Strategic Communication and Scientific Learning

Srijita Ghosh

September 2024

Abstract

In this paper we study the strategic communication problem for scientific information. The information is costly to process for both the sender and the receiver but costless to verify. We show that the sender cannot persuade the receiver to choose an action in all states,i.e., in some states following the recommendation of the sender, the receiver will choose to learn independently. We show that the persuasion strategy affects the learning strategy of the receiver. Under a high level of preference misalignment, the sender sends no recommendation, and by allowing the mixing of recommendations we ensure that the sender can still change the consideration set of the receiver which in turn affects his payoff. We also provide an alternate algorithm to find the consideration set of the agent facing uniform poster separable (UPS) cost of learning.

Keywords:: Strategic communication, Information design, Rational Inattention, Consideration set

JEL Code: D83, D91

1 Introduction

The Bayesian Persuasion literature considers strategic interaction problems under preference misalignment between a sender and a receiver. However, the sender needs to send scientific information The key feature of scientific information is that is costly to acquire, but easy to verify for all economic receivers. For example, consider an environmental protection agency recommending whether to use a new and improved seed variety for farmers or recommending a mining firm whether it is safe to use a new extraction technique in a vulnerable geographical location. In this paper, we analyze the optimal strategic communication policy for scientific information.

We consider a standard Bayesian Persuasion model where both the sender and the receiver have access to costly information. Upon learning the sender can costlessly send a recommendation to the receiver and the receiver can costlessly verify it. We assume both the sender and the receiver face the same Uniform Posterior Separable (UPS) cost function, parametrized by a single marginal cost of learning parameter. Furthermore, we show how the welfare changes if the receiver faces a higher marginal cost of learning than the sender.

Our main contribution is to show that the sender need not be able to directly manipulate the action of the receiver since receiver has access to the same learning technology and in such cases the sender manipulates the consideration set of the receiver's learning strategy. As an implication, we find that in states where the preference misaligned is the highest, the optimal strategy of the sender would be to not send any recommendation.

A secondary contribution is providing an alternate algorithm of finding consideration sets of receivers with UPS cost function, more specifically Shannon cost function. The use of this algorithm is not restricted to strategic communication problems only and can be applied to any decision problem where the decision maker faces a Shannon cost of learning.

Our first lemma shows that the learning strategy of any decision maker divides the simplex of prior beliefs into finitely many consideration sets. Using the Likelihood Invariant Posterior (LIP) property of the UPS cost function we show that for any prior belief in a given consideration set the posterior belief is identical and given by the extreme points of the consideration set, this can be found using the algorithm mentioned above.

Using this lemma we find the learning strategy of both the sender and the receiver and show if the sender truthfully recommends the action, the receiver will optimally choose to learn following the recommendation. Using this we show that in equilibrium the sender chooses to not truthfully communicate in every state and since verification is costless, mixing is achieved using no recommendation policy. Such mixing changes the interim belief of the receiver and puts him in a different consideration set. By choosing the consideration set optimally the sender can change the final action choice made by the receiver.

1.1 Literature Review

The strategic communication framework in this paper is built on the information design problem a la Kamenica and Gentzkow (2011)). Gentzkow and Kamenica (2014) consider a variation where only the sender faces a posterior separable cognitive cost function (refer Caplin et al. (2017)), where the main result from Kamenica and Gentzkow (2011) holds, i.e, the sender can manipulate the receiver's action. In Matyskova and Montes (2023) receiver can learn (costly) after obtaining the information from the sender whereas the sender faces no cost. Unlike this paper, they find that the receiver never learns on his own in equilibrium, because the sender can choose any information structure he wants. Bloedel and Segal (2018) also assume a cognitive cost for the receiver and with continuous state and two actions, they find the sender optimally partitions the state space into at most three intervals and provides either a simple recommendation strategy (*pooling*) or complex information strategy of disclosing the state(*separating*). Unlike their model disclosing the state is never optimal in this model since the receiver can learn on his own as well. Lipnowski et al (2020) show that if the receiver faces a cost of attention then even a benevolent sender (no preference misalignment) would not reveal full information and would prefer to restrict the choice set of the receiver, which improves welfare. In our model, the sender cannot directly restrict the choice set, but they do so with their optimal communication strategy.

The learning technology used in this paper is borrowed from Rational Inattention literature. Caplin and Dean (2015), Matějka and McKay (2015) have shown that solving the optimal choice of signals and the optimal posterior distribution is equivalent when Bayesian plausibility is satisfied. We use this formulation to solve the strategic interaction problem between the rationally inattentive receivers and a social sender with misalignment in preferences. In Caplin et al. (2019) related the RI learning problem to an endogenous consideration set problem. In this paper, we use a similar analysis and investigate how the sender can strategically change the consideration set of the receiver using the information policy.

The results obtained in this paper also relate to the consideration set literature. Manzini and Mariotti (2014) provide an axiomatic foundation of choice under consideration set formation. Lleras et al. (2017) characterize choice behavior under different consideration sets. In both cases, the formation of the consideration is a result of assumptions about the choice problem. However, in this paper even though we obtain a consideration set formulation, unlike the previous literature the formation of the consideration set is generated as an implication of a strategic communication problem. The formation of an endogenous choice set is thus not a result of choice axioms but an implication of the cost of learning. Since learning is costly, receivers exclude some actions from consideration based on prior beliefs. A similar intuition is present in Caplin et al. (2019). In this paper, we further show that strategic communication can manipulate this process.

The rest of the paper is organized as follows. Section 2 describes the model, section 3 presents the main results and section 4 concludes.

2 Model

2.1 Primitives

Let us consider a one-period model of an economy with a sender (e.g. the environment protection agency) and a receiver (e.g., the farmer). The sender wants to recommend an action (e.g., choosing new agricultural technology) to the receiver. The receiver's action (choice of the actual technology) affects the payoff of both the sender and the receiver.

The payoff relevant set of states is given by $\Omega = \{\omega_H, \omega_L, \omega_0\}$ where ω_H is the *high* state, ω_L represents a *low* state where caution is needed and ω_0 represents a *dire* state where the new technology is detrimental to the environment. The corresponding action set is given by $A = \{a_H, a_L, a_0\}$ where a_H, a_L, and a_0 denote the high, low and zero level of action chosen by the receiver.

The utility function $u : A \times \Omega \to \mathbb{R}$ of the receiver is as follows:

$$u(a,\omega) = \begin{cases} \alpha_H - i(\omega)\beta_H & \text{for } a = a_H \\ \alpha_L - i(\omega)\beta_L & \text{for } a = a_L \\ 0 & \text{for } a = a_0 \end{cases} \text{ where } \quad i(\omega) = \begin{cases} 0 & \text{for } \omega = \omega_H \\ i & \text{for } \omega = \omega_L \\ 1 & \text{for } \omega = \omega_0 \end{cases}$$

where $i \in (0,1)$. For simplicity we assume the positive component α_i to be state-independent. This is highest for action a_H, independent of the state. The negative component, denoted by parameter β_i is state-dependent and $i(\omega)$ denotes the marginal impact of the negative factors in state ω [1].

Note that, without perfect information a_0 denotes a riskless choice and a_H and a_L would be risky choices, where the risk is higher for the former action. We further

[1]For example, choosing the new technology always generates the highest yield, however, the impact of such technology on the environment (ie., soil quality, water table, biodiversity) depends on the state.

assume that

$$\alpha_H > \alpha_L > 0; \quad \beta_H > \beta_L > 0;$$
$$\alpha_L < \beta_L; \quad \alpha_L > i\beta_L; \quad \alpha_H < \beta_H$$
$$\alpha_H - \beta_H < \alpha_L - \beta_L < 0$$

Thus under full information, the receiver will optimally choose a_i is state ω_i for $i = H, L$ or 0.

The utility function of the sender is thus given by,

$$v(a_H, \omega) = u(a_H, \omega) - i(\omega)\nu$$
$$v(a_L, \omega) = u(a_L, \omega) - i(\omega)\nu$$
$$v(a_0, \omega) = u(a_0, \omega)$$

where $\nu \in (0, \frac{\alpha_L}{i} - \beta_L)$ is the additional cost ignored by the receiver, that measures the level of preference misalignment [2]. Since $\nu < \frac{\alpha_L}{i} - \beta_L$ implies $v(a_L, \omega_L) > v(a_0, \omega_L)$ and $i(\omega_H) = 0$ and a_0 maximizes $u(a, \omega_0)$, the sender will also choose action a_i in state ω_i under full information. Thus the preference misalignment is an artifact of the lack of full information.

2.2 Learning Technology

Both DMs (the sender and the receiver) enter the period with a common prior belief $\mu_0 \in \Delta(\Omega)$ and are Bayesian. Assume that $\mu_0 \in int(\Delta(\Omega))$, i.e., the true state is learnable and both the DMs have access to the same learning technology. We assume there is no inherent asymmetry in terms of access to information. We want to show under this assumption also the sender can manipulate the receiver's choice by appropriately choosing learning and communication strategies.

Let $\pi(s, \omega)$ denote the signal structure chosen by a DM to update his belief about the state ω, where $s \in S$ denotes a typical signal from the set of possible signals S. WLOG we can consider a set of signals as a set of possible actions, i.e., $S = A$ [3]. Let γ^i denote the posterior belief upon observing signal a_i,

$$\gamma^i(\omega_j) = Pr(\omega_j | s = a_i) = \frac{\pi(a_i, \omega_i)\mu_0(\omega_j)}{\sum_k \pi(a_i, \omega_k)\mu_0(\omega_k)}$$

As shown by Matejka and McKay (2015), Caplin, Dean, and Leahy (2018) we can ab-

[2]This denotes the level of negative externality of choosing the technology that the receiver does not consider

[3]If the DM chooses a signal structure that generates two separate signals for the same action then such a signal structure is equally Blackwell informative as the one where each signal generates a unique action.

stract away from the information structure $\pi(a_i, \omega)$ and consider directly the posterior distribution γ^i generated by the signal structure. This is because if two separate signal structures generate the same posterior distribution, the DM would choose the one with lower cost as they are equally Blackwell informative. Since $S = A$ this implies only one action is chosen at any posterior.

We can also define corresponding choice probabilities given any posterior belief over state as $P(a, \omega)$, i.e, the conditional (posterior) probability of choosing action a in state ω and $P(a)$, i.e., the unconditional (prior) probability of choosing action a. By Bayes plausibility,

$$\sum_k P(a_i, \omega_k)\mu_0(\omega_k) = P(a_i)$$

Following the tradition of the RI literature, we define the cost of learning function over choice probabilities $P(a, \omega)$ instead of γ^i directly. Since each action is chosen only at one posterior, the relation between the two objects is given by,

$$\gamma^i(\omega_j) = \frac{P(a_i, \omega_j)\mu_0(\omega_j)}{\sum_k P(a_i, \omega_k)\mu_0(\omega_k)}.$$

The cost of learning function is given by the Shannon mutual entropy between the conditional and unconditional choice probabilities,

$$K(\lambda, \mu_0) = \lambda D(P(a_i, \omega_j)||P(a_i)); \quad \text{where} \quad D(p||q) = \sum_x p(x) \ln \frac{p(x)}{q(x)}$$

and $\lambda \in (0, \infty)$ denotes the marginal cost of learning. From Matejka and McKay (2015) we know the optimal choice probabilities take the logistic form as follows,

$$P(a_i, \omega_j) = \frac{P(a_i)z(a_i, \omega)}{\sum_k P(a_k)z(a_k, \omega_j)}; \quad \text{where} \quad z(a_k, \omega_i) = \exp(u(a_k, \omega_k)/\lambda). \tag{logistic solution}$$

We assume that the payoff functions and the cost functions of both the receiver and the sender are common knowledge.

