COMMUNICATION WITH ENDOGENOUS INFORMATION ACQUISITION*

DI PEI

Feb. 14th, 2013

Abstract

I develop a theory of communication in which experts need to gather information at a cost before giving advice to decision makers. In contrast with the canonical sender-receiver game in which the sender's information is exogenous and costless, I show that first, the sender always communicates all her information to the receiver in any equilibrium; and second, her advice can be more informative when recommending a decision which is more favorable to her. By applying my model to study organizational design, I find that, paradoxically, both delegating decision rights to the sender as well as monitoring the sender's information acquisition process can decrease her incentive to acquire information.

Keywords: cheap talk, information acquisition, full communication, rational inattention

JEL Codes: D82, D83

*Mailing Address: Room MF 003, Manufacture des Tabacs, Toulouse School of Economics. Email: pd07tinghua@gmail.com. I am indebted to my advisors: Christian Hellwig and Jean Tirole for their guidance and encouragements. I am grateful to Patrick Bolton and Takuro Yamashita for their discussions and feedbacks. I thank Eduardo Azevedo, Jacques Crémer, Daniel Garrett, Paul Heidhues, Johannes Hörner, Alessandro Ispano, Fuhito Kojima, George Lukyanov, Lucas Maestri, Thomas Mariotti, Filip Matějka, Patrick Rey, Adrian Torchiana, Ruoying Wang and the seminar participates in St. Louis, Toulouse, the 2013 AEA Annual Conference for helpful comments. I am also grateful to the financial support from the European Research Council under the European Community's 7th Framework Programme FP7/2007-2013 grant agreement N263790 (InfoMacro). All remaining errors are mine.

1 INTRODUCTION

Experts often need to gather information and conduct costly research before giving advice to decision makers. For example, financial advisors analyze data before proposing investment strategies; doctors diagnose patients before recommending treatments and therapies.[1] In these situations, an expert's comparative advantage is on gathering and processing information, but not the superior information per se. When analyzing communication between experts and decision makers, the canonical sender-receiver game (Crawford and Sobel [1982], hereafter, CS) focuses on the informational advantage of the sender (she): if she is perfectly informed and has a conflict of interest with the receiver (he), then she can never fully communicate her information. Moreover, her advice is less informative when recommending a decision in which she enjoys a larger private benefit.[2]

In contrast, I emphasize the sender's ability to gather information by analyzing the CS game with an information acquisition stage, in which the sender acquires information at a cost. This interaction between *moral hazard in acquiring information* and *adverse selection in reporting information* overturns many conclusions of CS, in which the sender's information is exogenous and costless. Different from CS, I show that first, the sender *always* tells the receiver everything she knows, i.e. every equilibrium must achieve *full communication*; and second, her advice can be more informative when recommending an action which is more favorable to her (*communication informativeness*). Moreover, both *delegation* (delegate decision rights to the sender) and *monitoring* (monitor the sender's information acquisition process) can discourage the sender from acquiring information, which is in sharp contrast with Aghion and Tirole (1997) as well as Holmström (1979).

Formally speaking, my *full communication* result holds whenever the sender can always coarsen her information, and a coarser information structure costs less to acquire.[3] The intuition behind this result is simple: if the sender cannot transmit all her information ex post, then why does she acquire that information ex ante? By studying the problem in less detail, she saves time and enjoys more leisure. The same intuition applies when the receiver can consult multiple senders, or when

[1] Other examples include: a lobbyist studying a newly proposed regulation policy, a lawyer serving a new client, etc.

[2] There are many ways to interpret this conflict of interest. Financial advisors receive more commission fees from an increase in purchasing volume; doctors prefer their patients to select more expensive medical treatments; investment banks gain from issuing new equity stock and corporate mergers, etc. In these examples, an increase in purchasing volume, a more expensive treatment and issuing more new equity are decisions in which the sender enjoys a larger private benefit.

[3] Many other papers identify situations in which the sender's information can be fully communicated. This includes a multi-dimensional or a multi-sender setting (Krishna and Morgan [2001], Battaglini [2002], Ambrus and Takahashi [2008]), the receiver has private information or the sender has incomplete information (Watson [1996]), Ivanov [2010]), the sender has honesty concerns or she faces a convex lying cost (Olszewski [2004], Kartik et.al. [2007]), etc.

he can acquire some information himself.

This result establishes a benchmark in analyzing communication under conflict of interests: when the sender's information is acquired at a cost, all frictions caused by the misalignment of incentives must be fully absorbed at the information acquisition stage. Put it differently, the sender's incentive compatibility constraint never binds at the communication stage.

My communication informativeness result is shown when the state space is one-dimensional and the sender's ideal action is always strictly larger than the receiver's. For example, when a financial analyst (she) gives advice to an investor (he), she partially internalizes her customer's welfare, but also values her private benefit (a monetary reward, such as commission fees) which is increasing with the investor's purchasing volume. When the analyst needs to gather information, there are two effects influencing communication informativeness. One is the *adverse selection effect* in reporting information, which has been addressed by CS: when the analyst recommends investing more, the conflict of interest makes her advice less credible and her customer more vigilant. The other one is new, which is the *moral hazard effect* in acquiring information: when the analyst's preliminary research suggests that the investment opportunity is more profitable, she has larger incentives to study the situation in more detail since she expects to receive more private benefit. Thus, conveying more favorable information indicates that she is more knowledgeable and her advice is more valuable.[4] This moral hazard effect dominates when information acquisition costs are sufficiently large.

In firms, organizations and political institutions, the receiver (or principal) can also delegate decision rights to the sender (or agent). To overcome the moral hazard problem, he can also monitor the sender when she is acquiring information. To incorporate both elements, I allow the receiver to choose *publicly* whether or not to delegate and whether or not to monitor before the sender acquires information. I study the amount of information the sender acquires under each institutional arrangement. When information acquisition cost is high enough:

1. *Delegation undermines incentives:* The sender acquires less information when the decision right is delegated. This is because delegation gives the sender more freedom in choosing actions, and increases her expected payoff when she has coarser information.[5]

[4] In Crawford and Sobel (1982), every equilibrium is characterized by an interval partition. When the prior is uniform, the lengths of the intervals in any equilibrium partition must be increasing (from left to right). Thus, communication is more informative in smaller state values. Morgan and Stocken (2003) analyze a situation in which the receiver does not know the sender's preference and obtain similar insights. In contrast, I show that the lengths of the intervals in an equilibrium partition can be either increasing or decreasing in an endogenous information acquisition setting, depending on the information acquisition cost.

[5] Pioneered by Holmström (1977), many papers studied the impact of delegation on the agent's incentives and

2. *Monitoring backfires:* The sender acquires less information when the receiver can observe her information structure or her cost of information acquisition. Monitoring backfires because the receiver cannot commit not to change his equilibrium action set after detecting a deviation at the information acquisition stage.

Related Literature: In modeling information acquisition, some recent papers allow players to choose from a family of information structures, which can be parameterized by a finite-dimensional variable.[6] This approach is adopted by Argenziano et.al.(2011), which is the predecessor work of this paper. In their model, the sender chooses the preciseness of her information by deciding how many rounds of Bernoulli experiments to conduct. Their formulation has the advantage of explicitly describing the information acquisition process. They show under the uniform-quadratic assumption that communication may force the sender to overinvest in information acquisition; and sometimes, the receiver is better informed when communicating with the sender, comparing with acquiring the information himself, or delegating decision right to the sender.

Comparing with their paper, I adopt the idea of '*rational inattention*' (Sims [2003]), where the sender can flexibly allocate her '*attention*' between different ranges of the state space.[7] I establish benchmark results in a general framework. My model is also more tractable when analyzing communication informativeness, as well as doing applications on organizational design.[8]

Other recent papers also allow the sender to choose her information structure flexibly before communicating with the receiver. In Gentzkow and Kamenica (2012), the sender is allowed to choose from a rich set of partition information structures. They derive a full communication result when information is hard and the information acquisition process is observable. As indicated before, their paper differs from mine in the nature of information: they focus on hard verifiable information,

---

the principal's welfare. For example, Ottaviani (2000), Dessein (2002), Alonso and Matouschek (2008), Amador and Bagwell (2012) study the problem when the agent's information is exogenous. Aghion and Tirole (1997) allows the agent to acquire information, but does not address the decision flexibility problem associated with delegation. The papers closest to mine in explanation are Szalay (2005), Che and Kartik (2009). The main differences are: whether there is a conflict of interest between the two parties, and whether information is hard or soft.

[6] This approach has been broadly applied to non-strategic (Athey and Levin [2001]) as well as strategic settings (Hellwig and Veldkamp [2009], Amir and Lazzati [2011], Myatt and Wallace [2012], etc.). Comparing with the earliest contributions, which usually assume that an agent can only have two possible information structures after acquiring information (either she is perfectly informed, or she is completely ignorant, for example: Austen-Smith [1994], Aghion and Tirole [1997]), this formulation adds an intensive margin to the agent's informational choice, by allowing him to choose the preciseness of his information.

[7] Other papers adopting this idea include: Matějka and Sims (2010), Yang (2011), etc.

[8] The set of available information structures in Argenziano et.al.(2011) is not rich, making some of the equilibria in their model not fully communicative. Another application of my model is on the 'pandering' problem (Che et.al. [2013]), which is available upon request. I adopt their setup except assuming that the sender only knows the payoff of one project but needs to acquire information on the other. I show that the moral hazard problem in acquiring information can sometimes offset the distortions caused by asymmetric priors (pandering).

while mine is about soft unverifiable information. Kamenica and Gentzkow (2011) characterize the sender's optimal information structure when trying to persuade a decision maker. But in their model, information acquisition has no cost.

Ivanov (2010) studies the case in which the receiver can 'choose' the sender's information structure (informational control), and he shows that the optimal information structure must be an interval partition. Under informational control, communication informativeness can be improved upon the most informative CS equilibrium, i.e. the sender can sometimes convey more information when she knows less. In contrast, I show that when the sender acquires her information privately, the above comparison no longer holds — communication is always less informative than the optimal CS outcome.

2 THE GENERAL MODEL

In this section, I show that first, full communication is necessary for any equilibrium after defining '*coarseness*', '*richness*' and '*monotonicity*'; and second, the sender's *incentive constraint at the information acquisition stage* implies her *incentive constraint at the communication stage*, making the former a sufficient condition for an equilibrium, as long as the players' preferences satisfy a '*No Cycle Condition*'. Later, I extend my full communication result to a multi-sender setting as well as an informed-receiver setting.

2.1 Preferences and Priors

The receiver (he) needs to make a decision $a \in A = \mathbb{R}^k$ on an unfamiliar project, which is parameterized by the state variable $\theta \in \Theta \subset \mathbb{R}^k$. The receiver and the sender (she) are both risk neutral, and their gains from the project are $u^r(a, \theta)$ and $u^s(a, \theta)$ respectively. So each player's ex ante expected gain from the project is determined by the joint distribution of (θ, a).

Let $\Delta(\cdot)$ be the set of distributions on a given space. The two players share a common prior $\mu_0 \in \Delta(\Theta)$, which is absolutely continuous.

2.2 The Information Acquisition Stage

At the information acquisition stage, the sender chooses her information structure ψ from the set of '*available information structures*': Ψ. This choice is non-observable to the receiver. An *information structure* $\psi \in \Psi$ consists of a family of distributions $\{\psi(\cdot|\theta)\}_\theta$ over the *signal realization space* Ω, with a typical element $\omega \in \Omega$.

Every ω leads to a *posterior belief* on θ: $\mu^s_{\omega,\psi} \in \Delta(\Theta)$, which is called the sender's '*type*'. Every information structure induces a distribution over posterior beliefs. Two information structures are the *same* if and only if the distributions they induce are the same almost surely. An information structure is '*coarser*' than the other if it pools some sets of realizations together, and forms a new realization for each set.

Condition 2.1 (Coarseness). *ψ' is strictly coarser than ψ if $\psi' \neq \psi$ and there exists a set of realizations $\{\omega_\lambda\}_\lambda$ and a family of sets $\{\Omega_\lambda\}_\lambda$ ($\lambda \in \Lambda$, while Λ is a set of indexes) such that:*

1. *$\omega_\lambda \in \Omega_\lambda \subset \Omega$ for every λ;*

2. *$\Omega_\lambda \bigcap \Omega_{\lambda'} = \emptyset$ if $\lambda \neq \lambda'$.*

3. *For every $\theta \in \Theta$:*
$$\psi'(\omega|\theta) = \begin{cases} \psi(\omega|\theta) & \text{if } \omega \notin \bigcup_\lambda \Omega_\lambda \\ \int_{\omega \in \Omega_\lambda} \psi(\omega|\theta) d\omega & \text{if } \omega = \omega_\lambda \\ 0 & \text{otherwise} \end{cases}$$

The definition the '*richness*' of a set comes immediately after defining 'coarseness':

Condition 2.2 (Richness). *Ψ is rich if for any $\psi \in \Psi$ and ψ' strictly coarser than ψ, then $\psi' \in \Psi$.*

In a nutshell, a set of information structures is *rich* if the sender can always '*coarsen*' her information by disregarding the difference between any subset of signal realizations.

Acquiring information structure ψ brings an additive cost $C(\psi) \in \mathbb{R}^+$ to the sender. The '*monotonicity*' condition imposes an appropriate ranking among the costs of information structures:

Condition 2.3 (Monotonicity). *$C(\cdot)$ satisfies 'monotonicity' if for all $\psi, \psi' \in \Psi$:*

1. *$C(\psi) = C(\psi')$ if ψ and ψ' are the same;*

2. *$C(\psi) > C(\psi')$ if ψ' is strictly coarser than ψ.*

All my discussions are based on the following assumption:

Assumption 1. *Ψ is rich and $C(\cdot)$ satisfies monotonicity.*

2.3 The Communication Stage

Let M be the set of messages, which has sufficient number of elements. Under information structure ψ, the sender's mixed strategy after receiving ω is denoted by $\sigma_\omega^\psi \in \Delta(M)$. Her behavior strategy is composed of an information structure ψ, and a vector of distributions $\sigma^\psi = (\sigma_\omega^\psi)_\omega$. The two together induce a joint distribution of (θ, m).

Let $\mu_m^r \in \Delta(\Theta)$ be the receiver's posterior belief after receiving m. Vector $\mu^r = (\mu_m^r)_m$ is his *belief updating rule*. Let $\alpha_m \in \Delta(A)$ be his mixed strategy after receiving m. Vector $\alpha = (\alpha_m)_m$ is his *mixed action rule*.

