Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 160 (2015) 132–149

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Notes

Communication with endogenous information acquisition

Harry Di Pei

Department of Economics, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

Received 1 August 2014; final version received 20 August 2015; accepted 27 August 2015

Available online 12 September 2015

Abstract

I develop a theory of communication in which a sender gathers costly information before giving advice to a receiver. In a general setting, I show that the sender always communicates all her information to the receiver in every equilibrium. In the uniform-quadratic model in which the sender can choose any finite partition as her information structure, an upwardly biased sender can convey more precise information when recommending a larger action.

JEL classification: D82; D83

Keywords: Cheap talk; Information acquisition; Communication

1. Introduction

Experts often collect costly information before advising decision makers. Imagine an investment banker (she) persuading her customer (call him the CEO) on the acquisition of a firm, she has little idea about its value before mobilizing her research team to gather the relevant information. Needless to say, acquiring such information usually incurs cost. Similar situations arise when doctors diagnosing patients, lobbyists studying regulation policies, etc.

Motivated by these applications, I study the strategic information transmission problem in Crawford and Sobel (1982), with the innovation that the sender acquires her information en-

E-mail address: harrydp@mit.edu.

http://dx.doi.org/10.1016/j.jet.2015.08.011

dogenously before communicating with the receiver. I find that when information gathering is costly, the interaction between *moral hazard in acquiring information* and *adverse selection in reporting information* overturns many predictions in CS.

In my model, the sender chooses a partition of the state space as her information structure. This process is unobservable to the receiver and costly for the sender. My first result shows that the sender communicates everything she knows to the receiver in every equilibrium, with virtually no restriction on the state space as well as the players' preferences. Formally speaking, full communication is guaranteed whenever the sender can always coarsen her information partition, and a strictly coarser partition costs less to acquire. The intuition behind this result is simple: if the sender has an incentive to withhold information, then why does she acquire that information in the first place? By studying the problem in less detail, she achieves the same outcome at a lower cost.

Following this general result, I analyze the well-known '*uniform-quadratic*' model with an '*upwardly biased sender*' (her favorite action is always strictly larger than the receiver's), which has been the main focus of the strategic communication literature. When the cost of an information structure is proportional to its value (cubic cost),[1] I show that every equilibrium is characterized by an interval partition with *decreasing* interval lengths if information acquisition cost is large enough. This reverses the characterization result in CS, in which the interval lengths are increasing.

For some rough intuition, this '*reverse informativeness*' result is driven by the commitment effect of costly information acquisition. Our cubic cost function implies that given the number of elements in the partition, an information structure is more costly if the lengths of the intervals are more uniform. Since the sender is upwardly biased, and due to the covert nature of information acquisition, decreasing interval equilibria cannot be sustained when the cost of information acquisition is too small. This is because the sender strictly prefers the larger action at the partition point, so she has an incentive to move the partition point to the left, and to acquire an information structure which is more costly. However, when the cost of information acquisition is sufficiently large, this deviation is no longer profitable. Hence, higher information acquisition cost gives the sender more commitment power, which helps to sustain informative equilibria with decreasing interval lengths.

Related literature My work is closely related to a contemporaneous paper by Argenziano et al. (2014), in which the sender chooses the precision of her information by deciding how many rounds of Bernoulli Experiments to conduct.[2] They study equilibrium outcomes both when the sender's information structure is observable (overt) and when it is non-observable (covert), and apply their results to revisit the trade-off between delegation and communication. Both papers enrich the sender's informational choice comparing with earlier contributions, where the sender is either perfectly informed or completely ignorant.[3] When uncertainty is 1-dimensional, the

[1] Although this result is shown under the cubic cost function, the qualitative feature of the equilibrium is robust to more general cost functions. In Section 2 of the Online Appendix, I display a general cost functions under which '*reverse informativeness*' holds.

[2] In Argenziano et al. (2014), the state θ is uniformly distributed in [0, 1], the outcome of a Bernoulli experiment is binary: either 0 or 1, and outcome 1 occurs with probability θ. In their framework, the sender chooses how many rounds of independent experiments to conduct.

[3] For example, Aghion and Tirole (1997), Austen-Smith (1994), Hellwig and Veldkamp (2009), etc.

sender typically has a finite number of interim types, since she has no incentive to be fully informed when facing incentive problems in communication.

The main difference between the papers is the amount of flexibility allowed in the sender's informational choice. Their Bernoulli Experiment approach is better micro-founded by explicitly describing the information acquisition process, and fits better into applications where players cannot finetune their information structures.[4]

In contrast, my approach enables the sender to flexibly allocate her '*search effort*' over the entire state space. By allowing the sender to choose from the set of partitions, my model delivers clear comparisons between interval partition equilibria under costly information acquisition and CS equilibria. In practice, acquiring a partition information structure can be achieved by seeking answers for a set of deliberately chosen survey questions (for example, the customer's age; whether his income is below or above average, etc.), or examining a subset of the relevant attributes.

Flexible information acquisition stems from of idea of '*rational inattention*' (Sims, 2003) and has been adopted to study global games (Yang, 2014), security design (Yang, 2015) as well as disclosure games (Gentzkow and Kamenica, 2012; Kamenica and Gentzkow, 2011).

Yang (2014) shows that flexibility in players' informational choices leads to multiple equilibria and enables efficient coordination, which is in sharp contrast to the standard results in global games with rigid information. Analogously, my full communication and reverse informativeness results illustrate the impact of flexible information acquisition on the outcome of cheap talk games.

Layout The rest of the paper is organized as follows. Section 2 introduces a general setup and presents the full communication result. Section 3 focuses on the uniform-quadratic model and presents the reverse informativeness result. Section 4 concludes.

2. The general model

In this section, I setup a general model in which the sender acquires her information at a cost before communicating with the receiver (Fig. 1). I show that every equilibrium achieves full communication when the set of available information structures is '*rich*' and the cost of information acquisition satisfies '*monotonicity*'. I also discuss why full communication cannot be achieved when the receiver can observe the sender's information structure.

2.1. Setup

Primitives The receiver (he) needs to make a decision $a \in A$ on an unfamiliar project. The sender's and the receiver's gains from the project are $u^s(a,\theta)$ and $u^r(a,\theta)$ respectively, both depending on the receiver's action and the state of the world $\theta \in \Theta$. Both A and Θ are compact subsets of $\mathbb{R}^n$.

For any set X, I use $\Delta(X)$ to denote the set of distributions over X. Players share a common prior $\mu_0 \in \Delta(\Theta)$ over θ. For any $\Theta' \subset \Theta$, $\mu_0(\Theta')$ denotes the probability of Θ' under μ_0.

[4] Notice that the '*richness*' assumption in my model is violated in Argenziano et al. (2014), and neither is implied by the other.

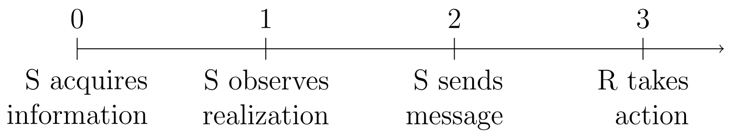

Fig. 1. Timeline.

Information acquisition An *information structure* ψ is a finite partition of Θ,[5] with $\psi(\theta)$ ($\subset \Theta$) the partition element containing θ. I assume that every partition element occurs with strictly positive probability, i.e. $\mu_0(\psi(\theta)) > 0$ for every $\theta \in \Theta$.