2.3 Strategic Communication

The communication problem is as follows:

1. All DMs enter with the common prior $\mu_0 \in \Delta(\Omega)$
2. The sender chooses a learning strategy
3. The sender chooses a communication strategy

4. receiver chooses optimal learning strategy
5. receiver chooses the optimal action
6. Payoffs are realized for all DMs

The learning technology available to both the DMs with common λ. The communication strategy of the sender consists of two types of actions, recommend an action a or not recommend at all. The separation of the learning and communication strategy of the sender allows that truth-telling is not necessary. Moreover, we assume that communication is not costly for the sender and any recommendation can be costlessly verified by the receiver. Thus the sender cannot lie but can hide information [4].

Note that, since learning is costly, it is without a loss that we can assume receivers choose their learning strategy after observing the information provided by the sender. This is true because in the state where after obtaining the recommendation from the sender if the receiver decided to obey the recommendation, then learning before the recommendation will generate sunk cost of learning (Ghosh, 2024).

2.4 Decision Problem

Given the cost of the learning function, the receiver wants to maximize the net expected utility conditional on information obtained from the sender. The one-to-one relationship between the posterior belief over states and conditional choice probabilities allows us to write the receiver's strategy as only choosing $P(a, \omega)$ optimally subject to the recommendation of the sender. Let μ denote the interim belief of the receiver upon observing the (no) recommendation of the sender. Thus the receiver's problem is as follows:

$$\max_{P(a,\omega)} E_\mu u(P(a, \omega)) - K(\lambda, \mu). \tag{1}$$

Given the decision problem of the receiver the sender chooses a learning strategy $\gamma : \Omega \to \Delta(\Omega)$ and a recommendation strategy $\sigma : \Delta(\Omega) \to A \cup \emptyset$ to maximize sender's payoff as follows:

$$\max_{\sigma,\gamma} v(P(a, \omega|\sigma(\gamma)), \omega) - K(\lambda, \mu_0) \tag{2}$$

where $P(a, \omega|\sigma)$ denote the optimal action chosen by the receiver where the interim belief μ is obtained using γ and $\sigma(\gamma)$.

[4]This can also be implemented by assuming a negligible cost of verification for the receiver.

2.5 Subgame Perfect Equuilibrium

To solve for the equilibrium we assume that before the state is realized the sender can commit to a learning and communication strategy. The subgame perfect equilibrium strategy of this comunication game is given by the triplet $(\gamma^*, \sigma^*, P^*(a, \omega))$ such that

1. $P^*(a, \omega)$ solves 1 for the receiver.
2. γ^*, σ^* jointly solve 2 for the sender.

For the rest of the paper we will only consider the SPNE of this communication game.

3 Results

3.1 Learning Problem

3.1.1 Receiver's Learning Startegy

To solve for the optimal recommendation policy $\sigma(\omega)$ for the sender we will approach the problem backward and solve the learning problem of the receiver first. Given the learning strategy, we will find the optimal recommendation and learning strategy of the sender.

Lemma 1. *receivers' optimal learning strategy divides the simplex over states,* $\Delta(\Omega)$ *into distinct consideration sets, and within each consideration set the optimal distribution of posterior beliefs is constant, i.e., independent of specific interim belief* μ*.*

The proof of the lemma is given in the appendix, here we describe the main intuition using figure 2. The proof of the lemma lies on a property of a class of cost function, namely, *uniform posterior separable* cost function, of which the Shannon entropy cost is a key example. The property, known as *likelihood invariant posterior*, or LIP, (refer to Caplin, Dean, and Leahy (2017)), indicates that for two decision problems the optimal posterior obtained in the first decision problem remains optimal in the second decision problem if the prior belief in the second decision problem keeps the optimal posterior from the first problem feasible.

In the two-state, two-action example, the optimal posterior is obtained by concavifying the net value function of the two actions. The LIP property means if the new prior lies in the interval joining the two optimal posteriors (for the two states) then the concavification process at the new prior would generate the same set of posteriors. The diagram below illustrates,

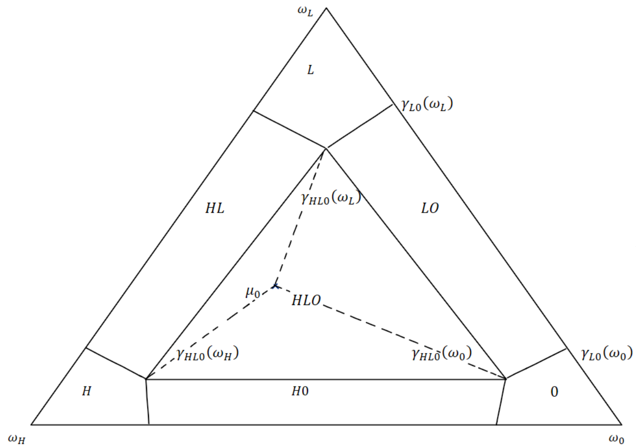

Figure 2: receiver's problem: consideration set

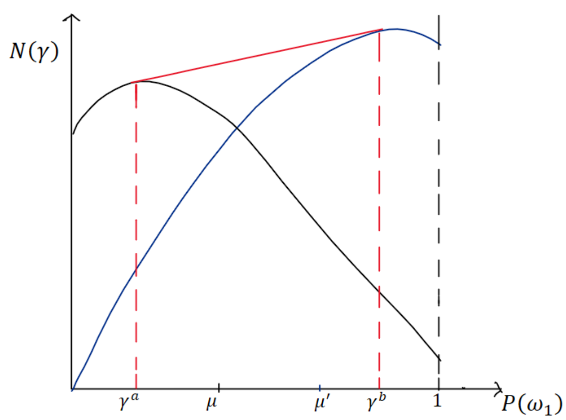

Figure 1: LIP: two-state, two-action problem

In figure 1, the black and the blue curve denote the net expected payoff from the two actions for different values of the probability of state ω_1. The red line concavifies the net expected payoff from the two actions. Suppose in the first problem, the prior is μ and the concavification generates the optimal posteriors as γ^a and γ^b. If in the second problem the prior μ' lies within $[\gamma^a, \gamma^b]$, i.e., the optimal posterior is *admissible* under the new prior, then the optimal posterior under μ' would also be given by γ^a and γ^b, since the concavification at μ' generates the same result at μ' as in μ.

We extend that in our example of three states and actions. In figure 2, consider any point in the interior of $\triangle HL0$ as the interim belief of the receiver given the sender's recommendation policy. Suppose the optimal posterior belief generated by this belief are the extreme points of this triangle, γ_{HL0}. By LIP this means for all interim belief $\mu \in int(\triangle HL0)$ the posterior belief thus generated is admissible and hence for all such beliefs, the optimal posterior would be generated by the same set of posteriors, namely the extreme points of $\triangle HL0$.

Since all beliefs in the interior of $\triangle HL0$ generate the same optimal posterior, we can find the extreme points of $\triangle HL0$ by assuming all three actions have equal unconditional choice probabilities. The rest of the proof shows, using LIP, a similar concavification argument is applicable for all beliefs in the interior of the triangle $HL0$, including the one that generates equal ex-ante choice probabilities. Also, for any belief on the boundary of $\triangle HL0$ or outside, at least one action is chosen with zero probability, generating the consideration set for all three actions.

Using similar logic we identify the consideration set for every pair of actions and residually identify the consideration set for single actions as well. Note that, the consideration sets need not be symmetric since the payoff structure is not symmetric. The process of generating the consideration set is general and can be applied to any finite set of states. Also, the consideration set that we identify here coincides with the description given in Caplin, Dean, and Leahy (2018), but our method is computationally easier.

For the rest of the paper, let us assume the common prior $\mu_0 \in \triangle HL0$, i.e., in absence of the sender's recommendation the receiver will choose all actions with positive probability. The model can solve all other cases, but this would be the most interesting case.

3.1.2 Sender's Learning Problem

Given the learning strategy of the receiver for a given interim belief $\mu : \sigma(\mu) \to \Delta(\Omega)$, we can solve for the sender's optimal learning strategy and the optimal recommendation strategy. We first show that learning is not optimal for the sender if the level of externality is substantially low.

Lemma 2. *Given any $\nu > 0$ there exists a $\lambda(\nu) > 0$ such that learning is preferred to not learning for $\lambda \leq \lambda(\nu)$, i.e., cost of learning is sufficiently small.*

The proof of the lemma is given in the appendix. For the rest of the paper, we will assume that for any ν, $\lambda \leq \lambda(\nu)$ such that learning is preferable by the sender.

3.2 Non-Strategic Communication

3.2.1 Optimal Strategy under Obedience

The following lemma shows that under misalignment of preferences, obedience is not an optimal response of the receiver. Obedience in this model refers to a strategy where the sender learns about each state and truthfully recommends the appropriate action and following this recommendation, the receiver chooses the recommended action without further learning. For example, if the sender's posterior belief suggests a_0 is the optimal action, he will recommend a_0, and the receiver will choose a_0 without further learning.

Lemma 3. *Obedience cannot be an equilibrium strategy.*

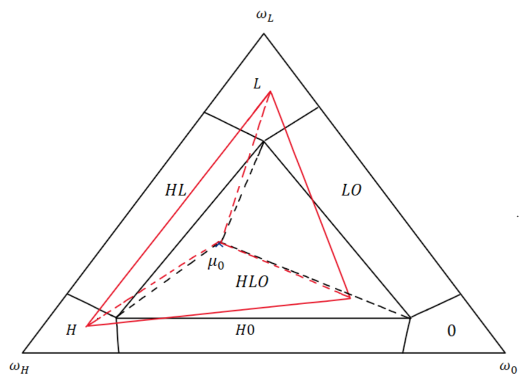

Figure 3: sender's problem: consideration set (in red) under complete compliance

The proof of the lemma is given in the appendix, here we explain the main intuition using figure 3. Since the learning technology for the sender and the receivers are identical, we can solve the sender's problem using a similar technique as described in Lemma 1. But due to the misalignment of preferences, the consideration set is different for the sender, ie., for certain beliefs when the receiver prefers to choose action a_j the sender may not recommend choosing action a_j.

In figure 3 the red triangle denotes the consideration set for all three actions for the sender under complete compliance. As can be seen from the diagram, if the sender truthfully reveals the state, the intermediate belief of the receiver would be such that following a_H or a_L recommendation the belief would lie in the consideration of H or L alone, resp. But for the recommendation of a_0, the optimal posterior belief lies outside the set □0, i.e., of choosing a_0 with probability 1, the receiver will choose to learn, following a_0 recommendation. Thus complete compliance cannot be an equilibrium here.