2.4 Equilibrium

A *Perfect Bayesian Equilibrium* (hereafter, equilibrium) is characterized by $(\psi, \sigma^\psi, \mu^r, \alpha)$, and satisfies:

1. Given the receiver's belief updating rule μ^r and his mixed action rule α, type $\mu_{\omega,\psi}^s$ sender chooses a message which maximizes her *interim expected gain* from the project.[9]

$$\int_\theta \int_a u^s(a,\theta) d\alpha_m(a) d\mu_{\omega,\psi}^s(\theta)$$

2. Given the receiver's belief updating rule and his mixed action rule, $\psi \in \Psi$ is chosen to maximize the sender's *ex ante expected gain* from the project minus the information acquisition cost:

$$\int_\Omega \Big(\underbrace{\int_\theta \psi(\omega|\theta) d\mu_0(\theta)}_{\text{Prob. of } \omega \text{ to occur}} \underbrace{\max_{m \in M} \{ \int_\theta \int_a u^s(a,\theta) d\alpha_m(a) d\mu_{\omega,\psi}^s(\theta) \}}_{\text{Sender's gain from the project under } \omega.} \Big) d\omega - C(\psi)$$

3. Under belief updating rule μ_m^r, almost all actions taken satisfy:

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

4. Both players' beliefs are updated according to Bayes Rule.

Furthermore, ψ is an *equilibrium information structure* if it is part of an equilibrium. $\Lambda^*(\psi) \equiv \{\alpha(m) | m \in M\}$ is the *equilibrium mix action set*. Each element in this set is an '*equilibrium mix*

[9] As a convention, all statements here hold with probability 1.

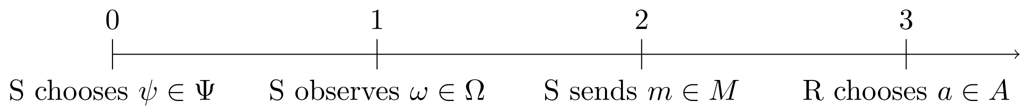

Figure I Timeline

action'. Let $A^*(\psi)$ be the *equilibrium action set*, with each of its element on the support of an equilibrium mixed action.

When $\Lambda^*(\psi)$ is a singleton, it is a *babbling equilibrium*, in which the sender acquires no information and the receiver's action is independent of her message. Obviously, the babbling equilibrium always exists.

2.5 Full Communication

First, I define '*full communication*':

Definition 1 (Full Communication)**.** *An equilibrium achieves full communication if the sender and the receiver's posterior beliefs are the same almost surely.*

Intuitively, the sender tells the receiver everything she knows in such an equilibrium. Now, I state my main result:

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

By '*pure strategy*', I only restrict the sender to use a pure strategy when acquiring information. I allow her to use mixed strategies when sending messages and the receiver to use mixed strategies when taking actions.

Proof of Proposition 1: I prove by contradiction.[10] Let $\psi \in \Psi$ be the equilibrium information structure. If there exists two types of senders with belief μ_1 and μ_2, both sending message m with positive probability, then:

$$m \in \arg\max_{\tilde{m}\in M} \int_\theta \int_a u^s(a,\theta) d\alpha_{\tilde{m}}(a) d\mu_j(\theta) \tag{2.1}$$

for $j = 1, 2$.

[10] Here, I focus on the case where each sender type occurs with positive probability. In Appendix I, I will discuss types occurring with 0 probability.

Let ω_1 and ω_2 be the signal realizations which induce μ_1 and μ_2 respectively. The sender can deviate by pooling ω_1 and ω_2 together at the information acquisition stage, which leads to an information structure coarser than ψ. When belonging to the new '*aggregate type*', the sender sends m. From (2.1), her expected payoff from the project remains unchanged.

This new information structure is available since Ψ is rich; and the sender can strictly reduces information acquisition cost since C is monotone. This is a profitable deviation, which leads to a contradiction.

In equilibrium, each message can be sent by at most one type of sender. From Bayes Rule, the receiver's can fully infer the sender's type. This is because he can correctly anticipate the sender's information structure when the sender is using a pure strategy, and each message can only be sent by 1 type of sender. So his posterior belief equals to the sender's. □

Proposition 1 shows that when the sender has no superior information ex ante, but enjoys sufficiently flexibility when choosing information structures, then her information must be fully transmitted in any equilibrium, i.e., there is no information loss in the strategic transmission process. This result not only provides useful benchmark analysis, but also stands significant as a necessary condition for any equilibrium.

Remark: My full communication result relies on the richness of Ψ and the monotonicity of C. Also, I focus on a subset of equilibria in which the sender uses a pure strategy to acquire information.

1. The concept of richness is closely related to the idea of '*rational inattention*': The sender can rationally allocate her '*attention*' between different pieces of information. At the information acquisition stage, when pooling a set of realizations together and disregarding the differences between them, she '*pays no attention*' to the realizations within this set. Some examples of a rich Ψ include:

    - Ψ is the set of information structures where $\psi(\theta|\omega)$ is non-singular for any $\omega \in \Omega$. This is the standard assumption in the rational inattention literature.[11]
    - Let x be a random variable which is independent with θ and has support X. Ψ is the set of partition information structures, where the sender partitions $\Theta \times X$ into several

[11] Each distribution function can be written as the convex combination of a singular distribution, an absolutely continuous distribution and a discrete distribution. By non-singular, I mean the weight of the singular part equals to 0.

disjoint subsets. After acquiring information, she knows which subset in the partition is (θ, x) actually in.

Next, I discuss how my formulation can accommodate the standard one in which the sender receives a signal realization which equals to θ plus a Gaussian white noise. This can be done by allowing for an additional '*inattention noise*' when the sender absorbs, understands and processes the signal.[12] For example, the CPI or the FED's announcement can be modeled as a noisy signal to the future inflation rate. Remembering these numbers only up to a certain digit brings a player-specific *inattention noise* when absorbing the signal.

2. The monotonicity condition imposes a restriction on the relative costs of two information structures which can be compared under the '*coarseness*' criteria. This criteria trivially satisfies the Blackwell's ordering. As a result, most 'appropriate' information acquisition cost functions where there is no free information satisfies monotonicity.

3. Finally, I examine when is it without loss of generality to focus on pure strategy equilibria. Let $\nu\psi \bigoplus (1-\nu)\psi'$ be a distribution over posterior beliefs which is equivalent to the following compound lottery: first we draw a lottery, and with probability ν, we get the distribution induced by ψ, and with probability $1-\nu$, we get the distribution induced by ψ'. The distribution over posterior beliefs defined by $\nu\psi \bigoplus (1-\nu)\psi'$ can be induced by an information structure since it satisfies the martingale property. The convexity condition is defined as follows:

Condition 2.4 (Strict Convexity)**.** *C and Ψ satisfy convexity if for any $\psi, \psi' \in \Psi$ $(\psi \neq \psi')$, $\nu \in (0,1)$: $\nu\psi \bigoplus (1-\nu)\psi' \in \Psi$, and*

$$C(\nu\psi \bigoplus (1-\nu)\psi') < \nu C(\psi) + (1-\nu)C(\psi') \tag{2.2}$$

For example, if the sender can choose any non-singular distribution as her information structure, then the information structure space satisfies convexity. If the cost of information acquisition is given by the *Shannon's Entropy*, then information acquisition cost satisfies monotonicity and convexity. I establish a sufficient condition for ruling out non-trivial mixed strategies:

[12] The '*inattention noise*' is the same as the '*receiver noise*' in Myatt and Wallace (2012). I change the phrasing to avoid confusion with the name of the player.

Lemma 2.1. *If C and Ψ satisfy convexity, the sender always uses a pure strategy to acquire information.*

The detailed proof can be seen in Appendix I. The main idea is that the sender can use a convex combination of the information structures on the support of that mixed strategy. This deviation is available from the convexity of Ψ. Under this deviation, the joint distribution of (θ, m) does not change. Since the receiver's action rule is already fixed under a given equilibrium, the joint distribution of (θ, a) will not change either. This implies that the sender's expected gain from the project remains the same. But she strictly saves information acquisition cost due to the convexity of C.

2.6 Equilibrium Information Structure

Every equilibrium which achieves full communication can be sufficiently characterized by an *equilibrium information structure*, which the sender acquires and fully transmits. This subsection explores when is the sender's incentive constraint at the information acquisition stage sufficient for an equilibrium information structure, i.e. when can her incentive constraint at the information acquisition stage imply her incentive constraint at the communication stage.

From now on, I assume that $u^r(a, \theta)$ is concave in a, so the receiver always chooses a deterministic action under any posterior belief. Let

$$a^*(\mu) \equiv \arg\max_{a \in A} \int_\theta u^r(a, \theta) d\mu(\theta)$$

The equilibrium action set is expressed as:

$$A^*(\psi) = \{a^*(\mu^r_m) | m \in M\}$$

For any $\psi \in \Psi$, let

$$A(\psi) \equiv \{a^*(\mu^s_{\omega,\psi}) | \omega \in \Omega\}$$

be the *action set induced by* ψ. The following claim is implied by Proposition 1:

Corollary 2.1. *If ψ is an equilibrium information structure, then $A(\psi) = A^*(\psi)$.*

I focus on the case where $A(\psi)$ is a finite set. Let $U^s(\psi'|\psi)$ be the sender's maximum ex ante expected gain from the project if she acquires information structure ψ' while she is free to induce any action in $A(\psi)$:

$$U^s(\psi'|\psi) \equiv \int_{\Omega} \Big[\max_{a \in A(\psi)} \Big\{ \int_{\theta} u^s(a,\theta) d\mu^s_{\omega,\psi'}(\theta) \Big\} \int_{\theta} \psi'(\omega|\theta) d\mu_0(\theta) \Big] d\omega - C(\psi') \tag{2.3}$$

Corollary 2.1 implies that ψ is an equilibrium information structure if and only if the following incentive constraints are satisfied:

1. The sender's incentive constraint at the information acquisition stage (IC-AC):

$$\psi \in \arg\max_{\psi' \in \Psi} \{U^s(\psi'|\psi) - C(\psi')\} \tag{2.4}$$

   This means, given the set of equilibrium actions the sender is allowed to induce, she has the incentive to choose the equilibrium information structure.

2. The sender's incentive constraint at the communication stage (IC-CO):

$$a^*(\mu) \in \arg\max_{a \in A(\psi)} \int_{\theta} u^s(a,\theta) d\mu(\theta) \tag{2.5}$$

   for any $\mu \in \{\mu^s_{\omega,\psi} | \omega \in \Omega\}$. This means, after the sender acquires the equilibrium information structure, she has the incentive to truthfully report her findings.

In the rest of this section, I show that when u^r and u^s satisfy the '*No Cycle Condition*', then (IC-AC) implies (IC-CO), making the former to be a sufficient condition for an equilibrium information structure.

Intuitively, the players preferences satisfy the 'No Cycle Condition' if there exists a ranking of the sender's types, such that a higher type never has the incentive to mimic a lower one. Formally speaking, let $A(\psi) = \{a_1^*, ..., a_n^*\}$. The type which induces a_i^* is called *type i*. The condition is defined as follows:

Definition 2. *u^s and u^r satisfy the 'No Cycle Condition' under information structure ψ if for any $2 \le k \le n$, any permutation $\{\tau_1, ..., \tau_n\}$ of set $\{1, 2, ..., n\}$, when type τ_i sender prefers $a^*_{\tau_{i+1}}$ to $a^*_{\tau_i}$ (for any $1 \le i \le k-1$), then type $a^*_{\tau_k}$ cannot prefer $a^*_{\tau_1}$ to $a^*_{\tau_k}$.*[13]

I will discuss the generality of this condition after presenting my result:

Proposition 2. *If u^r and u^s satisfy the No-Cycle condition under ψ, then ψ is an equilibrium information structure if and only if it satisfies (IC-AC).*

[13] Equivalently, there exists no '*Top Trading Cycle*' (Shapley and Scarf [1974]) with length greater than 1.

This means, under the 'No Cycle Condition', one can safely ignore the sender's incentives to report information truthfully and only needs to consider whether or not she has the incentive to acquire the correct information. This is because once the sender has a profitable deviation at the communication stage, then there must exist at least two types who have the same favorite equilibrium action, or there exists a '*cycle*', in which each type prefers its neighboring type's equilibrium action. The second case has been ruled out by assumption. In the first case, the sender can profitably deviate by pooling the two types together at the information acquisition stage. This result not only stands significant as a benchmark, but also simplifies the process of characterizing an equilibrium.

Proof: I prove by contradiction.

Suppose the sender has a profitable deviation at the communication stage but has no profitable deviation at the information acquisition stage, then under information structure ψ, there exists at least one type of sender: τ_1, whose favorite action in $A^*(\psi)$ is not $a^*_{\tau_1}$.

Let type τ_1's favorite action be $a^*_{\tau_2}$. If type τ_2's favorite action is $a^*_{\tau_2}$, then the sender has a profitable deviation at the information acquisition stage by pooling type τ_1 and τ_2 together, which is a contradiction. Let type τ_2's favorite action be $a^*_{\tau_3}$. Continue this process and define $\tau_1, ..., \tau_k, \tau_{k+1}....$ Since the number of types is finite and there exists no type in $\{\tau_1, ..., \tau_n\}$ such that τ_i's favorite action is $a^*_{\tau_i}$, so there exists $n' \leq n+1$ such that $\tau_{n'} = \tau_{n''}$ where $n'' < n'$. Then, $\tau_{n''}, ..., \tau_{n'}$ forms a cycle with length greater than 1, which contradicts the 'No Cycle Condition'. $\square$

Remark: The finiteness of $A^*(\psi)$ and the 'No Cycle Condition' are indispensable for Proposition 2. Both impose restrictions on the players' preferences. To illustrate the generality of this result, I discuss 4 examples. Example 1 is the canonical CS game, where $A^*(\psi)$ is necessarily finite and the No Cycle Condition is satisfied under any information structure. Example 2 extends this to multiple dimensions. Example 3 and 4 are counterexamples to Proposition 2 in which either the 'No Cycle Condition' or the finiteness condition fails.

Example 1: Let $\theta \in \Theta \in \mathbb{R}$, $a \in \mathbb{R}$ and Θ is compact. If

$$\frac{\partial u^s(a,\theta)}{\partial a} > \frac{\partial u^r(a,\theta)}{\partial a}$$

for any θ, the number of equilibrium actions must be finite in any equilibrium.[14]

The sender always chooses an information structure with a finite number of types. Let $\{\mu_1, ..., \mu_n\}$ be the set of possible posterior beliefs, and

$$a_i^* \equiv \arg\max_a \int_\theta u^r(a,\theta)d\mu_i(\theta)$$

$$a_i^s \equiv \arg\max_a \int_\theta u^s(a,\theta)d\mu_i(\theta)$$

Then, it is obvious that $a_i^* < a_i^s$ for any i. Furthermore, a_i^* must be different between each other (otherwise, the sender can pool two types with the same induced action together at the information acquisition stage). So it is without loss of generality to assume that $a_1^* < ... < a_n^*$. Apparently, u^s and u^r satisfy the 'No Cycle Condition' under any ψ. This is because a type with a higher index prefers her own equilibrium action than a lower type's. From Proposition 2, the sender's incentive constraint at the information acquisition stage is sufficient for an equilibrium information structure.