Let Ψ be the set of '*available information structures*'. At the information acquisition stage, the sender chooses $\psi \in \Psi$ at cost $C(\psi)$. This choice is *non-observable* to the receiver. The sender's total payoff equals to her expected gain from the project minus $C(\psi)$.

Every partition element (or signal realization) $\psi(\theta)$ leads to a *posterior belief*: $\mu^s_{\psi(\theta)} \in \Delta(\Theta)$, which I call the sender's '*type*'. The standard definition of partition coarseness provides a partial ordering of information structures:

Condition 1 *(Strictly coarser). ψ' is 'strictly coarser' than ψ if:*

1. $\psi(\theta) \subseteq \psi'(\theta)$ *for every* $\theta \in \Theta$.
2. *There exists* $\theta \in \Theta$ *such that* $\psi(\theta) \neq \psi'(\theta)$.

Next, I introduce the key assumptions of my model:

Assumption 1 *(Richness). If $\psi \in \Psi$ and ψ' is strictly coarser than ψ, then $\psi' \in \Psi$.*

Assumption 2 *(Monotonicity). If ψ' is strictly coarser than ψ, then $C(\psi') < C(\psi)$.*

In a nutshell, *richness* allows the sender to '*coarsen*' her information, and she strictly saves cost through coarsening when *monotonicity* holds.

Communication Let M be the set of messages, with cardinality greater or equal to that of Θ. Under information structure ψ, the sender's communication rule after receiving $\psi(\theta)$ is $\sigma_{\psi(\theta)} \in \Delta(M)$. Let $\mu^r_m \in \Delta(\Theta)$ be the receiver's posterior belief after receiving m. Let $\alpha_m \in \Delta(A)$ be his (mixed) action after receiving m.[6]

2.2. Full communication

Let $\mu^s \equiv (\mu^s_{\psi(\theta)})_{\psi(\theta)\in\psi}$, $\sigma \equiv (\sigma_{\psi(\theta)})_{\psi(\theta)\in\psi}$, $\mu^r \equiv (\mu^r_m)_{m\in M}$ and $\alpha \equiv (\alpha_m)_{m\in M}$. A *Perfect Bayesian Equilibrium* (hereafter, equilibrium) is characterized by $(\psi, \mu^s, \sigma, \mu^r, \alpha)$, and satisfies[7]:

[5] A partition is a set of subsets of Θ, i.e. $\psi = \{\Theta_1, \ldots, \Theta_n\}$ where $\Theta_i \subset \Theta$ for all i, $\cup_{i=1}^n \Theta_i = \Theta$, $\Theta_i \cap \Theta_j = \emptyset$ for any $i \neq j$, and $n = 1, 2, \ldots$. By definition, $\psi(\theta)$ is the partition element containing θ, so if $\psi = \{\Theta_1, \Theta_2, \ldots, \Theta_n\}$, there exists $k \in \{1, 2, \ldots, n\}$ such that $\psi(\theta) = \Theta_k$.

[6] Formally, for conditional probabilities to be well-defined, I require that $\sigma_{\psi(\theta)}$ and α_m to be Borel probability measures of Polish spaces M and A, respectively.

[7] For notation simplicity, I only write down the equilibrium conditions when the sender is using a pure strategy to acquire information.

1. Given μ^r and α, every type of sender sends messages which maximize her *interim expected gain* from the project, i.e. for every $\psi(\theta) \in \psi$, $\sigma_{\psi(\theta)}$ assigns probability 1 to the following subset of messages[8]:

$$\arg\max_{m \in M} \int_{\tilde{\theta}} u^s(\alpha_m, \tilde{\theta}) d\mu^s_{\psi(\theta)}(\tilde{\theta}).$$

2. Given μ^r and α, ψ is chosen to maximize the sender's expected payoff:

$$\max_{\psi \in \Psi} \Big\{ \sum_{\psi(\theta) \in \psi} \underbrace{\mu_0\Big(\psi(\theta)\Big)}_{\text{Prob. of } \psi(\theta)} \underbrace{\max_{m \in M} \{ \int_{\tilde{\theta}} u^s(\alpha_m, \tilde{\theta}) d\mu^s_{\psi(\theta)}(\tilde{\theta}) \}}_{\text{S's expected gain under } \psi(\theta).} - C(\psi) \Big\}.^{9}$$

3. The receiver chooses actions which maximize his expected payoff given his posterior belief, i.e. for every $m \in M$, α_m assigns probability 1 to the following subset of actions:

$$\arg\max_{a \in A} \int_{\tilde{\theta}} u^r(a, \tilde{\theta}) d\mu^r_m(\tilde{\theta}).$$

4. Players' beliefs are updated according to Bayes Rule whenever applicable.

An equilibrium achieves '*full communication*' if the sender's and the receiver's posterior beliefs are always the same. Intuitively, in such equilibria, the sender tells the receiver everything she knows. Now, I state my first result:

Proposition 1. *When Ψ is rich and C satisfies monotonicity, every pure strategy equilibrium achieves full communication.*

By '*pure strategy*', I only restrict the sender to use pure strategies when acquiring information, i.e. I allow her to use mixed strategies when sending messages and the receiver to use mixed strategies when taking actions. When the sender uses a mixed strategy to acquire information, then full communication cannot be guaranteed. However, in every equilibrium, the sender will never send the same message under two different signal realizations within one information structure. I show this claim in Section 3 of the Online Appendix, and will present an example of such mixed strategy equilibrium. I will also discuss cases when the receiver is consulting multiple senders, when the sender is communicating with multiple receivers, or when the receiver can also acquire information himself.

[8] For every $\alpha \in \Delta(A)$, $u^s(\alpha, \theta)$ and $u^r(\alpha, \theta)$ are defined in their natural ways:

$$u^s(\alpha, \theta) \equiv \int_{a \in A} u^s(a, \theta) d\alpha(a), \quad u^r(\alpha, \theta) \equiv \int_{a \in A} u^r(a, \theta) d\alpha(a).$$

[9] '$\sum_{\psi(\theta) \in \psi}$' means summing over all partition elements in ψ, i.e. if $\psi = \{\Theta_1, \ldots, \Theta_n\}$, then $\sum_{\psi(\theta) \in \psi}$ is equivalent to the sum over all Θ_i (i from 1 to n).

Proof of Proposition 1. The proof is done by contradiction. Suppose there exists an equilibrium $(\psi, \mu^s, \sigma, \mu^r, \alpha)$, such that there exist two partition elements of ψ, $\psi(\theta_1)$ and $\psi(\theta_2)$, with $\psi(\theta_1) \neq \psi(\theta_2)$, and a message $m \in M$, satisfying:

$$m \in \arg\max_{\tilde{m} \in M} \int_{\tilde{\theta}} u^s(\alpha_{\tilde{m}}, \tilde{\theta}) d\mu^s_{\psi(\theta_j)}(\tilde{\theta}), \tag{1}$$

for $j = 1, 2$.

Define $\tilde{\psi}$ to be the partition information structure where:

$$\tilde{\psi}(\theta) = \begin{cases} \psi(\theta) & \text{when } \theta \notin \psi(\theta_1) \cup \psi(\theta_2) \\ \psi(\theta_1) \cup \psi(\theta_2) & \text{when } \theta \in \psi(\theta_1) \cup \psi(\theta_2). \end{cases}$$

By definition, $\tilde{\psi}$ is strictly coarser than ψ. Define $\tilde{\sigma}$ to be the communication rule where[10]:

$$\tilde{\sigma}_{\tilde{\psi}(\theta)} = \begin{cases} \sigma_{\psi(\theta)} & \text{when } \theta \notin \psi(\theta_1) \cup \psi(\theta_2) \\ m & \text{when } \theta \in \psi(\theta_1) \cup \psi(\theta_2) \end{cases}$$

Fixing (μ^r, α), consider the following deviation of the sender: acquiring information structure $\tilde{\psi}$ at the information acquisition stage, and uses $\tilde{\sigma}$ at the communication stage.