3.2.2 Optimal Strategy under Truth-telling

Before we solve the optimal strategy for the sender, let us consider the restricted strategy space for the sender. Suppose the sender is forced to truthfully communicate the optimal action in each state. As shown in the previous section complete compliance is not possible for the recommendation of a_0. To find the equilibrium under truthful communication the sender incorporates this into his strategy.

Note that the restriction is imposed on the communication strategy, not the learning strategy of the sender. Thus it is possible for the sender to not learn at all and recommend all actions with probability μ_0. This would be an uninformative equilibrium since the communication would not affect the intermediate belief of the receiver. We assume away from such equilibrium and only concentrate on informative equilibrium, i.e, where $\mu(a_i) \neq \mu_0$ for any recommendation a_i.

The following lemma incorporates the insight from the previous section and shows that even under truthful communication the sender would not choose an informative enough learning strategy that ensures no learning by the receiver.

Lemma 4. *Under the assumption of truthful communication, every informative equilibrium induces learning by the receiver following only the recommendation* a_0.

Proof of the lemma is given in the appendix. The main intuition is as follows: the equilibrium under truthful communication can take one of three forms. In all cases, the receiver follows recommendation of a_H and a_L and the sender chooses the learning strategy such that $\mu(a_H) \in \square H$, $\mu(a_L) \in \square L$. But for a_0, there are three feasible strategies, either $\mu(a_0) \in \triangle HL0$ and receivers consider all three action with positive probability when learning, or $\mu(a_0) \in \square L0$ and receivers choose a_H with probability 0 following recommendation of a_0 or $\mu(a_0) \in \square H0$ and receivers choose a_L with probability 0 following recommendation of a_0. We show that each strategy generates fixed point mapping making them equilibrium strategies as well.

3.3 Strategic Communication

In this section, we consider the optimal strategy of a persuasive sender under commitment. We maintain our assumption that any recommendation can be costlessly verified but hiding information is possible. Thus the commitment in this framework reduces to committing to a mixed strategy under no recommendation since no verification is feasible in that case.

We find that under commitment the sender never optimally mixes a_H with a_0 but mixing a_L and a_0 can be optimal. Before we explore the optimal strategy of the sender let us write the expected payoff of the sender under commitment. Given $p \in [0,1]$ and actions $a_i, a_j \in A$ let $\gamma_{p,ij}, \sigma_{p,ij}$ denote a strategy whereupon observing signals other than a_i or a_j the receiver complies, but under no recommendation the receiver's belief would be such that the sender has observed a_i with probability p, and a_j with probability $1-p$. Let

$$V(\gamma_{p,ij}, \sigma_{p,ij}) = \max \gamma_{p,ij}, \sigma_{p,ij} v(P(a,\omega), \omega | \gamma_{p,ij}) - K(\lambda, \mu_0)$$

denote the expected social welfare following $\gamma_{p,ij}, \sigma_{p,ij}$. The following theorem outlines the structure of the equilibrium under commitment.

Theorem 1. *Under commitment, in equilibrium the sender*

1. *Recommend* a_H *upon observing signal* a_H
2. *Recommend* a_L *upon observing signal* a_L *with probability* $\hat{q} \in (0, 1)$
3. *No recommendation otherwise.*

Following this the receiver optimally chooses

1. *Follow recommendation for both* a_H *and* a_L
2. *Given no recommendation, update his belief to* $\gamma_{\hat{p},L0}$ *and learn using the consideration set* $\square L0$.

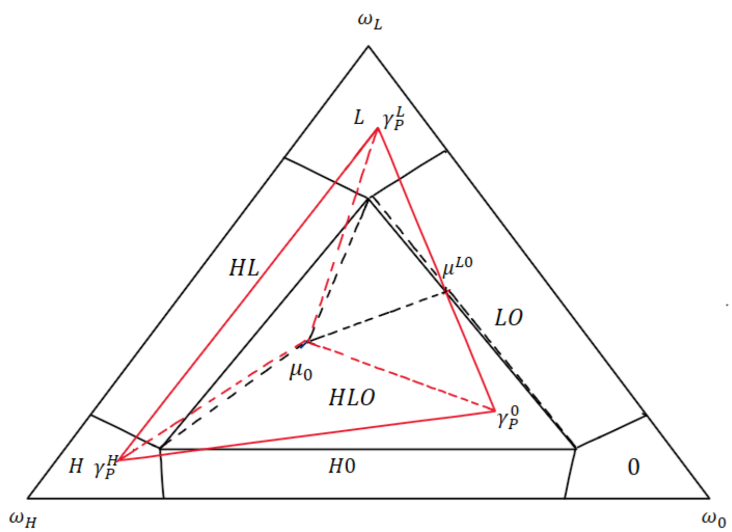

Figure 4: Equilibrium under Commitment

The proof of the proposition is given in the appendix. Here we explain the main intuition of the result using figure 4. The truthful communication shows that compliance does not occur only for a_0, thus in any equilibrium, under compliance, the sender would try to improve the payoff for a_0 signal. Since the learning and the communication strategy can be separated (through no recommendation), given commitment, the sender can be better off by mixing a_0 with either a_L.

To show that such an equilibrium exists, we assume that for any $q \in (0, 1)$ a posterior distribution γ_P with the above structure exists. We use this belief to find the optimal learning choice of the receiver and plug this into the sender's learning strategy. This formalizes as a fixed point problem for the sender's strategy γ. We find sufficient conditions in terms of q such that a fixed point exists. The optimal choice $\hat{q}$ maximizes the payoff for the sender.

3.4 Discussion of Assumptions

The crucial assumption in our model is that there is no *real* asymmetry of information since all DMs in the economy can learn about the underlying state if they pay the cost of learning. This reflects the nature of scientific information, i.e., such information is costly to obtain for any agent in the economy. However, our results hold true for a sufficiently small difference between the cost of learning parameters for the sender and the receiver.

We introduce a standard RI cost function, given by Shannon mutual entropy in an otherwise standard information design framework with three actions and three states. The characterization of the optimal policy relies heavily on the form of the cost function chosen here. However, the intuition behind the consideration set does not rely on the specific form of the cost function. As shown in lemma 1, we derive the consideration sets using the LIP property, which identifies a much larger set of cost functions, namely the UPS cost function, of which Shannon entropy is one example only. The concavification argument applies to any cost function in this class of functions.

But without the explicit form of the cost function, we would not be able to identify the extreme points of the consideration sets, even though we can show the existence of such points. Earlier papers in the literature (Gentzkow and Kamenica (2014) Matyskova and Montes (2023), Bloedel and Segal (2018)) have used the Shannon entropy function for obtaining closed-form analytical solutions. The properties of Shannon cost functions make it suitable for integrating the information design framework.

We continue with the assumption of commitment as is standard in the persuasion literature however, we weaken it by only considering commitment only applies for no recommendation. This requires assuming a negligible cost of verification, which may not be realistic in all cases. However, solving the optimization problem is significantly more difficult than verifying whether a solution is optimal, our assumption holds in most cases.

In this paper, we assume that both sides face the same cost of learning. If only the receiver faces the costs of learning, the sender will optimally choose the extreme points of the consideration set and ensure that the receiver never learns (see Matyskova and Montes (2023)). Whereas if the receiver does not have access to learning and the sender faces a cost of learning, then the optimal strategy would be given by the sender's consideration sets and obedience under commitment. Failure of obedience would be suboptimal since the posterior generated by the sender's learning is more informative than the prior and the receiver does not have access to information. Thus the both-sided cost is crucial in our framework and we find that even in this case the sender can manipulate the action of the receiver by strategically choosing the consideration set for the receiver.

4 Conclusion

We study a model of strategic communication between a sender and receivers in an economy in presence of externality. Under the assumption that learning is costly, we solve for the optimal information policy of the sender. We find that the sender truthfully reveals the signal when the action is less preferred by him. Otherwise, he strategically manipulates the consideration set of receivers leading to welfare improvement.

Our results lie on the assumption of the Shannon entropy cost function. One future direction would be to relax this assumption and study the optimal policy of the sender for a larger set of cost functions.

Also, we assume that the receivers in this model do not face any cost of verification following any recommendation. It would be interesting to explore what would happen if we relax the assumption and assume that the receiver pays a small cost of verification, smaller than the cost of learning.

Bibliography

Bloedel, A. W. and Segal, I. R. (2018). Persuasion with rational inattention. *Available at SSRN 3164033*.

Caplin, A. and Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *The American Economic Review*, 105(7):2183–2203.

Caplin, A., Dean, M., and Leahy, J. (2017). Rationally inattentive behavior: Characterizing and generalizing shannon entropy. Technical report, National Bureau of Economic Research.

Caplin, A., Dean, M., and Leahy, J. (2018). Rational Inattention, Optimal Consideration Sets, and Stochastic Choice. *The Review of Economic Studies*, 86(3):1061–1094.

Caplin, A., Dean, M., and Leahy, J. (2019). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies*, 86(3):1061–1094.

Gentzkow, M. and Kamenica, E. (2014). Costly persuasion. *American Economic Review*, 104(5):457–462.

Goldstein, I. and Leitner, Y. (2018). Stress tests and information disclosure. *Journal of Economic Theory*, 177:34–69.

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

Kolotilin, A., Mylovanov, T., Zapechelnyuk, A., and Li, M. (2017). Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964.

Leitner, Y. and Williams, B. (2020). Model secrecy and stress tests. *Available at SSRN 3606654*.

Lleras, J. S., Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2017). When more is less: Limited consideration. *Journal of Economic Theory*, 170:70–85.

Manzini, P. and Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica*, 82(3):1153–1176.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98.

Matyskova, L. and Montes, A. (2023). Bayesian persuasion with costly information acquisition. *Journal of Economic Theory*, 211:105678.

Taneva, I. (2019). Information design. *American Economic Journal: Microeconomics*, 11(4):151–85.

Wang, Y. (2015). Bayesian Persuasion with Multiple Receivers. Working Papers 2015-03-24, Wang Yanan Institute for Studies in Economics (WISE), Xiamen University.

A Appendix: Proofs

A.1 Proof of Lemma 1

Proof. To prove the lemma let us first note some important properties of the cost function. The cost function $K(.)$ is proportional to the Shannon mutual entropy cost function, which belongs to the class of uniform posterior separable (UPS) cost functions, as defined in Caplin, Dean, and Leahy (2017). The UPS cost functions are described as follows:

$$K(\mu, Q) = \sum_{\gamma \in \Gamma(Q)} Q(\gamma)T(\gamma) - T(\mu)$$

where $T(.)$ is convex. In the case of the Shannon cost function, the convex function is given by the Shannon entropy function. The defining characteristic for UPS cost functions is the Likelihood Invariant Posterior (LIP) property. LIP implies that for a given decision problem (μ, A) is a posterior distribution $\gamma \in \Delta(\Delta(\Omega))$ is obtained as the optimal learning strategy then it will remain an optimal posterior for any μ' such that γ is admissible under μ' using Bayes law.

Using the LIP property we want to argue that we can divide the simplex over the state space $\Delta(\Omega)$ into the consideration set. A consideration set for action i here defines a set of possible prior beliefs where only action i is chosen. In our 3-action example, we can thus find 7 consideration sets.