Example 2: Let $a \in \mathbb{R}^k$, $\theta \in \Theta \subset \mathbb{R}^k$. The sender and the receiver's utility functions are given by:

$$u^s = -\sum_{i=1}^{k}(a_{(i)} - \theta_{(i)} - s_{(i)})^2 \tag{2.6a}$$

$$u^r = -\sum_{i=1}^{k}(a_{(i)} - \theta_{(i)})^2 \tag{2.6b}$$

where $a_{(i)}$ denotes the ith coordinate of a and $s = (s_{(1)}, ..., s_{(k)})$ be the sender's bias vector. The same reasoning applies to other notations.

Corollary 2.2. *u^r and u^s satisfy the 'No Cycle Condition' under any information structure with finite number of types.*

The proof is in Appendix I. This example shows how the 'No Cycle Condition' can be applied to a multi-dimensional setting. The main problem is, $A^*(\psi)$ is not necessarily finite in a multi-dimensional setting, especially under the quadratic preferences shown above (Battaglini [2002], Ambrus and Takahashi [2008]). To ensure the validity of Proposition 2, one way is to adopt the formulation of Levy and Razin (2007), where the number of equilibrium actions must be necessarily finite even in a multi-dimensional setting.

[14] The detailed proof is in Lemma 1 of Crawford and Sobel (1982), which mainly uses the compactness of Θ.

Example 3: Let $\Theta = \{1, -1\}$, $A = [-1, 1]$, $u^r = -(a - \theta)^2$, $u^s = (a - \theta)^2$. The prior is $\theta = 1$ occurs with probability $\frac{1}{2}$. The sender can choose to learn θ at a cost $\varepsilon > 0$ or learn nothing at cost 0. If the equilibrium action set is $A^* = \{-1, 1\}$ and the message set is $M = \{-1, 1\}$ while $m = 1$ induces $a = 1$ and vice versa, the sender's incentive constraint at the information acquisition stage is satisfied when ε is small enough, but she has an incentive to misreport his information at the communication stage.

Example 4: θ is uniformly distributed on $\Theta = [0, 1]$ and $\theta_i = 2^{-i}$ $(i = 0, 1, ...)$. ψ is an interval partition information structure defined by the partition points $(\theta_0, \theta_1, ..., \theta_n, ..., 0)$, i.e. after acquiring information, the sender knows which interval $[\theta_i, \theta_{i-1}]$ is the true θ actually in. Let Ψ be the set of information structures coarser than ψ. The cost is given by the following cubic form:

$$c[1 - \sum_{i=1}^{\infty} P(\Theta_i)^3] \tag{2.7}$$

where P denotes the probability measure of a set. The sender and receiver's utility functions are given by:

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

If the sender is free to induce actions in $A(\psi)$, type $[2^{-n}, 2^{-n+1}]$ sender strictly prefers type $[2^{-n+1}, 2^{-n+2}]$'s equilibrium action than its own. So there is a profitable deviation at the communication stage once the sender acquires ψ. But there is no profitable deviation at the information acquisition stage when c is sufficiently small ($c < \frac{1}{2}$).

2.7 Discussions and Extensions

Proposition 1 and 2 contribute to the cheap talk communication literature by providing clear benchmarks in answering the following questions:

1. Is there any information loss in the strategic communication process?
2. Is the sender's incentive constraint at the communication stage binding?

I show that under the standard setup of cheap talk game, the answer to both questions are *no* unless one of the following conditions is violated:[15]

[15] Besides the origin of information, the higher order beliefs of the players also play a big role in determining whether the sender's information can be fully transmitted or not. As shown in Pei (2012), when the sender faces higher order

1. The sender has no superior information ex ante;

2. The sender has sufficient flexibility in acquiring information (captured by the richness of Ψ) and there is no free information.

Proposition 1 remains true when we allow the receiver to consult multiple senders, or to acquire some information himself. Suppose there are $I \in \mathbb{N}$ senders and let $\psi^i \in \Psi^i$ be the information structure the ith sender acquires, and ψ^r be the information structure the receiver acquires. Let $C^i : \Psi^i \to \mathbb{R}^+$ be sender i's information acquisition cost function. An equilibrium achieves *a full communication* if and only if each sender tells the receiver everything she knows.[16]

Corollary 2.3. *When Ψ^i is rich and C^i satisfies monotonicity for all $1 \leq i \leq I$, every pure strategy equilibrium is a full communication equilibrium.*

The proof is in Appendix I. The idea is that given the other senders' and the receiver's equilibrium behavior strategies, each message m sent by sender i induces a distribution over the receiver's posterior beliefs. This further leads to a joint distribution of (θ, a). At each posterior belief, sender i chooses her favorite joint distribution among the set of equilibrium joint distributions. Again, the intuition in Proposition 1 applies.

Remark: This result is general in the sense that it allows for any form of correlation between the senders' information structures. For example, several financial analysts can have access to the same source of public information, making the 'noise' of their individual signals correlated. But their individual interpretations can differ, which leads to idiosyncratic noise in their reports (Myatt and Wallace [2012]). Even so, every analyst will still fully communicate her information, as long as she uses a pure strategy to acquire information.

3 COMMUNICATION INFORMATIVENESS

In this section, I restrict the state space to be one-dimensional, and compare the informativeness of communication both within an equilibrium and between different equilibria. These are the questions I explore:

uncertainty, she still cannot fully transmit her information in any equilibrium, even if there is no conflict of interest between her and the receiver, and there exist no incongruent sender who sends the same message.

[16] For example, when every available information structure is a partition of Θ: $\psi^j \equiv \{\Theta^j_1, ..., \Theta^j_{n_j}\}$ where $j \in J \equiv \{1, 2, ..., I, r\}$. Then the receiver's posterior belief is characterized by the algebra induced by $\{\Theta^p_1, ..., \Theta^p_{n_p}\}$, where $\Theta^p_k = \bigcap_{j \in J} \Theta(j)$ and $\Theta(j) \in \psi^j$.

1. Within an equilibrium, in which part of the state space is communication more informative?

2. How informative are the equilibria in my model comparing with those in CS?

3. How communication informativeness changes with information acquisition cost?

All discussions in this section are based on the '*uniform-quadratic-cubic*' setup, which is introduced in subsection 3.1. More general settings as well as the proofs are shown in Appendix II.

3.1 The Uniform-Quadratic-Cubic Setting

Let $\Theta = [0,1]$ and the prior of θ is the uniform distribution on $[0,1]$. Let the players' utility functions be $u^s = -(a-\theta-b)^2$ and $u^r = -(a-\theta)^2$ respectively, where $0 < b < \frac{1}{4}$.

Let $m(\cdot)$ be the Lebesgue Measure of a set. A partition of Θ is denoted by:

$$\Theta = \cup_{i=1}^{n}\Theta_i, \quad n \in \mathbb{N} \cup \{\infty\}$$

such that $m(\Theta_i) > 0$ and $\Theta_i \cap \Theta_j = \emptyset$ if $i \neq j$. A *partition information structure* is characterized by a partition: after acquiring this information structure, the sender knows which Θ_i contains the true state. Let Ψ be the set of partition information structures. The cost of acquiring a partition information structure $\{\Theta_1, ..., \Theta_n\}$ is:

$$C(\Theta_1, ..., \Theta_n) = c\Big[1 - \sum_{i=1}^{n} m(\Theta_i)^3\Big] \tag{3.1}$$

in which $c \in \mathbb{R}^+$ is a cost parameter.

Lemma 3.1. *Ψ is rich, C satisfies monotonicity. (IC-AC) implies (IC-CO).*

Proof: Ψ is rich because merging several subsets in a partition together forms a new partition. The monotonicity of C is implied by the convexity of the cubic function. (IC-AC) implies (IC-CO) has been shown in section 2 (Example 1). □

An '*interval partition information structure*' is characterized by a partition where each Θ_i $(1 \leq i \leq n)$ is an interval. Let Ψ_0 $(\subset \Psi)$ be the set of interval partition information structures, we have:

Lemma 3.2. *Given any finite equilibrium action set, the sender will choose $\psi \in \Psi_0$.*

This is because u^s and u^r are both super-modular in (a, θ). So given the information acquisition cost function, it is not profitable to have a disjoint set in the partition. This Lemma also implies that all equilibria are *monotone*, in which the receiver's action is weakly increasing with θ. An equilibrium is an '*interval partition equilibrium*' if the equilibrium information structure belongs to Ψ_0.

Interval partition equilibria bring us convenience when comparing the informativeness of communication between different ranges of the state space: when the lengths of the intervals are increasing, communication is less informative when θ is large, and vice versa.

Remark: θ can also be interpreted as a signal to the payoff-relevant state: $\tilde{\theta}$. Each θ leads to a posterior belief over $\tilde{\theta}$. The sender knows a 'rough' range for the realization of θ, but is unsure about its exact value. For example, after reading the Wall Street Journal in the morning, an investor remembers that today's NASDAQ Composite Index is between 3130 and 3140 when being asked this question during lunch.

3.2 Comparison within an Equilibrium: Two Types of Equilibria

An interval partition is characterized by a set of partition points: $(\theta_0, \theta_1, ..., \theta_n)$, where $0 = \theta_0 < \theta_1 < ... < \theta_n = 1$. Fixing θ_{i-1} and θ_{i+1}, when moving θ_i closer to $\frac{1}{2}(\theta_{i-1}+\theta_{i+1})$, the sender acquires more information. The equilibrium partition points must satisfy the local incentive constraints at the information acquisition stage: given θ_{i-1} and θ_{i+1}, the sender has no incentive to change θ_i when choosing her information structure. As a result, θ_i must equalize the marginal benefit of acquiring more information with the marginal cost.

- When $\theta_i < \frac{1}{2}(\theta_{i-1}+\theta_{i+1})$, moving θ_i closer to $\frac{1}{2}(\theta_{i-1}+\theta_{i+1})$ makes the sender inducing action $a^*_{i-1} \equiv \frac{\theta_{i-1}+\theta_i}{2}$ more often when θ is around θ_i. The marginal benefit is:

$$MB(\theta_i) = \Big(\frac{\theta_i + \theta_{i+1}}{2} - \theta_i - b\Big)^2 - \Big(\frac{\theta_i + \theta_{i-1}}{2} - \theta_i - b\Big)^2 = \frac{1}{4}(\theta_{i+1} - \theta_{i-1})(\theta_{i-1} - 2\theta_i + \theta_{i+1} - 4b)$$

The marginal cost is:

$$MC(\theta_i) = 3c(\theta_{i+1} - \theta_{i-1})(\theta_{i+1} - 2\theta_i + \theta_{i-1})$$

- When $\theta_i > \frac{1}{2}(\theta_{i-1}+\theta_{i+1})$, moving θ_i closer to $\frac{1}{2}(\theta_{i-1}+\theta_{i+1})$ makes the sender inducing action

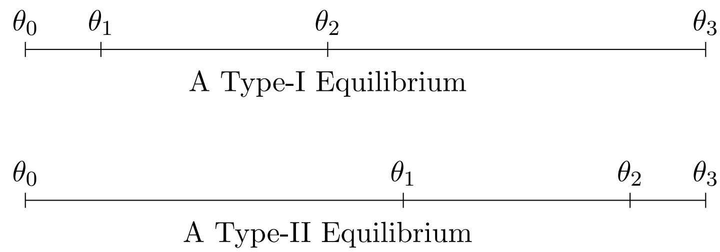

Figure II Two Types of Equilibria

$a_i^* \equiv \frac{\theta_{i+1}+\theta_i}{2}$ more often when θ is around θ_i. The marginal benefit is:

$$MB(\theta_i) = \Big(\frac{\theta_i + \theta_{i-1}}{2} - \theta_i - b\Big)^2 - \Big(\frac{\theta_i + \theta_{i+1}}{2} - \theta_i - b\Big)^2 = \frac{1}{4}(\theta_{i+1} - \theta_{i-1})(-\theta_{i-1} + 2\theta_i - \theta_{i+1} + 4b)$$

The marginal cost is:

$$MC(\theta_i) = 3c(\theta_{i+1} - \theta_{i-1})(-\theta_{i+1} + 2\theta_i - \theta_{i-1})$$

$|\theta_{i-1} - 2\theta_i + \theta_{i+1}|$ measures the difference in length between two adjacent intervals.

- When $\theta_i < \frac{1}{2}(\theta_{i-1} + \theta_{i+1})$, $\theta_{i-1} - 2\theta_i + \theta_{i+1}$ is positive and the lengths of the intervals are increasing. The term $\theta_{i-1} - 2\theta_i + \theta_{i+1} - 4b$ in the expression for marginal benefit implies that the benefit of information acquisition is *diminished* by the bias b.

- When $\theta_i > \frac{1}{2}(\theta_{i-1} + \theta_{i+1})$, $\theta_{i-1} - 2\theta_i + \theta_{i+1}$ is negative and the lengths of the intervals are decreasing. The term $-\theta_{i-1} + 2\theta_i - \theta_{i+1} + 4b$ in the expression for marginal benefit implies that the benefit of information acquisition is *amplified* by the bias b.

This implies that the sender's incentive to acquire information is larger when θ_i is above $\frac{\theta_{i-1}+\theta_{i+1}}{2}$. The driving force is just the direction of her bias ($b > 0$). The Lemma below is obtained by equalizing marginal benefit with marginal cost:

Lemma 3.3. *For any equilibrium information structure, the partition points must satisfy:*

$$\theta_{i-1} - 2\theta_i + \theta_{i+1} = \frac{4b}{1 - 12c} \tag{3.2}$$

for any $1 \leq i \leq n$.

The Lemma implies that the difference in length between any two adjacent intervals must be equal in any equilibrium. When $c < \frac{1}{12}$, $\frac{4b}{1-12c} > 0$, the lengths of the intervals in an equilibrium is strictly increasing. When $c > \frac{1}{12}$, $\frac{4b}{1-12c} < 0$, the lengths of the intervals in an equilibrium is strictly decreasing. I call the former '*Type-I Equilibrium*' and the latter '*Type-II Equilibrium*'. The following Proposition characterizes the ranges of parameters for the two types of informative equilibria to exist:

Proposition 3. *Type-I Equilibrium exists if and only if:* $c \in (0, \frac{1}{12}(1-4b))$*; Type-II Equilibrium exists if and only if:* $c \in (\frac{1}{12}(1+4b), \frac{1}{6}]$.

So when information acquisition cost is low, the equilibria are similar to those in CS, i.e. communication is more informative in smaller state values. But when information acquisition cost is high, the equilibria are different, i.e. communication is more informative in larger state values.[17]

To understand this result intuitively, I examine the two forces which affect the shape of equilibria: the *adverse selection effect in reporting information*, and the *moral hazard effect in acquiring information.*

- When information is cheap, the foremost concern is the *adverse selection problem.* Since the sender's favorite action is always larger than the receiver's, her recommendation becomes less credible when she claims that θ is large. This '*credibility effect*' makes her to acquire finer information in smaller state values and coarser information in larger ones. As a result, communication becomes more noisy when θ is large.

