninformative experiment cannot be optimal. In conclusion, the sender must choose the fully informative experiment $(t_L, t_H; 1 - \mu, \mu)$. □

Proof of Proposition 5

Proof. The first statement is obvious. To prove the second statement, we show when $\alpha^R(\mu) = \underline{a}$, the fully informative experiment $\bar{\boldsymbol{\tau}} = (t_L, t_H; 1 - \mu, \mu)$ is strictly better than the uninformative experiment $\boldsymbol{\tau_0}$. By Lemma A.1, α^R is weakly increasing, so $\hat{U}(t_L) = \hat{U}(\mu) = U(\underline{a})$. (Based on our model setting in Section 2, $\hat{U}(t_H) > U(\underline{a})$.) Let us consider the equilibrium induced by $\bar{\boldsymbol{\tau}}$ as follows.

1. When $\hat{U}(t_L) \geq \hat{U}(t_H) - c(t_L, \overline{m})$, $\bar{\boldsymbol{\tau}}$ induces a separating equilibrium.
   - If $\hat{U}(t_L) \geq \hat{U}(t_H) - c(t_L, m_c(t_H))$, t_L sends $m_c(t_L)$ and t_H sends $m_c(t_H)$. Thus, $E(\bar{\boldsymbol{\tau}}) > E(\boldsymbol{\tau_0})$.
   - Otherwise, $\hat{U}(t_L) = \hat{U}(t_H) - c(t_L, m_2(\bar{\boldsymbol{\tau}}))$. By Lemma A.3, $m_2(\bar{\boldsymbol{\tau}}) > m_c(t_H)$. Since t_H obtains a payoff
   $$\hat{U}(t_H) - c(t_H, m_2(\bar{\boldsymbol{\tau}})) > \hat{U}(t_H) - c(t_L, m_2(\bar{\boldsymbol{\tau}})) = \hat{U}(t_L),$$
   we have $E(\bar{\boldsymbol{\tau}}) > E(\boldsymbol{\tau_0})$.
2. When $\hat{U}(t_L) < \hat{U}(t_H) - c(t_L, \overline{m})$. Since $\hat{U}(t_L) > \hat{U}(\mu) - c(t_L, \overline{m})$, $\bar{\boldsymbol{\tau}}$ induces a partial-pooling equilibrium. Type t_L sends $\overline{m}$ with probability q, where q satisfies
   $$\hat{U}(\phi_1(\bar{\boldsymbol{\tau}}, q)) - c(t_L, \overline{m}) = \hat{U}(t_L).$$
   Then, type t_H obtains a payoff
   $$\hat{U}(\phi_1(\bar{\boldsymbol{\tau}}, q)) - c(t_H, \overline{m}) > \hat{U}(\phi_1(\bar{\boldsymbol{\tau}}, q)) - c(t_L, \overline{m}) = \hat{U}(t_L),$$
   so $E(\bar{\boldsymbol{\tau}}) > E(\boldsymbol{\tau_0})$.

□

Proof of Corollary 4

Proof. By the first condition, the optimal experiment is the uninformative experiment or a separating experiment that consists of two results.

For any experiment with two results $\boldsymbol{\tau} = (t_1, t_2; \tau_1, \tau_2)$, if it induces a separating equilibrium, by equation (A.4) and the conditions in Corollary 4, $\frac{\partial E(\boldsymbol{\tau})}{\partial t_1} > 0$. If $\boldsymbol{\tau}$ induces a partial-pooling equilibrium, by equation (A.5) and the conditions in Corollary 4, $\frac{\partial E(\boldsymbol{\tau})}{\partial t_1} > 0$. If $\boldsymbol{\tau}$ induces a total-pooling equilibrium, $E(\boldsymbol{\tau})$ increases with t_1 because c is strictly convex in t. Then, given any t_2, $E(\boldsymbol{\tau})$ always increases with t_1. Since $\lim_{t_1 \to \mu} E(\boldsymbol{\tau}) \leq \hat{U}(\mu)$, the sender chooses the uninformative experiment. □

Proof of Corollary 7

Proof. Since $c = K(m-t)^2$ is convex in t, by Proposition 1, the optimal experiment is separating, which either is uninformative or contains two results. Based on this, we first consider the sender's expected payoff from any separating experiment with two results, and then, figure out under which condition the sender chooses $\boldsymbol{\tau_0}$. Lastly, we derive how the sender's equilibrium payoff is affected by the cost intensity K.

First, let us consider any separating experiment with two results, denoted by $\boldsymbol{\tau} = (t_1, t_2; \tau_1, \tau_2)$. If $\underline{a} \leq t_1 < \mu$, we have $E(\boldsymbol{\tau}) \leq E(\boldsymbol{\tau_0})$ because the linear utility function implies $\boldsymbol{\tau_0}$ is optimal. If $t_1 < \underline{a}$, given any t_2, $t_1 = t_L$ is optimal. We then consider the optimal value of t_2.

1. When $\hat{U}(t_L) \geq \hat{U}(t_H) - c(t_L, \overline{m})$, that is, $K \geq 1 - \underline{a}$, any $\boldsymbol{\tau} = (t_L, t_2; \tau_1, \tau_2)$ induces a separating equilibrium. The sender chooses the fully informative experiment because it lets the sender obtain the highest expected utility without incurring any reporting cost.

2. When $\hat{U}(t_L) \leq \hat{U}(\mu) - c(t_L, \overline{m})$, that is, $K \leq \mu - \underline{a}$, any $\boldsymbol{\tau} = (t_L, t_2; \tau_1, \tau_2)$ cannot induce a separating equilibrium. Hence, the sender chooses $\boldsymbol{\tau_0}$.

