Information Design with Costly State Verification*

Lily Ling Yang[†]

Department of Economics

University of Mannheim

April 24, 2023

Abstract

We study the optimal information design when the receiver can probabilistically verify the state at a cost. The optimal mechanism trades off between influencing the receiver's action choice and state-verification choice. The optimal mechanism involves at most three messages, and is a combination of a cutoff structure and a negative assortative structure. For the action choice, the mechanism recommends the sender-preferred action when the state is above a threshold. Moreover, above this threshold, the mechanism only recommends state verification when the state is far enough from the receiver's outside option in either direction. In contrast to Matyskova and Montes (2023), the receiver acquires information with positive probability under the optimal mechanism. Moreover, the optimal mechanism reveals more information comparing to the case where state verification is exogenous. Finally, we apply our model to a monopoly pricing problem and show that making information more accessible to consumers has a nonmonotonic effect on the retail price.

1 Introduction

Different from the cheap-talk literature, the information-design literature assumes that the informed party (the sender) has commitment power and analyzes the optimal information structure

*Financial support from the German Research Foundation (DFG) through SFB-TR 224 is gratefully acknowledged. All remaining errors are my own.

†University of Mannheim, lily.yang@uni-mannheim.de

to influence the action of the other party (the receiver). In this paper, we study an information-design problem in which the receiver can verify the state at a cost.

By incorporating costly state verification in a model of Bayesian persuasion (Kamenica and Gentzkow, 2011; Kamenica, 2019), this paper studies how the optimal information structure trades off between influencing the receiver's choice of information acquisition and action, and provides full characterization of the optimal information structure.

Environment. We analyze the optimal disclosure mechanism in a stylized Bayesian persuasion problem with endogenous information acquisition. The sender would like to sell a product to the receiver, and the receiver decides whether to buy it. The receiver only wants to buy the product if the quality of the product is above a certain level, while the sender always wants to sell the product.

The sender commits to an information structure, and sends a message to the receiver according to the committed information structure and the actual quality. After receiving the message, the receiver can decide whether to incur a cost to acquire extra information about the quality before purchasing. We consider a stylized information acquisition technology, namely, probabilistic state verification. If the receiver decides to bare the cost, then she learns the quality perfectly with some probability. But the verification can also fail. In this case, the receiver does not learn anything in addition. Finally, based on the received message and the outcome of the state verification, the receiver chooses whether to buy the product or not.

The central innovation of our model, that the receiver could acquire further information using an imperfect technology after receiving information from the sender, is a natural assumption in many applications of Bayesian persuasion. The party who provides the information often does not have control over the other party's alternative information sources. For example, while E-commerce platforms can provide information about their products through reviews and recommendations, they cannot forbid consumers to go to other sources for further information on the products. Similarly, even though firms can provide information to potential investors, investors can also conduct their own research or purchase information from other sources.

Main result. The optimal information structure combines a cutoff structure and a negative assortative structure. It recommends the receiver to buy the product if and only if the quality is above a quality threshold. For quality above the threshold, the optimal information structure recommends not to verify the quality if and only if the quality is in an interval around the receiver's outside option. If it is recommended not to buy the product, the receiver does not buy the product and does not verify the quality. If it is recommended to buy the product and not to verify, the receiver buys the product and does not verify. If it is recommended to buy the product and verify the quality, the receiver verifies the quality and buys the product only if the quality is not found out to be lower than receiver's outside option.

We show that the optimal information structure only involves three messages, and can be characterized by three cutoffs. The lowest qualities are revealed, and moderate qualities are pooled and revealed to dissuade the receiver from state verification.

Related literature. This paper contributes to the literature on Bayesian persuasion initiated by Kamenica and Gentzkow (2011). Most of the literature assumes that the sender is the only source of information. Some notable exceptions consist of Au (2015), Kolotilin et al. (2017), Kolotilin (2018), and Guo and Shmaya (2019), which all consider a receiver with private information. Gentzkow and Kamenica (2016) and Li and Norman (2018, 2019) consider multiple senders.

This paper considers Bayesian persuasion with endogenous information acquisition. The paper is closest to Matyskova and Montes (2023), Bizzotto et al. (2020), and Ederer and Min (2022). Matyskova and Montes (2023) consider an environment with general payoff functions and uniformly posterior-separable information cost. One implication of such a cost function is that the receiver has access to information of any precision level and any form, which is not true in this paper. They show that the optimal information structure can be solved as a standard Bayesian persuasion problem under a receiver-never-learn constraint. In contrast, with limited access to information, the receiver still learns under the optimal information structure in this paper. Both Bizzotto et al. (2020) and Ederer and Min (2022) consider an environment with a binary state space, while this paper considers an environment with a continuous state space. The assumption of a rich enough state space is important for our results, as the non-unitary assortative structure we find in this paper cannot arise in a model with less than three states.

These authors also consider information technologies different from this paper. In Bizzotto et al. (2020), the receiver can run a binary test on the quality, the precision of which is symmetric in both states. Ederer and Min (2022) consider a probabilistic lie detection technology instead, which indicates whether the message is a lie or not.

In terms of our assumption on receiver's information technology, this paper is related to the recent literature on communication with detectable deception. The majority of the literature considers detectable deception in a cheap-talk setting. Dziuda and Salas (2018) and Balbuzanov (2019) consider lie detection similar to Ederer and Min (2022). Sadakane and Tam (2022) and Zhao (2018) consider state verification rather than lie detection, and Zhao (2018) also endogenizes state verification. Levkun (2021) considers fact checking provided by a third party, which is equivalent to state verification in the binary-state setting he considers.

In terms of the structure of our optimal information structure, our result shares some similarities with the negative assortative information structure in Goldstein and Leitner (2018), Guo and Shmaya (2019), Terstiege and Wasser (2020), and Kolotilin et al. (2022). Goldstein and Leitner (2018) consider the design of optimal stress test, and provide examples of nonmonotone rules. Guo and Shmaya (2019) consider a Bayesian persuasion problem with a privately informed receiver and find that the optimal disclosure mechanism has a nested-interval structure. Terstiege and Wasser (2020) study the buyer-optimal information structure under monopoly pricing with the constraint that the information structure be extensionproof: the seller must have no incentive to add information. They show that an optimal extensionproof information structure takes the form of unitary single-peaked p-pairwise disclosure. Kolotilin et al. (2022) provide simple conditions for the positive and negative assortative patterns of information disclosure, and according to their terminology, our optimal information structure has a non-unitary assortative pattern. Interestingly, even though our optimal mechanism is similar to those in Goldstein and Leitner (2018) and Guo and Shmaya (2019), the reasons behind are quite different. Goldstein and Leitner (2018) obtain the nonmonotone structure because the gain-to-cost ratio, which is crucial to their optimal stress test, is not monotone in the type. Guo and Shmaya (2019) obtain the nest-interval structure because, when the private information is precise enough, high types are easier to persuade, which means that a separating mechanism is optimal. In this paper, the negative assortative pattern among the recommended types arises because it is optimal for the sender to provide extra information to dissuade the receiver from state verification.

2 Model

We consider a Bayesian persuasion game with two players, Sender (S) and Receiver (R). Sender (he) sends Receiver (she) a message that depends on a state of nature that is unobserved by Receiver. Upon receiving the message, Receiver decides whether to verify the state. Receiver then chooses an action.

Disclosure. Let θ represents the state of nature, and θ is drawn from the cumulative distribution function (CDF) F over $[0,1]$, which admits the strictly positive probability density function (PDF) f. The sender chooses an information structure. After the state of nature is drawn, the receiver observes a message generated by the chosen information structure. An information structure is combination $(M, G(.))$ of a message set M and a function $G : [0,1] \rightarrow \Delta(M)$ such that if the state is θ, then a message $m \in M$ is drawn according to distribution $G(\theta)$ and observed by the receiver. A *no-disclosure* mechanism has $G(\theta) = G(\theta')$ for all $\theta \in [0,1]$. A *truth-telling* mechanism has distributions $G(\theta)$ whose supports are disjoint across θ, so that it fully reveals the state. A *cutoff* mechanism with two intervals has a cutoff θ^* such that 1) for $\theta < \theta^*$ and $\theta' > \theta^*$, the supports of $G(\theta)$ and $G(\theta')$ are disjoint, 2) for all $\theta, \theta' < \theta^*$, $G(\theta) = G(\theta')$, and 3) $G(\theta) = G(\theta')$ for all $\theta, \theta' > \theta^*$. A *deterministic* information structure has a degenerated $G(\theta)$ for each $\theta \in [0,1]$, and can be summarized by a function $\mathbf{m} : \theta \rightarrow M$.

State verification. Receiver can verify the state θ at a cost $c > 0$, and denote the state-verification effort by $e \in \{0,1\}$. The state verification generates a signal $s \in [0,1] \cup \{\phi\}$. The expertise level $q \in (0,1]$ determines the probability that Receiver learns the state. With probability q, $s = \theta$; and with probability $1-q$, $s = \phi$. The expertise q is independent of the state, so there is no updating of Receiver's belief when state verification is not successful. We refer to the vector (q,c) as the state-verification technology, and define $C = c/q$ as the quality-adjusted verification cost. For notation simplicity, we assume that Receiver still receives a signal $s = \phi$ if she chooses not to verify the state. We assume that Receiver does not verify the state whenever she is indifferent.

Payoffs. Receiver chooses a binary action $a \in \{0,1\}$. When $a = 1$, Sender gets a payoff of 1, and Receiver gets a payoff of θ. When $a = 0$, Sender gets zero payoff, and Receiver gets a constant payoff of $R \in (0,1)$. In addition, Receiver bears the state-verification cost c if she chooses to verify the state. Sender's payoff $u_S(a)$ is independent of the state θ, and Receiver's payoff $u_R(a, e, \theta)$ is summarized in Table 1.

	$e = 1$	$e = 0$
$a = 1$	$\theta - c$	θ
$a = 0$	$R - c$	R

Table 1. Receiver's payoff

Sender's objective is to maximize the probability of Receiver choosing $a = 1$. Receiver chooses $a = 1$ if $\mathbb{E}(\theta|s, m; G) > R$, and $a = 0$ if $\mathbb{E}(\theta|s, m; G) < R$. We further assume that Receiver always chooses $a = 1$ whenever she is indifferent.

Timeline. There are four stages. First, Sender chooses an information structure (M, G), and nature draws θ according to F. Second, Sender observes state θ, and sends message m according to $G(\theta)$. Third, Receiver observes message m, and decides whether to verify the state. Fourth, Receiver observes signal s, and chooses action a. The payoffs are realized.

2.1 An example

Consider an E-commerce platform (Sender) that receives commission fees according to the sales volume, so its payoff is independent of the product quality. The customer's payoff from purchasing the product depends on the product quality θ, which is uniformly distributed on $[0, 1]$. The product price p is 0.6, so the customer's payoff is $\theta - 0.6$ if she buys the product, and 0 otherwise. The platform's revenue is 1 if the customer buys the product, and 0 if not.