- When information is expensive, the foremost concern is the *moral hazard problem.* As shown before, the sender's marginal benefit to acquire information is larger when θ_i is above $\frac{\theta_{i-1}+\theta_{i+1}}{2}$. So, when she claims that θ is large, it signals that she has more detailed knowledge about θ. This is why communication can be more informative in larger state values, and an equilibrium is characterized by a partition of intervals with decreasing lengths.

The bottomline is that when the sender's information is costly acquired, the qualitative features of the equilibria not only depend on the players' preferences, but also on the cost of information acquisition.

[17] When allowing for more general utility functions, prior distributions as well as information acquisition cost functions, not every equilibrium belongs to those two types. There can exist equilibria where interval lengths are sometimes decreasing and sometimes increasing. I can only show that if there exists two adjacent intervals with increasing lengths in the equilibrium partition, then the cost must be below an upper bound; if there exists two adjacent intervals with decreasing lengths in the equilibrium partition, then the cost must be above a lower bound. Under the uniform-quadratic-cubic assumption, the 'lower bound' is greater than the 'upper bound', which rules out other equilibria beyond Type-I and Type-II.

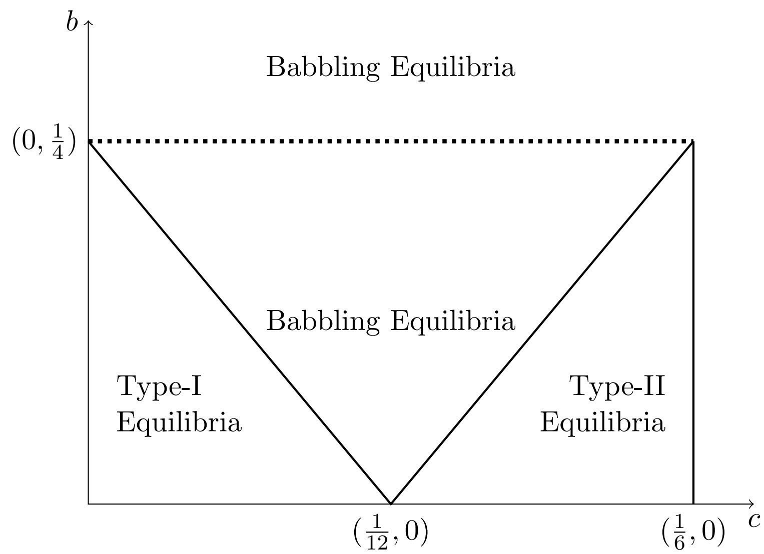

Figure III Existence of Informative Equilibria

3.3 Comparison with Crawford and Sobel (1982)

In this subsection, I compare the informativeness of our equilibria with those in CS. The comparison of informativeness between two equilibria is equivalent to compare the informativeness of their equilibrium information structures. I measure the 'informativeness' of ψ, $I(\psi)$, by the receiver's ex ante expected loss under that information structure. Formally speaking, when ψ is characterized by partition points $(\theta_0, \theta_1, ..., \theta_n)$, then $I(\psi)$ is a strictly increasing function of:

$$-\sum_{i=1}^{n}\int_{\theta_{i-1}}^{\theta_i} u^r(\overline{a}[\theta_{i-1},\theta_i],\theta)d\mu_0(\theta)$$

where μ_0 is the prior on θ and

$$\overline{a}[\theta_{i-1},\theta_i] \equiv \arg\max_a \int_{\theta_{i-1}}^{\theta_i} u^r(a,\theta)d\mu_0(\theta) \tag{3.3}$$

When the receiver's preference is quadratic and the prior of θ is uniform, it is without loss of generality to express informativeness as:[18]

$$I(\psi) \equiv 1 - 12\sum_{i=1}^{n}\int_{\theta_{i-1}}^{\theta_i}(\frac{\theta_{i-1}+\theta_i}{2}-\theta)^2 d\theta = 1-\sum_{i=1}^{n}(\theta_i-\theta_{i-1})^3 \tag{3.4}$$

[18]Note that when the sender's preference is also quadratic, her ex ante expected loss is just the receiver's plus b^2. So it does not matter whether to measure informativeness from the receiver's side or the sender's side

My first conclusion is directly implied by the first order conditions in information acquisition, which claims that communication informativeness cannot be improved upon CS.

Corollary 3.1. *Under any parameter values (b, c), every equilibrium is weakly less informative than the most informative CS equilibrium.*

This is because $|\theta_{i-1} - 2\theta_i + \theta_{i+1}| \geq 4b$ necessarily holds in any equilibrium. This implies that when the sender is ignorant ex ante, but is granted the discretion to acquire information, communication informativeness cannot be improved comparing with the case where the sender has perfect knowledge. This is because the conflict of interest discourages her from acquiring finer information.

Next, I show a convergence of our equilibria to a CS equilibria when $c \to 0$, which further implies a convergence of the informativeness level.

Proposition 4. *For any CS Equilibrium $\big(\theta_0(0), \theta_1(0), ..., \theta_n(0)\big)$, there exists $\varepsilon_0 > 0$, such that: for any sequence $\{c_i\}_{i=1}^{\infty}$ satisfying $\lim_{i\to\infty} c_i = 0$ and $c_i \in (0, \varepsilon_0)$, there exists*

$$\Big\{\big(\theta_0(c_i), \theta_1(c_i), ..., \theta_n(c_i)\big)\Big\}_{i=1}^{\infty}$$

such that $\big(\theta_0(c_i), \theta_1(c_i), ..., \theta_n(c_i)\big)$ is an equilibrium when $c = c_i$ and

$$\lim_{i\to\infty} \theta_j(c_i) = \theta_j(0)$$

for any $0 \leq j \leq n$.

The key step of the proof is to show that the local incentive constraints:

$$\theta_{i+1} - 2\theta_i + \theta_{i-1} = \frac{4b}{1 - 12c}$$

are sufficient when c is close to 0. An interval partition information structure is part of an equilibrium when the cost parameter is c if and only if the difference in length between two adjacent intervals is $\frac{4b}{1-12c}$. If this is the case, for any CS equilibrium, we can always find a sequence of equilibria under endogenous information acquisition which converge to that when $c \to 0$.

3.4 Comparative Statics

In this subsection, I fix the sender's bias and examine how communication informativeness changes when c increases. The Corollary below is directly implied by Proposition 3:

Corollary 3.2. *Communication informativeness is not monotone with respect to c.*

This result is obvious because when $c \in [\frac{1-4b}{12}, \frac{1+4b}{12}]$, no informative equilibrium exist, while there exists at least an informative equilibrium when $c \in (\frac{1+4b}{12}, \frac{1}{6}]$.

Next, I study the local comparative statics problem. I fix the number of partition points, and examine how communication informativeness changes when we increase c marginally.[19]

Corollary 3.3. *Communication informativeness decreases with c in a Type-I equilibrium, and increases with c in a Type-II equilibrium.*

The intuition is that an increasing c has two roles in the sender's incentive constraint at the information acquisition stage. First, it makes deviating to a less informative information structure more appealing, which we call the '*shirking effect*'. Secondly, it makes deviating to a more informative information structure less profitable, which we call the '*commitment effect*'. Whether a marginal increase in information acquisition cost increases or decreases communication informativeness depends on what is preventing communication from becoming more informative: Is it the marginal benefit is too large that the sender cannot commit herself not to acquire additional information; or is it the marginal cost is too large that the sender has an incentive to shirk.

- Under the 'Type-I Equilibrium', marginal cost is larger than marginal benefit when we move to a locally more informative information structure (moving θ_i rightwards). So an increase in marginal cost decreases informativeness.

- Under the 'Type-II Equilibrium', marginal benefit is larger than marginal cost when we move to a locally more informative information structure (moving θ_i leftwards). So an increase in marginal cost increases informativeness.

The marginal cost (MC, in blue) and marginal benefit (MB, in red) of information acquisition are depicted in Figure 4.

[19] Note that these are just local comparative statics, which may not hold globally.

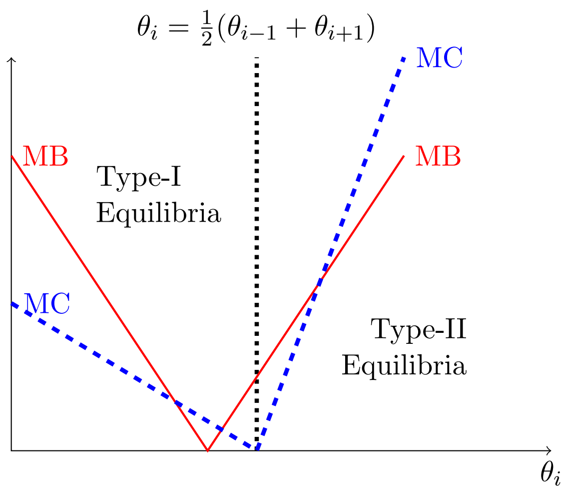

Figure IV Comparative Statics with respect to c

4 ORGANIZATIONAL DESIGN

In this section, I discuss the receiver's other alternatives when he lacks the knowledge for decision making. Specifically, I allow the him to delegate decision rights to the sender, or to monitor her information acquisition process. As in the previous literature, I replace the sender by the 'agent', and the receiver by the 'principal'. I study the effects of both institutional arrangements on the agent's incentive to acquire information as well as their welfare implications.

Throughout this section, I adopt the uniform-quadratic assumption and use:

$$I(\psi) \equiv 1 - \sum_{i=1}^{n} \left(\theta_i - \theta_{i-1}\right)^3$$

to measure the informativeness of an information structure. Let the cost of information structure be $C(I(\psi), c)$, while $C_1 > 0$, $C_2 > 0$, $C_{12} > 0$, $C(0, c) = 0$. The detailed specifications of C are different between subsections.

4.1 Delegation

In this subsection, I assume

$$C(I(\psi), c) = cI(\psi) \tag{4.1}$$

Also, contracts are incomplete and the principal has only two options ex ante:

1. Delegating decision right to the agent by letting her choose a after she acquires information;

2. Centralizing decision making and communicating with the agent through cheap talk.

I focus on the most informative equilibrium when multiple equilibria exist. The following Proposition examines how delegation affects the agent's incentives to acquire information.

Proposition 5. *The agent acquires strictly more information under delegation if and only if $c \leq \frac{1}{12}$. The agent acquires strictly less information under delegation if and only if a 'Type-II Informative Equilibrium' exists.*

Proof. Under delegation, the agent chooses:

$$a = b + \int \theta d\mu(\theta) \tag{4.2}$$

if her posterior belief is μ. Her optimization problem at the information acquisition stage is:

$$\min_{\psi \in \Psi_0} \left\{ \frac{1}{12} \sum_{i=1}^{n} (\theta_i - \theta_{i-1})^3 + c(1 - \sum_{i=1}^{n} (\theta_i - \theta_{i-1})^3) \right\} \tag{4.3}$$

So she acquires full information if $c \leq \frac{1}{12}$, and acquires no information otherwise. □

By comparing the ex ante expected welfare of the principal in the above two cases, I show that:

Corollary 4.1. *Delegation is optimal if and only if:*

$$c \leq \frac{1}{12} \quad \textit{and} \quad b \leq \sqrt{\frac{1}{12}}$$

Now I explain the intuition behind these results and compare them with the seminal contribution of Aghion and Tirole (1997). First, I discuss the incentives to acquire information. In Aghion and Tirole (1997), the agent is free to choose any action when she is being delegated the decision right; or she is the only player who has the information. Delegation does not increase the her flexibility in choosing actions. As a result, it only increases her payoff when acquiring finer information without simultaneously increases her payoff when acquiring coarser information. Therefore, delegation unambiguously increases her incentives to acquire information.

In my framework, however, the agent can only choose from a finite set of equilibrium actions under centralization, while she is free to choose any action under delegation. Under centralization, the principal's equilibrium action set will not change even if the agent deviates at the information acquisition stage. This is because her deviation is non-observable. I call this the '*rigid code effect*'.

The comparison between delegation with centralization is determined by how this effect affects the incentives to acquire information.

1. *When information acquisition cost is low*, the agent has the potential to acquire more information, but the 'rigid code effect' reduces her incentives to do that. This is because even if she acquires finer information, she cannot enrich the principal's equilibrium action set. So, her incentives to acquire information is higher under delegation.

2. *When information acquisition cost is high*, the agent has more incentives to deviate to a coarser information structure. The 'rigid code effect' makes her favorite action under this deviation not available. Hence her deviation becomes less profitable under centralization. Therefore, delegation decreases her incentives to acquire information.

Now I move on to discuss when is delegation optimal for the principal. If information acquisition cost is low, the agent acquires full information under delegation, but coarser information under centralization. If the bias is not too large ($b \leq \sqrt{\frac{1}{12}}$), the principal's welfare reduction due to '*loss of control*' is less than that caused by coarser information (Dessein [2002]). If information acquisition cost is high, delegation weakly decreases the agent's incentives to acquire information while also generating a '*loss of control*' effect. Obviously, centralization is the principal's optimal choice.

4.2 Monitoring Information Acquisition

In this subsection, I allow the principal to monitor the agent's information acquisition process. I obtain similar insights as Austen-Smith (1994), which shows that monitoring can backfire.

To facilitate the following discussions, I assume that:

Assumption 2. *$C(I(\psi), c)$ is strictly convex in $I(\psi)$.*

We compare the informativeness of equilibria under three monitoring structures:

1. *No Control:* The principal observes nothing about the agent's informational choice;

2. *Cost Control:* The principal can only observe the agent's information acquisition cost, but cannot directly observe the information structure;

3. *Direct Control:* The principal can observe the agent's information structure.

Notice that in all cases, the principal cannot observe the agent's signal realization, so the communication problem remains. To highlight the negative impact of monitoring, we assume that monitoring has no direct cost. The timeline of the game is as follows:

1. The principal publicly chooses the monitoring structure;
2. The agent observes the principal's choice and acquires information;
3. The principal observes a signal about the agent's informational choice according to the monitoring structure;
4. The agent sends a cheap talk message to the principal;
5. The principal updates his belief and takes an action.

Let ψ^n be the most informative equilibrium information structure under no control. The comparison between the informativeness of equilibria are summarized below:

Proposition 6. *Direct control weakly dominates cost control. No control strictly dominates direct control if and only if* $C(I(\psi^n), c) > \frac{1}{12} I(\psi^n)$.

Proof: I only prove the second part of the proposition and leave the first part to Appendix I.