3. When $\hat{U}(\mu) - c(t_L, \overline{m}) < \hat{U}(t_L) < \hat{U}(t_H) - c(t_L, \overline{m})$, that is, $\mu - \underline{a} < K < 1 - \underline{a}$. Since $\boldsymbol{\tau} = (t_L, t_2; \tau_1, \tau_2)$ induces a separating equilibrium,

$$\hat{U}(t_L) \geq \hat{U}(t_2) - c(t_L, \overline{m}),$$
$$t_2 \leq K + \underline{a}.$$

If $\hat{U}(t_L) \geq \hat{U}(t_2) - c(t_L, m_c(t_2))$, $E(\boldsymbol{\tau})$ increases with t_2. If

$$\hat{U}(t_L) = \hat{U}(t_2) - c(t_L, m_2(\boldsymbol{\tau})),$$
$$\underline{a} = t_2 - K(m_2(\boldsymbol{\tau}) - t_L)^2,$$

we have

$$m_2(\boldsymbol{\tau}) = \sqrt{\frac{t_2 - \underline{a}}{K}}.$$

Then,

$$\begin{aligned} E(\boldsymbol{\tau}) &= \hat{U}(t_L) + [\hat{U}(t_2) - c(t_2, m_2(\boldsymbol{\tau})) - \hat{U}(t_L)]\frac{\mu - t_L}{t_2 - t_L} \\ &= \underline{a} + [t_2 - K(\sqrt{\frac{t_2 - \underline{a}}{K}} - t_2)^2 - \underline{a}]\frac{\mu}{t_2} \\ &= \underline{a} + (-Kt_2 + 2\sqrt{K}\sqrt{t_2 - \underline{a}})\mu \end{aligned}$$

and

$$\frac{\partial E(\boldsymbol{\tau})}{\partial t_2} = \Big(-K + \sqrt{\frac{K}{t_2 - \underline{a}}}\Big)\mu.$$

Since $t_2 \le K + \underline{a}$, we have $\frac{\partial E(\boldsymbol{\tau})}{\partial t_2} > 0$. Therefore, $t_2 = K + \underline{a}$.

Second, we consider when $\boldsymbol{\tau_0}$ is optimal. Based on the analysis above, when $K \ge 1 - \underline{a}$, the fully informative experiment is optimal. When $K \le \mu - \underline{a}$, $\boldsymbol{\tau_0}$ is optimal. When $\mu - \underline{a} < K < 1 - \underline{a}$, the optimal separating experiment that consists of two results is $\tilde{\boldsymbol{\tau}} = (t_L, \tilde{t}_2; \tau_1, \tau_2)$, where $\tilde{t}_2 = K + \underline{a} < 1$. Since

$$\hat{U}(\tilde{t}_2) - c(t_L, m_c(\tilde{t}_2)) = K + \underline{a} - K(K + \underline{a})^2 > \underline{a} = \hat{U}(t_L),$$

$\tilde{t}_2$ separates with $m_2(\tilde{\boldsymbol{\tau}}) = 1$. Then,

$$\begin{aligned} E(\tilde{\boldsymbol{\tau}}) &= \hat{U}(t_L) + [\hat{U}(\tilde{t}_2) - c(\tilde{t}_2, m_2(\tilde{\boldsymbol{\tau}})) - \hat{U}(t_L)]\frac{\mu - t_L}{\tilde{t}_2 - t_L} \\ &= \underline{a} + [K + \underline{a} - K(1 - K - \underline{a})^2 - \underline{a}]\frac{\mu}{K + \underline{a}} \\ &= -\mu K^2 + \mu(2 - a)K + \underline{a}. \end{aligned}$$

If $\boldsymbol{\tau_0}$ is optimal, we have $E(\tilde{\boldsymbol{\tau}}) \le \hat{U}(\mu) = \mu$, implying $K \le 1 - \frac{1}{2}\underline{a} - \sqrt{\frac{1}{4}\underline{a}^2 - \underline{a} + \frac{\underline{a}}{\mu}}$. Therefore, $\tilde{\boldsymbol{\tau}}$ is optimal if $1 - \frac{1}{2}\underline{a} - \sqrt{\frac{1}{4}\underline{a}^2 - \underline{a} + \frac{\underline{a}}{\mu}} < K \le 1 - \underline{a}$, and $\boldsymbol{\tau_0}$ is optimal if $K \le 1 - \frac{1}{2}\underline{a} - \sqrt{\frac{1}{4}\underline{a}^2 - \underline{a} + \frac{\underline{a}}{\mu}}$ because $\mu - \underline{a} < 1 - \frac{1}{2}\underline{a} - \sqrt{\frac{1}{4}\underline{a}^2 - \underline{a} + \frac{\underline{a}}{\mu}}$ for any $\mu \in (\underline{a}, 1)$.

Lastly, we prove the sender's equilibrium payoff weakly increases with K. As K increases, the sender's equilibrium payoff changes from μ to $E(\tilde{\boldsymbol{\tau}})$ and then $\underline{a} + (1 - \underline{a})\mu$, the expected payoff from the fully informative experiment. When $K \in (1 - \frac{1}{2}\underline{a} - \sqrt{\frac{1}{4}\underline{a}^2 - \underline{a} + \frac{\underline{a}}{\mu}}, 1 - \underline{a})$, the sender's equilibrium payoff $E(\tilde{\boldsymbol{\tau}})$ increases with K because $\frac{\partial E(\tilde{\boldsymbol{\tau}})}{\partial K} = \mu(2 - \underline{a} - 2K) > 0$. Therefore, the sender's equilibrium payoff weakly increases with K. □

Appendix B. Existence of the Optimal Experiment

Built on Lemma 1, we have the following proposition.

Proposition B.1. *The optimal experiment exists.*

The proof is built on analysis and definitions (notations) in Section 3. Besides, we adopt the following notations.

- The set of the experiments that consist of n results is denoted by $X_n = \{(t_1, \dots, t_n; \tau_1, \dots, \tau_n) \in [0,1]^n \times (0,1]^n \mid t_1 < \dots < t_n, \sum_{i=1}^n t_i\tau_i = \mu, \sum_{i=1}^n \tau_i = 1\}$.
- The sender's expected payoff of experiments $\boldsymbol{\tau} \in X_n$ is a function $E_n(\boldsymbol{\tau}) : X_n \to \mathbb{R}$.