I show that this deviation is strictly profitable. First, $\tilde{\psi}$ is available by Assumption 1 (richness). Second, (1) implies that the sender's expected gain from the project under $(\tilde{\psi}, \tilde{\sigma})$ is the same as her expected gain from the project under (ψ, σ). Third, $C(\tilde{\psi}) < C(\psi)$ by Assumption 2 (monotonicity). Hence, I have shown that this is a profitable deviation, which leads to a contradiction.

This implies that when the sender uses a pure strategy to acquire information, each message is sent by at most one type of sender in any given equilibrium. From Bayes Rule, the receiver can fully infer the sender's type. Since they share a common prior, their posterior beliefs are equal. □

Proposition 1 shows that when the sender has *no superior knowledge ex ante*, *no alternative usage of information*, but enjoys *sufficient flexibility* in choosing information structures, she has no incentive to withhold her information. This result provides a necessary condition for equilibrium characterization. It also contrasts one of the key features of CS, that the sender's information is not fully transmitted due to the conflict of interests.

Remark Note that the sender's choice of information structure being unobserved by the receiver is crucial to this result. When the receiver observes the sender's informational choice, the sender's deviation at the information acquisition stage can change the receiver's action rule, which can lower her expected gain from the project. When the sender's expected loss from the project exceeds her gain from saving information acquisition cost, she will have an incentive to acquire information that cannot be fully transmitted.

[10] I abuse notation, and $\tilde{\sigma}_{\tilde{\psi}(\theta)} = m$ means that $\tilde{\sigma}_{\tilde{\psi}(\theta)}$ assigns probability 1 to message m.

3. The uniform-quadratic model

In this section, I adopt the framework in Section 2 to study the celebrated 1-dimensional '*uniform-quadratic*' model, which has been the main focus of the communication literature, and has been broadly applied in political economy as well as organizational design.[11]

Let $\Theta = A = [0, 1]$. θ is uniformly distributed on Θ, and players' payoffs are given by:

$$u^s(a,\theta) = -(a-\theta-b)^2, \quad u^r(a,\theta) = -(a-\theta)^2$$

with $0 < b < \frac{1}{4}$. I show that when information acquisition cost is large enough, an '*upwardly biased*' sender conveys more precise information when θ is large, which reverses the prediction in CS.

3.1. Information acquisition cost

Let Ψ be the set of finite partitions on Θ. Let $\Psi_0(\subset \Psi)$ be the set of *interval partitions*, where every partition element is an interval. I assume the cost of acquiring partition $\{\Theta_1, \dots, \Theta_n\}$ takes the following cubic form:

$$C(\Theta_1, \dots, \Theta_n) = c\Big[1 - \sum_{i=1}^{n} \mu_0(\Theta_i)^3\Big], \tag{2}$$

where $\mu_0(\cdot)$ is the Lebesgue Measure, and $c > 0$ measures the cost of information acquisition.

I will present my main result under the cubic cost function, and will discuss the reverse informativeness result under a more general class of cost functions in the Online Appendix.[12]

Remark This cubic function has a natural interpretation, that the cost of an *interval partition information structure* is proportional to its '*value*'[13] – the expected utility gain of a decision maker when he has access to this information. To see this, first, let us compute the receiver's expected payoff from the project when he has no information about θ:

$$\max_a \Big\{ \int_0^1 -(a-\theta)^2 d\theta \Big\} = -\frac{1}{12}.$$

Next, let us compute his expected payoff from the project when his information about θ is given by the following interval partition: $\{[\theta_0, \theta_1], (\theta_1, \theta_2], \dots, (\theta_{n-1}, \theta_n]\}$, where $0 = \theta_0 < \theta_1 < \dots < \theta_{n-1} < \theta_n = 1$, which is:

$$\max_{a_1,\dots,a_n} \Big\{ \sum_{i=1}^{n} \int_{\theta_{i-1}}^{\theta_i} -(a_i-\theta)^2 d\theta \Big\} = -\frac{1}{12}\sum_{i=1}^{n}(\theta_i - \theta_{i-1})^3.$$

Comparing with no information, the receiver's payoff is increased by:

$$\frac{1}{12} - \frac{1}{12}\sum_{i=1}^{n}(\theta_i - \theta_{i-1})^3.$$

11 For recent contributions based on the uniform-quadratic model, see Argenziano et al. (2014), Ivanov (2010), etc.

12 The general class of cost function I consider incorporates the Shannon's Entropy, which has been frequently seen in information economics, for example, Sims (2003), Yang (2014, 2015), etc.

13 Later on, I will show in Lemma 3.1 that only interval partition information structures can arise in equilibrium.

I call this difference: '*the value of interval partition information structure* $\{[\theta_0, \theta_1], (\theta_1, \theta_2], \dots, (\theta_{n-1}, \theta_n]\}$'.

Under the cubic cost function, the cost of that information structure is

$$c\Big[1 - \sum_{i=1}^{n} (\theta_i - \theta_{i-1})^3\Big],$$

which is proportional to its value.[14]

3.2. Some preliminary analysis

As in Section 2, I focus on equilibria in which the sender uses a pure strategy to acquire information. It is easy to check that Ψ is rich and C satisfies monotonicity.[15] I show in the next Lemma that the sender always acquires an interval partition information structure in equilibrium.

Lemma 3.1. *The sender chooses an interval partition information structure in every equilibrium.*

The proof is in Appendix A, which uses the fact that the sender's preference exhibits strictly increasing differences in a and θ,[16] as well as the cost of information acquisition only depends on the measure of each partition element.

Lemma 3.1 implies that it is without loss of generality to focus on interval partition information structures. Let $(\theta_0, \dots, \theta_n)$ be the partition points of an interval partition information structure with $0 = \theta_0 < \theta_1 < \dots < \theta_{n-1} < \theta_n = 1$. A necessary condition for $(\theta_0, \dots, \theta_n)$ to be an *equilibrium information structure* is:

$$(\theta_0, \theta_1, \dots, \theta_n) \in \arg \max_{(\theta'_0, \theta'_1, \dots, \theta'_n)} \Big\{ -\sum_{i=1}^{n} \int_{\theta'_{i-1}}^{\theta'_i} (a_i^* - \theta - b)^2 d\theta - c\Big(1 - \sum_{i=1}^{n} (\theta'_i - \theta'_{i-1})^3\Big)\Big\},$$

$$\text{subject to } 0 = \theta'_0 \le \theta'_1 \le \dots \le \theta'_{n-1} \le \theta'_n = 1 \text{ and } a_i^* = \frac{\theta_{i-1} + \theta_i}{2} \quad \text{for every } 1 \le i \le n.^{17} \tag{3}$$

This conditions says that given the receiver's strategy, the sender cannot strictly gain from the following class of deviations: acquiring a different interval partition information structure

[14] We can obtain the same conclusion by doing this exercise on the sender's preference.

[15] However, Ψ_0 does not satisfy the richness condition.

[16] u^s exhibits strictly increasing differences in a and θ if for every $a > a'$, $\theta > \theta'$

$$u^s(a, \theta) - u^s(a, \theta') > u^s(a', \theta) - u^s(a', \theta').$$

This condition is satisfied in our setting.