The platform designs a disclosure rule about the product quality and can commit to this rule. If the customer has no private information, the platform will reveal whether the quality is above or below 0.2. When the quality is revealed to be above 0.2, the average quality is 0.6, and the customer is indifferent between buying and not buying and buys the product. When the quality is revealed to be below 0.2, the average quality is 0.1, and the customer would not buy.

When the customer can acquire extra information after the platform's disclosure, the optimal disclosure rule is no longer a simple cutoff mechanism. Suppose the customer can learn the quality θ with probability 0.9 at a cost $c = 0.01$. The optimal disclosure rule is illustrated in Figure 1.

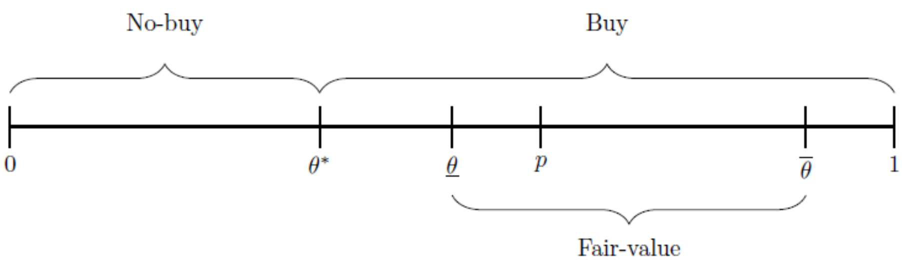

Figure 1. Optimal disclosure rule

In this numerical example, $\overline{\theta} = 0.92$, $\underline{\theta} = 0.50$ and $\theta^* = 0.34$. According to the optimal disclosure rule, the platform reveals whether the quality θ is below or above a quality threshold θ^* and whether the quality θ is close to the price 0.6. First, the platform recommends *No-buy* if quality θ is below θ^*, and the customer follows the recommendation and simply does not buy the product. The platform recommends *Buy* if θ is above θ^*. Second, on top of this No-buy/Buy recommendation, the platform further reveals whether θ is between $\underline{\theta}$ and $\overline{\theta}$, and indicates the product is *Fair-value* if so. In this case, the customer buys the product without further searching for product information. If a product is recommended as *Buy* but not indicated as *Fair-value*, the customer would further search for product information and buys the product only if she finds no bad news.

The average quality of *Fair-value* products is 0.71, while the average quality of *Buy* product is 0.67, which is still higher the price. We will show in our analysis later that a product indicated as *Fair-value* always has a higher expected value than a product indicated as *Buy* but not *Fair-value*, and more *No-Buy* recommendation is made.

3 Analysis

In this section, we first discuss the benchmark case with exogenous state verification. We then discuss the optimal information structure when state verification is endogenous.

3.1 Benchmark: exogenous state verification

When state verification is exogenous, then with probability q, Receiver is perfectly informed, and no information provided by Sender could influence her action choice a; with probability $1-q$, Receiver is uninformed, and chooses action based on the information received from Sender. Therefore, the optimal information structure is the same as when there is no private information. When there is no private information, the persuasion problem is a simple one: Sender, who wants to maximize the probability that $a = 1$, chooses an information structure and sends a message to Receiver accordingly. Receiver chooses $a = 1$ if and only if the expected value of θ given the message is at least R.

The optimal information structure is a simple cutoff mechanism with two messages, m_0 and m_1. There is a cutoff $\theta^e \in [0, R]$ such that $\mathbf{m}(\theta) = m_0$ for all $\theta < \theta^e$, and $\mathbf{m}(\theta) = m_1$ for all $\theta > \theta^e$. When Receiver has no private information, Receiver chooses $a = 0$ if $m = m_0$, Receiver chooses $a = 1$ if $m = m_1$. The cutoff θ^e is 0 if $\mathbb{E}(\theta) \geq R$ and given by

$$\mathbb{E}(\theta|\theta > \theta^e) = R$$

if $\mathbb{E}(\theta) < R$. Figure 2 illustrates the optimal information structure.

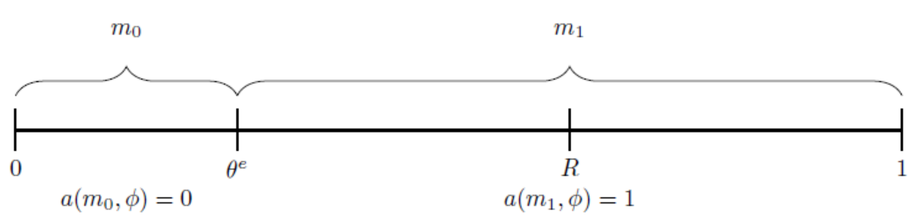

Figure 2. Optimal information structure when state verification is exogenous

When state verification is exogenous, given the perfect revealing nature of state verification, the optimal information structure does not distort disclosure to influence Receiver's belief updating. We will see in the next section that this is not true when state verification is endogenous or with other information technology (Guo and Shmaya 2019).

3.2 Preliminary analysis

The space of all information structures is the space of all functions that map the state space $[0,1]$ to the space of all distributions over the message space M. Our first result shows that we can reduce the search of the optimal information structure to information structures that involve only three messages. Denote the action choice given message m and signal s by $a(m,s)$, the state-verification choice given message m by $e(m)$.

The state-verification technology has two distinctive features: 1) when state verification is successful, Receiver receives perfect information about the state, that is, $s=\theta\in[0,1]$. The information content that Receiver receives from Sender becomes redundant. Thus, for any $m\in M$, $a(m,s)=1$ if $s\geq R$, and $a(m,s)=1$ if $s<R$. 2) When state verification is unsuccessful, Receiver receives no information, that is, $s=\phi$. Receiver's updating solely depends on information provided by Sender, which is the same as if no state verification has ever taken place. So, for any $m\in M$, $a(m,\phi)=1$ if $\mathbb{E}(\theta|m;G(\theta))\geq R$, and $a(m,\phi)=0$ if $\mathbb{E}(\theta|m;G(\theta))<R$.

Therefore, we do not need to keep track of Receiver's action when state verification is successful, but only when it is unsuccessful. Even though the optimal information structure can be very complicated, given that Receiver's action a and state-verification decision e are both binary, by Kamenica and Gentzkow (2011), any information structure is outcome-equivalent to an information structure with four types of messages: type-11, $a(m,\phi)=1$, $e(m)=1$; type-10, $a(m,\phi)=1$, $e(m)=0$; type-01, $a(m,\phi)=0$, $e(m)=1$; type-00, $a(m,\phi)=0$, $e(m)=0$. The message type fully determines Receiver's choices after receiving that message. For example, when Receiver receives a type-01 message, she verifies the state. When the state verification is successful, and the state is revealed to be above R, Receiver chooses $a=1$. When the state verification is unsuccessful, or the state is revealed to be below R, Receiver chooses $a=0$.

Now we show that type-01 messages are not a part of any optimal information structure.

Lemma 1 (No type-01) *The optimal information structure never sends message m such that $a(m,\phi)=0$, and $e(m)=1$.*

Proof. Suppose a type-01 message m is used in the optimal information structure. There must be a positive measure of $\theta \geq R$ such that $\Pr(m|\theta) > 0$. Otherwise, state verification always leads to $a = 0$ and it is suboptimal to verify the state. Therefore, we can consider another information structure that, for all θ, given which m is sent with positive probability, truthfully reveals the state with the same probability instead. For all $\theta \geq R$, $\Pr(a = 1) = q$ after m is sent in the original information structure and $\Pr(a = 1) = 1$ after θ is revealed. For all $\theta < R$, $\Pr(a = 1)$ remains unchanged. This is a strict improvement, which contradicts the optimality of the original information structure. Therefore, there cannot be any type-01 message in the optimal information structure. ■

We can further combine the other three types of messages into three messages. Denote the resulting three messages by m_ϕ, m_0, and m_1, which correspond to type-00, type-10, and type-11 messages, respectively.

Incentive compatibility implies that, for m_0,

$$\mathbb{E}(\theta|m_0) \geq R, \tag{A0}$$

$$q \Pr(\theta < R|m_0)(R - \mathbb{E}(\theta|\theta < R, m_0)) \leq c, \tag{C0}$$

and for m_1,

$$\mathbb{E}(\theta|m_1) \geq R, \tag{A1}$$

$$q \Pr(\theta < R|m_1)(R - \mathbb{E}(\theta|\theta < R, m_1)) \geq c. \tag{C1}$$

$(A0)$ and $(A1)$ imply that $a(m_0, \phi) = a(m_1, \phi) = 1$. $(C0)$ implies $e(m_0) = 0$, and $(C1)$ implies $e(m_1) = 1$. The left-hand side of $(C0)$ and $(C1)$ are the benefit of state verification. For $m \in \{m_0, m_1\}$, $a(m, \phi) = 1$. Therefore, state verification only changes Receiver's action a when $\theta < R$ and it happens with probability q. The payoff changes, when the state verification is successful, is $R - \theta$. Hence, the benefit of state verification is given by $q \Pr(\theta < R|m)(R - \mathbb{E}(\theta|\theta < R, m))$, which is increasing in $\Pr(\theta < R|m)$ and decreasing in $\mathbb{E}(\theta|\theta < R, m)$.

The presence of the constraints $(C0)$ and $(C1)$ indicate that Receiver's choice depends not only on the mean of the posterior, but also on the entire distribution of the posterior. This separates this paper from most other work on information design with a continuum of states (Kamenica and Gentzkow 2011; Dworczak and Martini 2019; Kleiner, Moldovanu and Strack 2021; Arieli et al. 2023).

Our next lemma shows that there is a cutoff θ^* below R such that, $a(m,\phi)=0$ for all $\theta<\theta^*$, and $a(m,\phi)=1$ for all $\theta\geq\theta^*$, which corresponds to the Buy/No-buy cutoff in the platform example.

Lemma 2 (Cutoff for $a(m,\phi)$) *There exists some $\theta^*<R$ such that $a(m,\phi)=0$ if and only if $\theta<\theta^*$.*

We prove Lemma 2 formally in the Appendix using calculus of variations. Here, we derive this result heuristically by focusing on deterministic information structures. Suppose first that there is some $\theta>R$, given which m_ϕ is sent in the optimal information structure. Then, consider another information structure that truthfully reveals these θ's. Since $\Pr(a=1|m_\phi)=0$, this strictly improves Sender's payoff. Therefore, the optimal information structure will only send m_ϕ when $\theta<R$.

Next, suppose that the optimal information structure sends m_ϕ for some $\theta'<R$ and m_0 for some $\theta''<\theta'$. Consider a new information structure that sends m_0 given θ' and m_ϕ given θ''. Given the new information structure, $\mathbb{E}(\theta|m_0)$ and $\mathbb{E}(\theta|\theta<R,m_0)$ are higher, while $\Pr(\theta<R|m_0)$ remains unchanged. Such a change relaxes both $(A0)$ and $(C0)$, which further enables the new information structure to send m_0 for some extra $\theta<R$ that sends m_ϕ in the original information structure. This strictly increase $\Pr(m_0|\theta<R)$, which implies $\Pr(a=1|\theta<R)$ is higher under the new information structure. This is a contradiction. The same logic applies to the message m_1.