Next, we prove the existence through three steps:

Step 1: For any finite $n \in \mathbb{N}$, the sender's expected payoff $E_n(\boldsymbol{\tau})$ is continuous in experiments $\boldsymbol{\tau} \in X_n$.

Step 2: For any finite $N \in \mathbb{N}$, the sender's expected payoff is upper semi-continuous in experiments $\boldsymbol{\tau} \in \bigcup_{n \le N} X_n$.

Step 3: The optimal experiment needs a finite number of results.

Step 1. Prove that given any finite n, $E_n(\boldsymbol{\tau})$ is continuous in experiments $\boldsymbol{\tau} \in X_n$.

The proof is via the following claims.

Claim 1: *$m_c(t) \equiv \arg\min_{m\in[0,\overline{m}]} c(t,m)$ is a continuous function of $t \in [0,1]$.*

Proof. Suppose $m_c(\cdot)$ is discontinuous at $\hat{t} \in [0,1]$. If $m_c(\cdot)$ is not right continuous at $\hat{t}$, there exists $\epsilon > 0$ such that for all $\delta > 0$, $\exists\ t(\delta)$ s.t. $0 < t(\delta) - \hat{t} < \delta$ and $|m_c(t(\delta)) - m_c(\hat{t})| \geq \epsilon$. Since $m_c(\cdot)$ weakly increases in t, $m_c(t(\delta)) \geq m_c(\hat{t}) + \epsilon > m_c(\hat{t})$. Because c is strictly quasi-convex in m, we have

$$c(t(\delta), m_c(\hat{t}) + \epsilon) < c(t(\delta), m_c(\hat{t})). \tag{A.6}$$

Let $(\delta_j)_{j\in\mathbb{N}}$ be a sequence such that $\delta_j > 0$ and $\lim_{j\to+\infty} \delta_j = 0$. Then, $\lim_{j\to+\infty} t(\delta_j) = \hat{t}$ and

$$\lim_{j\to+\infty} c(t(\delta_j), m_c(\hat{t}) + \epsilon) = c(\hat{t}, m_c(\hat{t}) + \epsilon), \lim_{j\to+\infty} c(t(\delta_j), m_c(\hat{t})) = c(\hat{t}, m_c(\hat{t})).$$

Because $c(\hat{t}, m_c(\hat{t}) + \epsilon) > c(\hat{t}, m_c(\hat{t}))$, then as $j \to +\infty$, $c(t(\delta_j), m_c(\hat{t}) + \epsilon) > c(t(\delta_j), m_c(\hat{t}))$, which violates inequality (A.6). Therefore, $m_c(\cdot)$ is right continuous at any $t \in [0,1)$. Similarly, we can also prove it is left continuous at any $t \in (0,1]$. □

Claim 2: *Denote $\boldsymbol{\tau'} = (t'_1,\ldots,t'_n;\tau'_1,\ldots,\tau'_n) \in X_n$. There exists $\epsilon > 0$, such that for any experiment $\boldsymbol{\tau'}$ that satisfies $\forall\ 1 \leq i \leq n$, $|t'_i - t_i| < \epsilon$ and $|\tau'_i - \tau_i| < \epsilon$, we have the conclusion that $m_i(\boldsymbol{\tau'})$ exists if for experiment $\boldsymbol{\tau}$, $m_i(\boldsymbol{\tau}) < \overline{m}$ exists. That is, for experiment $\boldsymbol{\tau} \in X_n$, if the separating message of t_i exists, then for an "arbitrarily close" experiment $\boldsymbol{\tau'} \in X_n$, the separating message for t'_i, $m_i(\boldsymbol{\tau'})$, exists. (The separating message is defined in Section 3.)*

Proof. Suppose for experiment $\boldsymbol{\tau}$, $m_i(\boldsymbol{\tau}) < \overline{m}$ exists. For $i = 1$, $m_1(\boldsymbol{\tau'}) = m_c(t'_1)$ must exist. For $i = 2$, we have

$$\hat{U}(t_1) - c(t_1, m_c(t_1)) \geq \hat{U}(t_2) - c(t_1, m_2(\boldsymbol{\tau})) > \hat{U}(t_2) - c(t_1, \overline{m}),$$

so as $\boldsymbol{\tau'} \to \boldsymbol{\tau}$ (i.e., $\forall\ 1 \leq i \leq n$, $t'_i \to t_i$ and $\tau'_i \to \tau_i$),

$$\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) > \hat{U}(t'_2) - c(t'_1, \overline{m}).$$

Also

$$\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) < \hat{U}(t'_2) - c(t'_1, m_c(t'_1)),$$

so $\exists\ \hat{m}_2 \in (m_c(t'_1), \overline{m})$ s.t. $\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) = \hat{U}(t'_2) - c(t'_1, \hat{m}_2)$. Since $c(t'_1, m)$ strictly increases in $m \in (m_c(t'_1), \overline{m})$,

$$\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) \geq \hat{U}(t'_2) - c(t'_1, m),$$

$\forall\ m \in [\hat{m}_2, \overline{m}]$. Therefore, $m_2(\boldsymbol{\tau'}) = \arg\min_{m\in[\hat{m}_2,\overline{m}]} c(t'_2, m)$ exists. Subsequently and similarly, we can show $\forall\ i$, $m_i(\boldsymbol{\tau'})$ exists if $m_i(\boldsymbol{\tau}) < \overline{m}$ exists. □

Claim 3: *For any $\boldsymbol{\tau} \in X_n$, $\forall\ 1 \leq j \leq n$, if $m_j(\boldsymbol{\tau}) < \overline{m}$ exists, $m_j(\cdot)$ is continuous at $\boldsymbol{\tau} \in X_n$.*

Proof. Let experiment $\boldsymbol{\tau} = (t_1,\ldots,t_n;\tau_1,\ldots,\tau_n)$ induce separating messages: $m_1(\boldsymbol{\tau}),\ldots,m_k(\boldsymbol{\tau})$, $k \leq n$. Denote $\boldsymbol{\tau'} = (t'_1,\ldots,t'_n;\tau'_1,\ldots,\tau'_n)$, where $\forall\ 1 \leq i \leq n$, $t'_i \to t_i$ and $\tau'_i \to \tau_i$. That is, $\boldsymbol{\tau'} \in X_n$ and $\boldsymbol{\tau'} \to \boldsymbol{\tau}$. Suppose

$m_k(\boldsymbol{\tau}) < \overline{m}$, and then, $\forall\ j \leq k$, $m_j(\boldsymbol{\tau'})$ exists. Next we prove $\lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} m_j(\boldsymbol{\tau'}) = m_j(\boldsymbol{\tau})$. Because $m_c(\cdot)$ is a continuous function,

$$\lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} m_1(\boldsymbol{\tau'}) = \lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} m_c(t'_1) = m_c(t_1) = m_1(\boldsymbol{\tau}).$$

The analysis for $m_2(\boldsymbol{\tau})$ is as follows.