[17] Notice that $a_i^* \equiv \frac{\theta_{i-1} + \theta_i}{2}$ is the receiver's equilibrium action after the sender tells him that $\theta \in (\theta_{i-1}, \theta_i]$. Since the receiver only has n equilibrium actions, partitions with more than n elements are never optimal for the sender, and hence, can be safely ignored.

If the sender chooses a partition with less than n elements, it can also be represented by $(\theta'_0, \dots, \theta'_n)$ with $0 = \theta'_0 \le \theta'_1 \le \dots \le \theta'_n = 1$. Of course, there will be degenerate partition elements, i.e. $\theta'_{i-1} = \theta'_i$, for some i. In this case, we should view the sender as acquiring an m-interval partition, where m is the number of non-degenerate elements in $\{[\theta'_0, \theta'_1], \dots, (\theta'_{n-1}, \theta'_n]\}$.

Also when referring to an interval in the partition, I will use $(\theta_{i-1}, \theta_i]$, although the reader should keep in mind that when $i = 1$, it should be $[\theta_0, \theta_1]$.

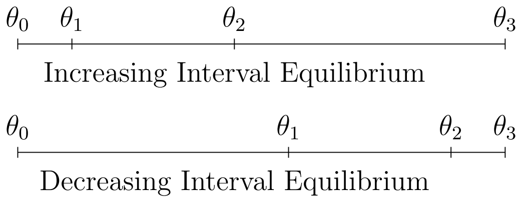

Fig. 2. Two types of equilibria.

$\{[\theta'_0, \theta'_1], \ldots, (\theta'_{n-1}, \theta'_n]\}$, and inducing a_i^* when $\theta \in (\theta'_{i-1}, \theta'_i]$. The next Lemma says that the above condition is also sufficient for an equilibrium.

Lemma 3.2. $(\theta_0, \ldots, \theta_n)$ *is an equilibrium information structure if and only if it satisfies (3).*

3.3. Reverse informativeness

I present my main result in this subsection, which examines the following feature of informative interval partition equilibria: whether the interval lengths are increasing or decreasing (from left to right). Fig. 2 shows the two types of equilibria, and intuitively, in '*increasing interval equilibria*', the sender conveys more precise information when θ is small; in '*decreasing interval equilibria*', the sender conveys more precise information when θ is large.

When $b > 0$, only increasing interval equilibria exist in CS. Hence, the implications on communication informativeness is reversed only when decreasing interval equilibria occur. Proposition 2 characterizes the range of parameter values, (b, c), under which each type of equilibrium exists:

Proposition 2 *(Reverse Informativeness). Increasing interval equilibria exist if and only if:* $c \in (0, \frac{1}{12}(1 - 4b))$*; Decreasing interval equilibria exist if and only if:* $c \in (\frac{1}{12}(1 + 4b), \frac{1}{6}]$.

Proposition 2 implies that when the sender's information is costly acquired, the qualitative feature of equilibria depends not only on the players' preferences, but also on the cost of information acquisition. When information acquisition cost is large enough, the sender's message is more informative when she recommends a larger action even though she is upwardly biased.

Proof of Proposition 2. The proof is done in three steps. First, I derive the first order condition for θ_i, which is necessary for any informative equilibrium. Second, I argue that no informative equilibrium exists when $c > \frac{1}{6}$ by exhibiting an incentive constraint that is violated. Third, I show by construction that informative equilibria exist when $c \in (0, \frac{1}{12}(1 - 4b)) \cup (\frac{1}{12}(1 + 4b), \frac{1}{6}]$.

Step 1: Consider maximization program (3). If $(\theta_0, \ldots, \theta_n)$, with $0 = \theta_0 < \theta_1 < \ldots < \theta_n = 1$ and $n \geq 2$, are the partition points of an equilibrium information structure, then conditional on the sender choosing $\theta'_0 = \theta_0, \ldots, \theta'_{i-1} = \theta_{i-1}$ and $\theta'_{i+1} = \theta_{i+1}, \ldots, \theta'_n = \theta_n$, she has no incentive to deviate from θ_i when choosing θ'_i. Since $\theta_i \in (\theta_{i-1}, \theta_{i+1})$, i.e. the optimum is interior, hence, the first order condition with respect to θ_i has to be satisfied:

Lemma 3.3. *In every informative equilibrium ($n \geq 2$), the difference in length between any two adjacent intervals equals to $\frac{4b}{1-12c}$, that is, for every $1 \leq i \leq n-1$:*

$$(\theta_{i+1} - \theta_i) - (\theta_i - \theta_{i-1}) = \frac{4b}{1-12c}. \tag{4}$$

Proof of Lemma. Take the first order condition with respect to θ_i' on the sender's objective function, and evaluate the derivative at $(\theta_0', \ldots, \theta_n') = (\theta_0, \ldots, \theta_n)$, we have:

$$-(a_i^* - \theta_i - b)^2 + (a_{i+1}^* - \theta_i - b)^2 + 3c(\theta_i - \theta_{i-1})^2 - 3c(\theta_{i+1} - \theta_i)^2 = 0.$$

Rearranging terms and plugging in $a_i^* = \frac{\theta_{i-1}+\theta_i}{2}$ and $a_{i+1}^* = \frac{\theta_i + \theta_{i+1}}{2}$, we have:

$$-\frac{\theta_{i+1} - \theta_{i-1}}{2}\left(\frac{\theta_{i+1} - 2\theta_i + \theta_{i-1}}{2} - 2b\right) + 3c(\theta_{i+1} - 2\theta_i + \theta_{i-1})(\theta_{i+1} - \theta_{i-1}) = 0.$$

Canceling out $\theta_{i+1} - \theta_{i-1}$ and rearranging terms, we get: $\theta_{i+1} - 2\theta_i + \theta_{i-1} = \frac{4b}{1-12c}$. □

Since $\theta_{i+1} - 2\theta_i + \theta_{i-1}$ is the difference in length between two adjacent intervals, this Lemma also implies:

- When $\frac{4b}{1-12c} \in [0, 1)$ (or $c < \frac{1-4b}{12}$), only increasing interval equilibria may exist;
- When $\frac{4b}{1-12c} \in (-1, 0]$ (or $c > \frac{1+4b}{12}$), only decreasing interval equilibria may exist;
- When $|\frac{4b}{1-12c}| \geq 1$ (or $c \in [\frac{1-4b}{12}, \frac{1+4b}{12}]$), no informative equilibrium exists.

Moreover, as in Crawford and Sobel (1982), this necessary condition also uniquely pins down the partition point(s) given the size of the partition.

Step 2: I use the necessary condition derived in Step 1 and show that no informative equilibrium exists when $c > \frac{1}{6}$. The proof is done by contradiction. If $(\theta_0, \ldots, \theta_n)$ is an equilibrium, then the following deviation cannot be profitable for the sender:

- Pooling intervals $(\theta_{i-1}, \theta_i]$ and $(\theta_i, \theta_{i+1}]$ together at the information acquisition stage, and inducing action $a_i^* = \frac{\theta_{i-1}+\theta_i}{2}$ when $\theta \in (\theta_{i-1}, \theta_{i+1}]$.