Let us start with the consideration set where all three actions are chosen. One member of such a consideration set would be a belief, say μ_{HL0} such that $P(a_H) = P(a_L) = P(a_0) = 1/3$, i.e., ex-ante all actions are chosen with equal probabilities. Plugging these values in *logisticsolution* we get,

$$P(a, \omega) = \frac{z(a, \omega)}{\sum_k z(a_k, \omega)},$$

This implies for an receiver with prior μ_{HL0} it is optimal to choose a learning strategy where upon observing signal i the receiver will choose action i and the probability of obtaining signal i is state ω (this is equivalent to the probability of choosing action i in state ω) is described by the above equation. Note that, $\mu_{HL0} \neq \mu_0$, thus to find the μ_{HL0} that generates an equal probability of choosing any action we solve the following set of three equations,

$$
\begin{aligned}
\mu_{HL0}(\omega_0)P(a_0,\omega_0)+\mu_{HL0}(\omega_L)P(a_0,\omega_L)+\mu_{HL0}(\omega_H)P(a_0,\omega_H)&=P(a_0)=1/3\\
\mu_{HL0}(\omega_0)P(a_L,\omega_0)+\mu_{HL0}(\omega_L)P(a_L,\omega_L)+\mu_{HL0}(\omega_H)P(a_L,\omega_H)&=P(a_L)=1/3\\
\mu_{HL0}(\omega_0)P(a_H,\omega_0)+\mu_{HL0}(\omega_L)P(a_H,\omega_L)+\mu_{HL0}(\omega_H)P(a_H,\omega_H)&=P(a_H)=1/3
\end{aligned}
$$

Plugging in the values of $P(a,\omega)$ we solve for μ_{HL0}. Given μ_{HL0} we can find the posterior belief upon observing signal i as follows,

$$\gamma^i_{HL0}(\omega_j)=\frac{P(a_i,\omega_j)\mu_{HL0}(\omega_j)}{\sum_k P(a_i,\omega_k)\mu_{HL0}(\omega_k)}.$$

This would be the optimal strategy of the DM only if the prior μ_0 makes the γ_{HL0} admissible. To prove that let us consider the convex set generated by the three posterior beliefs γ^k_{HL0} for $k=H,L,$ and 0. This forms a triangle within the simplex $\Delta(\Omega)$, which we will denote as $\triangle HL0$. Consider any arbitrary prior belief $\mu'\in int(\triangle HL0)$. For any such μ' the posteriors γ^i_{HL0} are feasible, i.e., can be obtained using the Bayes rule for some unconditional choice probability distribution $P(a_i)$.

Furthermore, since μ_{HL0} can be written as a convex combination of γ^k_{HL0} we know that $\mu_{HL0}\in int(\triangle HL0)$. Hence, for all $\mu'\in int(\triangle HL0)$ the optimal learning strategy is given by γ^k_{HL0}. Thus $\triangle HL0$ generates the consideration set of all three actions. If the common prior $\mu_0\notin int(\triangle HL0)$, then μ_0 cannot be written as a convex combination of γ^i. But this would imply at μ_0 at least one action is chosen with zero probability, contradicting our assumption that at μ_0 all three actions are chosen with strictly positive probability.

Next, we consider the consideration set for two actions only. WLOG let us investigate the consideration set for actions a_H and a_L, this includes all prior beliefs μ' where $P(a_0)=0$. To construct this set let us start with the extreme points of $\triangle HL0$. Consider prior beliefs $\mu_1=\gamma^H_{HL0}$ and $\mu_2=\gamma^L_{HL0}$. We know the receivers choose a_H with probability 1 if his belief is ar μ_1 and similarly a_L with probability 1 for μ_2. Thus for any arbitrary belief $\mu'\in int(\overline{\gamma^H_{HL0}\gamma^L_{HL0}})$ would belong to the consideration set of a_H and a_L only.

Note that, if $P(a_0)=0$ then the DM only obtains two signals a_H and a_L. Thus the resultant posterior belief over states would have only two beliefs in the support, one for each signal a_H and a_L. Both these two belief would assign a probability of zero to state $\omega_0$since $P(a_0,\omega_H)=P(a_0,\omega_L)=0$. Thus we can conclude that the optimal posterior belief would lie on the boundary of the simplex where $\mu(\omega_0)=0$, where $P(a_0)=0$.

On this line segment, we can find the posterior beliefs for choosing only a_H and

a_L using the same method as before and assuming $P(a_H) = P(a_L) = 1/2$. This generates the other boundary of the consideration set for a_H and a_L. Joining the two sets of extreme points we find the consideration set HL which takes the shape of a trapezoid. Similarly, we find the consideration set of all other sets of two actions. The consideration sets for single actions can then be characterized residually defined by the boundary of the two-action consideration sets.

Caplin, Dean, and Leahy found that the consideration set can be characterized by finding the convex set at the intersection of a set of linear equations. For example, to find γ_{HL0}^{H} we need to consider the intersection of the two following equations,

$$f(\gamma, a_H, a_L) = 1 \quad \text{and} \quad f(\gamma, a_H, a_0) = 1$$

where

$$f(\gamma, a, b) = \sum_{\omega} \frac{z(b,\omega)}{z(a,\omega)} \gamma(\omega).$$

It is straightforward to verify that at γ_{HL0}^{H} indeed $f(\gamma, a_H, a_L) = f(\gamma, a_H, a_0) = 1$. Thus both the methods generate the same set of consideration sets but instead of using a set of linear equations we derive the extreme points of the consideration by applying LIP and solving for the optimal posterior when all actions are chosen with equal probabilities. $\square$

A.2 Proof of Lemma 2

For any $\nu > 0$ if $\lambda \to 0$, learning is not costly and by learning the sender can get a better payoff. The net payoff function under learning is given by,

$$V_L = E\hat{u}(P(a,\omega)|\mu_0) - K(\lambda, \mu_0)$$

As λ increases since $K(.) = \lambda D(P(a,\omega)||P(a))$, the cost of learning increases. Also, as λ increases the probability of mismatching state increases, which implies $E\hat{u}(P(a,\omega)|\mu_0)$ decreases with λ. However, for high λ the benefit from not learning and leaving it on the receiver is also smaller since the receiver doesn't learn as much as well. This implies, that there exists λ sufficiently small such that learning is better than not learning for the sender.

A.3 Proof of Lemma 3

Proof. Let us consider the learning problem of the sender under complete compliance, i.e., the receivers obediently follow the recommendation of the sender. Complete compliance thus implies if the sender recommends a_i the receivers will choose a_i with probability 1. The learning problem of the sender and the receivers is identical except $z(a, \omega)$. Let us denote $z_C(a, \omega) = \exp(v(a, \omega)/\lambda)$, i.e., the $z(.)$ counterpart for the sender under compliance. Since there is no strategic interaction, the sender always recommends the signal he obtains, and the optimal posterior belief distribution is given by,

$$\gamma_C(a_i, \omega_j) = \frac{z_C(a_i, \omega_j)\mu(\omega_j)}{P_C(a_h)z_P(a_H, \omega_j) + P_C(a_L)z_P(a_L, \omega_j) + P_C(a_0)z_P(a_0, \omega_j)}$$

If complete compliance is an equilibrium then we will get,

$$\begin{aligned}
\gamma_C^H &= (\gamma_C(a_H, \omega_0), \gamma_C(a_H, \omega_L), \gamma_C(a_H, \omega_0)) \in \Box H \\
\gamma_C^L &= (\gamma_C(a_L, \omega_0), \gamma_C(a_L, \omega_L), \gamma_C(a_L, \omega_0)) \in \Box L \\
\gamma_C^0 &= (\gamma_C(a_0, \omega_0), \gamma_C(a_0, \omega_L), \gamma_C(a_0, \omega_0)) \in \Box 0
\end{aligned}$$

If any belief $\gamma \in \Box H$ then γ will satisfy the following two inequalities that outline the boundary of $\Box H$,

$$\begin{aligned}
f(\gamma^H, a_H, a_L) \leq 1 \\
f(\gamma^H, a_H, a_0) \leq 1
\end{aligned}$$

where

$$f(\gamma, a, b) = \sum_{\omega} \frac{z(a, \omega)}{z(b, \omega)} \gamma(\omega).$$

We can rewrite the above inequalities as follows,

$$\begin{aligned}
f(\gamma_C^H, a_H, a_L) = \frac{z(a_L, \omega_0)}{z(a_H, \omega_0)} \frac{z_C(a_H, \omega_0)\mu(\omega_0)}{\sum_{a \in A} P(a) z_C(a, \omega_0)} + \frac{z(a_L, \omega_L)}{z(a_H, \omega_L)} \frac{z_C(a_H, \omega_L)\mu(\omega_0)}{\sum_{a \in A} P(a) z_C(a, \omega_L)} \\
+ \frac{z(a_L, \omega_H)}{z(a_H, \omega_H)} \frac{z_C(a_H, \omega_H)\mu(\omega_H)}{\sum_{a \in A} P(a) z_C(a, \omega_H)} \leq 1 \\
f(\gamma_C^H, a_H, a_0) = \frac{z(a_0, \omega_0)}{z(a_H, \omega_0)} \frac{z_C(a_H, \omega_0)\mu(\omega_0)}{\sum_{a \in A} P(a) z_C(a, \omega_0)} + \frac{z(a_0, \omega_L)}{z(a_H, \omega_L)} \frac{z_C(a_H, \omega_L)\mu(\omega_0)}{\sum_{a \in A} P(a) z_C(a, \omega_L)} \\
+ \frac{z(a_0, \omega_H)}{z(a_H, \omega_H)} \frac{z_C(a_H, \omega_H)\mu(\omega_H)}{\sum_{a \in A} P(a) z_C(a, \omega_H)} \leq 1
\end{aligned}$$

Rearranging terms we get,

$$\begin{aligned}\frac{z(a_L,\omega)}{z(a_H,\omega)}\frac{z_C(a_H,\omega)\mu(\omega)}{\sum_{a\in A}P(a)z_C(a,\omega)} &= \frac{z_C(a_H,\omega)}{z(a_H,\omega)}\frac{z(a_L,\omega)}{z_C(a_L,\omega)}\frac{z_C(a_L,\omega)\mu(\omega)}{\sum_{a\in A}P(a)z_C(a,\omega)} \\ &= \frac{z_C(a_H,\omega)}{z(a_H,\omega)}\frac{z(a_L,\omega)}{z_C(a_L,\omega)}\gamma_C(a_L,\omega)\end{aligned}$$

This implies,

$$\begin{aligned}f(\gamma_C^H,a_H,a_L) &= \exp\frac{-\nu}{\lambda}\exp\frac{\nu}{\lambda}\gamma_C(a_L,\omega_0)+\exp\frac{-i\nu}{\lambda}\exp\frac{i\nu}{\lambda}\gamma_C(a_L,\omega_L)+\gamma_C(a_L,\omega_H) \\ &= \gamma_C(a_L,\omega_0)+\gamma_C(a_L,\omega_L)+\gamma_C(a_L,\omega_H)=1,\end{aligned}$$

hence, the first inequality is satisfied. Similarly, we can show,

$$\begin{aligned}\frac{z(a_0,\omega)}{z(a_H,\omega)}\frac{z_C(a_H,\omega)\mu(\omega)}{\sum_{a\in A}P(a)z_C(a,\omega)} &= \frac{z_C(a_H,\omega)}{z(a_H,\omega)}\frac{z(a_0,\omega)}{z_C(a_0,\omega)}\frac{z_C(a_0,\omega)\mu(\omega)}{\sum_{a\in A}P(a)z_C(a,\omega)} \\ &= \frac{z_C(a_H,\omega)}{z(a_H,\omega)}\gamma_C(a_0,\omega)\end{aligned}$$

since $z(a_0,\omega)=z_C(a_0,\omega)=1$. This implies

$$f(\gamma_C^H,a_H,a_0)=\exp\frac{-\nu}{\lambda}\gamma_C(a_0,\omega_0)+\exp\frac{-i\nu}{\lambda}\gamma_C(a_0,\omega_L)+\gamma_C(a_0,\omega_H)<1$$

since $\nu>0$ and $\sum_{\omega\in\Omega}\gamma_P(a_0,\omega)=1$. Thus $\gamma_P^H\in\Box H$.