1. If $\hat{U}(t_1) - c(t_1, m_c(t_1)) > \hat{U}(t_2) - c(t_1, m_c(t_2))$, then $m_2(\boldsymbol{\tau}) = m_c(t_2)$.

   Since $\hat{U}$, c, and m_c are all continuous functions,

   $$\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) > \hat{U}(t'_2) - c(t'_1, m_c(t'_2)).$$

   Therefore, $m_2(\boldsymbol{\tau'}) = m_c(t'_2)$ and $\lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} m_2(\boldsymbol{\tau'}) = \lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} m_c(t'_2) = m_c(t_2) = m_2(\boldsymbol{\tau})$.

2. If $\hat{U}(t_1) - c(t_1, m_c(t_1)) = \hat{U}(t_2) - c(t_1, m_c(t_2))$, we also have $m_2(\boldsymbol{\tau}) = m_c(t_2)$.

   If $\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) \geq \hat{U}(t'_2) - c(t'_1, m_c(t'_2))$, we have the same argument as above. If $\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) < \hat{U}(t'_2) - c(t'_1, m_c(t'_2))$, we have $\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) = \hat{U}(t'_2) - c(t'_1, m_2(\boldsymbol{\tau'}))$. Then,

   $$\begin{aligned}\lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} \left[\hat{U}(t'_2) - c(t'_1, m_2(\boldsymbol{\tau'}))\right] &= \lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} \left[\hat{U}(t'_1) - c(t'_1, m_c(t'_1))\right] \\ &= \hat{U}(t_1) - c(t_1, m_c(t_1)) \\ &= \hat{U}(t_2) - c(t_1, m_c(t_2)),\end{aligned}$$

   implies $\lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} c(t'_1, m_2(\boldsymbol{\tau'})) = c(t_1, m_c(t_2))$.[21]

   Next we prove $\lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} m_2(\boldsymbol{\tau'}) = m_c(t_2)$. Suppose not. Then, there exists $\epsilon > 0$ such that for all $\delta > 0$, $\exists\ \boldsymbol{\tau}(\delta) = (t_1(\delta), \ldots, t_n(\delta); \tau_1(\delta), \ldots, \tau_n(\delta)) \in X_n$, s.t. $\forall\ i$, $|t_i(\delta) - t_i| < \delta$, $|\tau_i(\delta) - \tau_i| < \delta$ and $|m_2(\boldsymbol{\tau}(\delta)) - m_c(t_2)| \geq \epsilon$. Then, $m_2(\boldsymbol{\tau}(\delta)) \leq m_c(t_2) - \epsilon$ or $m_2(\boldsymbol{\tau}(\delta)) \geq m_c(t_2) + \epsilon$. Since $\lim_{\delta\to 0} m_c(t_2(\delta)) = m_c(t_2)$, then as $\delta \to 0$, $m_2(\boldsymbol{\tau}(\delta)) \geq m_c(t_2(\delta)) > m_c(t_2) - \epsilon$. Thus, we have $m_2(\boldsymbol{\tau}(\delta)) \geq m_c(t_2) + \epsilon$. Because

   $$\lim_{\delta\to 0} \left[c(t_1(\delta), m_2(\boldsymbol{\tau}(\delta))) - c(t_1(\delta), m_c(t_2) + \epsilon)\right] = c(t_1, m_c(t_2)) - c(t_1, m_c(t_2) + \epsilon) < 0,$$

   then as $\delta \to 0$, $c(t_1(\delta), m_2(\boldsymbol{\tau}(\delta))) < c(t_1(\delta), m_c(t_2) + \epsilon)$, which contradicts that c is strictly quasi-convex in m.

3. If $\hat{U}(t_1) - c(t_1, m_c(t_1)) < \hat{U}(t_2) - c(t_1, m_c(t_2))$, then $m_2(\boldsymbol{\tau}) > m_c(t_2)$ and

   $$\hat{U}(t_1) - c(t_1, m_c(t_1)) = \hat{U}(t_2) - c(t_1, m_2(\boldsymbol{\tau})).$$

   Then,

   $$\begin{aligned}\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) &< \hat{U}(t'_2) - c(t'_1, m_c(t'_2)), \\ \hat{U}(t'_1) - c(t'_1, m_c(t'_1)) &= \hat{U}(t'_2) - c(t'_1, m_2(\boldsymbol{\tau'})),\end{aligned}$$

   implying $\lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} c(t'_1, m_2(\boldsymbol{\tau'})) = c(t_1, m_2(\boldsymbol{\tau}))$. Using the proof method in the second scenario, we can prove $\lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} m_2(\boldsymbol{\tau'}) = m_2(\boldsymbol{\tau})$. Subsequently and similarly, we can prove $\forall\ 1 \leq j \leq k$, $\lim_{\boldsymbol{\tau'}\to\boldsymbol{\tau}} m_j(\boldsymbol{\tau'}) = m_j(\boldsymbol{\tau})$, so $m_j(\cdot)$ is continuous at experiment $\boldsymbol{\tau} \in X_n$.