Since the equilibrium outcome remains unchanged for $\theta \notin (\theta_{i-1}, \theta_{i+1}]$ under this deviation, we only examine her expected loss when $\theta \in (\theta_{i-1}, \theta_{i+1}]$. Her expected loss in equilibrium is smaller than her expected loss after deviating if:

$$\sum_{j=i}^{i+1} \left\{ \int_{\theta_{j-1}}^{\theta_j} (a_j^* - \theta - b)^2 d\theta - c(\theta_j - \theta_{j-1})^3 \right\} \leq \int_{\theta_{i-1}}^{\theta_{i+1}} (a_i^* - \theta - b)^2 d\theta - c(\theta_{i+1} - \theta_{i-1})^3.$$

Re-arranging terms, we get:

$$c(\theta_{i+1} - \theta_{i-1})^3 - c(\theta_i - \theta_{i-1})^3 - c(\theta_{i+1} - \theta_i)^3 \leq \int_{\theta_i}^{\theta_{i+1}} (a_i^* - \theta - b)^2 d\theta - \int_{\theta_i}^{\theta_{i+1}} (a_{i+1}^* - \theta - b)^2 d\theta.$$

The left hand side of this inequality equals to $3c(\theta_{i+1}-\theta_i)(\theta_i-\theta_{i-1})(\theta_{i+1}-\theta_{i-1})$, while the right hand side equals to $(a_i^*-a_{i+1}^*)(\theta_{i+1}-\theta_i)(a_{i+1}^*+a_i^*-2b-\theta_{i+1}-\theta_i)$.

Plugging $a_i^*=\frac{\theta_i+\theta_{i-1}}{2}$ and $a_{i+1}^*=\frac{\theta_i+\theta_{i+1}}{2}$ into the right hand side, the original inequality is equivalent to:

$$3c(\theta_{i+1}-\theta_i)(\theta_i-\theta_{i-1})(\theta_{i+1}-\theta_{i-1})\le\frac{1}{2}(\theta_{i+1}-\theta_i)(\theta_{i+1}-\theta_{i-1})(2b+\frac{\theta_{i+1}-\theta_{i-1}}{2}).$$

Since $\theta_{i+1}>\theta_i>\theta_{i-1}$, we can cancel $(\theta_{i+1}-\theta_i)(\theta_{i+1}-\theta_{i-1})$ on both sides, which gives:

$$6c(\theta_i-\theta_{i-1})\le 2b+\frac{\theta_{i+1}-\theta_{i-1}}{2}=2b+\frac{\theta_{i+1}-\theta_i}{2}+\frac{\theta_i-\theta_{i-1}}{2}.$$

The necessary condition we have derived in Step 1 requires $(\theta_{i+1}-\theta_i)-(\theta_i-\theta_{i-1})=\frac{4b}{1-12c}$, which means that

$$6c\Big(\theta_{i+1}-\theta_i-\frac{4b}{1-12c}\Big)\le 2b+\frac{1}{2}(\theta_{i+1}-\theta_i)+\frac{1}{2}\Big(\theta_{i+1}-\theta_i-\frac{4b}{1-12c}\Big).$$

This is equivalent to:

$$(1-6c)(\theta_{i+1}-\theta_i)+2b-\frac{2b}{1-12c}+6c\frac{4b}{1-12c}\ge 0.$$

When $c>\frac{1}{6}$, $2b-\frac{2b}{1-12c}+6c\frac{4b}{1-12c}=0$, and $(1-6c)(\theta_{i+1}-\theta_i)<0$, this is a contradiction. Hence, no informative equilibrium exists when $c>\frac{1}{6}$.

Step 3: I construct 2-partition equilibria when $c\in(0,\frac{1-4b}{12})\cup(\frac{4b+1}{12},\frac{1}{6}]$. Consider the following strategy profile:

- The sender acquires interval partition information structure $\{[0,\frac{1}{2}-\frac{2b}{1-12c}],(\frac{1}{2}-\frac{2b}{1-12c},1]\}$. She claims that the state is low when $\theta\in[0,\frac{1}{2}-\frac{2b}{1-12c}]$, and claims that the state is high when $\theta\in(\frac{1}{2}-\frac{2b}{1-12c},1]$.
- The receiver chooses $a_1^*\equiv\frac{1}{4}-\frac{b}{1-12c}$ when the sender claims that the state is low, and chooses $a_2^*\equiv\frac{3}{4}-\frac{b}{1-12c}$ when the sender claims that the state is high.

I show that this is an equilibrium. First, conditional on the sender's strategy, the receiver's strategy is sequentially rational since $\frac{\theta_1}{2}=a_1^*$ and $\frac{\theta_1+1}{2}=a_2^*$. Second, according to Lemmas 3.1 and 3.2, conditional on the receiver's strategy, the sender's incentive constraint is satisfied if and only if $\theta_1=\frac{1}{2}-\frac{2b}{1-12c}$ is the solution to the following optimization program:

$$\max_{\theta_1\in[0,1]}\Big\{-\int_0^{\theta_1}(a_1^*-\theta-b)^2d\theta-\int_{\theta_1}^{1}(a_2^*-\theta-b)^2d\theta-c\big(1-\theta_1^3-(1-\theta_1)^3\big)\Big\}.$$

Take the first order condition with respect to θ_1, we have:

$$-(a_1^*-\theta_1-b)^2+(a_2^*-\theta_1-b)^2+3c\theta_1^2-3c(1-\theta_1)^2=0.$$

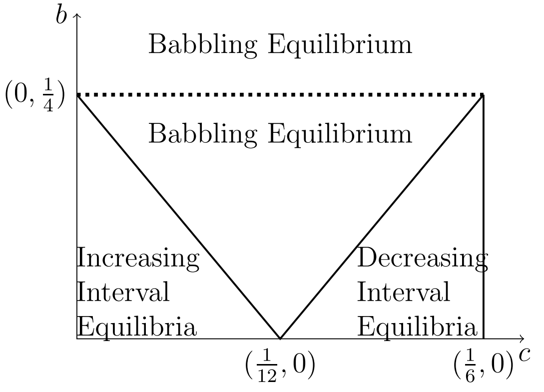

Fig. 3. Existence of informative equilibria.

The above equation is equivalent to:

$$3c(2\theta_1 - 1) + (a_2^* - a_1^*)(a_1^* + a_2^* - 2\theta_1 - 2b) = 0$$

Plugging in the value of a_1^* and a_2^*, we get a linear equation of θ_1, with solution:

$$\theta_1 = \frac{1}{2} - \frac{2b}{1-12c} = \frac{1}{2}\left(1 - \frac{4b}{1-12c}\right),$$

and $\frac{1}{2} - \frac{2b}{1-12c} \in (0, 1)$ when $c \in (\frac{4b+1}{12}, \frac{1}{6}] \bigcup (0, \frac{1-4b}{12})$.

Next, the second order derivative of the objective function with respect to θ_1 equals to:

$$\begin{aligned} &-2(\theta_1 + b - a_1^*) + 2(\theta_1 + b - a_2^*) + 6c\theta_1 + 6c(1-\theta_1) \\ &= -2(a_2^* - a_1^*) + 6c = -1 + 6c \leq 0, \end{aligned}$$

and hence, the objective function is concave when $\theta_1 \in [0, 1]$, which shows that the first order condition is sufficient for a global maximum. Hence, $\theta_1 = \frac{1}{2} - \frac{2b}{1-12c}$ solves the optimization problem, and thus, the sender's strategy is a best response to the receiver's strategy. □

This result is shown graphically in Fig. 3, which displays the set of parameter values under which informative equilibria exist, and whether they take the form of increasing or decreasing intervals. Fixing $b \in (0, \frac{1}{4})$, when $c \in (0, \frac{1-4b}{12}) \cup (\frac{1+4b}{12}, \frac{1}{6}]$, the equilibrium partition point in a 2-partition equilibrium is given by:

$$\theta_1 = \frac{1}{2} - \frac{2b}{1-12c}.$$

This relationship is depicted in Fig. 4.