The optimal information structure is equivalent to an information structure that truthful reveals all $\theta<\theta^*$. For low θ, the optimal information structure truthfully reveals the state and dissuades Receiver both from choosing $a=1$ and investing in the state verification. Figure 3 illustrates the cutoff structure for $a(m,\phi)$.

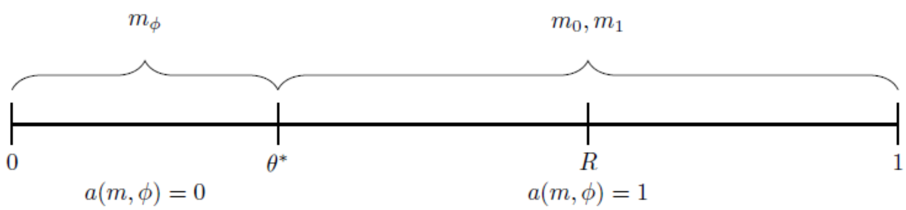

Figure 3. Cutoff structure for $a\,(m,\phi)$

Given $(A0)$ and $(A1)$, $E\,(\theta|m_0) \geq R$ and $E\,(\theta|m_1) \geq R$. Therefore, $\theta^* \geq \theta^e$. This means that the optimal information structure reveals more low states when state verification is endogenous. Moreover, the optimal information structure sends only one message above θ^e when state verification is exogenous and up to two messages above θ^* when state verification is endogenous. This means that the optimal information structure with endogenous state verification is more informative above the cutoff as well. Thus, we have

Corollary 1 *The optimal information structure reveals more information when state verification is endogenous than when state verification is exogenous.*

3.3 Optimal information design

In this section, we consider a constrained information design problem. We show that, given any θ^*, the optimal information structure has a negative assortative structure on the interval $[\theta^*, 1]$. That is, there exists two thresholds $\underline{\theta}$ and $\overline{\theta}$, such that $\mathbf{m}\,(\theta) = m_0$ for all $\theta \in \left(\underline{\theta}, \overline{\theta}\right)$, and $\mathbf{m}\,(\theta) = m_1$ for all $\theta \in (\theta^*, \underline{\theta}) \cup \left(\overline{\theta}, 1\right]$. Hence, we have the main characterization result of this paper.

Theorem 1 *Any optimal information structure is outcome-equivalent to an information structure with at most three messages that is characterized by three cutoffs θ^*, $\underline{\theta}$, and $\overline{\theta}$ such that $0 \leq \theta^* \leq \underline{\theta} \leq \overline{\theta} \leq 1$ and satisfies*

1. $$\mathbf{m}\,(\theta) = \begin{cases} m_\phi & \text{if } \theta < \theta^* \\ m_0 & \text{if } \underline{\theta} < \theta < \overline{\theta} \\ m_1 & o.w. \end{cases};$$

2. $a\,(m_\phi, \phi) = 0,\ a\,(m_0, \phi) = a\,(m_1, \phi) = 1,\ e\,(m_\phi) = e\,(m_0) = 0,$ *and* $e\,(m_1) = 1$*;*

3. for all $m \in \{m_\phi, m_0, m_1\}$*,* $a(m,s) = 1$ *if* $s = \theta > R$*, and* $a(m,s) = 0$ *if* $s = \theta < R$*.*

The optimal information structure is a combination of a cutoff structure and a negative assortative structure. For the action recommendation, the optimal information structure has a cutoff structure. Below θ^*, the recommended action is $a = 0$. For this interval, sending m_ϕ is equivalent to truthful revealing, because, for all $\theta < R$, Receiver's optimal action is $a = 0$ and $\theta^* < R$. Above θ^*, the recommended action is $a = 1$. For the state-verification recommendation, the optimal information structure has a negative assortative structure for $\theta > \theta^*$. State verification is only recommended for θ in between θ^* and $\underline{\theta}$, and in between $\overline{\theta}$ and 1. A typical three-message optimal information structure is illustrated in Figure 4.

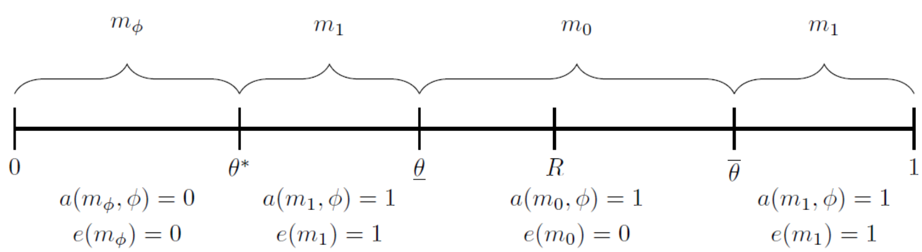

Figure 4. Three-message optimal information structure

To understand Theorem 1, we formulate Sender's choice between messages m_0 and m_1 as an infinite-dimensional maximization problem. We show that, for any feasible and incentive compatible information structure with a specific cut-off θ^*, it is not optimal if it does not have a negative assortative structure on $[\theta^*, 1]$. The proof is in Appendix, and we explain the result heuristically by focusing on deterministic information structure.

Consider the following constrained maximization problem with a fixed $\theta^* > 0$.

$$\max_{\substack{x_0(\theta), x_1(\theta) \in [0,1] \\ x_0(\theta)+x_1(\theta)=1}} \int_{\theta^*}^{R} [x_0(\theta) + x_1(\theta)(1-q)] f(\theta)\, d\theta + \int_R^1 f(\theta)\, d\theta$$

s.t.

$$\int_{\theta^*}^{1} (\theta - R) x_0(\theta) f(\theta)\, d\theta \geq 0, \quad \text{(A0)}$$

$$\int_{\theta^*}^{1} (\theta - R) x_1(\theta) f(\theta)\, d\theta \geq 0, \quad \text{(A1)}$$

$$C \int_R^1 x_0(\theta) f(\theta)\, d\theta + \int_{\theta^*}^{R} (\theta - R + C) x_0(\theta) f(\theta)\, d\theta \geq 0, \quad \text{(C0)}$$

where $x_0(\theta)$ and $x_1(\theta)$ are the probabilities that m_0 and m_1 are sent given θ, respectively. Note that we have ignored the constraint $(C1)$. This is because when $(C1)$ is violated, Receiver would not verify the state. In this case, we can simply replace the message m_1 with m_0. This new solution will be a solution of the original problem with the constraint $(C1)$.

We can see from the constraints $(A0)$-$(C1)$ the trade-off between sending m_0 and m_1: for $\theta > R$, sending either signal leads to certain acceptance and relaxes $(A0)$ or $(A1)$, but sending m_0 also relaxes $(C0)$; for $\theta < R$, sending m_1 induces investigation and a positive probability of rejection, but it only tightens $(A1)$, while sending m_0 induces no investigation, but it affects $(C0)$ in additional to tightening $(A0)$.

Moreover, note that the extent to which assigning $\theta > R$ to either signal relaxes $(A0)$ or $(A1)$ depends on the distance between θ and R, $\theta - R$. High θ brings more to the table than low θ. However, to $(C0)$, quantity matters rather than quality, that is, the extent how assigning $\theta > R$ to m_0 relaxes $(C0)$ depends only on $f(\theta)$, but not $\theta - R$.

Rewrite $(A0)$ and $(A1)$ as

$$\underbrace{\int_R^1 (\theta - R)\, x_0(\theta)\, f(\theta)\, d\theta}_{CR_0} \geq \underbrace{\int_{\theta^*}^R (R - \theta)\, x_0(\theta)\, f(\theta)\, d\theta}_{CC_0},$$

and

$$\underbrace{\int_R^1 (\theta - R)\, x_1(\theta)\, f(\theta)\, d\theta}_{CR_1} \geq \underbrace{\int_{\theta^*}^R (R - \theta)\, x_1(\theta)\, f(\theta)\, d\theta}_{CC_1}.$$

The left-hand side can be viewed as credibility resources (CR_0 and CR_1), and the right-hand side can be viewed as credibility costs (CC_0 and CC_1). Each unit of $\theta > R$ brings in $\theta - R$ units of credibility resources to either m_0 or m_1, and the effect is the same. On the other hand, each unit of $\theta < R$ incurs $\theta - R$ units of credibility cost to either m_0 or m_1, and the effect is the same as well. If the maximization problem is subject to only $(A0)$ and $(A1)$, then $x_0(\theta) = 1$ for all $\theta \geq \theta^*$. The only reason to send m_1 instead of m_0 is to satisfy $(C0)$.

Note that

$$CR_0 + CR_1 = \int_R^1 (\theta - R)\, f(\theta)\, d\theta,$$

which is a constant. Thus, the total credibility resource are fixed and an information structure simply distributes them across m_0 and m_1.

Furthermore, rewrite $(C0)$ as

$$\underbrace{C\int_R^1 x_0(\theta) f(\theta)\, d\theta}_{VC_0} \geq \underbrace{\int_{\theta^*}^R (R-\theta-C)\, x_0(\theta) f(\theta)\, d\theta}_{VB_0}.$$

The left-hand side can be viewed as the net verification cost (VC_0), and the right-hand side the net verification benefit (VB_0). State verification does not bring any benefit for $\theta > R$, because $a = 1$ regardless of the outcome of state verification. Each unit of $\theta > R$ raises C units of net verification cost, and each unit of $\theta < R$ brings $R - \theta - C$ units of net verification benefit.

To show that the optimal information structure must have a cutoff structure above R, consider the following maximization problem

$$\max_{x_0(\theta)\in[0,1]} C\int_R^1 x_0(\theta) f(\theta)\, d\theta$$

s.t.

$$\int_R^1 (\theta - R)\, x_0(\theta) f(\theta)\, d\theta = CR_0,$$

which requires the information structure to provide a fixed amount of credibility resource CR_0 while maximizing the net verification cost. The solution to this maximization problem has a cutoff structure, which is characterized by a cutoff $\overline{\theta} \in (R, 1)$ such that

$$\int_R^{\overline{\theta}} (\theta - R) f(\theta)\, d\theta = CR_0.$$

Intuitively, since high θ counts more toward the credibility resource, keeping the total credibility resource constant at CR_0, sending m_0 for low θ would increase the total probability that m_0 is sent.

For any information structure, if we replace the part of the original information structure above R with the resulting cutoff structure such that $x_0(\theta) = 1$ if $\theta \in \left(R, \overline{\theta}\right)$ and $x_1(\theta) = 1$ if $\theta \in \left(\overline{\theta}, 1\right)$. The resulting information structure is also the unique maximizer of $\int_R^1 x_0(\theta) f(\theta)\, d\theta$, which is equivalent to relaxing ($C0$). If the resulting information induces more net verification cost than the original one, then it is possible to further improve Sender's payoff by replacing some m_1 with m_0. Thus, the optimal structure must have a cutoff structure on $[R, 1]$.