□

[21] More precisely, the limit as $\boldsymbol{\tau'} \to \boldsymbol{\tau}$ is under the condition that $\boldsymbol{\tau'}$ satisfies $\hat{U}(t'_1) - c(t'_1, m_c(t'_1)) < \hat{U}(t'_2) - c(t'_1, m_c(t'_2))$.

Claim 4: *Given any finite n, $E_n(\boldsymbol{\tau})$ is continuous in experiments $\boldsymbol{\tau} \in X_n$.*

Proof. In this proof, we prove the continuity of E_n at $\boldsymbol{\tau}$ if $\boldsymbol{\tau}$ induces a partial-pooling equilibrium in which the threshold type takes a mixed strategy. Based on that, if $\boldsymbol{\tau}$ induces other kinds of D1 equilibria, the proof is similar and thus omitted.

Suppose experiment $\boldsymbol{\tau} = (t_1, \dots, t_n; \tau_1, \dots, \tau_n)$ induces a D1 equilibrium in which type $t_i \le t_{p-1}$ sends $m_i(\boldsymbol{\tau})$, type $t_i \ge t_{p+1}$ sends $\overline{m}$, and type t_p sends $\overline{m}$ with probability q and $m_p(\boldsymbol{\tau}) < \overline{m}$ with probability $1-q$, where $q \in (0,1)$. Then, the following IC conditions are satisfied:

$$
\begin{aligned}
&\hat{U}(t_i) - c(t_i, m_i(\boldsymbol{\tau})) > \hat{U}(\phi_{i+1}(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m}),\ i < p,\\
&\hat{U}(t_p) - c(t_p, m_p(\boldsymbol{\tau})) < \hat{U}(\phi_{p+1}(\boldsymbol{\tau}, 1)) - c(t_p, \overline{m}),\\
&\hat{U}(t_p) - c(t_p, m_p(\boldsymbol{\tau})) > \hat{U}(\phi_{p}(\boldsymbol{\tau}, 1)) - c(t_p, \overline{m}),\\
&\hat{U}(t_p) - c(t_p, m_p(\boldsymbol{\tau})) = \hat{U}(\phi_{p}(\boldsymbol{\tau}, q)) - c(t_p, \overline{m}).
\end{aligned}
$$

The sender's expected utility from experiment $\boldsymbol{\tau}$ is

$$
E_n(\boldsymbol{\tau}) = \sum_{i=1}^{p-1} \tau_i \big[\hat{U}(t_i) - c(t_i, m_i(\boldsymbol{\tau}))\big] + \sum_{i=p}^{n} \tau_i \big[\hat{U}(\phi_p(\boldsymbol{\tau}, q)) - c(t_i, \overline{m})\big].
$$

Denote $\boldsymbol{\tau'} = (t'_1, \dots, t'_n; \tau'_1, \dots, \tau'_n)$, where $\forall\ 1 \le i \le n$, $t'_i \to t_i$ and $\tau'_i \to \tau_i$. That is, $\boldsymbol{\tau'} \in X_n$ and $\boldsymbol{\tau'} \to \boldsymbol{\tau}$. According to Claim 2, we have $m_i(\boldsymbol{\tau'})$, $i \le p$, exists because $m_i(\boldsymbol{\tau}) < \overline{m}$, $i \le p$, exists. Since $\hat{U}$, c, $m_i(\cdot)$, $\phi_i(\cdot, 1)$ are all continuous in $\boldsymbol{\tau} \in X_n$, we have

$$
\begin{aligned}
&\hat{U}(t'_i) - c(t'_i, m_i(\boldsymbol{\tau'})) > \hat{U}(\phi_{p}(\boldsymbol{\tau'}, 1)) - c(t'_i, \overline{m}),\ i = 1, \dots, j,\\
&\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'})) < \hat{U}(\phi_{p+1}(\boldsymbol{\tau'}, 1)) - c(t'_p, \overline{m}),\\
&\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'})) > \hat{U}(\phi_{p}(\boldsymbol{\tau'}, 1)) - c(t'_p, \overline{m}).
\end{aligned}
$$

Then, $\exists\ q' \in (0,1)$ s.t.

$$
\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'})) = \hat{U}(\phi_p(\boldsymbol{\tau'}, q')) - c(t'_p, \overline{m}).
$$

Thus,

$$
\begin{aligned}
\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} \big[\hat{U}(\phi_p(\boldsymbol{\tau'}, q')) - c(t'_p, \overline{m})\big] &= \lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} \big[\hat{U}(t'_p) - c(t'_p, m_p(\boldsymbol{\tau'}))\big]\\
&= \hat{U}(t_p) - c(t_p, m_p(\boldsymbol{\tau}))\\
&= \hat{U}(\phi_p(\boldsymbol{\tau}, q)) - c(t_p, \overline{m}),
\end{aligned}
$$

implying $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} \hat{U}(\phi_p(\boldsymbol{\tau'}, q')) = \hat{U}(\phi_p(\boldsymbol{\tau}, q))$.

Next, we prove $q' \to q$, as $\boldsymbol{\tau'} \to \boldsymbol{\tau}$. Suppose not. Then, there exists $\epsilon > 0$ such that for all $\delta > 0$, $\exists\ \boldsymbol{\tau}(\delta) = (t_1(\delta), \dots, t_n(\delta); \tau_1(\delta), \dots, \tau_n(\delta)) \in X_n$, s.t. $\forall\ 1 \le i \le n$, $|t_i(\delta) - t_i| < \delta$, $|\tau_i(\delta) - \tau_i| < \delta$, $|q(\delta) - q| \ge \epsilon$ and $\lim_{\delta \to 0} U\big(\alpha^R(\phi_p(\boldsymbol{\tau}(\delta), q(\delta)))\big) = \hat{U}(\phi_p(\boldsymbol{\tau}, q))$. Since $\phi(\boldsymbol{\tau}, q)$ is continuous and strictly decreases in q, $\hat{U}(\phi_p(\boldsymbol{\tau}, q))$ is continuous and strictly decreases in q. Thus,

$$
\lim_{\delta \to 0} \big[\hat{U}(\phi_p(\boldsymbol{\tau}(\delta), q(\delta))) - \hat{U}(\phi_p(\boldsymbol{\tau}(\delta), q + \epsilon))\big] = \hat{U}(\phi_p(\boldsymbol{\tau}, q)) - \hat{U}(\phi_p(\boldsymbol{\tau}, q + \epsilon)) > 0,
$$

implying when $\delta \to 0$, $\hat{U}(\phi_p(\boldsymbol{\tau}(\delta), q(\delta))) > \hat{U}(\phi_p(\boldsymbol{\tau}(\delta), q + \epsilon))$. Similarly, when $\delta \to 0$, $\hat{U}(\phi_p(\boldsymbol{\tau}(\delta), q(\delta))) <$