3.4. Discussion

I explain the logic behind the reverse informativeness result using 2-partition equilibria. Since the sender's information acquisition process is covert, and she cannot commit to an information structure, her marginal benefit and her marginal cost of *increasing the partition point*, θ_1, must be equal in equilibrium. In what follows, I compute this marginal benefit and marginal cost when the receiver's equilibrium actions are $a_1^* \equiv \frac{\theta_1}{2}$ and $a_2^* \equiv \frac{1+\theta_1}{2}$ respectively.

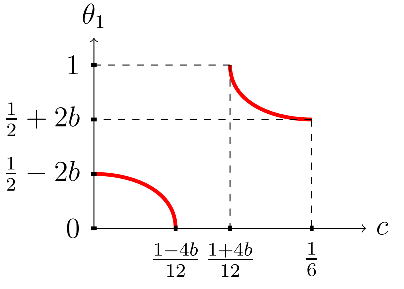

Fig. 4. Equilibrium partition point in 2-partition equilibria ($b = \frac{1}{16}$).

- The sender's expected payoff from the project is: $-\int_0^{\theta_1}(a_1^* - \theta - b)^2 d\theta - \int_{\theta_1}^1 (a_2^* - \theta - b)^2 d\theta$. Taking the derivative with respect to θ_1, her marginal benefit (MB) is given by:

$$MB = (a_2^* - \theta_1 - b)^2 - (a_1^* - \theta_1 - b)^2 = (\underbrace{a_2^* - a_1^*}_{=\frac{1}{2}})(\underbrace{a_1^* + a_2^*}_{=\theta_1+\frac{1}{2}} - 2\theta_1 - 2b)$$

$$= -\frac{1}{2}\theta_1 + \frac{1}{4}(1-4b)$$

- Her cost of information acquisition is $c\Big(1 - \theta_1^3 - (1-\theta_1)^3\Big)$. Her marginal cost (MC) is:

$$MC = \frac{\partial c\Big(1 - \theta_1^3 - (1-\theta_1)^3\Big)}{\partial \theta_1} = 3c(1 - 2\theta_1) = 6c(\frac{1}{2} - \theta_1).$$

  Since the cost of information acquisition is maximized at $\theta_1 = \frac{1}{2}$, so the marginal cost is positive when $\theta_1 < \frac{1}{2}$, and negative when $\theta_1 > \frac{1}{2}$. The marginal cost at $\theta_1 = \frac{1}{2}$ is always 0.

Fig. 5 depicts the sender's marginal benefit and the marginal cost as functions of θ_1. When c increases, the marginal cost curve becomes steeper and rotates clockwise around the point $(\frac{1}{2}, 0)$. The equilibrium partition point is where the two lines intersect. From the left panel to the right, we have increasing interval equilibrium, non-existence of informative equilibrium, and decreasing interval equilibrium.

For some intuition, since the sender is upwardly biased ($b > 0$), when $c = 0$, θ_1 must be below $\frac{1}{2}$, because otherwise, the sender will have an incentive to reduce θ_1. When c is low ($c \in (0, \frac{1-4b}{12})$), an increase in c makes this deviation more tempting, because when $\theta_1 < \frac{1}{2}$, reducing θ_1 also reduces the cost of information acquisition. As a result, θ_1 decreases with c. However, when c becomes sufficiently large ($c \in (\frac{1+4b}{12}, \frac{1}{6}]$), a high information acquisition cost helps to sender to commit not to acquire a more expensive information structure. This commitment effect is stronger when c increases, and the 2-partition equilibrium becomes more informative when c increases.

The existence of informative equilibria is not monotone with respect to c because when $c \in [\frac{1-4b}{12}, \frac{1+4b}{12}]$, the information acquisition cost is too large to support equilibria where $\theta_1 < \frac{1}{2}$ (since the sender is too tempted to reduce θ_1), while it is not large enough to commit the sender to acquire an information structure with $\theta_1 > \frac{1}{2}$ (the cost is too small such that the sender is tempted to reduce θ_1 and acquire a more expensive information structure, since her marginal

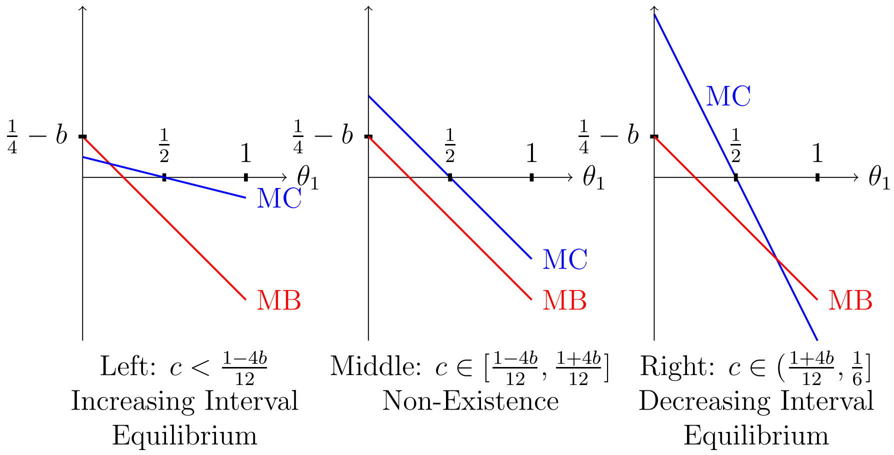

Fig. 5. Marginal benefit & marginal cost of increasing θ_1.

benefit of doing so is very large). Thus, no informative equilibrium exists.[18] When c is too large ($c > \frac{1}{6}$), the sender has a profitable '*non-local*' deviation by pooling the two intervals together and acquiring no information, thus the only equilibrium in this case is the babbling equilibrium.

4. Conclusion

I introduce endogenous information acquisition into cheap talk games and find that contrary to the conventional wisdom in Crawford and Sobel (1982), conflict of interest between players never causes the sender to withhold information; and their paper's implication on communication informativeness is reversed when information acquisition cost is sufficiently large. However, my benchmark result is sensitive to the flexibility of the sender's informational choice and the receiver's knowledge about the sender's information structure.[19]

Acknowledgments

I am indebted to Christian Hellwig, Jean Tirole, Robert Gibbons and Juuso Toikka for guidance and support. I thank an associate editor, a referee, Eduardo Azevedo, Patrick Bolton, Jacques Crémer, Glenn Ellison, Paul Heidhues, Johannes Hörner, Fuhito Kojima, George Lukyanov, Alan Olivi, Heikki Rantakari, Joel Sobel, Birger Wernerfelt, Nils Wernerfelt, Takuro Yamashita, seminar participants at MIT, Toulouse, St. Louis, AEA 2013 for helpful discussions. I am grateful to the financial support from the European Research Council under the European Community's 7th Framework Programme FP7/2007-2013 grant agreement N263790.

[18] The non-existence of informative equilibrium with more than two partition elements follows from a similar intuition, i.e. for any two adjacent intervals, $(\theta_{i-1}, \theta_i]$ and $(\theta_i, \theta_{i+1}]$, the sender always has an incentive to change θ_i.