Next, we consider γ_P^L which will lie in $\Box L$ if

$$\begin{aligned}f(\gamma_C^L,a_L,a_H)&\le 1\\ f(\gamma_C^L,a_L,a_0)&\le 1\end{aligned}$$

Using similar adjustments we can write,

$$\begin{aligned}f(\gamma_C^L,a_L,a_H) &= \sum_{\omega\in\Omega}\frac{z_C(a_L,\omega)}{z(a_L,\omega)}\frac{z(a_H,\omega)}{z_C(a_H,\omega)}\frac{z_C(a_H,\omega)\mu(\omega)}{\sum_{a\in A}P(a)z_P(a,\omega)} \\ &= \exp\frac{-\nu}{\lambda}\exp\frac{\nu}{\lambda}\gamma_C(a_H,\omega_0)+\exp\frac{-i\nu}{\lambda}\exp\frac{i\nu}{\lambda}\gamma_P(a_H,\omega_L)+\gamma_C(a_H,\omega_H) \\ &= \gamma_C(a_H,\omega_0)+\gamma_C(a_H,\omega_L)+\gamma_C(a_H,\omega_H)=1.\end{aligned}$$

Comparing a_L and a_0 we get,

$$\begin{aligned}f(\gamma^L,a_L,a_0) &= \sum_{\omega\in\Omega}\frac{z_C(a_L,\omega)}{z(a_L,\omega)}\frac{z_C(a_0,\omega)\mu(\omega)}{\sum_{a\in A}P(a)z_C(a,\omega)} \\ &= \exp\frac{-\nu}{\lambda}\gamma_C(a_0,\omega_0)+\exp\frac{-i\nu}{\lambda}\gamma_C(a_0,\omega_L)+\gamma_C(a_0,\omega_H)<1\end{aligned}$$

since $\nu > 0$ and $\sum_{\omega\in\Omega} \gamma_C(a_0, \omega) = 1$. This implies $\gamma_C^L \in \Box L$.

Now we consider γ_P^0 which will lie in $\Box 0$ if

$$f(\gamma_C^0, a_0, a_H) \leq 1$$
$$f(\gamma_C^0, a_0, a_L) \leq 1$$

Using similar adjustments as before we get,

$$\begin{aligned} f(\gamma_C^0, a_0, a_H) &= \sum_{\omega\in\Omega} \frac{z(a_H, \omega)}{z_C(a_H, \omega)} \gamma_C(a_H, \omega) \\ &= \exp\frac{\nu}{\lambda}\gamma_C(a_H, \omega_0) + \exp\frac{i\nu}{\lambda}\gamma_C(a_H, \omega_L) + \gamma_C(a_H, \omega_H) > 1 \end{aligned}$$

since $\nu, i > 0$ and $\sum_{\omega\in\Omega} \gamma_C(a_H, \omega) = 1$. Also,

$$\begin{aligned} f(\gamma_C^0, a_0, a_L) &= \sum_{\omega\in\Omega} \frac{z(a_L, \omega)}{z_C(a_L, \omega)} \gamma_C(a_L, \omega) \\ &= \exp\frac{\nu}{\lambda}\gamma_C(a_L, \omega_0) + \exp\frac{i\nu}{\lambda}\gamma_C(a_L, \omega_L) + \gamma_C(a_L, \omega_H) > 1 \end{aligned}$$

since $\nu, i > 0$ and $\sum_{\omega\in\Omega} \gamma_C(a_L, \omega) = 1$. Thus $\gamma_C^0 \notin \Box 0$. This implies if the sender communicates the signal truthfully the receiver will be better off by deviating and learning to follow a_0 recommendation. Hence, proved. □

A.4 Proof of lemma 4

Proof. In this lemma, we explore whether the receiver chooses to learn when the sender truthfully recommends the action based on the observed signal. Since learning is costly for the receiver, he only chooses to learn by himself if his interim belief is not in a consideration set that contains only one action, e..g, $\Box L$. The previous lemma shows even when receivers perfectly comply with the sender's recommendation sender chooses a learning strategy such that the posterior belief following the recommendation of a_H and a_L would lie in $\Box H$ and $\Box L$ respectively, i.e., the receiver would not have any incentive to learn further. However, the posterior belief following the recommendation of a_0 does not lie in $\Box 0$, where the receiver would optimally choose to learn.

Under truthful communication, since the receiver has the opportunity to learn after the recommendation, it would always be optimal for the sender to choose a learning strategy such that the posterior belief following a_H or a_L lies in $\Box H$ and $\Box L$ respectively. We would show the following strategy would be the optimal strategy under truthful communication:

- receiver complies with a_H and a_L

- receiver does not comply with a_0 recommendation

Let us consider a strategy as described above. Let v_T denote the expected payoff of the sender under truthful communication. Since following a_0 the receiver learns by himself, the sender's payoff needs to incorporate the receiver's choice. Suppose post learning the receiver's belief belongs to the consideration set $CS(JK)$, then the receiver's optimal posterior is given by the extreme points of the consideration sets, i.e.,

$$v_T(a,\omega) = P_{JK}(a_J)v(a_J,\omega) + P_{JK}(a_K)v(a_K,\omega))$$

where $P_{JK}(a)$ is unconditional probability of choosing action a_J in consideration set $CS(JK)$ given interim belief $\mu(\omega)$ is equal to sender's optimal posterior choice γ_T.

Under the strategy of compliance in a_H and a_L only, we will consider four possible strategies, namely,

- $\gamma_T^0 \in \triangle HL0$
- $\gamma_T^0 \in \square L0$
- $\gamma_T^0 \in \square H0$
- $\gamma_T^0 \in \square 0$

where γ_T^i denote the posterior belief of the sender following a signal a_i in the truthful communication game. We will show that only the first three are feasible strategies and will characterize the equilibrium strategy.

Let us consider the following strategy where a_H and a_L recommendations are followed but after a_0 the interim belief lies in $\triangle HL0$. Under this strategy γ_T, $v_T(a_H,\omega) = v(a_H,\omega)$ and $v_T(a_L,\omega) = v(a_L,\omega)$ for any $\omega \in \Omega$ but

$$v_T(a_0,\omega) = P_{HL0}(a_0)v(a_0,\omega) + P_{HL0}(a_L)v(a_L,\omega) + P_{HL0}(a_H)v(a_H,\omega) \quad \forall\omega \in \Omega.$$

where $P(a)_{HL0}$ denotes the probability of action a being chosen following a recommendation of a_0 and it is a function of γ_T. Following a similar exercise as of lemma 1, we can find the optimal learning strategy $\hat{\gamma}_T$ given $v_T(a,\omega)$ and prior belief μ_0. To show this is a feasible strategy, we will first check indeed $\gamma_T^H \in \square H$, $\gamma_T^L \in \square L$ and $\gamma_T^0 \in \triangle HL0$.

The necessary and sufficient condition for $\gamma_T^H \in \square H$ is given as follows

$$f(\gamma_T^H, a_H, a_L) \leq 1$$
$$f(\gamma_T^H, a_H, a_0) \leq 1$$

Using similar rearrangement of terms as in lemma 3 we get,

$$f(\gamma_T^H, a_H, a_L) = \sum_{\omega\in\Omega} \frac{z_T(a_H,\omega)}{z(a_H,\omega)} \frac{z(a_L,\omega)}{z_T(a_L,\omega)} \gamma_T(a_L,\omega)$$
$$= \sum_{\omega\in\Omega} \gamma_T(a_L,\omega) = 1$$

and,

$$\begin{aligned} f(\gamma_T^H, a_H, a_0) &= \sum_{\omega\in\Omega} \frac{z_T(a_H,\omega)}{z(a_H,\omega)} \frac{1}{z_T(a_0,\omega)} \gamma_T(a_0,\omega) \\ &= \exp \frac{-\nu - P_{HL0}(a_H)(\alpha_H - \beta_H - \nu) - P_{HL0}(a_L)(\alpha_L - \beta_L - \nu)}{\lambda} \gamma_T(a_0,\omega_0) \\ &+ \exp \frac{-i\nu - P_{HL0}(a_H)(\alpha_H - i\beta_H - i\nu) - P_{HL0}(a_L)(\alpha_L - i\beta_L - i\nu)}{\lambda} \gamma_T(a_0,\omega_L) \\ &+ \exp \frac{-P_{HL0}(a_H)\alpha_H - P_{HL0}(a_L)\alpha_L}{\lambda} \gamma_T(a_0,\omega_H) \end{aligned}$$

Note that, $\lim_{P_{HL0}(a_0)} f(\gamma_T^H, a_H, a_0) < 1$ and $f'(\gamma_T^H, a_H, a_L) > 0$ in $P_{HL0}(a_0)$. This implies the sender can choose a learning strategy such that $\gamma_T^H \square H$. It can be shown, under this condition $\gamma_T^L \in \square L$ and $\gamma_T^0 \in \triangle HL0$

Similarly, for $\gamma_T^L \in \square L$ the necessary and sufficient condition is

$$f(\gamma_T^L, a_L, a_0) \leq 1$$
$$f(\gamma_T^L, a_L, a_0) \leq 1$$

Using similar rearrangement of terms we get,

$$f(\gamma_T^L, a_L, a_H) = \sum_{\omega\in\Omega} \frac{z_T(a_L,\omega)}{z(a_L,\omega)} \frac{z(a_H,\omega)}{z_T(a_H,\omega)} \gamma_T(a_L,\omega)$$
$$= \sum_{\omega\in\Omega} \gamma_T(a_L,\omega) = 1$$

and

$$\begin{aligned} f(\gamma_T^L, a_L, a_0) &= \sum_{\omega\in\Omega} \frac{z_T(a_L,\omega)}{z(a_L,\omega)} \frac{1}{z_T(a_0,\omega)} \gamma_T(a_0,\omega) \\ &= \exp \frac{-\nu - P_{HL0}(a_H)(\alpha_H - \beta_H - \nu) - P_{HL0}(a_L)(\alpha_L - \beta_L - \nu)}{\lambda} \gamma_T(a_0,\omega_0) \\ &+ \exp \frac{-i\nu - P_{HL0}(a_H)(\alpha_H - i\beta_H - i\nu) - P_{HL0}(a_L)(\alpha_L - i\beta_L - i\nu)}{\lambda} \gamma_T(a_0,\omega_L) \\ &+ \exp \frac{-P_{HL0}(a_H)\alpha_H - P_{HL0}(a_L)\alpha_L}{\lambda} \gamma_T(a_0,\omega_H) \end{aligned}$$

which will also be less than 1 if $\gamma_T^H \in \square H$.