$\hat{U}(\phi_p(\boldsymbol{\tau}(\delta), q-\epsilon))$. Then, $q-\epsilon < q(\delta) < q+\epsilon$, which leads to a contradiction. Thus, $q' \to q$, as $\boldsymbol{\tau'} \to \boldsymbol{\tau}$.

Therefore, the sender's expected utility from experiment $\boldsymbol{\tau'}$

$$E_n(\boldsymbol{\tau'}) = \sum_{i=1}^{p-1} \tau_i' \big[\hat{U}(t_i') - c(t_i', m_i(\boldsymbol{\tau'}))\big] + \sum_{i=p}^{n} \tau_i' \big[\hat{U}(\phi_p(\boldsymbol{\tau'}, q')) - c(t_i', \overline{m})\big]$$

approaches $E_n(\boldsymbol{\tau})$, i.e., $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} E_n(\boldsymbol{\tau'}) = E_n(\boldsymbol{\tau})$. Hence, $E_n(\boldsymbol{\tau})$ is continuous at $\boldsymbol{\tau} \in X_n$. Similarly, the continuity can also be proved if $\boldsymbol{\tau}$ induces other kinds of D1 equilibria. □

Step 2. *We prove if the sender chooses from the experiments that contain $n \leq N$ results, the optimal experiment exists.*

Proof. Though we have proved $E_n(\boldsymbol{\tau})$ is continuous in $\boldsymbol{\tau} \in X_n$, we cannot derive the existence of the optimal experiment directly because X_n is not a closed set, for any $n \leq N$. To construct a closed set for experiments, we define the set

$$Q_N \equiv \Big\{(t_1, \ldots, t_N; \tau_1, \ldots, \tau_N) \in [0,1]^N \times [0,1]^N \bigm| t_1 \leq \cdots \leq t_N; \sum_{i=1}^{N} t_i \tau_i = \mu; \sum_{i=1}^{N} \tau_i = 1\Big\}.$$

Q_N is closed and $X_N \subset Q_N$. We still denote each element of Q_N as $\boldsymbol{\tau}$. $\boldsymbol{\tau} \in X_N$ represents an experiment with N results, while $\boldsymbol{\tau} \in Q_N \setminus X_N$ can represent an experiment with less than N results. Then, we construct a new payoff function $F(\cdot) : Q_N \to \mathbb{R}$ as follows.

- If $\boldsymbol{\tau} \in X_N$, $F(\boldsymbol{\tau}) = E_N(\boldsymbol{\tau})$.
- If $\boldsymbol{\tau} \in Q_N \setminus X_N$, $F(\boldsymbol{\tau})$ is the expected payoff from experiment $\beta(\boldsymbol{\tau})$, where $\beta(\boldsymbol{\tau})$ is derived after deleting t_j if $\tau_j = 0$, and after combining t_j and t_{j+1} if $t_j = t_{j+1}$.

  For example, if $t_1 < \cdots < t_N$ and $\forall\ i$, $\tau_i > 0$ except $\tau_j = 0$, then $\boldsymbol{\tau}$ can be considered as experiment $\beta(\boldsymbol{\tau}) = (t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_N; \tau_1, \ldots, \tau_{j-1}, \tau_{j+1}, \ldots, \tau_N)$. If $t_1 < \cdots < t_j = t_{j+1} < \cdots < t_N$ and $\forall\ i$, $\tau_i > 0$, then $\boldsymbol{\tau}$ can be considered as experiment $\beta(\boldsymbol{\tau}) = (t_1, \ldots, t_j, t_{j+2}, \ldots, t_N; \tau_1, \ldots, \tau_{j-1}, \tau_j + \tau_{j+1}, \tau_{j+2}, \ldots, \tau_N)$.

Any $\boldsymbol{\tau} \in Q_N$ represents an experiment and any experiment in $\bigcup_{n=1}^{N} X_n$ can be represented by a $\boldsymbol{\tau} \in Q_N$. Then, the existence of the optimal experiment in $\bigcup_{n=1}^{N} X_n$ is equivalent to the existence of the maximum of $F(\boldsymbol{\tau})$, $\boldsymbol{\tau} \in Q_N$:

$$\max_{n=1,\ldots,N} \big[\max_{\boldsymbol{\tau} \in X_n} E_n(\boldsymbol{\tau})\big] \Leftrightarrow \max_{\boldsymbol{\tau} \in Q_N} F(\boldsymbol{\tau})$$

By Claim 4, $E_N(\cdot)$ is continuous in $\boldsymbol{\tau} \in X_N$. Next, we **prove $\boldsymbol{F(\cdot)}$ is upper semi-continuous in $\boldsymbol{\tau} \in \boldsymbol{Q_N}$.**

<u>*Step 1.1.*</u> *Prove: $F(\cdot)$ is upper semi-continuous at $\boldsymbol{\tau} \in Q_N \setminus X_N$, where $\boldsymbol{\tau} = (t_1, \ldots, t_N; \tau_1, \ldots, \tau_N)$ satisfies $t_1 < \cdots < t_N$, $\tau_j = 0$ and $\forall\ i \neq j$, $\tau_i > 0$.*

In this proof, though $m_j(\boldsymbol{\tau})$ is undefined, we still let $m_{j+1}(\boldsymbol{\tau})$ denote the separating message of type t_{j+1}. Denote $\boldsymbol{\tau'} = (t_1', \ldots, t_N'; \tau_1', \ldots, \tau_N')$, where $\forall\ 1 \leq i \leq N$, $t_i' \to t_i$ and $\tau_i' \to \tau_i$. That is, $\boldsymbol{\tau'} \to \boldsymbol{\tau}$, $\boldsymbol{\tau'} \in Q_N$. If $\tau_j' = 0$, the proof is obvious based on Claim 4. Next, we show the proof for $\tau_j' > 0$.