Second, consider the following minimization problem

$$\min_{x_0(\theta)\in[0,1]} \int_{\theta^*}^{R} (R-\theta-C)\, x_0(\theta)\, f(\theta)\, d\theta$$

s.t.

$$\int_{\theta^*}^{R} (R-\theta)\, x_0(\theta)\, f(\theta)\, d\theta = CC_0.$$

Notice that

$$\begin{aligned} \int_{\theta^*}^{R} (R-\theta-C)\, x_0(\theta)\, f(\theta)\, d\theta &= \int_{\theta^*}^{R} (R-\theta)\, x_0(\theta)\, f(\theta)\, d\theta - C\int_{\theta^*}^{R} x_0(\theta)\, f(\theta)\, d\theta \\ &= CC_0 - C\int_{\theta^*}^{R} x_0(\theta)\, f(\theta)\, d\theta. \end{aligned}$$

The solution to this maximization problem also has a cutoff structure, which is characterized by a cutoff $\underline{\theta} \in (\theta^*, R)$ such that

$$\int_{\underline{\theta}}^{R} (R-\theta)\, f(\theta)\, d\theta = CC_0.$$

Replace the part of the original information structure below R with the resulting cutoff structure, such that $x_0(\theta) = 1$ if $\theta \in (\underline{\theta}, R)$ and $x_1(\theta) = 1$ if $\theta \in (\theta^*, \underline{\theta})$. The resulting information structure also uniquely maximizes $\int_{\theta^*}^{R} x_0(\theta)\, f(\theta)\, d\theta$, which is equivalent to maximizing $\Pr(a = 1 | \theta \in (\theta^*, R))$. Therefore, the optimal structure has a cutoff structure on $[\theta^*, R]$.

In Proposition 1, we further classify the optimal information structure into four groups.

Proposition 1 *Any optimal information structure is outcome-equivalent to one of the following mechanism:*

1. *No-disclosure mechanism: Sender reveals no information and always recommends $a = 1$, and state verification is never recommended;*

2. Cutoff mechanism: Sender reveals whether θ is above or below a cutoff and recommends $a = 1$ if and only if θ is above the cutoff, and state verification is never recommended;

3. Negative assortative mechanism: Sender reveals whether θ is in an interval around R and recommends state verification if θ is not in this interval, and $a = 1$ is always recommended;

4. Three-message mechanism.

Figure 5 illustrates all four types of optimal information structures.

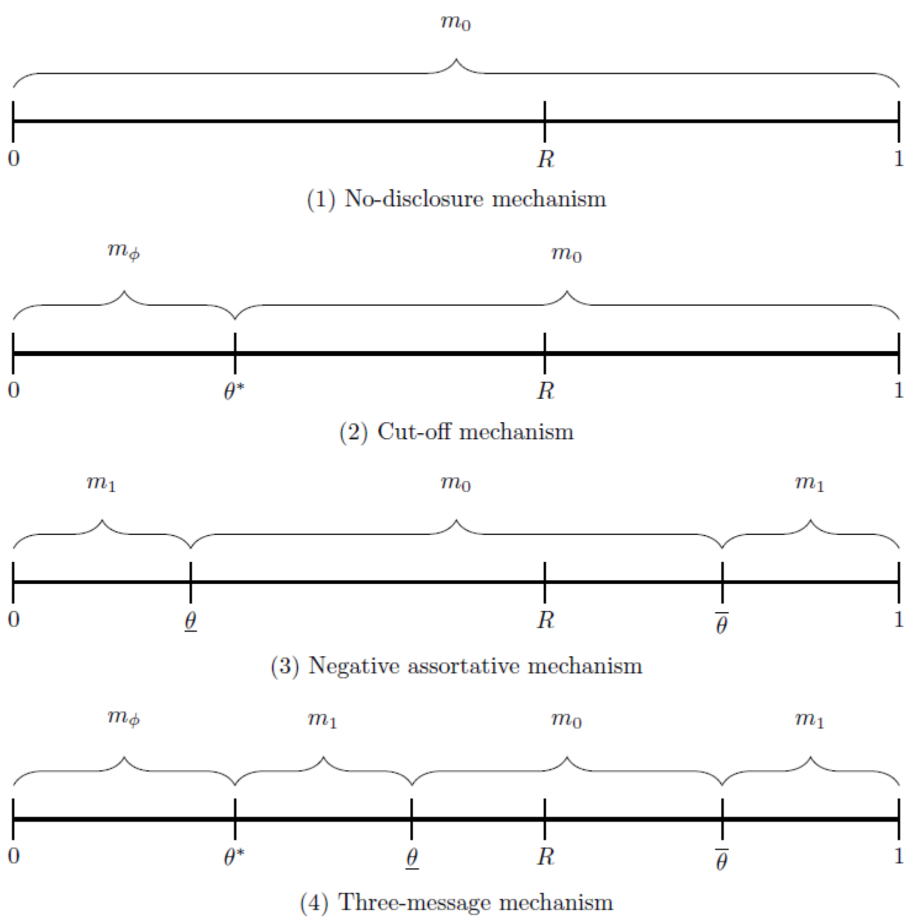

Figure 5. Classification of optimal information structures

The no-disclosure mechanism corresponds to the information structure that only sends one message. Since m_0 is sent with positive probability in any optimal information structure, then in any no-disclosure mechanism, m_0 is the only message sent by Sender. This is also the most preferred information structure by Sender, and Sender achieves the highest payoff possible when the optimal information structure is a no-disclosure mechanism. It occurs when the prior belief is optimistic and state verification is expensive.

The cutoff mechanism corresponds to the information structure that does not send m_1. Sender reveals enough information such that Receiver does not verify the state. It occurs when the prior belief is less optimistic and state verification is expensive.

The negative assortative mechanism corresponds to the information structure that does not send m_ϕ. Sender always recommends his preferred action, but it is too costly to dissuade state verification completely. It occurs when the prior belief is optimistic but state verification is less expensive.

The three-message mechanism is the one discussed in Theorem 1, and all three messages are sent. It occurs when prior belief is less optimistic and state verification is less expensive.

In the next section, we solve the optimal information structure for the environment with uniform distribution, and provide explicit conditions for each type of optimal information structures.

4 Optimal information structure under uniform distribution

In this section, we consider the uniform distribution, i.e., for all $\theta \in (0,1)$, $f(\theta) = 1$, and illustrate how our characterization can be used to solve for the optimal information structure.

We first provide the necessary and sufficient condition under which the optimal information structure is a no-disclosure mechanism. In a no-disclosure mechanism, by Proposition 1, only m_0 is sent. When Sender always sends m_0, which recommends $a = 1$ and $e = 0$, Receiver finds it optimal to follow the recommendation that $a = 1$ if and only if

$$R \leq \mathbb{E}(\theta) = \frac{1}{2}. \tag{1}$$

Furthermore, for no state verification to be optimal, the benefit of state verification must be small enough, that is, $(C0)$ must be satisfied. Therefore,

$$C \int_R^1 d\theta + \int_0^R (\theta - R + C)\, d\theta = C(1-R) + \left(C - \frac{R}{2}\right) R \geq 0,$$

which can be simplified to

$$R \leq \sqrt{2C}. \tag{2}$$

Given (1) and (2), under the no-disclosure mechanism, action $a = 1$ is chosen with probability 1, so Sender has achieved the highest possible payoff. As a result, (1) and (2) are necessary and sufficient for the optimal mechanism to be a no-disclosure mechanism.

Proposition 2 *Suppose the state is uniformly distributed. The optimal information structure is a no-disclosure mechanism if and only if*

$$R \leq \frac{1}{2} \text{ and } R \leq \sqrt{2C}.$$

Second, we provide the sufficient and necessary condition, under which the optimal information structure induces state verification, that is, m_1 is sent with positive probability. We start with a useful lemma that partially characterizes optimal information structures that recommend state verification with positive probability, which does not rely on the uniform distribution assumption.

Lemma 3 *If, in the optimal information structure, there is a positive measure of θ such that $\mathbf{m}(\theta) = m_1$, then $(C0)$ is binding.*

To see why Lemma 3 must hold, suppose $(C0)$ is not binding in the optimal information structure. If there is some positive measure of θ such that $\mathbb{E}(\theta) \geq R$ and $\mathbf{m}(\theta) = m_1$. We can always pick a subset of these θ's such that the expected value equals R and set $\mathbf{m}(\theta) = m_0$ instead of m_1. By construction, $(A0)$ and $(A1)$ still hold. Since $(C0)$ is not binding, if the measure of these θ's is small enough, $(C0)$ holds as well. This strictly improves Sender's payoff and therefore is a contradiction.

Given $\overline{\theta} \geq R$, if $(C0)$ is binding, $\underline{\theta}$ is uniquely pinned down by

$$C\left(\overline{\theta} - R\right) - \left(R - \underline{\theta}\right)\left(R - \frac{R + \underline{\theta}}{2} - C\right) = 0.$$

Denote the unique solution to the above equation that satisfies $\underline{\theta} \leq R$ by

$$\underline{\theta}_{C0}\left(\overline{\theta}\right) := R - C - \sqrt{2C\left(\overline{\theta} - R\right) + C^2}.$$

Suppose $(A0)$ is binding, $\underline{\theta}$ is uniquely pinned down by

$$\underline{\theta}_{A0}\left(\overline{\theta}\right) := 2R - \overline{\theta},$$

which is the solution to the equation

$$\left(\overline{\theta} - R\right)\left(\frac{\overline{\theta} + R}{2} - R\right) + \left(R - \theta\right)\left(\frac{R + \theta}{2} - R\right) = 0.$$

By Lemma 3, if $(C0)$ is not binding, m_1 is not sent. In that case, $\overline{\theta} = 1$ and $\underline{\theta} = \theta^*$. Suppose $(C0)$ binds. Since $(A0)$ requires that $\underline{\theta}$ is not smaller than $\underline{\theta}_{A0}\left(\overline{\theta}\right)$, we must have $\underline{\theta}_{C0}\left(\overline{\theta}\right) \geq \underline{\theta}_{A0}\left(\overline{\theta}\right)$. Moreover, $\frac{\partial}{\partial \theta}\left(\underline{\theta}_{C0}\left(\theta\right) - \underline{\theta}_{A0}\left(\theta\right)\right) > 0$, and $\underline{\theta}_{C0}\left(\theta\right) = \underline{\theta}_{A0}\left(\theta\right)$ if $\theta = R + 4C$. Therefore, $\overline{\theta} \geq R + 4C$ whenever m_1 is sent with positive probability. When $R + 4C \geq 1$, no such $\overline{\theta} < 1$ exists. As a result, the optimal information structure does not send m_1. Together with the condition $R > \frac{1}{2}$, which guarantees $\theta^* > 0$, $1 \leq R + 4C$ forms a sufficient condition for the optimality of the cutoff mechanism. When $q \leq \frac{2}{3}$, these two conditions are also necessary. Since Proposition 2 has already identified the condition for a no-disclosure mechanism, in the remaining region in the parameter space, i.e., $\sqrt{2C} < R < 1 - 4C$, the optimal information structure must recommend state verification with positive probability. Proposition 3 further identifies conditions for the optimal information structure to recommend state verification with positive probability when $q > \frac{2}{3}$.