Finally we consider if $\gamma_T^0 \in \triangle HL0$. However, there is no necessary and sufficient condition based on $f(x,a,b)$. This is why we start with the necessary condition, $f(x,a,b) > 1$ for all $a,b \in A$ and $x = \gamma_T^0$.

$$f(\gamma_T^0, a_0, a_H) = \sum_{\omega\in\Omega} z_T(a_0,\omega)\frac{z(a_H,\omega)}{z_T(a_H,\omega)}\gamma_T(a_H,\omega) > 1$$

since $f(\gamma_T^H, a_H, a_0) < 1$. By similar logic $f(\gamma_T^0, a_0, a_L) > 1$. Additionally,

$$f(\gamma_T^0, a_H, a_0) = \sum_{\omega\in\Omega} \frac{z(a_0,\omega)}{z(a_H,\omega)}\frac{z_T(a_0,\omega)}{z_T(a_H,\omega)}\gamma_T(a_H,\omega) > 1$$

since $f(\gamma_T^H, a_H, a_0) < 1$. By similar logic $f(\gamma_T^0, a_L, a_0) > 1$.

$$f(\gamma_T^0, a_H, a_L) = \sum_{\omega\in\Omega} \frac{z(a_H,\omega)}{z_T(a_H,\omega)}\frac{z_T(a_0,\omega)}{z(a_L,\omega)}\gamma_T(a_H,\omega) > 1$$

since $f(\gamma_T^H, a_H, a_0) < 1$. By similar logic $f(\gamma_T^0, a_L, a_H) > 1$. However, since this is not the sufficient condition, we need to check if $\gamma_T(a_0,\omega_L) > t\gamma(a_0,\omega_L) + (1-t)\gamma(a_H,\omega_L)$ for any $t \in [0,1]$, i.e., γ^0 lies above $\overline{BC}$, $\gamma_T(a_0,\omega_H) > t\gamma(a_0,\omega_H) + (1-t)\gamma(a_L,\omega_H)$ for any $t \in [0,1]$, i.e., i.e., γ^0 lies to the left of $\overline{AB}$, and $\gamma_T(a_0,\omega_0) > t\gamma(a_H,\omega_H) + (1-t)\gamma(a_L,\omega_H)$ for any $t \in [0,1]$, i.e., i.e., γ^0 lies to the right og $\overline{AC}$. To determine this let us compare γ and γ_T. To begin with,

$$\begin{aligned}
\gamma_T(a_0,\omega_L) &= \frac{z_T(a_0,\omega_L)\mu(\omega_L)}{P_T(a_H)z_T(a_H,\omega_L) + P_T(a_L)z_T(a_L,\omega_L) + P_T(a_0)z_T(a_0,\omega_L)} \\
\gamma(a_0,\omega_L) &= \frac{z(a_0,\omega_L)\mu(\omega_L)}{P(a_H)z(a_H,\omega_L) + P(a_L)z(a_L,\omega_L) + P(a_0)z(a_0,\omega_L)} \\
\gamma(a_H,\omega_L) &= \frac{z(a_H,\omega_L)\mu(\omega_L)}{P(a_H)z(a_H,\omega_L) + P(a_L)z(a_L,\omega_L) + P(a_0)z(a_0,\omega_L)}.
\end{aligned}$$

Since $z_T(a_0,\omega_L) > z(a_0,\omega_L)$, $z_T(a_H,\omega_L) < z(a_H,\omega_L)$, $z_T(a_L,\omega_L) < z(a_L,\omega_L)$ and $P_T(a_0) < 1$ we get $\gamma_T(a_0,\omega_L) > \gamma(a_0,\omega_L)$. Since $z_T(a_0,\omega_L) > z(a_H,\omega_L)$, we also get $\gamma_T(a_0,\omega_L) > \gamma(a_H,\omega_L)$, which implies $\gamma_T(a_0,\omega_L) > t\gamma(a_0,\omega_L) + (1-t)\gamma(a_H,\omega_L)$ for any $t \in [0,1]$. Furthermore,

$$\begin{aligned}
\gamma_T(a_0,\omega_H) &= \frac{z_T(a_0,\omega_H)\mu(\omega_H)}{P_T(a_H)z_T(a_H,\omega_H) + P_T(a_L)z_T(a_L,\omega_H) + P_T(a_0)z_T(a_0,\omega_H)} \\
\gamma(a_0,\omega_H) &= \frac{z(a_0,\omega_H)\mu(\omega_H)}{P(a_H)z(a_H,\omega_H) + P(a_L)z(a_L,\omega_H) + P(a_0)z(a_0,\omega_H)} \\
\gamma(a_L,\omega_H) &= \frac{z(a_L,\omega_H)\mu(\omega_L)}{P(a_H)z(a_H,\omega_H) + P(a_L)z(a_L,\omega_H) + P(a_0)z(a_0,\omega_H)}.
\end{aligned}$$

Since $z_T(a_0,\omega_H) > z(a_0,\omega_H)$, $z_T(a_H,\omega_H) = z(a_H,\omega_L)$, $z_T(a_L,\omega_H) = z(a_L,\omega_H)$ and

$P_T(a_0) < 1$ we get $\gamma_T(a_0, \omega_H) > \gamma(a_0, \omega_L)$. Since $z_T(a_0, \omega_H) > z(a_L, \omega_H)$, we also get $\gamma_T(a_0, \omega_H) > \gamma(a_L, \omega_H)$, which implies $\gamma_T(a_0, \omega_H) > t\gamma(a_0, \omega_H) + (1-t)\gamma(a_L, \omega_H)$ for any $t \in [0, 1]$. Finally,

$$\gamma_T(a_0, \omega_0) = \frac{z_T(a_0, \omega_0)\mu(\omega_0)}{P_T(a_H)z_T(a_H, \omega_0) + P_T(a_L)z_T(a_L, \omega_0) + P_T(a_0)z_T(a_0, \omega_0)}$$
$$\gamma(a_H, \omega_0) = \frac{z(a_H, \omega_0)\mu(\omega_0)}{P(a_H)z(a_H, \omega_0) + P(a_L)z(a_L, \omega_0) + P(a_0)z(a_0, \omega_0)}$$
$$\gamma(a_L, \omega_0) = \frac{z(a_L, \omega_0)\mu(\omega_0)}{P(a_H)z(a_H, \omega_0) + P(a_L)z(a_L, \omega_0) + P(a_0)z(a_0, \omega_0)}.$$

Since $z_T(a_0, \omega_0) > z(a_H, \omega_0)$, $z_T(a_H, \omega_0) < z(a_H, \omega_0)$, $z_T(a_L, \omega_0) < z(a_L, \omega_0)$ and $P_T(a_0) < 1$ we get $\gamma_T(a_0, \omega_0) > \gamma(a_L, \omega_0)$. Since $z_T(a_0, \omega_0) > z(a_H, \omega_0)$, we also get $\gamma_T(a_0, \omega_0) > \gamma(a_H, \omega_0)$, which implies $\gamma_T(a_0, \omega_0) > t\gamma(a_L, \omega_0) + (1-t)\gamma(a_H, \omega_0)$ for any $t \in [0, 1]$. Thus $\gamma_T^0 \in \triangle HL0$ and does not lie on either of the boundaries.

To show this strategy can be an equilibrium strategy, we need to check whether there exists any $\gamma_T^0 \in \triangle HL0$ such that it is consistent with the learning strategy of the receiver given in the lemma 1. The following equation denotes the relationship between any interim belief $\mu' \in \triangle HL0$ and $P(a)$

$$\gamma(a_0, \omega) = \frac{P'(a, \omega)}{P'(a)}\mu'(\omega)$$

where $P'(a, \omega)$ and $P'(\omega)$ denote the conditional and unconditional probabilities resp. for belief, $\mu' \in \triangle HL0$ and γ denote the optimal choice of the receiver characterized in lemma 1. Thus $P(a)$ is a continuous function of μ', let us denote this function as g_1. Also, γ_T^0 is a continuous function in $P(a)$ for $\gamma_T^0 \in \triangle HL0$ since all three actions are chosen with strictly positive probability. Let us denote this function as g_2. The function $g_1 \circ g_2 : int(\triangle HL0) \to int(\triangle HL0)$ and is continuous. Note that, we can use the $\int(\triangle HL0)$ because the boundary points of $\triangle HL0$ are not contained in the consideration set of $HL0$ and we have already shown γ_T^0 does not lie on any boundary of $\triangle HL0$. This implies there exists a fixed point of the composite mapping $g_1 \circ g_2$ which would be the equilibrium belief γ_T^0.

Using similar arguments we can show that there exists an equilibrium where $\gamma_T^H \in \square H$, $\gamma_T^L \in \square L$ and $\gamma_T^0 \in \square L0$. The updated payoff function, in this case, would be,

$$v_T(a_0, \omega) = P_{L0}(a_0)v(a_0, \omega) + P_{L0}(a_L)v(a_L, \omega)$$

and the necessary and sufficient condition would be,

$$f(\gamma_T^H, a_H, a_0) = \exp\frac{-\nu - P_{L0}(\alpha_L)(\alpha_L - \beta_L - \nu)}{\lambda}\gamma_T(a_0, \omega_0) + \exp\frac{-i\nu - P_{L0}(\alpha_L)(\alpha_L - i\beta_L - i\nu)}{\lambda}\gamma_T(a_0, \omega_L) + \gamma_T(a_0, \omega_H) \leq 1.$$

Similarly one can show there exists an equilibrium where $\gamma_T^H \in \square H$, $\gamma_T^L \in \square L$ and $\gamma_T^0 \in \square H0$. The updated payoff function, in this case, would be,

$$v_T(a_0, \omega) = P_{H0}(a_0)v(a_0, \omega) + P_{H0}(a_H)v(a_H, \omega)$$

and the necessary and sufficient condition would be,

$$f(\gamma_T^H, a_H, a_0) = \exp\frac{-\nu - P_{H0}(\alpha_H)(\alpha_H - \beta_H - \nu)}{\lambda}\gamma_T(a_0, \omega_0) + \exp\frac{-i\nu - P_{H0}(\alpha_H)(\alpha_H - i\beta_H - i\nu)}{\lambda}\gamma_T(a_0, \omega_L) + \gamma_T(a_0, \omega_H) \leq 1.$$

Finally, we consider the strategy $\gamma_T^H \in \square H$, $\gamma_T^L \in \square L$ and $\gamma_T^0 \in \square 0$ the payoff function becomes,

$$v_T(a_0, \omega) = v(a_0, \omega).$$

By lemma 2, $\gamma_T^0 \notin \square 0$, thus such an equilibrium cannot exist.

Hence, proved.