Under cost control and direct control, the agent always has the following outside option: She acquires no information, and after the principal observes this, he always chooses $a = \frac{1}{2}$. The agent's expected loss is $\frac{1}{12} + b^2$. When the agent acquires information structure ψ, which is characterized by partition points $(\theta_0, ..., \theta_n)$, her expected loss is no less than:

$$C(I(\psi), c) + \frac{1}{12} \sum_{i=1}^{n} (\theta_i - \theta_{i-1})^3 + b^2$$

ψ can be an equilibrium information structure under cost control or direct control only if

$$C(I(\psi), c) + \frac{1}{12} \sum_{i=1}^{n} (\theta_i - \theta_{i-1})^3 + b^2 \leq \frac{1}{12} + b^2 \tag{4.4}$$

When $C(I(\psi^n), c) > \frac{1}{12} I(\psi^n)$, an information structure which is weakly more informative than ψ^n cannot be acquired when the principal controls. Hence, the no control outcome, ψ^n, is optimal.

When $C(I(\psi^n), c) < \frac{1}{12} I(\psi^n)$, the principal can threat the agent that if he does not acquire information structure ψ^* $(I(\psi^*) > I(\psi^n))$, then they switch to the babbling equilibrium. Then the agent has to choose between acquiring information structure ψ^* and acquiring no information. Obviously, ψ^* gives her a higher expected payoff when c is small enough. Hence, the most informative outcome under direct control dominates that under no control when c is small, and the reverse is true when c is large. □

The main idea is that when information acquisition is non-observable, even if the agent acquires no information, the equilibrium action set remains unchanged; when information acquisition is observable, however, the principal cannot commit not to change the equilibrium action set after observing a deviation at the information acquisition stage. If the equilibrium action set does not contain any action within $[\frac{1}{2}, \frac{1}{2} + 2b]$ (this is indeed the case in any Type-II Equilibrium), then the agent's payoff when acquiring no information is higher when the principal monitors. Monitoring backfires when the task is challenging due to this *commitment effect* (Crémer [1995]). But the comparison between direct control and cost control coincides with the classical result in Grossman and Hart (1983), where having more precise information about the agent provides better incentives. So the optimal monitoring structure is either monitoring extensively, or no monitoring at all.

Next, I discuss when will the most informative equilibria in Ivanov (2010) or Crawford and Sobel (1982) can be chosen in an endogenous information acquisition setting. The most informative *Ivanov Information Structure* is defined as:

Definition 3. *The most informative Ivanov Information Structure* $(\theta_0, ..., \theta_n)$ *is the solution to the following optimization problem:*

$$\min_{n,\theta_0,\ldots,\theta_n} \sum_{i=1}^{n} (\theta_i - \theta_{i-1})^3$$

subject to

$$\theta_{i+1} - \theta_{i-1} \geq 4b, \quad 0 = \theta_0 < ... < \theta_n = 1, \quad n \in \mathbb{N}$$

for any $1 \leq i \leq n-1$.[20]

Similarly, the *most informative CS Information Structure* is obtained by minimizing the same objective subject to $\theta_{i+1} - 2\theta_i + \theta_{i-1} = 4b$ for any $1 \leq i \leq n-1$.

Corollary 4.2. *If c is low enough: the most informative Ivanov Information Structure is chosen by the agent under direct control; the most informative CS Information Structure is chosen by the agent under cost control.*

[20] The solution to this problem is characterized in Lemma 3 of Ivanov (2010).

This establishes a foundation for the most informative Ivanov Information Structure as well as most informative CS Information Structure, which says that the most informative equilibria in these papers can be acquired in an endogenous information acquisition setting when a certain monitoring structure is chosen, and when the cost of information acquisition is not too large.

5 ALLEYS FOR FUTURE RESEARCH

Endogenous and Exogenous Information: Comparing with the other polar case in the literature, where the sender is endowed with perfect knowledge, I assume that the sender has no superior information ex ante, i.e. all her information is costly to acquire. Most applications involve both features. Incorporating both superior knowledge and the possibility of acquiring new information will significantly enhance the realism of the model, although at the cost of sacrificing tractability.

Communicating with Multiple Receivers: When the sender gives advice to multiple receivers with heterogenous preferences (Goltsman and Pavlov 2011), she may not be able communication all her information to everyone of them. Whether communicating with a second receiver increases or decreases communication informativeness with the first one depends on whether adverse selection effect or moral hazard effect dominates. When information acquisition cost is low and adverse selection effect dominates, then the information acquired '*specific*' for the second receiver makes the sender '*know too much*' — this hinders her communication with the first one; but when information acquisition cost is high and moral hazard effect dominates, then the '*scale effect*' of multiple-receivers increases the sender's benefit from acquiring information — this alleviates the moral hazard problem and increases communication informativeness.

TOULOUSE SCHOOL OF ECONOMICS

A Appendix I: Remaining Proofs

Proof of Proposition 1. I finish the proof of Proposition 1 by discussing types occurring with 0 probability.

For a given ψ, the sender's type is entirely determined by ω. To reduce notations, use ω as a representation of her type. Let Ω_1 be the set of types which occurs with positive probability. Let $\Lambda_1 \equiv \{\alpha_i\}_{i=1}^{\infty}$ be the set of equilibrium mixed actions induced with positive probability. Obviously, Ω_1 and Λ_1 are at most countable.

Let Ω_0 be the set of types occurring with 0 probability. Let $\Lambda_0 \equiv \{\alpha_\lambda\}_\lambda$ be the set of equilibrium mixed actions induced with 0 probability. Let

$$p_0 \equiv \Pr(\omega \in \Omega_0)$$

When $p_0 = 0$, the conclusion in Proposition 1 holds trivially. So I focus on the case when $p_0 > 0$. Let $\Omega_{01} \subset \Omega_0$ be the set of types which induces a mixed action in Λ_1 in equilibrium. $\Omega_0(\alpha_\lambda) \subset \Omega_0$ be the set of types which induces α_λ.

The proof proceeds as the following:

1. The measure of Ω_{01} is 0.

   If not, there exists $i \in \mathbb{N}$ such that there is a positive measure of types in Ω_{01} which induces α_i. Then, there is a profitable deviation at the information acquisition stage by pooling these types together and induce α_i, which is a contradiction.

2. The measure of $\Omega_0(\alpha_\lambda)$ is 0 for every λ.

   If not, the sender can pool $\Omega_0(\alpha_\lambda)$ into one type at the information acquisition stage induce α_λ. This is a profitable deviation, which is a contradiction.

3. There exists a 0 measure set Ω_{00} ($\subset \Omega_0$) such that, in Ω_0/Ω_{00}, no two types induce the same action.

   If not, we can proceed as follows. For a given $\omega \in \Omega_0$, find out one type in Ω_0 which induces the same action as ω. If it exists, denote it by ω'. Remove ω and ω' from Ω_0 and do the same for another element, until such a process is not possible for any element in Ω_0. So, we will stop at a time when we remove a positive measure of elements. Pool ω with ω' together at the information acquisition stage, and induce their mutual action if she belongs to that new aggregate type. This is a strictly profitable deviation, which is a contradiction.

So in $\Omega - \Omega_{00}$, the receiver can fully deduce the sender's type. Since Ω_{00} is proved to be a 0 measure set, so every equilibrium is a full communication equilibrium. □

Proof of Lemma 2.1: I construct a profitable deviation of the sender if she uses a non-trivial mixed strategy to acquire information.

At the communication stage, she chooses $m \in M$ such that:

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

Let $\psi(\mu)$ be the probability of posterior belief μ to occur under information structure ψ. Let $m(\mu)$ be a message which satisfies the above equation.

Let $\mathcal{P}$ be a probability measure on Ψ, whose mass is not concentrated on any single element. When the sender acquires information structure $\psi' \equiv \int \psi \mathcal{P}(\psi)$ and sends message $m(\mu)$ when she has belief μ, the joint distribution of (θ, m), and hence (θ, a) remains unchanged.

From the convexity assumption on Ψ and C, we know that $\psi' \in \Psi$ and

$$C(\psi') < \int_\psi C(\psi) d\mathcal{P}(\psi) \tag{A.2}$$

which is just Jensen's Inequality. This means, the sender strictly saves her information acquisition cost, which is a profitable deviation given the receiver's equilibrium mixed action set. □

Proof of Corollary 2.2: Let $\{\mu_1, ..., \mu_n\}$ be the set of equilibrium posterior beliefs under information structure ψ. Let $\mathbb{E}^i \equiv \int_\Theta \theta d\mu_i(\theta)$. Without loss of generality, I assume $\mathbb{E}^i \neq \mathbb{E}^{i'}$ if $i \neq i'$.

The receiver takes action $\mathbb{E}^i$ when he has belief μ_i. Type i sender's favorite equilibrium action is the one closest to her ideal point: $s + \mathbb{E}^i$.

Now, I prove the corollary by contradiction.

Suppose type $\{1, 2, ..., j\}$ $(2 \leq j \leq k)$ forms a cycle: type i prefers the equilibrium action of type $i+1$ for $i = 1, 2, ..., j-1$; type j prefers the equilibrium action of type 1, then $\mathbb{E}^{i+1}$ lies in the k-dimensional sphere with center $s + \mathbb{E}^i$ and radius s.

Reset and re-scale the coordination system, such that $\mathbb{E}^1 = 0$, $s = (1, 0, ..., 0)$, we have:

1. $s + \mathbb{E}^1 = (1, 0, ..., 0)$;

2. If $[s + \mathbb{E}^i]_{(1)} > 1$, then $[\mathbb{E}^{i+1}]_{(1)} > 0$. So $[s + \mathbb{E}^{i+1}]_{(1)} > 1$

By induction, we know that $\mathbb{E}^1$ cannot be in the sphere with radius 1 and center $s + \mathbb{E}^j$, which is a contradiction. $\square$

Proof of Corollary 2.3: Given the other senders' as well as the receiver's behavior strategies, and $\psi^i \in \Psi_i$. For each realization $\omega^i \in \Omega^i$, let $\psi^i(\theta|\omega^i)$ be her belief on the on the true state conditional on her own signal realization, and $\psi^i(\omega^{-i}|\omega^i, \theta)$ be her belief on the other players' signal realizations conditional on her own as well as the true state of the world. Since for a given message m as well as the other players' behavior strategies, the distribution of the receiver's action depends only on the state θ as well as the other players' signal realizations, which we denote by $\mu_m(a|\omega^{-i}, \theta)$ Let $p(\omega^i)$ be the probability of ω^i to occur, and $p(\omega^i|\theta)$ be that probability conditional on θ.

If sender i's optimal message under ω_1^i and ω_2^i are both m, then her expected payoff from the project by separating ω_1^i and ω at the information acquisition stage is:

$$V_1 \equiv \sum_{k=1}^{2} p(\omega_k^i) \int_\theta \int_{\omega^{-i}} \Big\{ \int_a u^{s,i}(a, \theta) d\mu_m(a|\omega^{-i}, \theta) \Big\} d\psi^i(\omega^{-i}|\omega_k^i, \theta) d\psi^i(\theta|\omega_k^i) \tag{A.3}$$

Her expected payoff by merging them together, forming new type ω^i, and sending m under ω^i is:

$$V_2 \equiv p(\omega^i) \int_\theta \int_{\omega^{-i}} \Big\{ \int_a u^{s,i}(a, \theta) d\mu_m(a|\omega^{-i}, \theta) \Big\} d\psi^i(\omega^{-i}|\omega^i, \theta) d\psi^i(\theta|\omega^i)$$

where $p(\omega^i) = p(\omega_1^i) + p(\omega_2^i)$,

$$\psi^i(\theta|\omega^i) = \frac{\psi^i(\theta|\omega_1^i)p(\omega_1^i) + \psi^i(\theta|\omega_2^i)p(\omega_2^i)}{p(\omega_1^i) + p(\omega_2^i)}$$

and

$$\psi^i(\omega^{-i}|\omega^i, \theta) = \frac{p(\omega_1^i|\theta)\psi^i(\omega^{-i}|\omega_1^i, \theta) + p(\omega_2^i|\theta)\psi^i(\omega^{-i}|\omega_2^i, \theta)}{p(\omega_1^i|\theta) + p(\omega_2^i|\theta)}$$

Let $\Lambda(\omega^{-i},\theta) \equiv \int_A u^{s,i}(a,\theta)d\mu_m(a|\omega^{-i},\theta)$ (we will use Λ for short). We have:

$$\begin{aligned} V_2 - V_1 &= \sum_{k=1}^{2} p(\omega_k^i) \int_{\Theta\times\Omega^{-i}} \Lambda d\psi^i(\omega^{-i}|\omega^i,\theta)d\psi^i(\theta|\omega_k^i) - \sum_{k=1}^{2} p(\omega_k^i) \int_{\Theta\times\Omega^{-i}} \Lambda d\psi^i(\omega^{-i}|\omega_k^i,\theta)d\psi^i(\theta|\omega_k^i) \\ &= \int_{\Theta\times\Omega^{-i}} \frac{\Lambda p(\omega_1^i)p(\omega_2^i)}{\mu_0(\theta)\left[p(\omega_1^i|\theta)+p(\omega_2^i|\theta)\right]} \\ &\quad \Big\{p(\theta|\omega_2^i)d\psi^i(\omega^{-i}|\omega_2^i,\theta)d\psi^i(\theta|\omega_1^i) + p(\theta|\omega_1^i)d\psi^i(\omega^{-i}|\omega_1^i,\theta)d\psi^i(\theta|\omega_2^i) \\ &\quad - p(\theta|\omega_2^i)d\psi^i(\omega^{-i}|\omega_1^i,\theta)d\psi^i(\theta|\omega_1^i) - p(\theta|\omega_1^i)d\psi^i(\omega^{-i}|\omega_2^i,\theta)d\psi^i(\theta|\omega_2^i)\Big\} \\ &= 0 \end{aligned} \tag{A.4}$$

The above equation makes use of Bayes Rule:

$$p(\omega_k^i|\theta) = \frac{p(\theta|\omega_k^i)p(\omega_k^i)}{\mu_0(\theta)}$$

and $d\psi^i(\theta|\omega_k^i) = p(\theta|\omega_k^i)$ for $k = 1, 2$. □

Proof of Proposition 3: Let $l_i \equiv \theta_i - \theta_{i-1}$ be the length of the ith interval in the partition.

First, a necessary condition for $\{l_1, l_2, ..., l_n\}$ to be an equilibrium is

$$l_i - l_{i-1} = \frac{4b}{1-12c} \tag{A.5}$$

for all i. So

$$\frac{4b}{1-12c} < 1$$

must hold as long as an informative equilibrium exists. This implies that $c < \frac{1-4b}{12}$ or $c > \frac{1+4b}{12}$.