Proposition 3 *Suppose the state is uniformly distributed. If $q \leq \frac{2}{3}$, then the optimal information structure recommends state verification with positive probability if and only if*

$$\sqrt{2C} < R < 1 - 4C.$$

If $\frac{2}{3} < q < 1$, then the optimal information structure recommends state verification with positive *probability if and only if*

$$\sqrt{2C} < R < 1 - Q(q)\,C,$$

where

$$Q(q) := \frac{4q}{3 - 2q - q^2 - (1-q)^{\frac{3}{2}}\sqrt{9-q}}.$$

If $q = 1$, then the optimal information structure never recommends state verification with positive probability.

Moreover, $Q\left(\frac{2}{3}\right) = 4$, $\lim_{q\to 1} Q(q) = \infty$, and, for all $q \in (0,1)$, $Q'(q) > 0$, which implies that as q gets larger, the set of R such that the optimal information structure recommends state verification gets smaller. Intuitively, as q increases, while C remains fixed, it becomes more likely the true state is discovered by Receiver when the state is verified. This means that it is relatively more costly to send message m_1. As a result, Sender reduces the use of message m_1. The proposition is illustrated in Figures 6(a) and 6(b).

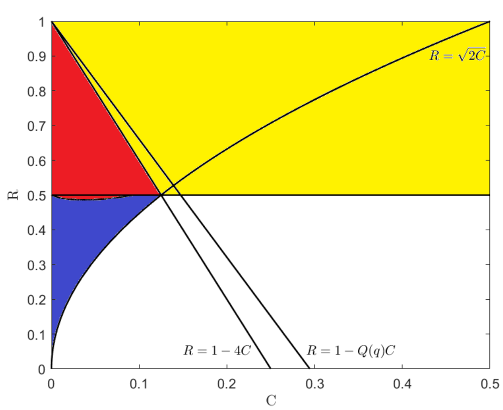

Figure 6(a). Optimal mechanism ($q = 0.6$)

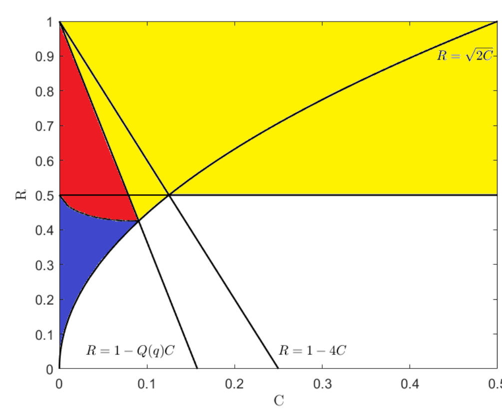

Figure 6(b). Optimal mechanism ($q = 0.8$)

For both figures, the optimal information structure is a no-disclosure mechanism in the white region, and a cutoff mechanism in the yellow region. Both of these mechanisms recommend no state verification. The optimal mechanism is a three-message mechanism in the red region and a negative assortative mechanism in the blue region. Both of these mechanisms involve state verification. Figure 6(a) represents the situation when $q < \frac{2}{3}$, when the state verification technology is relatively imprecise. In this case, the line $R = 1 - Q(q)C$ is always to the right of the line $R = 1 - 4C$. The total area of the blue and red regions is always the same, which corresponds to the first part of Proposition 3. When $q = \frac{2}{3}$, the two lines coincide. Figure 6(b) represents the situation when $q > \frac{2}{3}$, when the state verification technology is relatively precise. In this case, the line $R = 1 - Q(q)C$ is always to the left of the line $R = 1 - 4C$ and it determines the boundary between the yellow region and the (combined) blue and red region, which corresponds to the second part of Proposition 3. The boundary between the blue and red regions does not admit a closed form expression and is plotted here numerically.

Moreover, Figures 6(a) and 6(b) also show that, a sufficient condition for no state verification in the optimal information structure is $C \geq \frac{1}{8}$. Intuitively, when state verification is costly enough, it is possible to prevent Receiver from verifying. On the other hand, given c and q, the optimal information structure induces state verification only for intermediate level of R. This is because, for low R, it is easier to support no disclosure and for high R, that is, it is easier to satisfy $(A0)$ than $(C0)$, so the optimal information structure allocates all the credibility resources to m_0 rather than m_1.

5 Information and price setting

A seller can supply information to consumers to segment the market and charge higher prices to high-value buyers. Suppose in our model Sender is a seller of a good and Receiver a buyer. In this case, the outside option R can be interpreted as the price set by the seller and the seller sets the price and chooses the information revealed to the buyer at the same time. Lewis and Sappington (1994) consider a monopoly pricing problem with a seller who chooses how much information potential buyers would get. In their setting, the seller chooses how likely potential buyers would have access to an exogenously given signal and they show in a variety of settings that the seller would either choose to provide the maximal amount of information or no information. In this section, we show in our setting that no information is only optimal when the verification cost is high enough, and full information is never optimal. Moreover, the optimal price is not monotone in the verification cost.

For simplicity, consider the uniform distribution case, i.e., for all $\theta \in (0, 1)$, $f(\theta) = 1$, and assume the verification technology is perfect, i.e., $q = 1$. We can solve the seller's problem in two steps: step 1, given any R, choose the optimal information structure; step 2, choose the optimal R^*. By Proposition 3, when $q = 1$, for any price R, the optimal information structure is either a no-information mechanism or a cutoff mechanism. As a result, no state verification would take place in equilibrium. Figures 6(a) and (b) reduce to Figure 7, which has only white (no-disclosure mechanism) and yellow (cutoff mechanism) regions.

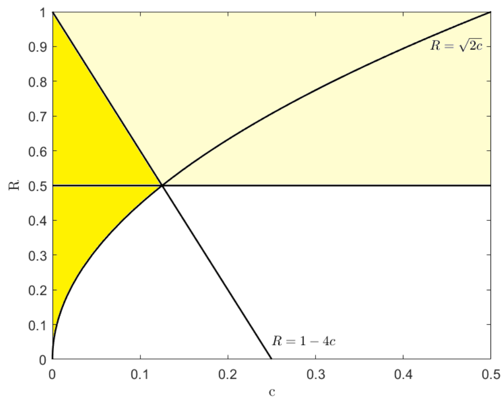

Figure 7. Optimal mechanism ($q = 1$)

For all combinations of (c, R) in the white region, the optimal information structure is a no-disclosure mechanism, no state verification takes place, and all buyers buy the product. For all combinations of (c, R) in the (light and dark) yellow region, the optimal information structure is a cutoff mechanism, no state verification takes place, and all buyers with θ above the cutoff θ^* buy the product.

Next, we would like to find the optimal price R^* given c. For each c, we proceed by looking for the most profitable price region by region.

Consider first the white region. When buyers has no information about their θ's, the seller sets the price and sell to all buyers. Therefore, the seller would set a price that as close to an uninformed buyer's willingness to pay, $\mathbb{E}(\theta) = \frac{1}{2}$, as possible. Thus, for a given c, the most profitable price in this region must lie in the upper boundary.

The trade-off faced by the seller in the yellow region is more interesting. Setting a higher price has two opposite effects. A higher price leads to lower sales. However, the extent how the cutoff θ^* increases as the price increases depends on which of the two constraints, $(A0)$ or

($C0$), binds. When ($A0$) is the binding constraint (the light yellow region), the seller's problem is equivalent to the monopoly pricing problem when the buyers know their θ's. As a result, the price that maximizes the seller's payoff in this region is the price closest to the monopoly price $\frac{1}{2}$. Thus, for a given c, the most profitable price in the light yellow region must lie in the lower boundary.

When ($C0$) is binding (the dark yellow region), the seller's problem is different from the monopoly pricing problem, and we can show that the most profitable price in this region is higher than the monopoly price, and increasing in the verification cost.

Because of the presence of two different concerns–to guarantee that the buyers are willing to buy the product in light yellow region, and to guarantee that the buyers don't have enough incentive to verify the state in the dark yellow region–the optimal price is non-monotone in the verification cost.

Proposition 4 *Suppose the state is uniformly distributed, the verification technology is perfect and the seller chooses both the information structure and price. The optimal price R^* is given by*

$$R^* = \begin{cases} \frac{1}{2} & \text{if } c \geq \frac{1}{8}, \\ 1 - 4c & \text{if } \frac{1}{10} < c < \frac{1}{8}, \\ \frac{1}{2} + \frac{3}{16}c + \frac{1}{16}\sqrt{c\,(9c+16)} & \text{if } c \leq \frac{1}{10}. \end{cases} \tag{3}$$

Moreover, the optimal information structure recommends Buy *if $\theta > \theta^*$ and* No-buy *if $\theta < \theta^*$, where θ^* is given by*

$$\theta^* = \begin{cases} 0 & \text{if } c \geq \frac{1}{8}, \\ 1 - 8c & \text{if } \frac{1}{10} < c < \frac{1}{8}, \\ R^* - c - \sqrt{2c\,(1-R^*) + c^2} & \text{if } c \leq \frac{1}{10}. \end{cases} \tag{4}$$

Figure 8 illustrates the optimal price R^* as a function of c.

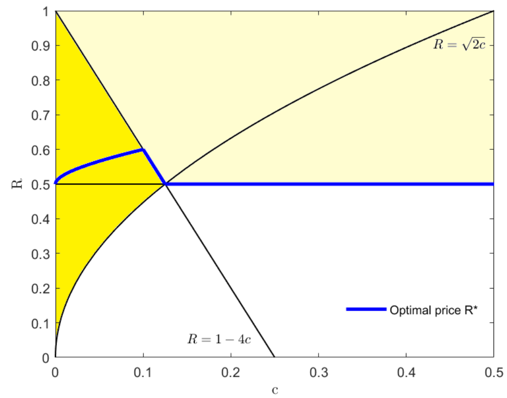

Figure 8. Optimal price R^* $(q = 1)$

Note that the effect of decreasing the verification cost on the optimal price is not monotone. Therefore, making information more accessible to consumers could *increase* the retail price in equilibrium. To understand why this is the case, note that when $\frac{1}{10} < c < \frac{1}{8}$, the optimal price is such that both $(A0)$ and $(C0)$ bind ((c, R) lies in the boundary of the dark and light yellow regions). Moreover, the seller would prefer to set a even higher price if $(A0)$ can be ignored. Suppose now c decreases. This means that $(A0)$ is no longer binding and thus the seller can increase the price.