$\square$

A.5 Proof of Theorem 1

Proof. Lemma 4 has established there are three types of equilibrium when the sender truthfully communicates the obtained signal. They are as follows

- Strategy $HL0$: $\gamma_P^H \in \square H$, $\gamma_P^L \in \square L$, and $\gamma_P^H \in \triangle HL0$
- Strategy $L0$: $\gamma_P^H \in \square H$, $\gamma_P^L \in \square L$, and $\gamma_P^H \in \square L0$
- Strategy $H0$: $\gamma_P^H \in \square H$, $\gamma_P^L \in \square L$, and $\gamma_P^H \in \square H0$,

where the subscript P denotes the sender's choice. Since the least amount of learning is needed for strategy $HL0$, this would be the cheapest strategy to implement. Also, since $\alpha_H - \beta_H < \alpha_L - \beta_L$, strategy $L0$ yields the highest expected gross payoff. In this strategic communication environment, we want to find a strategy such that the learning strategy resembles strategy $HL0$ but the communication strategy improves it to something akin to strategy $L0$.

Consider the following strategy. The sender's learning strategy is such that $\gamma_P^H \in \Box H$, $\gamma_P^L \in \Box L$, and $\gamma_P^H \in \triangle HL0$. However, the sender does not always truthfully recommend it. The communication strategy is as follows:

- Recommend a_H following signal a_H
- Recommend a_L with probability q following a_L
- No recommendation otherwise

The expected payoff of the sender following this strategy would be given by,

$$
\begin{aligned}
v_q(a_H, \omega) &= v(a_H, \omega) \\
v_q(a_0, \omega) &= P_q(a_L)v(a_L, \omega) + P_q(a_0)v(a_0, \omega) \\
v_q(a_L, \omega) &= qv(a_L, \omega) + (1-q)P_q(a_L)v(a_L, \omega) + P_q(a_0)v(a_0, \omega)
\end{aligned}
$$

where $P_q(a)$ denote the unconditional probability of choosing a following no recommendation. Let p denote the posterior probability that the no recommendation message is derived from the a_L signal. Then we can write,

$$
p = \frac{(1-q)P_P(a_L)}{(1-q)P_P(a_L) + P_P(a_0)}.
$$

Let $\mu_{p,\emptyset} \in \Delta(\Omega)$ denote the interim belief of the receiver following no recommendation. Then,

$$
\mu_{p,\emptyset}(\omega) = p\gamma_P(a_L, \omega) + (1-p)\gamma_P(a_0, \omega)
$$

The proof comprises the following steps:

- Step 1: verify $\gamma_P^H \in \Box H$, $\gamma_P^L \in \Box L$, and $\gamma_P^H \in \triangle HL0$ and $\mu_{p,\emptyset}(\omega) \in \Box L0$.
- Step 2: the composite function that maps any interim belief following no recommendation into $\mu_{p,\emptyset}(\omega)$ has a fixed point
- Step 3: q is chosen optimally.

As before, the following conditions are necessary and sufficient for $\gamma_P^H \in \Box H$,

$$
\begin{aligned}
f(\gamma_P^H, a_H, a_0) &\leq 1 \\
f(\gamma_P^H, a_H, a_L) &\leq 1
\end{aligned}
$$

By rearranging terms we get,

$$
\begin{aligned}
f(\gamma_P^H, a_H, a_0) &= \sum_{\omega\in\Omega} \frac{z_P(a_H,\omega)}{z(a_H,\omega)} \frac{1}{z_P(a_0,\omega)} \gamma_P(a_0,\omega) \\
&= \exp\frac{-\nu - P_q(a_L)(\alpha_L - \beta_L - \nu))}{\lambda}\gamma_P(a_0,\omega_0) \\
&+ \exp\frac{-i\nu - P_q(a_L)(\alpha_L - i\beta_L - i\nu))}{\lambda}\gamma_P(a_0,\omega_L) + \exp\frac{-P_q(a_L)\alpha_L}{\lambda}\gamma_P(a_0,\omega_H).
\end{aligned}
$$

For $P_q(a_0) \to 1$ we get $f(\gamma_P^H, a_H, a_0) < 1$ and also at $P_q(a_0) = 0$ the inequality holds by lemma 3. Thus this inequality always holds. Furthermore,

$$
\begin{aligned}
f(\gamma_P^H, a_H, a_L) &= \sum_{\omega\in\Omega} \frac{z_P(a_H,\omega)}{z(a_H,\omega)} \frac{z(a_L,\omega)}{z_P(a_L,\omega)} \gamma_P(a_L,\omega) \\
&= \exp\frac{(1-q)(1-P_q(a_L))(\alpha_L - \beta_L - \nu)}{\lambda}\gamma_P(a_L,\omega_0) \\
&+ \exp\frac{(1-q)(1-P_q(a_L))(\alpha_L - i\beta_L - i\nu)}{\lambda}\gamma_P(a_L,\omega_L) \\
&+ \exp\frac{(1-q)(1-P_q(a_L))(\alpha_L)}{\lambda}\gamma_P(a_L,\omega_H)
\end{aligned}
$$

Let us note that as $q \to 1$ we have $f(\gamma_P^H, a_H, a_L) = 1$ and

$$
\begin{aligned}
\frac{\partial\gamma_P(a_L,\omega_0)}{\partial q} &< 0 \\
\frac{\partial\gamma_P(a_L,\omega_L)}{\partial q} &> 0 \\
\frac{\partial\gamma_P(a_L,\omega_H)}{\partial q} &\geq 0.
\end{aligned}
$$

At $q = 0$, we have

$$\gamma_P(a_L,\omega_0) > \gamma_P(a_L,\omega_L) > \gamma_P(a_L,\omega_H)$$

since $z_P(a_L,\omega) = z_P(a_0,\omega)$ for all $\omega \in \Omega$ in this case but at $q = 1$

$$\gamma_P(a_L,\omega_0) < \gamma_P(a_L,\omega_H) < \gamma_P(a_L,\omega_L).$$

This implies as q decreases from 1, $f(\gamma_P^H, a_H, a_L)$ initially decreases ($\gamma_P(a_L,\omega_0)$ dominates and $\alpha_L - \beta_L - \nu < 0$), then increases ($\gamma_P(a_L,\omega_L)$ dominates and $\alpha_L - \beta_L - i\nu > 0$). This implies there exists a q small enough such that $f(\gamma_P^L, a_H, a_L) \leq 1$. This generates a lower bound on q, say, $q \leq q_1 \leq 1$ where the inequality holds.

The necessary and sufficient condition for $\gamma_P^L \in \square L$ would be

$$f(\gamma_P^L, a_L, a_0) \leq 1$$
$$f(\gamma_P^H, a_H, a_L) \leq 1$$

This can be written as,

$$\begin{aligned}
f(\gamma_P^L, a_L, a_0) &= \sum_{\omega\in\Omega} \frac{z_P(a_L,\omega)}{z(a_L,\omega)} \frac{1}{z_P(a_0,\omega)} \gamma_T(a_0,\omega) \\
&= \exp \frac{-((1-q)(1-P_q(a_L)) + P_q(a_L))(\alpha_L - \beta_L - \nu)}{\lambda} \gamma_P(a_0,\omega_0) \\
&+ \exp \frac{-((1-q)(1-P_q(a_L)) + P_q(a_L))(\alpha_L - i\beta_L - i\nu)}{\lambda} \gamma_P(a_0,\omega_L) \\
&+ \exp \frac{-((1-q)(1-P_q(a_L)) + P_q(a_L))\alpha_L}{\lambda} \gamma_P(a_0,\omega_H)
\end{aligned}$$

Similar to the last inequality at $q = 1$ we get $f(\gamma_P^L, a_L, a_0) < 1$ and $f'(\gamma_P^L, a_L, a_0) < 0$ in q. This implies for sufficiently large q the inequality holds. This generates a lower bound, say $q \geq q_2 \geq 0$ where the inequality holds. . Similarly,

$$\begin{aligned}
f(\gamma_P^L, a_L, a_H) &= \sum_{\omega\in\Omega} \frac{z_P(a_L,\omega)}{z(a_L,\omega)} \frac{z(a_H,\omega)}{z_P(a_H,\omega)} \gamma_P(a_H,\omega) \\
&= \exp \frac{-(1-q)(1-P_q(a_L))(\alpha_L - \beta_L - \nu)}{\lambda} \gamma_P(a_H,\omega_0) \\
&+ \exp \frac{-(1-q)(1-P_q(a_L))(\alpha_L - i\beta_L - i\nu)}{\lambda} \gamma_P(a_H,\omega_L) \\
&+ \exp \frac{-(1-q)(1-P_q(a_L))(\alpha_L)}{\lambda} \gamma_P(a_H,\omega_H) \leq 1
\end{aligned}$$

for $q \geq q_2$ this inequality. We can verify that since $\gamma_P(a_L, \omega_0)$ decreases and $\gamma_P(a_0, \omega_0)$ increases in q, we get, $q_2 \leq q_1$. Thus there exists $q \in [q_2, q_1]$ such that all the necessary and sufficient conditions hold.

Finally, for γ_P^0 we check,

$$f(\gamma_P^0, a_0, a_H) = \sum_{\omega\in\Omega} z_P(a_0,\omega) \frac{z(a_H,\omega)}{z_P(a_H,\omega)} \gamma_P(a_H,\omega) > 1$$

since $f(\gamma_P^0, a_H, a_0) < 1$. By similar logic $f(\gamma_P^0, a_0, a_L) > 1$. Additionally,

$$f(\gamma_P^0, a_H, a_0) = \sum_{\omega\in\Omega} \frac{z(a_0,\omega)}{z(a_H,\omega)} \frac{z_P(a_0,\omega)}{z_P(a_H,\omega)} \gamma_T(a_H,\omega) > 1$$

by lemma 4. By similar logic $f(\gamma_T^0, a_L, a_0) > 1$.

$$f(\gamma_P^0, a_H, a_L) = \sum_{\omega\in\Omega} \frac{z(a_H,\omega)}{z_P(a_H,\omega)} \frac{z_T(a_0,\omega)}{z(a_L,\omega)} \gamma_P(a_H,\omega) > 1$$

since $f(\gamma_P^H, a_H, a_0) < 1$. Also, since $z_P(a_0,\omega_L) > z(a_0,\omega_L)$, $z_P(a_H,\omega_L) < z(a_H,\omega_L)$, $z_P(a_L,\omega_L) < z(a_L,\omega_L)$ and $P_q(a_0) < 1$ we get $\gamma_P(a_0,\omega_L) > \gamma(a_0,\omega_L)$. Since $z_P(a_0,\omega_L) > z(a_H,\omega_L)$, we also get $\gamma_P(a_0,\omega_L) > \gamma(a_H,\omega_L)$, which implies $\gamma_P(a_0,\omega_L) > t\gamma(a_0,\omega_L) + (1-t)\gamma(a_H,\omega_L)$ for any $t \in [0,1]$.

Furthermore, since $z_P(a_0,\omega_0) > z(a_H,\omega_0)$, $z_P(a_H,\omega_0) < z(a_H,\omega_0)$ but $z_P(a_L,\omega_0) \lesseqgtr z(a_L,\omega_0)$ but the difference in numerator dominates, hence we get $\gamma_T(a_0,\omega_0) > \gamma(a_L,\omega_0)$. Since $z_P(a_0,\omega_0) > z(a_H,\omega_0)$, we also get $\gamma_T(a_0,\omega_0) > \gamma(a_H,\omega_0)$, which implies $\gamma_T(a_0,\omega_0) > t\gamma(a_L,\omega_0) + (1-t)\gamma(a_H,\omega_0)$ for any $t \in [0,1]$.