Suppose $\boldsymbol{\tau}$ induces a pure strategy partial-pooling equilibrium, and the proofs for other kinds of D1 equilibria are similar and omitted. Assume all the types smaller than t_p separates and other types pool. As $\tau_j' \to 0$, the limit of the sender's expected payoff from $\boldsymbol{\tau'}$ depends on the payoff of $t_1', \ldots, t_{j-1}', t_{j+1}', \ldots, t_N'$. By the logic in the proof of Claim 4, when $t_j > t_{p-1}$, experiment $\boldsymbol{\tau'}$ also induces $t'_1, \ldots, t'_{p-1}$ to separate and $t'_p, \ldots, t'_N$ to pool. t'_{p-1} may pool with probability $\epsilon \to 0$ and t'_p may separate with probability $\epsilon \to 0$. Thus, $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} f_{\boldsymbol{\tau'}}(t'_i) = f_{\boldsymbol{\tau}}(t_i)$, $i \neq j$, that implies $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} F(\boldsymbol{\tau'}) = F(\boldsymbol{\tau})$.

Next, we prove $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} F(\boldsymbol{\tau'}) < F(\boldsymbol{\tau})$ if $t_j < t_{p-1}$. For $\boldsymbol{\tau}$, it induces a D1 equilibrium in which type $t_i < t_p$, $i \neq j$, sends $m_i(\boldsymbol{\tau}) < \overline{m}$ and $t_i \geq t_p$ sends $\overline{m}$. Though t_j is not a type in $\boldsymbol{\tau}$ (because $\tau_j = 0$), to compare with experiment $\boldsymbol{\tau'}$, we can consider the "separating message" for t_j, denoted by $\widetilde{m}_j$, that is the least costly message for t_j to separate from any t_i, $i < j$. Because $\boldsymbol{\tau}$ induces $t_1, \ldots, t_{j-1}, t_{j+1}$ to separate,

$$\begin{aligned}\hat{U}(t_{j-1}) - c(t_{j-1}, m_{j-1}(\boldsymbol{\tau})) &\geq \hat{U}(t_{j+1}) - c(t_{j-1}, m_{j+1}(\boldsymbol{\tau})) \\ &> \hat{U}(t_j) - c(t_{j-1}, \overline{m}).\end{aligned}$$

Thus, $\widetilde{m}_j$ exists. Since $\boldsymbol{\tau'} \to \boldsymbol{\tau}$, we have $\hat{U}(t'_{j-1}) - c(t'_{j-1}, m_{j-1}(\boldsymbol{\tau'})) > \hat{U}(t'_j) - c(t'_{j-1}, \overline{m})$. Then, $m_j(\boldsymbol{\tau'})$ exists and $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} m_j(\boldsymbol{\tau'}) = \widetilde{m}_j$.

We have the following analysis of t_j.

- If $\hat{U}(t_j) - c(t_j, \widetilde{m}_j) < \hat{U}(\phi_{j+1}(\boldsymbol{\tau}, 1)) - c(t_j, \overline{m})$, we have $\hat{U}(t'_j) - c(t'_j, m_j(\boldsymbol{\tau'})) < \hat{U}(\phi_j(\boldsymbol{\tau'}, 1)) - c(t'_j, \overline{m})$ because $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} \phi_{j+1}(\boldsymbol{\tau'}, 1) = \lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} \phi_j(\boldsymbol{\tau'}, 1)$. Thus, in the D1 equilibrium induced by $\boldsymbol{\tau'}$, $t'_1, \ldots, t'_{j-1}$ separate, and $t'_j, \ldots, t'_N$ pool. Then, for $i < j$, $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} f_{\boldsymbol{\tau'}}(t'_i) = f_{\boldsymbol{\tau}}(t_i)$. For $j < i < p$, since

$$\hat{U}(t_i) - c(t_i, m_i(\boldsymbol{\tau})) \geq \hat{U}(\phi_{i+1}(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m}) > \hat{U}(\phi_{j+1}(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m}),$$

  we have $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} f_{\boldsymbol{\tau'}}(t'_i) < f_{\boldsymbol{\tau}}(t_i)$. For $i \geq p$,

$$f_{\boldsymbol{\tau}}(t_i) = \hat{U}(\phi_p(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m}) > \hat{U}(\phi_{j+1}(\boldsymbol{\tau}, 1)) - c(t_i, \overline{m}) = \lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} f_{\boldsymbol{\tau'}}(t'_i).$$

  Therefore, $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} F(\boldsymbol{\tau'}) < F(\boldsymbol{\tau})$.

- If $\hat{U}(t_j) - c(t_j, \widetilde{m}_j) \geq \hat{U}(\phi_{j+1}(\boldsymbol{\tau}, 1)) - c(t_j, \overline{m})$. We calculate the "separating message" based on $\widetilde{m}_j$ for type t_{j+1} as

$$\begin{aligned}\widetilde{m}_{j+1} &= \underset{m \in [0, \overline{m}]}{\arg\min}\, c(t_{j+1}, m) \\ \text{s.t. } &\hat{U}(t_j) - c(t_j, \widetilde{m}_j) \geq \hat{U}(t_{j+1}) - c(t_j, m), m \geq \widetilde{m}_j.\end{aligned}$$