Moreover, in contrast to Lewis and Sappington (1994), in which the optimal information structure is characterized by two extremes: either no information or maximal information, an intermediate level of information is provided by the seller in our model when the verification cost is low. Intuitively, in order to keep the verification incentive sufficiently low when the verification cost is low, the seller would have to set a very low price, which would make selling to all buyers unattractive. Instead, the seller sets a higher price and only sells to buyers with valuation high enough to justify buying at the price without verification.

6 Discussion

In this section, we discuss one extension of the model, and compare the optimal information structure to the equilibrium communication strategy when there is no commitment power.

6.1 Unlimited round of state verification

One constraint on our information acquisition technology is that Receiver can only verify the state once. State verification technology itself is not most flexible, and the one-round assumption further limits the flexibility of the information acquisition technology.

When Receiver can choose as many rounds of state verification as she wants, then if she strictly prefers verifying the state for one round, she would verify the state until the verification is successful.

Lemma 4 *Receiver either does not learn the state or always learns the state.*

When the previous round of state verification is not successful, Receiver does not update her belief, and the incentive for state verification remains the same as in the beginning of the previous round. Therefore, given a message, if Receiver has no incentive to verify the state for one round, then she does not learn the state; if Receiver has incentive to verify the state, then she learns the state perfectly.

This means that our model with unlimited rounds of state verification is equivalent to our baseline model when $q = 1$. In Proposition 3, we show that the optimal information structure never recommends state verification when the state is uniformly distributed and $q = 1$. This result also holds for other distributions.

Proposition 5 (No state-verification mechanism) *Suppose unlimited rounds of state verification. The optimal information structure is a no-disclosure mechanism or a cutoff mechanism, and no state verification takes place.*

To see why Proposition 5 holds, suppose in the optimal information structure, m_1 is sent with positive probability. For each θ that sends m_1, since Receiver always learns θ by Lemma 4, $\Pr(a = 1)$ if and only if $\theta \geq R$, when unlimited rounds of state verification is allowed. Therefore, an information structure that sends m_1 for θ above R cannot be optimal, as Sender can pool these θ's with some θ's below R and sends m_0 instead. On the other hand, if the information structure only sends m_1 below R, Receiver would not choose to verify. Therefore, the optimal information structure would never send m_1 with positive probability.

This result is reminiscent of the receiver-never-learn result in Matyskova and Montes (2023). Since the sender can provide any information that the receiver may acquire, it is without loss to assume that the sender provides information that prevents the receiver from acquiring information. Notice, however, that our state verification technology and cost do not satisfy their assumptions and, therefore, their result does not apply to our setting.

6.2 Role of commitment power

Communication with detectable deception has been studied in a cheap-talking setting (Dziuda and Salas, 2018; Balbuzanov, 2019; Levkun, 2021; Sadakane and Tam, 2022; Zhao, 2018). The resulting equilibrium communication strategy has a negative assortative strategy. In a cheap talk game, low types can always mimic high types, and there is no way to truthfully reveal the lowest types. Such feature can be observed in both environments with exogenous and endogenous deception detection. For example, 1) with exogenous lie detection, Dziuda and Salas (2018) show that the moderate types and the highest types tell the truth, and the lowest types pretend to be the highest types; 2) with endogenous state verification, Zhao (2018) shows that, in the sender-preferred equilibrium, the lowest types and the highest types pool. Moreover, Zhao (2018) shows that the receiver never chooses the sender-preferred action without verification, otherwise every type would lie. In equilibrium, the low types never benefit from lying. The fewer the lying low types, the higher the state verification effort, because the purpose of state verification is to find out the high types.

In contrast, an information designer can always truthfully reveal the lowest types and the optimal information structure does so. Hence, the receiver would chooses the sender-preferred action without verification for moderate types, which is the key difference between the communication strategy with and without commitment power. The lowest types are revealed to induce the sender to choose the sender-preferred action even without successful state verification, and the moderate types are further revealed to persuade the receiver to choose the sender-preferred action without going through the costly state verification. In other words, the better the lying moderate types, the lower the state verification effort, because the purpose of state verification is to find out the low types.

References

[1] Arieli, I., Babichenko, Y., Smorodinsky, R., & Yamashita, T. (2023). Optimal persuasion via bi-pooling. Theoretical Economics, 18(1), 15-36.

[2] Au, P. H. (2015). Dynamic information disclosure. *The RAND Journal of Economics, 46(4)*, 791-823.

[3] Balbuzanov, I. (2019). Lies and consequences: The effect of lie detection on communication outcomes. *International Journal of Game Theory, 48,* 1203-1240.

[4] Bizzotto, J., Rüdiger, J., & Vigier, A. (2020). Testing, disclosure and approval. *Journal of Economic Theory*, *187*, 105002.

[5] Dworczak, P., & Martini, G. (2019). The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5), 1993-2048.

[6] Dziuda, W., & Salas, C. (2018). Communication with detectable deceit. Available at SSRN 3234695.

[7] Ederer, F., & Min, W. (2022). Bayesian persuasion with lie detection (No. w30065). National Bureau of Economic Research.

[8] Gentzkow, M., & Kamenica, E. (2016). Competition in persuasion. *The Review of Economic Studies, 84(1)*, 300-322.

[9] Goldstein, I., & Leitner, Y. (2018). Stress tests and information disclosure. *Journal of Economic Theory, 177*, 34-69.

[10] Guo, Y., & Shmaya, E. (2019). The interval structure of optimal disclosure. *Econometrica, 87(2)*, 653-675.

[11] Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review, 101(6)*, 2590-2615.

[12] Kleiner, A., Moldovanu, B., & Strack, P. (2021). Extreme points and majorization: Economic applications. *Econometrica*, 89(4), 1557-1593.

[13] Kolotilin, A. (2018). Optimal information disclosure: A linear programming approach. *Theoretical Economics, 13(2),* 607-635.

[14] Kolotilin, A., Corrao, R., & Wolitzky, A. (2022). Persuasion with non-linear preferences.

[15] Kolotilin, A., Mylovanov, T., Zapechelnyuk, A., & Li, M. (2017). Persuasion of a privately informed receiver. *Econometrica, 85(6)*, 1949-1964.

[16] Lewis, T. R., & Sappington, D. E. (1994). Supplying information to facilitate price discrimination. *International Economic Review*, 35(2), 309-327.

[17] Li, F., & Norman, P. (2018). On Bayesian persuasion with multiple senders. *Economics Letters, 170,* 66-70.

[18] Li, F., & Norman, P. (2021). Sequential persuasion. *Theoretical Economics, 16(2),* 639-675.

[19] Levkun, A. (2021). Communication with strategic fact-checking. Working paper.

[20] Matyskova, L. & Montes, A. (2023). Bayesian persuasion with costly information acquisition. Working paper.

[21] Sadakane, H., & Tam, Y. C. T. (2022). Cheap talk and lie detection.

[22] Terstiege, S., & Wasser, C. (2020). Buyer-optimal extensionproof information. *Journal of Economic Theory, 188*, 105070.

[23] Zhao, X. (2018). How to persuade a group: Simultaneously or sequentially.

7 Appendix

Proof of Lemma 2. Our information design problem can be formulated as

$$\max_{x_0(\theta),x_1(\theta)\in[0,1]} \int_0^R \left[x_0\left(\theta\right)+x_1\left(\theta\right)\left(1-q\right)\right] f\left(\theta\right) d\theta + \int_R^1 \left(x_0\left(\theta\right)+x_1\left(\theta\right)\right) f\left(\theta\right) d\theta$$

s.t.

$$\begin{aligned}
\int_0^1 (\theta - R) x_0(\theta) f(\theta) d\theta &\geq 0, && \text{(A0)} \\
\int_0^1 (\theta - R) x_1(\theta) f(\theta) d\theta &\geq 0, && \text{(A1)} \\
C \int_R^1 x_0(\theta) f(\theta) d\theta + \int_0^R (\theta - R + C) x_0(\theta) f(\theta) d\theta &\geq 0, && \text{(C0)} \\
\forall \theta \in [0,1], \quad 1 - x_0(\theta) - x_1(\theta) &\geq 0, && \text{(TP)}
\end{aligned}$$

where $x_0(\theta)$ and $x_1(\theta)$ are probabilities that m_0 and m_1 are sent, respectively, and the last inequality follows from the law of total probability, i.e., the probabilities that m_ϕ, m_0 and m_1 are sent must sum up to 1. Let λ_0, λ_1, μ and $\eta(\theta)$ be the Lagrange multipliers corresponding to the constraints $(A0)$, $(A1)$, $(C0)$ and (TP), respectively, and $x_0^*(\theta)$ and $x_1^*(\theta)$ be the maximizers. The Euler–Lagrange equations are given by

$$\forall \theta > R, \quad \begin{array}{l} \lambda_0(\theta - R) + 1 + \mu C - \hat{\eta}(\theta) > 0 \Rightarrow x_0^*(\theta) = 1, \\ \lambda_0(\theta - R) + 1 + \mu C - \hat{\eta}(\theta) < 0 \Rightarrow x_0^*(\theta) = 0, \end{array} \tag{5}$$

$$\forall \theta > R, \quad \begin{array}{l} \lambda_1(\theta - R) + 1 - \hat{\eta}(\theta) > 0 \Rightarrow x_1^*(\theta) = 1, \\ \lambda_1(\theta - R) + 1 - \hat{\eta}(\theta) < 0 \Rightarrow x_1^*(\theta) = 0, \end{array} \tag{6}$$

$$\forall \theta < R, \quad \begin{array}{l} (\lambda_0 + \mu)(\theta - R) + 1 + \mu C - \hat{\eta}(\theta) > 0 \Rightarrow x_0^*(\theta) = 1, \\ (\lambda_0 + \mu)(\theta - R) + 1 + \mu C - \hat{\eta}(\theta) < 0 \Rightarrow x_0^*(\theta) = 0. \end{array} \tag{7}$$

$$\forall \theta < R, \quad \begin{array}{l} \lambda_1(\theta - R) + 1 - q - \hat{\eta}(\theta) > 0 \Rightarrow x_1^*(\theta) = 1, \\ \lambda_1(\theta - R) + 1 - q - \hat{\eta}(\theta) < 0 \Rightarrow x_1^*(\theta) = 0, \end{array} \tag{8}$$

where $\hat{\eta}(\theta) := \eta(\theta) / f(\theta)$. For $\theta > R$, define the marginal benefits of sending messages m_0 and m_1, $\overline{B}_0(\theta)$ and $\overline{B}_1(\theta)$, by

$$\begin{aligned}
\overline{B}_0(\theta) &: = \lambda_0(\theta - R) + 1 + \mu C, \\
\overline{B}_1(\theta) &: = \lambda_1(\theta - R) + 1.
\end{aligned}$$

For all $\theta > R$, $\overline{B}_0(\theta), \overline{B}_1(\theta) > 0$. This means that we cannot have $x_0^*(\theta) + x_1^*(\theta) < 1$. This is because, in that case, (5) and (6) imply $\hat{\eta}(\theta) \geq \max\left\{\overline{B}_0(\theta), \overline{B}_1(\theta)\right\} > 0$, which in turn implies that (TP) is binding. Thus, m_ϕ is not sent for any $\theta > R$.