Finally, since $z_P(a_0,\omega_H) > z(a_0,\omega_H)$, $z_P(a_H,\omega_H) = z(a_H,\omega_L)$, $z_T(a_L,\omega_H) < z(a_L,\omega_H)$ and $P_T(a_0) < 1$ we get $\gamma_T(a_0,\omega_H) > \gamma(a_0,\omega_H)$. But $z_P(a_0,\omega_H) < z(a_L,\omega_H)$ and $z_P(a_L,\omega_L) < z(a_L,\omega_L)$, we get $\gamma_T(a_0,\omega_H) < \gamma(a_L,\omega_H)$, which implies for every q there exists a t' such that $\gamma_T(a_0,\omega_H) > t\gamma(a_0,\omega_H) + (1-t)\gamma(a_L,\omega_H)$ for all $t > t'$.

Moreover and higher q induces a higher t but a higher q also implies $\gamma_T(a_0,\omega_H) \approx \gamma(a_0,\omega_H)$. Combine this with the observation that a higher q implies $\mu_{p,\emptyset}$ would more likely be inside $\triangle HL0$ and the distance between $\mu_{p,\emptyset}$ and γ_P^0 increases with q. At $q = 1$ both $\gamma_P^0, \mu_{p,\emptyset} \in \triangle HL0$ and at $q = 0$, $\gamma_P^0 \in \triangle HL0$ but $\mu_{p,\emptyset} \in \square L0$. This implies there exists $q \geq q_3$ such that all the conditions hold. Note that, since $\gamma_P(a_L,\omega|\gamma^0 \in \triangle HL0) < \gamma_P(a_L,\omega|\gamma^0 \in \square L0)$ we also get $q_3 < q_1$. Thus there exists $q \in [\min\{q_2,q_3\}, q_1]$ such that $\gamma_P^0 \in \triangle HL0$ and $\mu_{p,\emptyset} \in \square L0$.

Consider the following mapping g_1 that maps $q \in (\min\{q_2,q_3\}, q_1)$ to an interim belief $\mu \in int(\square L0)$. This mapping would be continuous in q by Bayes Law. Consider the mapping g_2 that takes interim belief $\mu \in int(\square L0)$ to a unconditional probability distribution $P_p(a)$. This mapping is continuous in μ by lemma 4. Consider the mapping g_3 that maps $P(a_p)(a)$ to a learning strategy γ_P^0, which is also continuous in $P_p(a)$ by lemma 4. Consider a final mapping g_4 from γ_P^0 to $q \in (\min\{q_2,q_3\}, q_1)$ such that $\mu_{p,\emptyset} \in \square L0$. This mapping will also be continuous in γ_P^0 by Bayes law. Thus $g_1 \circ g_2 \circ g_3 \circ g_4$ is continuous and a fixed point exists by Brouwer's FP theorem. In the case where $(\min\{q_2,q_3\}, q_1)$ is an empty set, we get $q_2 = q_1$. Setting $q = q_1 = q_2$ would generate an equilibrium value of q.

To find the equilibrium under this strategy we need to solve for the optimal q. Let us rewrite the expected payoff of the sender in terms of q as follows:

$$V(q) = E_{\mu_0}(v(a,\omega,q)) - K(\lambda,\mu_0,q)$$

Thus $V(q)$ is twice continuously differentiable and bounded for all $q \in [\min\{q_2, q_3\}, q_1]$. As q increases two opposing effect takes place, namely, the expected payoff from a_L increases, increasing V but the cost of learning for a_0 also increases, decreasing V. Given λ, if either effect dominates for the entire range of q then we obtain a corner solution, if not there exists an interior $\hat{q}$ that maximizes $V(q)$.

Hence, proved.

□

A.6 Proof of Theorem 2

Proof. We will prove the theorem in three steps. First, we will compare the payoff function of the sender under strategic communication and full compliance. Second, we will rewrite the relationship between posterior probability and payoff function. Third, we will consider various possible cases under different parameters and show in which state the sender is more likely to make mistakes under strategic communication.

Step 1: Let us rewrite the payoff of the sender under strategic communication and compare it with the full compliance payoff. Under full compliance, the payoff of the sender is given by the sender's utility function $v(a, \omega)$ since there is no distortion due to the receiver's action. However, under strategic communication when the sender does not send any recommendation the receiver learns on their own which changes the sender's payoff. The following denotes the payoff of the sender under strategic communication,

$$\begin{aligned} v_S(a_H, \omega) &= v(a_H, \omega); \\ v_S(a_L, \omega) &= qv(a_L, \omega); \quad \text{where } q = \hat{q} + (1-\hat{q})q' \\ v_S(a_0, \omega) &= q'v(a_L, \omega) \end{aligned}$$

where q' denotes the unconditional probability with which the receiver chooses a_L following no recommendation.

Since, strategic communication does not change the payoff from choosing action a_H in any state, the only two diiferences we will study are $\Delta_0 = v(a_0, \omega_0) - v(a_L, \omega_0)$ and $\Delta_L = v(a_L, \omega_L) - v(a_0, \omega_L)$.

Step 2: Let $P_S(a, \omega)$ and $P_C(a, \omega)$ denote the conditional probability of choosing a in state ω under strategic communication and full compliance resp. We can write them as

$$\frac{P_S(a_i,\omega)}{1-P_S(a_i,\omega)} = \frac{P_S(a_i)z_S(a_i,\omega)}{\sum_{j\neq i} P_S(a_j)z_S(a_j,\omega)}; \quad z_S(a_i,\omega) = \exp(v_S(a_i,\omega)/\lambda);$$
$$\frac{P_C(a_i,\omega)}{1-P_C(a_i,\omega)} = \frac{P_C(a_i)z_C(a_i,\omega)}{\sum_{j\neq i} P_C(a_j)z_C(a_j,\omega)}; \quad z_C(a_i,\omega) = \exp(v(a_i,\omega)/\lambda).$$

Note that,

$$z_S(a_0,\omega_0) < z_C(a_0,\omega_0); z_S(a_0,\omega_L) > z_C(a_0,\omega_L); z_S(a_0,\omega_H) > z_C(a_0,\omega_H); \quad \forall q' \in (0,1)$$
$$z_S(a_L,\omega_L) < z_C(a_L,\omega_L); z_S(a_L,\omega_0) > z_C(a_L,\omega_0); z_S(a_L,\omega_H) < z_C(a_L,\omega_H); \quad \forall q \in (0,1)$$

But,

$$z_S(a_H,\omega) = z_C(a_H,\omega) \quad \forall \omega \in \Omega$$

Step 3: Let us consider possible values of $\hat{q}$ and q' that determine the learning strategy under strategic communication. Let μ_N denote the intermediate belief following no recommendation. A higher q' is generated by higher $P(a_L|\mu_N)$, which can happen if Δ_0 is sufficiently low and Δ_L is sufficiently high, i.e., the mistake from choosing a_0 in ω_L outweighs the mistake of choosing a_L in ω_0. Also, a higher $\hat{q}$ reduces $P(a_l|\mu_N)$, since it denotes a lower probability of the optimal action being a_L following no recommendation. Combining these two we explore the following possibilities.

Case 1: $P_S(a_0) > P_C(a_0)$ and $P_S(a_L) > P_C(a_L)$ such that,

$$\frac{P_S(a_0,\omega_0)}{1-P_S(a_0,\omega_0)} > \frac{P_C(a_0,\omega_0)}{1-P_C(a_0,\omega_0)}; \frac{P_S(a_L,\omega_L)}{1-P_S(a_L,\omega_L)} > \frac{P_C(a_L,\omega_0)}{1-P_C(a_L,\omega_0)}.$$

The first inequality implies $z_S(a_0,\omega_L) \gg z_C(a_0,\omega_L)$, and $z_S(a_0,\omega_H) \gg z_C(a_0.\omega_H)$. The second inequality implies $z_S(a_L,\omega_0) \gg z_C(a_L,\omega_0)$. These two conditions together imply high q' and low q, which cannot be true.

Case 2: $P_S(a_0) < P_C(a_0)$ and $P_S(a_L) < P_C(a_L)$ such that,

$$\frac{P_S(a_0,\omega_0)}{1-P_S(a_0,\omega_0)} < \frac{P_C(a_0,\omega_0)}{1-P_C(a_0,\omega_0)}; \frac{P_S(a_L,\omega_L)}{1-P_S(a_L,\omega_L)} < \frac{P_C(a_L,\omega_0)}{1-P_C(a_L,\omega_0)}.$$

This first inequality implies $z_S(a_0,\omega_L) \approx z_C(a_0,\omega_L)$ and $z_S(a_0,\omega_H) \approx z_C(a_0,\omega_H)$. The second inequality implies $z_S(a_L,\omega_0) \approx z_S(a_L,\omega_0)$. These two conditions together imply low q' and high q, which can be achieved with high $\hat{q}$. In this case, strategic communication generates more mistakes in learning in both ω_0 and ω_L.

Case 3: $P_S(a_0) > P_C(a_0)$ and $P_S(a_L) \gtreqqless P_C(a_L)$ such that

$$\frac{P_S(a_0,\omega_0)}{1-P_S(a_0,\omega_0)} \geq \frac{P_C(a_0,\omega_0)}{1-P_C(a_0,\omega_0)}; \frac{P_S(a_L,\omega_L)}{1-P_S(a_L,\omega_L)} < \frac{P_C(a_L,\omega_0)}{1-P_C(a_L,\omega_0)}.$$

This first inequality implies $z_S(a_0,\omega_L) \gg z_C(a_0,\omega_L)$, and $z_S(a_0,\omega_H) \gg z_C(a_0.\omega_H)$. The second inequality implies $z_S(a_L,\omega_0) \geq z_S(a_L,\omega_0)$.These two conditions together imply high q' and intermediate q, which can be achieved by a sufficiently high $\hat{q}$. This implies $\Delta_L \gg \Delta_0$, i.e., it is more costly to make mistake in ω_L than in ω_0. In this case under strategic communication, the sender makes more mistakes in state ω_L.

Case 4: $P_S(a_0) \gtreqqless P_C(a_0)$ and $P_S(a_L) > P_C(a_L)$ such that

$$\frac{P_S(a_0,\omega_0)}{1-P_S(a_0,\omega_0)} < \frac{P_C(a_0,\omega_0)}{1-P_C(a_0,\omega_0)}; \frac{P_S(a_L,\omega_L)}{1-P_S(a_L,\omega_L)} \geq \frac{P_C(a_L,\omega_0)}{1-P_C(a_L,\omega_0)}.$$

The first inequality requires $z_S(a_0,\omega_L) \geq z_C(a_0,\omega_L)$ and $z_S(a_0,\omega_H) \geq z_C(a_0,\omega_H)$. The second inequality requires $z_S(a_L,\omega_0) \gg z_C(a_L,\omega_0)$. These two conditions together imply intermediate q' and low q, which can be achieved with low $\hat{q}$. This implies $\Delta_L \ll \Delta_0$, i.e., it is more costly to make mistake in ω_0 than in ω_L. In this case under strategic communication, the sender makes more mistakes in state ω_0.

Hence, proved. $\square$