Next, I construct 2-partition informative equilibria when $c \in (\frac{4b+1}{12}, \frac{1}{6}) \bigcup (0, \frac{1-4b}{12})$. Denote the equilibrium partition point to be θ_1^*, the sender's expected loss in equilibrium is

$$(\frac{1}{12} - c)[(1-\theta_1^*)^3 + \theta_1^{*3}] + c + b^2$$

Since I have already proved that this is a local maximum, the only possibility that it is not a global maximum is when $\theta_1 = 0$ or $\theta_1 = 1$ (the constraint $\theta_1 \geq \theta_0$ or $\theta_2 \geq \theta_1$ binds.). So $\theta_1 = \theta_1^*$ is a global maximum if and only if:

$$(\frac{1}{12} - c)[(1-\theta_1^*)^3 + \theta_1^{*3}] + c + b^2 \leq \frac{1}{3} + (b - a_1^*)^2 + (b - a_1^*) \tag{A.6a}$$

$$(\frac{1}{12}-c)[(1-\theta_1^*)^3+\theta_1^{*3}]+c+b^2 \leq \frac{1}{3}+(b-a_2^*)^2+(b-a_2^*) \tag{A.6b}$$

These inequalities reduce to:

$$c \leq \frac{1}{6} \tag{A.7}$$

So $l_i - l_{i-1} = \frac{4b}{1-12c}$ is both sufficient and necessary for a 2-partition equilibrium as long as $c \leq \frac{1}{6}$. Hence, we can construct a 2-partition equilibrium when $|\frac{4b}{1-12c}| < 1$ and $c < \frac{1}{6}$.

$c \leq \frac{1}{6}$ is also necessary for any informative equilibrium. This is because for an n-partition equilibrium, if $c > \frac{1}{6}$, then the sender can gain by pool two intervals: $[\theta_{i-1}^*, \theta_i^*]$ and $[\theta_i^*, \theta_{i+1}^*]$ together at the information acquisition stage and induces a_i^* on the new aggregate interval $[\theta_{i-1}^*, \theta_{i+1}^*]$, which is a profitable deviation according to the above inequality. □

Proof of Corollary 4.1: Obviously, when $c > \frac{1}{12}$, centralization is optimal.

When $c \leq \frac{1}{12}$ and $b > \frac{1}{4}$, only babbling equilibrium exists under centralization. The principal's expected loss is $\frac{1}{12}$ under centralization and b^2 under delegation. Delegation is optimal if and only if $b^2 < \frac{1}{12}$.

When $c \leq \frac{1}{12}$ and $b \leq \frac{1}{4}$, the expected loss of the receiver under delegation is:

$$L_d = b^2$$

The expected loss the the receiver under centralization satisfies the following inequality:

$$L_c \geq \frac{1}{12}\sum_{i=0}^{n-1}(l_1+i\delta)^3 \tag{A.8}$$

where l_1 is the length of the shortest interval and $\delta = |\frac{4b}{1-12c}| > 4b$. Also δ, l_1 and n satisfy:

$$\delta\frac{n(n-1)}{2}+nl_1 = 1$$

$$\delta\frac{n(n+1)}{2} \geq 1$$

So, when $c \leq \frac{1}{12}$, we have:

$$\frac{1}{12}\sum_{i=1}^{n}(l_1+i\delta)^3 \geq \frac{1}{12}\sum_{i=1}^{n}(i\delta)^3 \geq \frac{\delta^2}{16}\frac{2(n-1)^2 n}{3(n+1)} \tag{A.9}$$

The last term is greater than b^2 when $n \geq 3$. When $n = 1$, $L_c = \frac{1}{12} > b^2$. When $n = 2$, $L_c = \frac{1}{96}((1-\delta)^3 + (1+\delta)^3) > \frac{1}{48}(1 + 48b^2) > b^2$.

To summarize, delegation dominates centralization when $c \leq \frac{1}{12}$ and $b \leq \sqrt{\frac{1}{12}}$. $\square$

Proof of Proposition 6: The proof is done through the following steps. First, I estimate an upper bound of informativeness under cost control. Then, I estimate a lower bound of informativeness under direct control.

Cost Control: Since the agent can always increase information acquisition cost by acquiring wasteful information, so the principal can only impose a 'cap' on the agent's information acquisition cost. First, I show that it is without loss of generality to consider the following class of equilibria under cost control. The principal specifies an information structure ψ^c, which satisfies:

$$C(I(\psi^c), c) \leq \frac{1}{12} I(\psi^c)$$

He threats the agent that if her information acquisition cost does not equal to $C(\psi^c)$, he will ignore her message in the communication stage (babbling equilibrium). The agent acquires ψ^c and fully communicates her information to the principal. This is called a '*fully communicative equilibrium*' under cost control.

Obviously, for all other equilibria where the sender acquires ψ and uses a pure strategy in communication, there always exists $\psi' \in \Psi$ which is strictly coarser than ψ, such that the agent can fully communicates ψ' in that equilibrium. Then, it is equivalent to the fully communicative equilibrium with equilibrium information structure ψ'.

If the agent is using a mixed strategy in communication, rank the types of the agent by their posterior beliefs of θ. Let type i be the smallest type sender to use a mixed strategy, then she must be indifferent between a_i^* and a_{i+1}^*. Since there are no other types sending a_i^*, the principal's sequential rationality condition implies that $\mathbb{E}_i[\theta] = a_i^*$. Then from type i agent's indifference condition, $a_{i+1}^* = a_i^* + 2b$. There must exist another type higher than i who also induces a_{i+1}^* with strictly positive probability. Without loss of generality, let us call this type $i+1$. So a_{i+1}^* is both type i and type $i+1$'s favorite equilibrium action. If the agent pools type i and $i+1$ and induce action a_{i+1}^*, her expected payoff from the project remains unchanged. Now I construct a profitable deviation of the agent. She divides $[\theta_{i-1}, \theta_{i-1} + \varepsilon]$ from the rest of $\Theta_i \bigcup \Theta_{i+1}$ at the information acquisition stage and induce a_i^* when $\theta \in [\theta_{i-1}, \theta_{i-1} + \varepsilon]$, and a_{i+1}^* when $\theta \in \Theta_i \bigcup \Theta_{i+1}$

and $\theta > \theta_{i-1} + \varepsilon$. We can always make ε small enough to make the information acquisition cost lower than the original information structure. Under such a deviation, her expected payoff from the project strictly increases while making her information acquisition cost weakly below the '*controlling cap*'. This leads to a contradiction. So she never uses a mixed strategy at the communication stage under cost control.

Now I show that communication informativeness under cost control is bounded by two bounds: $C(I(\psi^c), c) \leq \frac{1}{12} I(\psi^c)$ and the most informative CS equilibrium. The first one is obvious. The second one is shown in the Lemma below:

Lemma A.1. *Let $(\theta_0^*, ..., \theta_n^*)$ be the partition points. For any ψ^c which is fully communicative, and any $i \neq j$:*

$$|\theta_{i+1}^* - 2\theta_i^* + \theta_{i-1}^*| = |\theta_{j+1}^* - 2\theta_j^* + \theta_{j-1}^*| \geq 4b \tag{A.10}$$

Proof of Lemma A.1: If ψ^c satisfies:

$$C(\psi^c, c) \leq \frac{1}{12}\big[1 - \sum_{i=1}^{n}(\theta_i^* - \theta_{i-1}^*)^3\big]$$

then the agent's maximization problem at the information acquisition stage is:

$$\max_{(\theta_0, \ldots, \theta_n)} -\sum_{i=1}^{n} \int_{\theta_{i-1}}^{\theta_i} (\frac{\theta_{i-1}^* + \theta_i^*}{2} - \theta - b)^2 d\theta \tag{A.11}$$

s.t.

$$\sum_{i=1}^{n}(\theta_i - \theta_{i-1})^3 \geq \sum_{i=1}^{n}(\theta_i^* - \theta_{i-1}^*)^3 \tag{A.12}$$

and

$$0 = \theta_0 \leq ... \leq \theta_n = 1 \tag{A.13}$$

Ignore the second constraint and solve the relaxed problem. Let λ be the Lagrange multiplier of the first constraint, the Lagrangian can be written as:

$$\mathcal{L} = -\sum_{i=1}^{n} \int_{\theta_{i-1}}^{\theta_i} (\frac{\theta_{i-1}^* + \theta_i^*}{2} - \theta - b)^2 d\theta + \lambda\big[\sum_{i=1}^{n}(\theta_i - \theta_{i-1})^3 - \sum_{i=1}^{n}(\theta_i^* - \theta_{i-1}^*)^3\big] \tag{A.14}$$

Take the FOC with respect to θ_i^*:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \theta_i}\Big|_{\theta_i=\theta_i^*} &= -\frac{\theta_{i+1}^* - \theta_{i-1}^*}{4}\left[2\theta_i^* - \theta_{i-1}^* - \theta_{i+1}^* + 4b - 12\lambda(2\theta_i^* - \theta_{i-1}^* - \theta_{i+1}^*)\right] \\ &= -\frac{\theta_{i+1}^* - \theta_{i-1}^*}{4}[4b - (1-12\lambda)(l_{i+1} - l_i)] \\ &= 0 \end{aligned} \tag{A.15}$$

Let $l_i \equiv \theta_i^* - \theta_{i-1}^*$, so:

$$l_{i+1} - l_i = \frac{4b}{1-12\lambda} \tag{A.16}$$

Since $\lambda \geq 0$, so when $\lambda \in [0, \frac{1}{6}]$, $|l_{i+1} - l_i| \geq 4b$.

I complete the proof by showing that $l_{i+1} - l_i \notin (-4b, 0)$. If $l_{i+1} - l_i \in (-4b, 0)$, then:

$$\frac{\theta_{i-1}^* + \theta_{i+1}^*}{2} < \theta_i^* < \frac{\theta_{i-1}^* + \theta_{i+1}^*}{2} + 2b$$

The agent is indifferent between a_i^* and a_{i+1}^* at $\theta = \frac{1}{4}(\theta_{i-1}^* + 2\theta_i^* + \theta_{i+1}^*) - b$. Since

$$|\frac{1}{4}(\theta_{i-1}^* + 2\theta_i^* + \theta_{i+1}^*) - b - (\theta_{i-1}^* + \theta_{i+1}^* - \theta_i^*)| < |\frac{1}{4}(\theta_{i-1}^* + 2\theta_i^* + \theta_{i+1}^*) - b - \theta_i^*| \tag{A.17}$$

□

Since $|\theta_{i+1}^* - 2\theta_i^* + \theta_{i-1}^*|$ is the difference of length between two adjacent intervals, this Lemma implies that the informativeness of an equilibrium cannot exceed the most informative CS Equilibrium.

Direct Control: The solution to the following maximization problem can be implemented under direct control:

$$\max_{\psi \in \Psi_0} I(\psi) \quad s.t. \quad \theta_{i+1} - \theta_{i-1} \geq 4b, \quad C(I(\psi), c) \leq \frac{1}{12} I(\psi)$$

In this equilibrium, the principal threats the agent that if she does not acquire information structure ψ, then he will ignore her message at the communication stage. The constraint $\theta_{i+1} - \theta_{i-1} \geq 4b$ ensures that the agent fully communicates her information at the communication stage and $C(I(\psi), c) \leq \frac{1}{12} I(\psi)$ ensures that she prefers to acquire information structure ψ rather than deviating and switching to the babbling equilibrium.

If constraint $C(I(\psi), c) \leq \frac{1}{12} I(\psi)$ is not binding, then the solution to this maximization problem is the most informative Ivanov Equilibrium, which is shown to be more informative than the most informative CS equilibrium. If constraint $C(I(\psi), c) \leq \frac{1}{12}I(\psi)$ is binding, the agent cannot acquire an information structure with informativeness level more than $I(\psi)$.

To conclude, cost control is always being dominated by direct control. $\square$

B Appendix II: General Formulation of Interval Partition Equilibria

I generalize my findings in section 3 by adopting a more general framework. Again, I assume that $\Theta = [0, 1]$.[21] Let $F(\theta)$ be the cdf of the prior on θ and $f(\theta)$ be its pdf. Let $P(\Theta_i) \equiv \int_{\Theta_i} dF(\theta)$ be the probability of Θ_i to occur under F.

B.1 Information Acquisition

Let Ψ to be set of partition information structures. Let $g : [0, 1] \to \mathbb{R}$ be a strictly increasing and convex function on $[0, 1]$ with $g \in \mathcal{C}^2$, $g(0) = 0$ and $\lim_{x \to 0^+} g'(x) = 0$. I use

$$-\sum_{i=1}^{n} g(P(\Theta_i))$$

to measure the amount of information acquired. Let $c \in \mathbb{R}_+$ be the cost parameter. I assume that the cost of information takes the following form:

$$C(-\sum_{i=1}^{n} g(P(\Theta_i)), c)$$

where C is $\mathcal{C}^2$ on $\mathbb{R} \times \mathbb{R}_+$ and satisfies $C_1 > 0$, $C_2 > 0$, $C_{12} > 0$ and $C(-g(1), c) = 0$ for any c. From the Jensen's Inequality, we know that C satisfies monotonicity. For simplicity, I focus on linear cost function in our following discussions:

$$C(-\sum_{i=1}^{n} g(P(\Theta_i)), c) = -c\sum_{i=1}^{n} g\big(P(\Theta_i)\big) + cg(1) \tag{B.1}$$

Without further notice, my results are robust to the more general form of cost function.

[21] This is without loss of generality as long as Θ is convex and compact.

Remark: This generalized form of cost function is consistent with the idea that an information structure is more costly if it reduces more uncertainty. The '*Shannon's Entropy*' is just a specific parametric form of this generalized family.[22]

B.2 Preferences

Let the sender and receiver have utility functions $u^s(\theta, a)$ and $u^r(\theta, a)$. For any $i \in \{s, r\}$, $\theta \in [0,1]$ and $a \in \mathbb{R}$: $u^i \in \mathcal{C}^3$, $u^i_{12} > 0$, $u^i_{22} < 0$, and there exists $a^i(\theta)$ for every $\theta \in [0,1]$ such that:

$$a^i(\theta) \in \arg\max_a u^i(\theta, a), \quad u^i_2(\theta, a^i(\theta)) = 0$$

From the concavity and continuity of u^i, $a^i(\theta)$ is unique and differentiable. Under the above conditions on the preferences, I show that if an equilibrium has finite number of equilibrium actions, then it must be an interval partition equilibrium.