[19] Flexibility of the sender's informational choice refers to the assumption that the sender always has the flexibility to "coarsen" her information structure by acquiring a strictly coarser partition, and by doing so, she can strictly save her information acquisition cost. Under this degree of flexibility, all pure strategy equilibria achieve full communication.

Appendix A. Remaining proofs

Proof of Lemma 3.1. I show by contradiction that partitions which are not interval partitions cannot be equilibrium information structures. The proof is done by constructing a profitable deviation for the sender.

Formally, let $A^* = \{a_1^*, \ldots, a_n^*\}$ be the equilibrium action set,[20] with $a_1^* < \ldots < a_n^*$. Suppose the sender's equilibrium strategy is:

- Acquiring partition information structure $\{\Theta_1, \ldots, \Theta_n\}$ at the information acquisition stage, and inducing action a_i^* when $\theta \in \Theta_i$,

where there exists i such that Θ_i is not an interval.

Let $p_i \equiv \mu_0(\Theta_i)$, $\Theta_1' \equiv [0, p_1]$, and $\Theta_i' \equiv (\sum_{j=1}^{i-1} p_j, \sum_{j=1}^{i} p_j]$ for every $i \geq 2$. Since $\sum_{i=1}^{n} p_i = 1$, $\{\Theta_1', \ldots, \Theta_n'\}$ is an interval partition of [0, 1].

I show that the sender's strategy that I have just described is dominated by the following strategy:

- Acquiring interval partition information structure $\{\Theta_1', \ldots, \Theta_n'\}$ at the information acquisition stage, and inducing action a_i^* when $\theta \in \Theta_i'$.

Furthermore, this domination is strict if there exists $i \in \{1, 2, \ldots, n\}$, such that either $\Theta_i \backslash (\Theta_i \cap \Theta_i')$ or $\Theta_i' \backslash (\Theta_i \cap \Theta_i')$ has strictly positive measure.

Since θ is uniformly distributed on [0, 1], $\mu_0(\Theta_i') = p_i = \mu_0(\Theta_i)$ for all i. Since the cost of an information structure only depends on the measures of its partition elements, the two information structures cost the same. Hence, we only need to show that:

$$\underbrace{-\sum_{i=1}^{n} \int_{\Theta_i} (a_i^* - \theta - b)^2 d\theta}_{\text{payoff under non-interval partition information structure}} \leq \underbrace{-\sum_{i=1}^{n} \int_{\Theta_i'} (a_i^* - \theta - b)^2 d\theta}_{\text{payoff under interval partition information structure}} .$$

Since

$$\sum_{i=1}^{n} \int_{\Theta_i} (a_i^* - \theta - b)^2 d\theta$$
$$= \sum_{i=1}^{n} \mu_0(\Theta_i) a_i^{*2} + \int_0^1 \theta^2 d\theta + \int_0^1 b^2 d\theta + 2b \int_0^1 \theta d\theta - 2b \sum_{i=1}^{n} \mu_0(\Theta_i) a_i^*$$
$$- 2\sum_{i=1}^{n} \int_{\Theta_i} a_i^* \theta d\theta,$$

and

[20] From the conclusion of Proposition 1 as well as the assumption that the number of partition element is finite in every available information structure, it is without loss of generality to focus on finite equilibrium action sets.

$$\sum_{i=1}^{n}\int_{\Theta_i'}(a_i^*-\theta-b)^2 d\theta$$
$$=\sum_{i=1}^{n}\mu_0(\Theta_i')a_i^{*2}+\int_0^1\theta^2 d\theta+\int_0^1 b^2 d\theta+2b\int_0^1\theta d\theta-2b\sum_{i=1}^{n}\mu_0(\Theta_i')a_i^*$$
$$-2\sum_{i=1}^{n}\int_{\Theta_i'}a_i^*\theta d\theta,$$

as well as $\mu_0(\Theta_i)=\mu_0(\Theta_i')=p_i$ for all i, the two expressions differ only in their last terms. Hence what we need to show is:

$$\sum_{i=1}^{n}\int_{\Theta_i}a_i^*\theta d\theta \le \sum_{i=1}^{n}\int_{\Theta_i'}a_i^*\theta d\theta. \tag{A.1}$$

To show this inequality, let

$$Q_i \equiv \int_{\theta\in\Theta_i}\theta d\theta, \quad Q_i' \equiv \int_{\theta\in\Theta_i'}\theta d\theta.$$

Since Θ_i and Θ_i' have the same Lebesgue measure, a useful observation is that:

$$\sum_{i=k}^{n}Q_i \le \sum_{i=k}^{n}Q_i' = \int_{p_1+\ldots+p_{k-1}}^{1}\theta d\theta,$$

for all $k\in\{1,2,\ldots,n\}$. Using summation by parts, we have:

$$\sum_{i=1}^{n}\int_{\Theta_i}a_i^*\theta d\theta = \sum_{i=1}^{n}a_i^*Q_i$$
$$=a_1^*\sum_{j=1}^{n}Q_j+\sum_{i=2}^{n}(a_i^*-a_{i-1}^*)\Big(\sum_{j=i}^{n}Q_j\Big)$$
$$\le a_1^*\sum_{j=1}^{n}Q_j'+\sum_{i=2}^{n}(a_i^*-a_{i-1}^*)\Big(\sum_{j=i}^{n}Q_j'\Big)$$
$$=\sum_{i=1}^{n}a_i^*Q_i'=\sum_{i=1}^{n}\int_{\Theta_i'}a_i^*\theta d\theta.$$

The inequality in the third line is true since $a_1^*<\ldots<a_n^*$ as well as $\sum_{i=k}^{n}Q_i\le\sum_{i=k}^{n}Q_i'$ for all k, which proves (A.1). □

Proof of Lemma 3.2. The '*only if*' part ((3) is necessary) is obvious. I show the '*if*' part ((3) is sufficient) by showing that the following strategy profile constitutes a Perfect Bayesian Equilibrium if (3) holds:

- The sender acquires interval partition information structure

$$\{[\theta_0, \theta_1], (\theta_1, \theta_2], \ldots, (\theta_{n-1}, \theta_n]\}$$

  and induces $a_i^* \equiv \frac{\theta_{i-1}+\theta_i}{2}$ when $\theta \in (\theta_{i-1}, \theta_i]$;
- The receiver chooses a_i^* after the sender telling him that $\theta \in (\theta_{i-1}, \theta_i]$;
- Every message the sender sends induces an action in $\{a_1^*, \ldots, a_n^*\}$.

First, fixing the sender's strategy, the receiver's action rule is sequentially rational. Hence, we only need to show that given the receiver's action rule, the sender's strategy is a best response.[21]

Let $A^* \equiv \{a_1^*, \ldots, a_n^*\}$ with a_i^* defined as above. Whenever she acquires an interval partition information structure, a strategy of the sender can be characterized by $\{\theta_0', \ldots, \theta_n', a_1, \ldots, a_n\}$ where $0 = \theta_0' \le \theta_1' \le \ldots \le \theta_n' = 1$ is the information structure she acquires and $a_i \in A^*$ is the equilibrium action the sender induces after knowing that $\theta \in (\theta_{i-1}', \theta_i']$.[22]

I show by contradiction that $\{\theta_0, \ldots, \theta_n, a_1^*, \ldots, a_n^*\}$ is the sender's best response. Suppose the sender's best response is $\{\theta_0', \ldots, \theta_n', a_1, \ldots, a_n\}$, and it delivers her a strictly higher expected payoff comparing with $\{\theta_0, \ldots, \theta_n, a_1^*, \ldots, a_n^*\}$. I introduce three useful observations.