Similarly, for $\theta < R$, define the marginal benefits of sending messages m_0 and m_1, $\underline{B}_0(\theta)$ and $\underline{B}_1(\theta)$, by

$$\begin{aligned}\underline{B}_0(\theta) &: = (\lambda_0 + \mu)(\theta - R) + 1 + \mu C, \\ \underline{B}_1(\theta) &: = \lambda_1(\theta - R) + 1 - q.\end{aligned}$$

Define

$$\theta^* := \max\left\{\min\left\{R - \frac{1-q}{\lambda_1}, R - \frac{1+\mu C}{\lambda_0 + \mu}\right\}, 0\right\}.$$

For all $\theta < \theta^*$, we must have $x_0^*(\theta) = x_1^*(\theta) = 0$ and m_ϕ is sent with probability 1. This is because, by construction, $\theta^* > 0$ implies that for all $\theta < \theta^*$, $\underline{B}_0(\theta), \underline{B}_1(\theta) < 0$. Since $\hat{\eta}(\theta) \geq 0$, (7) and (8) imply that $x_0^*(\theta) = x_1^*(\theta) = 0$. For all $\theta \in (\theta^*, R)$, at least one of $\underline{B}_0(\theta)$ and $\underline{B}_1(\theta)$ is strictly larger than 0. Suppose $\underline{B}_0(\theta) > 0$. We must have $x_0^*(\theta) = 1$ or $\hat{\eta}(\theta) \geq \underline{B}_0(\theta) > 0$. In both cases, (TP) is binding. Similarly, (TP) must be binding when $\underline{B}_1(\theta) > 0$. Thus, for all $\theta \in (\theta^*, R)$, m_ϕ is not sent. ■

Proof of Theorem 1. By Lemma 2, our information design problem can be re-formulated as

$$\max_{\substack{x_0(\theta)\in[0,1] \\ \theta^*\in[0,R]}} \int_{\theta^*}^{R} \left[x_0(\theta) + (1 - x_0(\theta))(1-q)\right] f(\theta)\, d\theta + \int_R^1 f(\theta)\, d\theta$$

s.t.

$$\int_{\theta^*}^{1} (\theta - R)\, x_0(\theta) f(\theta)\, d\theta \geq 0, \tag{A0}$$

$$\int_{\theta^*}^{1} (\theta - R)(1 - x_0(\theta)) f(\theta)\, d\theta \geq 0, \tag{A1}$$

$$C\int_R^1 x_0(\theta) f(\theta)\, d\theta + \int_{\theta^*}^{R} (\theta - R + C)\, x_0(\theta) f(\theta)\, d\theta \geq 0, \tag{C0}$$

Let λ_0, λ_1 and μ be the Lagrange multipliers corresponding to the constraints $(A0)$, $(A1)$, and

$(C0)$, respectively, and let $x_0^*(\theta)$ be the maximizer. The Euler–Lagrange equations are given by

$$\forall \theta > R, \quad \begin{array}{l} (\lambda_0 - \lambda_1)(\theta - R) + \mu C > 0 \Rightarrow x_0^*(\theta) = 1, \\ (\lambda_0 - \lambda_1)(\theta - R) + \mu C < 0 \Rightarrow x_0^*(\theta) = 0, \end{array} \tag{9}$$

$$\forall \theta < R, \quad \begin{array}{l} (\lambda_0 - \lambda_1 + \mu)(\theta - R) + q + \mu C > 0 \Rightarrow x_0^*(\theta) = 1, \\ (\lambda_0 - \lambda_1 + \mu)(\theta - R) + q + \mu C < 0 \Rightarrow x_0^*(\theta) = 0. \end{array} \tag{10}$$

Moreover, the optimality of θ^* implies

$$\begin{array}{ll} \begin{array}{c} [x_0^*(\theta) + (1 - x_0^*(\theta))(1-q)] + [\lambda_0 x_0^*(\theta) + \lambda_1(1 - x_0^*(\theta))](\theta^* - R) \\ +\mu(\theta^* - R + C)\, x_0^*(\theta) = 0 \end{array} & \text{if } \theta^* > 0, \\ \begin{array}{c} [x_0^*(\theta) + (1 - x_0^*(\theta))(1-q)] + [\lambda_0 x_0^*(\theta) + \lambda_1(1 - x_0^*(\theta))](\theta^* - R) \\ +\mu(\theta^* - R + C)\, x_0^*(\theta) \geq 0 \end{array} & \text{if } \theta^* = 0. \end{array} \tag{11}$$

Suppose first $\lambda_0 > \lambda_1$. (9) implies that $x_0^*(\theta) = 1$ for all $\theta > R$. $(A1)$ then implies that $x_0^*(\theta) = 1$ for all $\theta > \theta^*$. Similarly, if $\lambda_0 = \lambda_1$ and $\mu > 0$, then $x_0^*(\theta) = 1$ for all $\theta > R$. As a result, $x_0^*(\theta) = 1$ for all $\theta > \theta^*$.

If $\lambda_0 \leq \lambda_1$ and $\mu = 0$, then $x_0^*(\theta) = 1$ for all $\theta \in (\theta^*, R)$. Since m_1 is not sent below R, if m_1 is sent for a positive measure of θ above R, then $(C1)$ is violated and Receiver would not verify the state. This means that $x_0^*(\theta) = 1$ for all $\theta > R$ as well.

Next, suppose $\lambda_0 < \lambda_1$ and $\mu > 0$. Consider

$$\overline{\theta} := R + \frac{\mu C}{\lambda_1 - \lambda_0}.$$

If $\overline{\theta} \geq 1$, then m_1 is not sent with positive probability above R. $(A1)$ then implies m_1 is not sent with positive probability. If $\overline{\theta} < 1$, then $x_0^*(\theta) = 1$ for all $\theta \in \left(R, \overline{\theta}\right)$ and $x_0^*(\theta) = 0$ for all $\theta \in (\overline{\theta}, 1]$. Since $\mu > 0$, $(A1)$ implies $x_0^*(\theta) = 0$ for some $\theta \in (\theta^*, R)$. Since $q > 0$, (10) implies that $x_0^*(\theta) = 1$ for θ smaller than but close enough to R. Moreover, by (10), in order to have $x_0^*(\theta) = 0$ for some $\theta < R$, we must have $\lambda_0 - \lambda_1 + \mu > 0$. Define

$$\underline{\theta} := R - \frac{q + \mu C}{\lambda_0 - \lambda_1 + \mu}.$$

In this case, $x_0^*(\theta) = 1$ for all $\theta \in (\underline{\theta}, R)$ and $x_0^*(\theta) = 0$ for all $\theta \in (\theta^*, \underline{\theta})$.

In all cases, $\mathbf{m}(\theta)$ satisfies the descriptions in Theorem 1. $\blacksquare$

Proof of Lemma 3. From the proof of Theorem 1, the only case in which m_1 is sent with positive probability in the optimal information structure is when $\lambda_0 < \lambda_1$ and $\mu > 0$. Since $\mu > 0$, $(C0)$ is binding. $\blacksquare$

Proof of Proposition 3. Consider $\sqrt{2C} < R < 1 - 4C$. By Proposition 2, the optimal information structure is not a no-disclosure mechanism. This means that one of $(A0)$ and $(C0)$ must bind. Otherwise, Sender can send more m_0 and increase his payoff. By Lemma 3, if $(C0)$ is not binding, then $\overline{\theta} = 1$ under the optimal information structure. But since $1 > R + 4C$, we have $\underline{\theta}_{C0}(1) \geq \underline{\theta}_{A0}(1)$. Therefore, under the optimal information structure, $(C0)$ must be binding and $\overline{\theta} \in [R + 4C, 1]$. We would like to identify conditions under which $\overline{\theta} \neq 1$ under the optimal information structure. To do so, we first write down Sender's problem as an optimization problem over $\overline{\theta}$ and show the first order condition at $\overline{\theta} = 1$ is not satisfied under the stated conditions. Then, we show the first order condition is necessary and sufficient for $\overline{\theta} = 1$ to be the maximizer.

Define

$$\theta_{A1}^*\left(\overline{\theta}, \underline{\theta}\right) := R - \sqrt{(R - \underline{\theta})^2 + (1 - R)^2 - \left(\overline{\theta} - R\right)^2}.$$

If $\theta^* = \theta_{A1}^*\left(\overline{\theta}, \underline{\theta}\right)$, then θ^* is the solution to

$$\left(1 - \overline{\theta}\right)\left(\frac{1 + \overline{\theta}}{2} - R\right) = (\underline{\theta} - \theta)\left(R - \frac{\underline{\theta} + \theta}{2}\right),$$

which means that $(A1)$ binds. Thus, given $\underline{\theta}$ and $\overline{\theta}$, $\theta_{A1}^*\left(\overline{\theta}, \underline{\theta}\right)$ identifies θ^* under the assumption that $(A1)$ binds.

Suppose that $(C0)$ is binding and $\theta^* = 0$, the sender's objective function becomes

$$\Pr(a = 1) = 1 - \underline{\theta}_{C0}\left(\overline{\theta}\right) + (1 - q)\,\underline{\theta}_{C0}\left(\overline{\theta}\right) = 1 - q\left(R - C - \sqrt{2C\left(\overline{\theta} - R\right) + C^2}\right),$$

which is strictly increasing in $\overline{\theta}$.

Next, suppose $(A1)$ and $(C0)$ are binding and $\theta^* > 0$, the sender's objective function becomes

$$\begin{aligned}
& \Pr(a=1) \\
= \; & \Pr(\theta \geq R) + \Pr(\theta < R)\left[\Pr(m=m_0|\theta < R) + (1-q)\Pr(m=m_1|\theta<R)\right] \\
= \; & 1-R+R-\underline{\theta}_{C0}\left(\overline{\theta}\right) + (1-q)\left(\underline{\theta}_{C0}\left(\overline{\theta}\right) - \theta^*_{A1}\left(\overline{\theta}, \underline{\theta}_{C0}\left(\overline{\theta}\right)\right)\right) \\
= \; & 1-R+\left(C+\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\right) \\
& +(1-q)\left(R-C-\sqrt{2C\left(\overline{\theta}-R\right)+C^2} - \left(R-\sqrt{\left(R-\underline{\theta}\right)^2+(1-R)^2-\left(\overline{\theta}-R\right)^2}\right)\right) \\
= \; & 1-R+q\left(C+\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\right) \\
& +(1-q)\left(\sqrt{\left(C+\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\right)^2+(1-R)^2-\left(\overline{\theta}-R\right)^2}\right).
\end{aligned}$$

Differentiating the objective function yields

$$\frac{d\Pr(a=1)}{d\overline{\theta}} = \frac{qC+(1-q)\dfrac{\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\left(C+R-\overline{\theta}\right)+C^2}{\sqrt{\left(C+\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\right)^2+(1-R)^2-\left(\overline{\theta}-R\right)^2}}}{\sqrt{2C\left(\overline{\theta}-R\right)+C^2}}.$$

Suppose $q=1$. Then, $\frac{d\Pr(a=1)}{d\overline{\theta}} > 0$, which implies that $\overline{\theta}=1$ is optimal. The optimal information structure would never recommend state verification.