  $\widetilde{m}_{j+1}$ exists and is the least costly message for t_{j+1} to separate from $t_1, \ldots, t_j$. Since

$$\begin{aligned}\hat{U}(t_{j-1}) - c(t_{j-1}, m_{j-1}(\boldsymbol{\tau})) &\geq \hat{U}(t_j) - c(t_{j-1}, \widetilde{m}_j) \\ &> \hat{U}(t_{j+1}) - c(t_{j-1}, \widetilde{m}_{j+1}),\end{aligned}$$

  then $\widetilde{m}_{j+1} > m_{j+1}(\boldsymbol{\tau})$ or $\widetilde{m}_{j+1} = m_{j+1}(\boldsymbol{\tau}) = m_c(t_{j+1})$. We then repeat the above analysis of t_j to have the analysis of t_{j+1}. Recursively, we have the analysis of $t_{j+2}, \ldots, t_{p-1}$. Similarly, we can prove $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} F(\boldsymbol{\tau'}) < F(\boldsymbol{\tau})$ until $\hat{U}(t_{p-1}) - c(t_{p-1}, \widetilde{m}_{p-1}) \geq \hat{U}(\phi_p(\boldsymbol{\tau}, 1)) - c(t_{p-1}, \overline{m})$, in which case, $\widetilde{m}_p \geq m_p(\boldsymbol{\tau})$ exists. Since $\hat{U}(t_p) - c(t_p, \widetilde{m}_p) \leq \hat{U}(t_p) - c(t_p, m_p(\boldsymbol{\tau})) \leq \hat{U}(\phi_p(\boldsymbol{\tau}, 1)) - c(t_p, \overline{m})$, we have $\lim_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} F(\boldsymbol{\tau'}) \leq F(\boldsymbol{\tau})$.

Step 1.2. Prove: $F(\cdot)$ is upper semi-continuous at $\boldsymbol{\tau} \in Q_N \setminus X_N$, where $\boldsymbol{\tau} = (t_1, \ldots, t_N; \tau_1, \ldots, \tau_N)$ satisfies $t_1 < \cdots < t_j = t_{j+1} < \cdots < t_N$, and $\forall\ i$, $\tau_i > 0$.

$\boldsymbol{\tau}$ can be considered as the experiment $(t_1, \ldots, t_j, t_{j+2}, \ldots, t_N; \tau_1, \ldots, \tau_{j-1}, \tau_j + \tau_{j+1}, \tau_{j+2}, \ldots, \tau_N)$. Denote $\boldsymbol{\tau'} = (t'_1, \ldots, t'_N; \tau'_1, \ldots, \tau'_N)$, where $\forall\, 1 \leq i \leq N$, $t'_i \to t_i$ and $\tau'_i \to \tau_i$. That is, $\boldsymbol{\tau'} \to \boldsymbol{\tau}$, $\boldsymbol{\tau'} \in Q_N$. When $t'_j = t'_{j+1}$, by Claim 4, the proof is obvious. When $t'_j < t'_{j+1}$, $\lim\limits_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} m_{j+1}(\boldsymbol{\tau'}) = \lim\limits_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} m_j(\boldsymbol{\tau'}) = m_j(\boldsymbol{\tau})$ if $m_j(\boldsymbol{\tau}) < \overline{m}$ exists. Therefore, according to the proof of Claim 4, we have $\lim\limits_{\boldsymbol{\tau'} \to \boldsymbol{\tau}} F(\boldsymbol{\tau'}) = F(\boldsymbol{\tau})$.

By analogy, $F(\cdot)$ is upper semi-continuous at any $\boldsymbol{\tau} \in Q_N \setminus X_N$. Thus, $F(\cdot)$ is upper semi-continuous in $\boldsymbol{\tau} \in Q_N$. Because Q_N is closed and bounded, the maximum of $F(\cdot)$ for $\boldsymbol{\tau} \in Q_N$ exists. □

Step 3: *The optimal experiment exists.*

Proof. By the proof of Proposition 1, for any experiment with $n \geq 4$ results, we can find another experiment with fewer than or equal to three results, that induces weakly higher expected payoff for the sender ex ante. Then, the sender's optimal experiment choice from the experiments that contain $n \leq 3$ results is optimal among all experiments. Therefore, based on Step 1, the optimal experiment exists.

□

Lemma B.1. *The best separating experiment $\boldsymbol{\tau^s}$ exists.*

Proof. According to Proposition 1, the best separating experiment needs at most two results, thus, based on the proof of Proposition B.1, we only need to prove the existence of the maximum of $F(\boldsymbol{\tau})$, where $\boldsymbol{\tau}$ induces a separating equilibrium and

$$\boldsymbol{\tau} \in Q_2 = \big\{(t_1, t_2; \tau_1, \tau_2) \in [0,1]^2 \times [0,1]^2 \;\big|\; t_1 \leq t_2, t_1\tau_1 + t_2\tau_2 = \mu, \tau_1 + \tau_2 = 1\big\}.$$

$\boldsymbol{\tau} \in Q_2 \setminus X_2$ represents the uninformative experiment and induces a separating equilibrium. $\boldsymbol{\tau} \in X_2$ is an experiment with two results: $t_1 < t_2$, and induces a separating equilibrium if and only if

$$\hat{U}(t_1) - c(t_1, m_c(t_1)) \geq \hat{U}(t_2) - c(t_1, \overline{m}).$$

By the implicit function theorem, there exists a unique continuous function of t_1, $l(\cdot) : [0, \mu] \to [0, +\infty)$, that satisfies

$$\hat{U}(t_1) - c(t_1, m_c(t_1)) = \hat{U}(l(t_1)) - c(t_1, \overline{m}).$$

Then, if $t_1 < t_2 \leq l(t_1)$, $\boldsymbol{\tau} \in X_2$ is a separating experiment with two results. Thus, the set of the separating experiments with at most two results can be represented as

$$\boldsymbol{\tau} \in Q_2^s = \big\{(t_1, t_2; \tau_1, \tau_2) \in [0,1]^2 \times [0,1]^2 \;\big|\; t_1 \leq t_2 \leq l(t_1), t_1\tau_1 + t_2\tau_2 = \mu, \tau_1 + \tau_2 = 1\big\}.$$

From the proof of Proposition B.1, $F(\cdot)$ is upper semi-continuous in $\boldsymbol{\tau} \in Q_2$. Q_2^s is a closed and bounded subset of Q_2, so the best separating experiment exists. □