- **Observation 1:** Since $u^s(a, \theta)$ exhibits strictly increasing differences in a and θ, by Topkis Theorem (Topkis, 1998), $a_1 \le a_2 \le \ldots \le a_n$.
- **Observation 2:** Every equilibrium action can be induced in at most one non-degenerate interval, or formally, for every $1 \le k \le n$, there exists **at most one** j ($1 \le j \le n$), such that $\theta_{j-1} < \theta_j$ and $a_j = a_k^*$.
  The proof shares the same idea as Proposition 1. Suppose towards a contradiction, that there exists such i and j, such that $1 \le i < j \le n$ with $\theta_{i-1}' < \theta_i'$, $\theta_{j-1}' < \theta_j'$ and $a_i = a_j = a_k^*$. Then the sender's expected payoff strictly increases when she pools intervals $(\theta_{i-1}', \theta_i']$ and $(\theta_{j-1}', \theta_j']$ together at the information acquisition stage (she can do this since the set of information structures is rich), and induces action a_k^* when $\theta \in (\theta_{i-1}', \theta_i'] \cup (\theta_{j-1}', \theta_j']$. This is because the alternative strategy strictly saves information acquisition cost (by monotonicity) while her expected gain from the project remains unchanged.
- **Observation 3:** Eq. (3) implies that the sender's payoff is (weakly) higher under $\{\theta_0, \ldots, \theta_n, a_1^*, \ldots, a_n^*\}$ comparing with $\{\theta_0^*, \ldots, \theta_n^*, a_1^*, \ldots, a_n^*\}$ for every $0 = \theta_0^* \le \theta_1^* \le \ldots \le \theta_{n-1}^* \le \theta_n^* = 1$.

In what follows, I use these three observations to show that $\{\theta_0, \ldots, \theta_n, a_1^*, \ldots, a_n^*\}$ is the sender's best response (given the receiver's strategy). Specifically, I show that if an alternative strategy, $\{\theta_0', \ldots, \theta_n', a_1, \ldots, a_n\}$, is the sender's best response, then there exists another strategy $\{\theta_0'', \ldots, \theta_n'', a_1^*, \ldots, a_n^*\}$, which achieves the same expected payoff for the sender as in $\{\theta_0', \ldots, \theta_n', a_1, \ldots, a_n\}$, but this payoff is weakly smaller than the sender's expected payoff under $\{\theta_0, \ldots, \theta_n, a_1^*, \ldots, a_n^*\}$. Thus, the sender has no profitable deviation.

[21] In this proof, I only consider the sender's ex ante incentives. Her incentive constraints at the interim stage has been implied since $\theta_{i-1} < \theta_i$ for all i, and hence, $\theta \in (\theta_{i-1}, \theta_i]$ happens with strictly positive probability.

[22] When $\theta_{i-1}' = \theta_i'$, the interval $(\theta_{i-1}', \theta_i']$ is degenerate, meaning that the sender is acquiring an information structure with fewer than n partition elements. Also, since we are interested in the sender's best responses, it is without loss of generality to focus on pure strategies.

Suppose $\{\theta'_0, \dots, \theta'_n, a_1, \dots, a_n\}$ is the sender's best response, define set $\mathcal{I}$ as:

$$\mathcal{I} \equiv \left\{ i \middle| \exists j, \text{ s.t. } \theta'_{j-1} \neq \theta'_j, \text{ and } a_j = a^*_i \right\}.$$

By definition, $\{a^*_i | i \in \mathcal{I}\}$ is the set of actions induced with strictly positive probability under $\{\theta'_0, \dots, \theta'_n, a_1, \dots, a_n\}$. From Observations 1 and 2, and because $\{\theta'_0, \dots, \theta'_n, a_1, \dots, a_n\}$ is a best response, for every $i \in \mathcal{I}$, there exists a unique j such that $\theta'_{j-1} \neq \theta'_j$ and $a_j = a^*_i$. Let $k(i) = j$, and hence, $k(i)$ is uniquely defined for all $i \in \mathcal{I}$. This implies that:

$$\bigcup_{i \in \mathcal{I}} \left(\theta'_{k(i)-1}, \theta'_{k(i)} \right] = [0, 1].$$

Next, let us define $(\theta''_0, \dots, \theta''_n)$. For every $i \in \mathcal{I}$, let

$$\theta''_{i-1} \equiv \theta'_{k(i)-1}, \quad \theta''_i \equiv \theta'_{k(i)}.$$

From Observation 1, for every $i \in \mathcal{I}$ or $i + 1 \in \mathcal{I}$, θ''_i is well-defined and is weakly increasing in its index i.

For all other $0 \leq i \leq n$, i.e. $i \notin \mathcal{I}$ and $i + 1 \notin \mathcal{I}$, θ''_i is uniquely defined by the constraint: $0 = \theta''_0 \leq \theta''_1 \leq \dots \leq \theta''_{n-1} \leq \theta''_n = 1$. Since

$$\bigcup_{i \in \mathcal{I}} \left(\theta'_{k(i)-1}, \theta'_{k(i)} \right] = [0, 1].$$

$\{[\theta''_0, \theta''_1], \dots, (\theta''_{n-1}, \theta''_n]\}$ is an interval partition information structure.

Consider strategy $\{\theta''_0, \dots, \theta''_n, a^*_1, \dots, a^*_n\}$. By definition, it delivers her the same expected payoff as strategy $\{\theta'_0, \dots, \theta'_n, a_1, \dots, a_n\}$. From Observation 3, the sender's expected payoff is weakly greater under $\{\theta_0, \dots, \theta_n, a^*_1, \dots, a^*_n\}$ comparing with $\{\theta''_0, \dots, \theta''_n, a^*_1, \dots, a^*_n\}$, which contradicts the assumption that strategy $\{\theta'_0, \dots, \theta'_n, a_1, \dots, a_n\}$ gives the sender a strictly higher payoff comparing to $\{\theta_0, \dots, \theta_n, a^*_1, \dots, a^*_n\}$.

Hence, there exists no strictly profitable deviation, which implies that (3) is also sufficient. □

Appendix B. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jet.2015.08.011.

References

Aghion, P., Tirole, J., 1997. Formal and real authority in organizations. J. Polit. Econ. 105, 1–29.

Argenziano, R., Severinov, S., Squintani, F., 2014. Strategic information acquisition and transmission. Working paper. University of Essex.

Austen-Smith, D., 1994. Strategic transmission of costly information. Econometrica 62, 955–963.

Crawford, V., Sobel, J., 1982. Strategic information transmission. Econometrica 50, 1431–1451.

Gentzkow, M., Kamenica, E., 2012. Disclosure of endogenous information. Working paper. University of Chicago.

Hellwig, C., Veldkamp, L., 2009. Knowing what others know: coordination motives in information acquisition. Rev. Econ. Stud. 76, 223–251.

Ivanov, M., 2010. Informational control and organizational design. J. Econ. Theory 145, 721–751.

Kamenica, E., Gentzkow, M., 2011. Bayesian persuasion. Am. Econ. Rev. 101, 2590–2615.

Sims, C., 2003. Implications of rational inattention. J. Monet. Econ. 50, 665–690.

Topkis, D., 1998. Supermodularity and Complementarity. Princeton University Press.

Yang, M., 2014. Coordination with flexible information acquisition. J. Econ. Theory. Forthcoming.

Yang, M., 2015. Optimality of debt under flexible information acquisition. Working paper. Duke University.