Lemma B.1. *If the sender is allowed to induce any action in a finite action set, it is optimal for him to choose an interval partition information structure.*

Proof of Lemma B.1. I prove by contradiction.
Let $\{a^*_1, ..., a^*_n\}$ be the receiver's equilibrium action set, where $a^*_1 < ... < a^*_n$. If $\psi \notin \Psi_0$, then, there exists $0 \leq i < j \leq n$ and $\Theta^0_i \subset \Theta_i$, $\Theta^0_j \subset \Theta_j$ such that:

$$P(\Theta^0_i) = P(\Theta^0_j) > 0$$

[22]The entropy of prior F is:

$$H(prior) = -\int_0^1 f(\theta) \ln f(\theta) d\theta$$

The expected entropy of the posterior distribution induced by partition structure $(\theta_0, \theta_1, ..., \theta_n)$ is given by:

$$\begin{aligned} H(\theta_1, ..., \theta_{n-1}) \qquad &= -\sum_{i=1}^n \Big[\big(F(\theta_i) - F(\theta_{i-1})\big) \int_{\theta_{i-1}}^{\theta_i} \frac{f(\theta)}{F(\theta_i) - F(\theta_{i-1})} \ln \frac{f(\theta)}{F(\theta_i) - F(\theta_{i-1})} d\theta\Big] \\ &= -\sum_{i=1}^n \Big[\int_{\theta_{i-1}}^{\theta_i} f(\theta) \ln f(\theta) d\theta - \int_{\theta_{i-1}}^{\theta_i} f(\theta) \ln \big(F(\theta_i) - F(\theta_{i-1})\big) d\theta\Big] \\ &= -\int_0^1 f(\theta) \ln f(\theta) d\theta + \sum_{i=1}^n P[\theta_{i-1}, \theta_i] \ln P[\theta_{i-1}, \theta_i] \end{aligned}$$

The amount of information acquired is given by:

$$I(\theta_1, ..., \theta_{n-1}) = -\sum_{i=1}^n P[\theta_{i-1}, \theta_i] \ln P[\theta_{i-1}, \theta_i]$$

with $g(x) = x \ln x$ a convex function.

and for any $\theta \in \Theta_i^0$, $\theta' \in \Theta_j^0$, $\theta > \theta'$.

Let ψ' be a partition structure such that the underlying partition is the same as ψ except that:

$$\Theta_i' = (\Theta_i - \Theta_i^0) \bigcup \Theta_j^0, \quad \Theta_j' = (\Theta_j - \Theta_j^0) \bigcup \Theta_i^0$$

Then, $C(\psi) = C(\psi')$. The sender's expected loss is reduced by:

$$\begin{aligned} &\int_{\Theta_i^0} u^s(\theta, a_j^*) dF(\theta) - \int_{\Theta_i^0} u^s(\theta, a_i^*)^2 dF(\theta) \\ &+ \int_{\Theta_j^0} u^s(\theta, a_i^*) dF(\theta) - \int_{\Theta_j^0} u^s(\theta, a_j^*) dF(\theta) \\ &= \int_{\Theta_i^0} \Big[\int_{a_i^*}^{a_j^*} u_2^s(\theta, a) da \Big] dF(\theta) - \int_{\Theta_i^0} \Big[\int_{a_i^*}^{a_j^*} u_2^s(\theta, a) da \Big] dF(\theta) \end{aligned} \tag{B.2}$$

Since $a_j^* > a_i^*$ and $u_{12} > 0$, so the sender can strictly gain by deviating to ψ', which is a contradiction. □

To facilitate further discussions, I define two functions: $\overline{a}$ and W. For each pair of $0 \le \theta_{i-1} < \theta_i \le 1$:[23]

$$\overline{a}[\theta_{i-1}, \theta_i] \equiv \arg\max_a \int_{\theta_{i-1}}^{\theta_i} u^r(\theta, a) dF(\theta) \tag{B.3}$$

For any $0 \le \theta_{i-1} \le \theta_i \le \theta_{i+1} \le 1$ and $c \in \mathbb{R}^+$:[24]

$$W(\theta_{i-1}, \theta_i, \theta_{i+1}, c) \equiv u^s(\theta_i, \overline{a}[\theta_i, \theta_{i+1}]) - u^s(\theta_i, \overline{a}[\theta_{i-1}, \theta_i]) + cg'(P[\theta_i, \theta_{i+1}]) - cg'(P[\theta_{i-1}, \theta_i]) \tag{B.4}$$

$\theta_0, ..., \theta_n$ is a '*solution*' under c if and only if $W(\theta_{i-1}, \theta_i, \theta_{i+1}, c) = 0$ for any $1 \le i \le n-1$. A solution is '*Locally Type-I*' at i if $P[\theta_{i-1}, \theta_i] < P[\theta_i, \theta_{i+1}]$; a solution is '*Locally Type-II*' at i if $P[\theta_{i-1}, \theta_i] > P[\theta_i, \theta_{i+1}]$. If a solution is Locally Type-I for all i, then it is called a '*Type-I Solution*', the same reasoning apply to '*Type-II Solution*'.

Now I introduce some regularity conditions on the primitives which will be useful in the discussions below:

---

[23]This is well-defined for $\theta_i \to \theta_{i-1}^-$. Since the bounded convergence theorem implies that:

$$\lim_{\theta_i \to \theta_{i-1}^+} \overline{a}[\theta_{i-1}, \theta_i] = a^r(\theta_{i-1})$$

[24]Notice that when $c = 0$, W is just the 'V function' in Crawford and Sobel (1982).

Condition B.1. *For any θ and a,*

$$u_2^s(\theta, a) - u_2^r(\theta, a) > 0 \tag{B.5}$$

Condition B.2. *For any θ_{i-1}, θ_i and θ_{i+1} such that $P[\theta_{i-1}, \theta_i] = P[\theta_i, \theta_{i+1}]$:*

$$u^s(\theta_i, \overline{a}[\theta_{i-1}, \theta_i]) < u^s(\theta_i, \overline{a}[\theta_i, \theta_{i+1}]) \tag{B.6}$$

Condition B.3 (M-ε)**.** *There exists $\varepsilon > 0$ such that for any $c \in [0, \varepsilon]$, and two forward solutions θ and θ' under c such that $\theta_0 = \theta'_0$, $\theta_1 < \theta'_1$, then: $\theta_i < \theta'_i$ for all $2 \leq i \leq n$.*

Below are remarks on the conditions:

1. Condition B.1 means that the sender is right hand biased. The usual assumption $a^s(\theta) > a^r(\theta)$, which is weaker, is not sufficient to imply that the number of actions in any equilibrium must be finite. This is because when the sender has incomplete information on θ, not only does his favorite action at each state matters, but also how fast does utility declines when we move away from $a^s(\theta)$ and $a^r(\theta)$.

2. Condition B.2 imposes a restriction on the prior distribution, which is satisfied when F is close to uniform distribution. Under this restriction, all solutions are *Type-I* when $c = 0$. This facilitates our discussions in the next subsection, where we prove how does the solution change from Type-I to Type-II when c increases.

3. Condition B.3 is a stronger version of the M-condition in Crawford and Sobel (1982), which says that the M property still holds for small perturbations. This ensures that for $c < \varepsilon$, fixing θ_0 and θ_n, there can be at most one forward solution to the problem. I will use this assumption later to prove the equilibrium convergence result.

I extend the CS finiteness result to a setting where the sender has incomplete information under Condition B.1:

Lemma B.2. *The number of equilibrium actions must be finite in any equilibrium.*

Proof of Lemma B.2. I show that there exists $\varepsilon > 0$ such that for any two equilibrium actions $a_1^* < a_2^*$, $|a_2^* - a_1^*| > \varepsilon$.

For this, I only need to show for any a_1^*, a_2^* such that $a_1^* < a_2^* < a_1^* + \varepsilon$, then for any distribution $\psi(\theta)$ such that:

$$\int u^s(\theta, a_2^*) d\psi(\theta) < \int u^s(\theta, a_1^*) d\psi(\theta) \tag{B.7}$$

then:

$$\arg\max_a \int u^r(\theta, a) d\psi(\theta) < a_1^* \tag{B.8}$$

Suppose it is not true, then:

$$\int_\theta \Big[\int_{a_1^*}^{a_2^*} u_2^s(\theta, a) da \Big] d\psi(\theta) < 0 \tag{B.9}$$

and

$$\int u_2^r(\theta, a_1^*) d\psi(\theta) \geq 0 \tag{B.10}$$

Since $u_{22}^s < 0$ and $[0,1] \times [a^r(0), a^s(1)]$ is compact, so there exists $0 < \zeta_1 < \zeta_2$, such that $|u_{22}^s| \in [\zeta_1, \zeta_2]$. Then we have:

$$\begin{aligned} 0 \quad & > \int_\theta \Big[\int_{a_1^*}^{a_2^*} u_2^s(\theta, a) da \Big] d\psi(\theta) \\ & > (a_2^* - a_1^*) \int u_2^s(\theta, a_2^*) d\psi(\theta) \\ & > (a_2^* - a_1^*) \int \Big[u_2^s(\theta, a_1^*) - \zeta_2 (a_2^* - a_1^*) \Big] d\psi(\theta) \end{aligned}$$

So:

$$\int u_2^s(\theta, a_1^*) d\psi(\theta) < \zeta_2 \varepsilon \tag{B.11}$$

Also, since $u_2^s - u_2^r > 0$, so there exists $\eta > 0$ such that $|u_2^s - u_2^r| \geq \eta$, so:

$$\zeta_2 \varepsilon > \int \Big[u_2^s(\theta, a_1^*) - u_2^s(\theta, a_1^*) \Big] d\psi(\theta) > \eta \tag{B.12}$$

Take $\varepsilon < \frac{\eta}{\zeta_2}$ leads to a contradiction. $\square$

B.3 Equilibrium

Given the cost parameter c, the necessary conditions for $(\theta_0, \theta_1, ..., \theta_n)$ to be an equilibrium information structure are:

1. It is a solution under c
2. The initial value conditions $\theta_0 = 0$ and $\theta_n = 1$ are satisfied.

If an equilibrium can be characterized by a Type-I Solution, then it is called a '*Type-I Equilibrium*'; If it is characterized by a Type-II solution, then it is called a '*Type-II Equilibrium*'. I show the following Proposition under Condition B.2:

Proposition 7. *There exists $\overline{c}$, $\underline{c} > 0$ such that:*

1. *A solution is Locally Type-I only if $c \leq \overline{c}$;*
2. *A solution is Locally Type-II only if $c \geq \underline{c}$;*
3. *When $\overline{c} < \underline{c}$,*[25] *only Type-I and Type-II equilibria exist.*

Proof of Proposition 7: The proof proceeds as follows:

1. For each θ_{i-1}, θ_i there exists $\overline{c}(\theta_{i-1}, \theta_i) \geq 0$ such that there exists $\theta_{i+1}(c) \leq 1$, $W(\theta_{i-1}, \theta_i, \theta_{i+1}(c), c) = 0$ if and only if $c \leq \overline{c}(\theta_{i-1}, \theta_i)$;
2. For each θ_{i+1}, θ_i there exists $\underline{c}(\theta_i, \theta_{i+1}) \geq 0$ such that there exists $\theta_{i-1}(c) \geq 0$, $W(\theta_{i-1}(c), \theta_i, \theta_{i+1}, c) = 0$ if and only if $c \geq \underline{c}(\theta_i, \theta_{i+1})$;
3. $\overline{c}(\theta_{i-1}, \theta_i)$ is continuous in $\{(\theta_{i-1}, \theta_i) | 0 \leq \theta_{i-1} \leq \theta_i \leq 1\}$;
4. $\underline{c}(\theta_i, \theta_{i+1})$ is continuous in $\{(\theta_i, \theta_{i+1}) | 0 \leq \theta_i \leq \theta_{i+1} \leq 1\}$;
5. From the compactness of $\{(\theta_{i-1}, \theta_i) | 0 \leq \theta_{i-1} \leq \theta_i \leq 1\}$ and $\{(\theta_i, \theta_{i+1}) | 0 \leq \theta_i \leq \theta_{i+1} \leq 1\}$, and the continuity of $\overline{c}$ and $\underline{c}$, we know that $\inf \underline{c}(\theta_i, \theta_{i+1})$ and $\sup \overline{c}(\theta_{i-1}, \theta_i)$ exists.

Type-I Solutions: Let us consider the following 'local' problem: For fixed θ_{i-1}, θ_i, what is the value of θ_{i+1}. Let $\theta_{i+1}(c)$ be the solution to $W(\theta_{i-1}, \theta_i, \theta_{i+1}, c) = 0$, we can prove the following Lemma:

Lemma B.3. *In any Type-I solution, $\theta_{i+1}(c) > \theta_{i+1}(0)$ for any $c > 0$. Furthermore, there exists $\overline{c}$, such that a 'Type-I' solution exists if and only if $c < \overline{c}$.*

Proof. From the concavity of u^s,

$$u^s(\theta_i, \overline{a}[\theta_i, \theta_{i+1}]) > u^s(\theta_i, \overline{a}[\theta_{i-1}, \theta_i]) \tag{B.13}$$

[25] This is satisfied under the uniform-quadratic-cubic specification.

for any $\theta_{i+1} < \theta_{i+1}(0)$. Since $g'(P[\theta_i, \theta_{i+1}]) > g'(P[\theta_{i-1}, \theta_i])$ in any Type-I solution, so $\theta_{i+1}(c) > \theta_{i+1}(0)$.

When $\theta_{i+1} = \theta_{i+1}(0)$, $W = c[g'(P[\theta_i, \theta_{i+1}]) - g'(P[\theta_{i-1}, \theta_i])] > 0$. The derivative of W with respect to θ_{i+1} is given by:

$$\frac{\partial W}{\partial \theta_{i+1}} = -\frac{u_2^s(\theta_i, \overline{a}) u_2^r(\theta_{i+1}, \overline{a}) f(\theta_{i+1})}{\displaystyle\int_{\theta_i}^{\theta_{i+1}} u_{22}^r(\theta, \overline{a}) dF(\theta)} + cg''(P[\theta_i, \theta_{i+1}]) f(\theta_{i+1}) \tag{B.14}$$

There exists a Type-I solution if and only if there exists $\theta_{i+1}(0) < \theta_{i+1}(c) < 1$, such that:

$$-c[g'(P[\theta_i, \theta_{i+1}(0)]) - g'(P[\theta_{i-1}, \theta_i])] = \int_{\theta_{i+1}(0)}^{\theta_{i+1}(c)} \frac{\partial W(\theta_{i-1}, \theta_i, \theta_{i+1}, c)}{\partial \theta_{i+1}} d\theta_{i+1} \tag{B.15}$$

Since

$$u_2^s(\theta_i, \overline{a}) < 0, \quad u_2^r(\theta_{i+1}, \overline{a}) > 0, \quad \int_{\theta_i}^{\theta_{i+1}} u_{22}^r(\theta, \overline{a}) dF(\theta) < 0, \quad g''(P) > 0$$

So for any $c_1 < c_2$, if a Type-I solution exists when $c = c_2$, then

$$-c_1[g'(P[\theta_i, \theta_{i+1}(c_2)]) - g'(P[\theta_{i-1}, \theta_i])] > \int_{\theta_{i+1}(0)}^{\theta_{i+1}(c_2)} \frac{\partial W(\theta_{i-1}, \theta_i, \theta_{i+1}, c_1)}{\partial \theta_{i+1}} d\theta_{i+1} \tag{B.16}$$

From the continuity of both sides, there exists $\theta_{i+1}(c_1) \in (\theta_{i+1}(0), \theta_{i+1}(c_2))$ such that $\theta_{i+1} = \theta_{i+1}(c_1)$ is a Type-I solution when $c = c_1$. $\square$

Furthermore, define $\overline{c}(\theta_{i-1}, \theta_i) = 0$ if $\theta_{i+1}(0) > 1$. Now I prove the continuity of $\overline{c}$. First, I show that it is continuous when $\theta_{i-1} \neq \theta_i$; later, I show that it is continuous at $\theta_{i-1} = \theta_i$.