Consider next $q \in (0,1)$. Let $A := \frac{1}{1-q}$. We have

$$\begin{aligned}
& \tfrac{d\Pr(a=1)}{d\overline{\theta}}|_{\overline{\theta}=1} < 0 \\
\Leftrightarrow \; & (A-1)C\left(C+\sqrt{2C(1-R)+C^2}\right)+\sqrt{2C(1-R)+C^2}\,(C+R-1)+C^2<0 \\
\Leftrightarrow \; & AC^2-\sqrt{2C(1-R)+C^2}\,(1-R-AC)<0
\end{aligned}$$

If $R+AC \geq 1$, then $\frac{d\Pr(a=1)}{d\overline{\theta}}|_{\overline{\theta}=1} > 0$. Thus, to identify the condition for $\frac{d\Pr(a=1)}{d\overline{\theta}}|_{\overline{\theta}=1} < 0$, we only need to focus on the case when $R + AC < 1$. Thus,

$$\begin{aligned}
&\tfrac{d\Pr(a=1)}{d\overline{\theta}}|_{\overline{\theta}=1} < 0 \\
\Leftrightarrow\quad & A^2C^4 - \left(2C(1-R) + C^2\right)(1-R-AC)^2 < 0 \\
\Leftrightarrow\quad & -C(1-R)\left(2A(A-1)C^2 - (1-R)(4A-1)C + 2R^2 - 4R + 2\right) < 0 \\
\Leftrightarrow\quad & 2A(A-1)C^2 - (1-R)(4A-1)C + 2R^2 - 4R + 2 > 0
\end{aligned}$$

Let

$$H(R) := 2A(A-1)C^2 - (1-R)(4A-1)C + 2R^2 - 4R + 2.$$

The function $H(.)$ is quadratic and $H(\infty) = H(-\infty) = \infty$. Moreover,

$$H(1-AC) = -AC^2 < 0$$

and for $R \in (0, 1-AC)$ and $q \in (0,1)$,

$$H'(R) = (4A-1)C + 4R - 4 < (4A-1)C - 4AC = -C < 0.$$

Thus, there is at most one root of $H(.)$ on $[0, 1-AC]$ and it is given by

$$R^*(C) = 1 - \frac{4A(A-1)}{4A-1-\sqrt{8A+1}}C.$$

Thus, $\frac{d\Pr(a=1)}{d\overline{\theta}}|_{\overline{\theta}=1} < 0$ if and only if $R < R^*(C)$.

Finally, we show that the first order condition is sufficient. Let

$$\Omega\left(\overline{\theta}\right) := \frac{\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\left(C+R-\overline{\theta}\right)+C^2}{\sqrt{\left(C+\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\right)^2+(1-R)^2-\left(\overline{\theta}-R\right)^2}}.$$

Note that $\frac{d\Pr(a=1)}{d\overline{\theta}} \lesseqgtr 0$ if and only if $\Omega\left(\overline{\theta}\right) \lesseqgtr -\frac{qC}{1-q}$. The first order condition is sufficient if for all $\overline{\theta} \in [R+4C, 1]$, $\Omega'\left(\overline{\theta}\right) < 0$.

$$\begin{aligned} & \Omega'\left(\overline{\theta}\right) \\ = & -\frac{C}{\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\left(\left(C+\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\right)^2+(1-R)^2-\left(\overline{\theta}-R\right)^2\right)^{\frac{3}{2}}} \\ & \times\{2C^3+\left(\overline{\theta}-R\right)\left(6C^2-3CR+3C\overline{\theta}+2R^2+2R\overline{\theta}-6R-\overline{\theta}^2+3\right) \\ & +2C\sqrt{2C\left(\overline{\theta}-R\right)+C^2}\left(C+2\left(\overline{\theta}-R\right)\right)\}. \end{aligned}$$

which is negative if

$$\Phi\left(\overline{\theta}, C\right) = 6C^2 - 3CR + 3C\overline{\theta} + 2R^2 + 2R\overline{\theta} - 6R - \overline{\theta}^2 + 3 > 0.$$

We have

$$\begin{aligned} \frac{d\Phi\left(\overline{\theta}, C\right)}{d\overline{\theta}} &= 3C + 2R - 2\overline{\theta} < 3C + 2R - 2\left(R+4C\right) = -5C < 0, \\ \frac{d\Phi\left(\overline{\theta}, C\right)}{dC} &= 12C + 3\left(\overline{\theta} - R\right) > 0. \end{aligned}$$

Thus, $\Phi\left(\overline{\theta}, C\right)$ is minimized at $\left(\overline{\theta}, C\right) = (1, 0)$. Since

$$\Phi(1,0) = 2(1-R)^2 > 0,$$

for all $\overline{\theta} \in [R+4C, 1]$, $\Phi\left(\overline{\theta}, C\right) > 0$. ■

Proof of Proposition 4. Suppose $q = 1$. Then by Proposition 3, there are at most two messages that sent with positive probability in the optimal information structure. One induces acceptance without verification, and one induces rejection without verification. The seller chooses both the information available and price. The maximization problem becomes

$$\max_{R,\theta\in[0,1]} R(1-\theta)$$

s.t.

$$\theta \geq 2R - 1, \tag{12}$$

$$\theta \geq R - c - \sqrt{2c(1-R) + c^2}, \tag{13}$$

where the inequality (12) follows from ($A0$) and (13) follows from ($C0$).

We use R^* and θ^* to denote the maximizers. The maximization problem can be decomposed into two steps: step 1, given any R, choose the optimal θ^*; step 2, given $\theta^*(R)$, choose the optimal R^*.

Consider step 1. When $R < \frac{1}{2}$ and $R < \sqrt{2c}$ (white region of Figure 7), the right-hand sides of the two constraints (12) and (13) are less than 0. It is thus optimal to set $\theta^* = 0$. Suppose next $R \geq \frac{1}{2}$ or $R \geq \sqrt{2c}$. Since the objective is decreasing in θ, one of the constraints must bind in this case. Comparing (12) and (13), we find that when $R \geq 1 - 4c$ (light yellow region of Figure 7), (12) is binding; when $R \leq 1 - 4c$ (dark yellow region of Figure 7), (13) binds.

Consider next step 2. Given c, since the objective function is increasing in R and the two constraints do not bind in the white region, the most profitable price of this region lies in the upper boundary. Consider next the light yellow region, where (12) is binding. Ignore (13) and use (12) to eliminate θ, the maximization problem becomes

$$\max_{R\in[0,1]} 2R(1-R).$$

The unconstrained problem has the maximizer $R^* = \frac{1}{2}$. Thus, the most profitable price of this region must lie in the lower boundary.

Finally, consider the dark yellow region, where (13) is binding. Ignore (12) and use (13) to eliminate θ, the maximization problem becomes

$$\max_{R\in[0,1]} R\left(1 - R + c + \sqrt{2c(1-R) + c^2}\right).$$

The F.O.C. is given by

$$c - 2R + 1 + \frac{1}{\sqrt{2c(1-R) + c^2}} \left(2c - 3Rc + c^2\right) = 0,$$

which has solution

$$R^*_{C0} = \frac{1}{2} + \frac{3}{16}c + \frac{1}{16}\sqrt{c(9c+16)}.$$

Notice that the objective is strictly concave in R.

$$\begin{aligned}
& \frac{\partial^2}{\partial R^2}\left(R\left(1 - R + c + \sqrt{2c(1-R)+c^2}\right)\right) \\
= \; & -2 - \frac{c^2}{\left(2c(1-R)+c^2\right)^{\frac{3}{2}}}(2c - 3R + 4) \\
\leq \; & -2 - \frac{c^2}{\left(2c(1-R)+c^2\right)^{\frac{3}{2}}}(2c + 1) \\
< \; & 0,
\end{aligned}$$

where we have used the fact that $R \leq 1$ in the first inequality. As a result, the most profitable price of this region is the price closest to R^*_{C0}. Note also that when $c < \frac{1}{10}$, R^*_{C0} lies in the dark yellow region, i.e., $R^*_{C0} \in (\sqrt{2C}, 1 - 4C)$. Putting the results for different regions together, we find that R^* is given by (3). To see that θ^* is given by (4), note first that, by Proposition 2, if $c \geq \frac{1}{8}$ and $R^* = \frac{1}{2}$, then $\theta^* = 0$. When $\frac{1}{10} < c < \frac{1}{8}$ and $R^* = 1 - 4c$, (12) is binding and thus $\theta^* = 1 - 8c$. When $c \leq \frac{1}{10}$ and $R^* = \frac{1}{2} + \frac{3}{16}c + \frac{1}{16}\sqrt{c(9c+16)}$, (13) is binding and thus $\theta^* = R^* - c - \sqrt{2c(1-R^*)+c^2}$. ■

Proof of Proposition 5. When Receiver can choose unlimited rounds of state verification, the information design problem reduces to

$$\max_{x_0(\theta) \in [0,1]} \int_0^1 x_0(\theta) f(\theta) \, d\theta$$

s.t.

$$\int_0^1 (\theta - R)\, x_0(\theta)\, f(\theta)\, d\theta \;\geq\; 0, \tag{A0}$$

$$C\int_R^1 x_0(\theta)\, f(\theta)\, d\theta + \int_0^R (\theta - R + C)\, x_0(\theta)\, f(\theta)\, d\theta \;\geq\; 0. \tag{C0}$$

Let λ_0 and μ be the Lagrange multipliers corresponding to the constraints $(A0)$ and $(C0)$, respectively and let $x_0^*(\theta)$ be the maximizer. The Euler–Lagrange equations are given by

$$\forall\theta > R, \quad \begin{array}{l} \lambda_0(\theta - R) + 1 + \mu C > 0 \Rightarrow x_0^*(\theta) = 1, \\ \lambda_0(\theta - R) + 1 + \mu C < 0 \Rightarrow x_0^*(\theta) = 0, \end{array} \tag{14}$$

$$\forall\theta < R, \quad \begin{array}{l} (\lambda_0 + \mu)(\theta - R) + 1 + \mu C > 0 \Rightarrow x_0^*(\theta) = 1, \\ (\lambda_0 + \mu)(\theta - R) + 1 + \mu C < 0 \Rightarrow x_0^*(\theta) = 0. \end{array} \tag{15}$$

Since $\lambda_0, \mu \geq 0$, (14) and (15) imply immediately that there exists $\theta^* \in [0, R)$ such that m_0 is sent for $\theta > \theta^*$ and m_ϕ is sent for $\theta < \theta^*$. ■