When $\theta_{i-1} \neq \theta_i$, let θ'_{i-1} and θ'_i satisfies: $|\theta_{i-1} - \theta'_{i-1}| < \xi$, $|\theta_i - \theta'_i| < \xi$. Taylor expand $W(\theta'_{i-1}, \theta'_i, \theta_{i+1}(c), c)$ around (θ_{i-1}, θ_i) and ignore higher order terms, we have:

$$W(\theta'_{i-1}, \theta'_i, \theta_{i+1}(c), c) \leq \xi\Big(\Big|\frac{\partial W}{\partial \theta_{i-1}}\Big| + \Big|\frac{\partial W}{\partial \theta_i}\Big|\Big) \tag{B.17}$$

As shown above, $\frac{\partial W}{\partial \theta_i}$ and $\frac{\partial W}{\partial \theta_{i-1}}$ are bounded when $\theta_{i-1} \neq \theta_i$. Since $\frac{\partial W}{\partial c} > 0$, reduce c by $A(\xi)$ such that $\lim_{\xi \to 0} A(\xi) = 0$, to make $W(\theta'_{i-1}, \theta'_i, \theta_{i+1}(c), c - A(\xi)) = 0$.

Next, I prove that $\overline{c}$ is continuous at $\theta_{i-1} = \theta_i$. Since

$$\lim_{\theta_i \to \theta_{i-1}^-} \overline{a}[\theta_{i-1}, \theta_i] = a^r(\theta_{i-1})$$

and u_1^s, u_2^s, g'' are bounded, we get the conclusion.

Type-II Solutions: Let $\theta_{i-1}(c)$ be the solution to $W(\theta_{i-1}, \theta_i, \theta_{i+1}, c) = 0$. Let θ'_{i-1} be the point satisfies $P[\theta'_{i-1}, \theta_i] = P[\theta_i, \theta_{i+1}]$. Obviously, if a Type-II solution exists under c, then $\theta_{i-1}(c) \in [0, \theta'_{i-1})$. I begin with showing the Lemma below:

Lemma B.4. *For any (θ_i, θ_{i+1}), there exists $\underline{c}$, such that a Type-II solution exists if and only if $c > \underline{c}$.*

Proof. When $\theta_{i-1} = \theta'_{i-1}$, from Assumption 1, we have:

$$W(\theta'_{i-1}, \theta_i, \theta_{i+1}, c) = u^s(\theta_i, \overline{a}[\theta_i, \theta_{i+1}]) - u^s(\theta_i, \overline{a}[\theta'_{i-1}, \theta_i]) > 0 \tag{B.18}$$

Take the derivative of W with respect to θ_{i-1}:

$$\frac{\partial W}{\partial \theta_{i-1}} = -\frac{u_2^s(\theta_i, \overline{a}) u_2^r(\theta_{i-1}, \overline{a}) f(\theta_{i-1})}{\displaystyle\int_{\theta_{i-1}}^{\theta_i} u_{22}^r(\theta, \overline{a}) dF(\theta)} + c g''(P[\theta_{i-1}, \theta_i]) f(\theta_{i-1}) \tag{B.19}$$

There exists a Type-II solution if and only if there exists $0 \leq \theta_{i-1}(c) < \theta'_0$ such that:

$$u^s(\theta_i, \overline{a}[\theta_i, \theta_{i+1}]) - u^s(\theta_i, \overline{a}[\theta'_{i-1}, \theta_i]) = \int_{\theta_{i-1}(c)}^{\theta'_{i-1}} \frac{\partial W(\theta_{i-1}, \theta_i, \theta_{i+1}, c)}{\partial \theta_{i-1}} d\theta_{i-1} \tag{B.20}$$

Since

$$u_2^s(\theta_i, \overline{a}) > 0, \quad u_2^r(\theta_{i-1}, \overline{a}) < 0, \quad \int_{\theta_{i-1}}^{\theta_i} u_{22}^r(\theta, \overline{a}) dF(\theta) < 0, \quad g''(P) > 0$$

Similar to the previous proof, we get the conclusion. $\square$

The continuity of $\underline{c}$ is proved similarly as the continuity of $\overline{c}$. $\square$

Remark: Under the uniform-quadratic-cubic assumption, $\underline{c} > \overline{c}$. This implies that there exists no equilibria in which the lengths of the intervals are sometimes increasing and sometimes decreasing. In general, this conclusion is not true. When $\underline{c} < \overline{c}$, there exists c such that there exists an equilibrium where the lengths of the intervals is sometimes increasing, sometimes decreasing. I call this a '*compound equilibrium*'.

B.4 Equilibrium Convergence

When the Crawford-Sobel M-Property is robust to small perturbations (Condition B.3), the local incentive constraints are sufficient when c is small.

Lemma B.5. *There exists $\varepsilon_0 > 0$ such that when $c < \varepsilon_0$, the local incentive constraints at the information acquisition stage are sufficient.*

This result leads to a convergence in the equilibrium partition points when $c \to 0$.

Proposition 8. *The conclusion in Proposition 4 holds under Condition B.1, B.2 and B.3.*

Proof of Lemma B.5 and Proposition 8: The proof relies on the following Theorem:

Theorem 1 (The Maximum Theorem)**.** *Let $f : S \times \Theta \to \mathbb{R}$ be a continuous function, and $\mathcal{D} : \Theta \to P(S)$ be a compact-valued, continuous correspondence. Let*

$$\mathcal{D}^*(\theta) = \arg\max\{f(x, \theta)|x \in \mathcal{D}(\theta)\}$$

Then $\mathcal{D}^*$ is a compact valued, upper-semi-continuous correspondence on Θ.

Let $(a_1^*, ..., a_n^*)$ be the equilibrium actions in a CS Equilibrium. When c is small enough, let $(a_1^*(c), ..., a_n^*(c))$ be the actions defined by:

$$a_j^*(c) = \overline{a}[\theta_{j-1}(c), \theta_j(c)]$$

where $(0, \theta_1(c), ..., \theta_{n-1}(c), 1)$ is a forward solution to $W(\theta_{i-1}, \theta_i, \theta_{i+1}, c) = 0$.
I prove the proposition through 3 steps.

In the first step, I show that $a_j^*(c) \equiv \overline{a}[\theta_{j-1}(c), \theta_j(c)]$ is right-continuous at $c = 0$, which is implied by the right continuity of $\theta_j(c)$ at $c = 0$ $(j = 1, 2, ..., n-1)$. The $M - \varepsilon$ condition ensures that $\theta_j(c)$ is well defined for any $c < \varepsilon$.

For a fixed c, if $W(\theta_{i-1}, \theta_i, \theta_{i+1}, c) = 0$, then θ_{i+1} changes continuously with θ_{i-1} and θ_i. First I show the following Lemma:

Lemma B.6. *If $(\theta_0, \theta_1, \theta_2, ..., \theta_n)$ is a solution at $c = c_0$, $(\theta_0, \theta_1, \theta_2', ..., \theta_n')$ is a solution at $c = c_1$, s.t. $c_1 < c_0$, $\theta_2' < \theta_2$. Then $\theta_i' < \theta_i$ for any $i \geq 2$.*

Proof. Since $c_1 < c_0$, and we are under a Type-I Solution, it is obvious that $\theta_2' < \theta_2$.

Let $(\theta_1, \theta_2', \theta_{3,1}, ..., \theta_{n,1})$ be a forward solution under $c = c_0$, from the $M - \varepsilon$ condition, $\theta_{i,1} < \theta_i$ for $i \geq 3$. Let $(\theta_1, \theta_2', \theta_3')$ be a solution at $c = c_1$, so $\theta_3' < \theta_{3,1} < \theta_3$.

Let $(\theta_2', \theta_3', \theta_{4,2}, ..., \theta_{n,2})$ be a forward solution under $c = c_0$, from the $M - \varepsilon$ condition, $\theta_{i,2} < \theta_{i,1}$ for $i \geq 4$. Let $(\theta_2', \theta_3', \theta_4')$ be a solution at $c = c_1$, so $\theta_4' < \theta_{4,2} < \theta_{4,1} < \theta_4$.

Continue this process, we have: $\theta_i' < \theta_i$ for all $i \geq 2$. □

Also, for fixed θ_{i-1} and θ_i, θ_{i+1} is right continuous with respect to c. Since $\theta_n' < \theta_n$, I can increase θ_1 to make $\theta_n'(\theta_1) = \theta_n$. From the continuity conditions we mentioned above, $\theta_j(c)$ is right-continuous with respect to c.

In the second step, I show there exists a solution to the following maximization problem that converges to $(\theta_1(0), \theta_2(0), ..., \theta_{n-1}(0))$ when $c \to 0$.

$$\max_{\theta_1,...,\theta_{n-1}} f(\theta_1, ..., \theta_{n-1}) \tag{B.21}$$

s.t. $0 \leq \theta_1 \leq ... \leq \theta_{n-1} \leq 1$, where

$$f(\theta_1, ..., \theta_{n-1}) \equiv \sum_{i=1}^{n} \int_{\theta_{i-1}}^{\theta_i} u^s(\theta, a_i^*(c)) dF(\theta) + c \sum_{i=1}^{n} g\big(P[\theta_{i-1}, \theta_i]\big) \tag{B.22}$$

Since $\{(\theta_1, ..., \theta_{n-1}) | 0 \leq \theta_1 \leq ... \leq \theta_{n-1} \leq 1\}$ is compact, and the maximization problem is continuous, according to the *Maximum Theorem*, the correspondence

$$\mathcal{D}^*(c) \equiv \arg\max\{f(\theta_1, ..., \theta_{n-1})\} \tag{B.23}$$

is upper-semi-continuous. Since $\mathcal{D}^*(0) = \{(\theta_1(0), ..., \theta_{n-1}(0))\}$, which is a singleton, there exists $0 < \varepsilon_1 < \varepsilon$, such that when $c < \varepsilon_1$, the solution to the maximization problem is unique. Let the solution to be $(\theta_1'(c), ..., \theta_{n-1}'(c))$, so $(\theta_1'(c), ..., \theta_{n-1}'(c)) \to (\theta_1(0), ..., \theta_{n-1}(0))$ when $c \to 0$.

For the last step, I show that $(\theta_1(c), ..., \theta_{n-1}(c)) = (\theta_1'(c), ..., \theta_{n-1}'(c))$ when c is small enough. Since $\theta_j(c) \to \theta_j(0)$, $\theta_j'(c) \to \theta_j(0)$, so for any $\xi > 0$, there exists $0 < \varepsilon_2 < \varepsilon_1$, such that for any $c < \varepsilon_2$, $|\theta_j(c) - \theta_j'(c)| < \xi$ for any j. So $\theta_1'(c) < ... < \theta_{n-1}'(c)$ when ξ is small enough. From the Kuhn-Tucker complementary slackness condition, we know that none of the constraints $\theta_j \leq \theta_{j+1}$

is binding. So both $(\theta'_1(c), ..., \theta'_{n-1}(c))$ and $(\theta_1(c), ..., \theta_{n-1}(c))$ are solutions to:

$$u^s(\theta_i, \overline{a}[\theta_i(c), \theta_{i+1}(c)]) - u^s(\theta_i, \overline{a}[\theta_{i-1}(c), \theta_i(c)]) + cg'\big(P[\theta_i, \theta_{i+1}]\big) - cg'\big(P[\theta_{i-1}, \theta_i]\big) = 0 \tag{B.24}$$

Lemma B.6 is proved.

Let us continue the proof of Proposition 8. Let

$$\big|\theta_i(c) - \theta'_i(c)\big| = \max\Big\{\big|\theta_k(c) - \theta'_k(c)\big|\Big|k = 1, 2, ..., n-1\Big\}$$

Assume $\big|\theta_i(c) - \theta'_i(c)\big| = \xi_0 < \xi$. If $\xi_0 > 0$, since

$$u^s(\theta_i(c), a^*_{i+1}(c)) - u^s(\theta_i(c), a^*_i(c)) + cg'\big(P[\theta_i(c), \theta_{i+1}(c)]\big) - cg'\big(P[\theta_{i-1}(c), \theta_i(c)]\big) = 0 \tag{B.25a}$$

$$u^s(\theta'_i(c), a^*_{i+1}(c)) - u^s(\theta'_i(c), a^*_i(c)) + cg'\big(P[\theta'_i(c), \theta'_{i+1}(c)]\big) - cg'\big(P[\theta'_{i-1}(c), \theta'_i(c)]\big) = 0 \tag{B.25b}$$

Expand the second equation at $\theta_{i-1}(c), \theta_i(c), \theta_{i+1}(c)$

$$\begin{aligned}
&0 = u^s(\theta'_i(c), a^*_{i+1}(c)) - u^s(\theta'_i(c), a^*_i(c)) + cg'\big(P[\theta'_i(c), \theta'_{i+1}(c)]\big) - cg'\big(P[\theta'_{i-1}(c), \theta'_i(c)]\big)\\
&= \int_{a^*_{i-1}}^{a^*_i} u^s_2(\theta'_i(c), a)da + cg'\big(P[\theta'_i(c), \theta'_{i+1}(c)]\big) - cg'\big(P[\theta'_{i-1}(c), \theta'_i(c)]\big)\\
&= \xi_0 \int_{a^*_{i-1}}^{a^*_i} u^s_{12}(\theta_i(c), a)da + c\Big[g''(P_i)f(\theta_{i+1}(c))(\theta'_{i+1}(c) - \theta_{i+1}(c)) - g''(P_i)f(\theta_i(c))(\theta'_i(c) - \theta_i(c))\\
&- g''(P_{i-1})f(\theta_i(c))(\theta'_i(c) - \theta_i(c)) + g''(P_{i-1})f(\theta_{i-1}(c))(\theta'_{i-1}(c) - \theta_{i-1}(c))\Big]
\end{aligned} \tag{B.26}$$

Since $u^s_{12} > 0$, so it has a strictly positive lower bound. Since $|\theta_{i+1}(c) - \theta'_{i+1}(c)| < \xi_0$, $|\theta_{i-1}(c) - \theta'_{i-1}(c)| < \xi_0$, and g'', f are all bounded from above, we know that there exists c small enough such that:

$$\begin{aligned}
&\xi_0 \int_{a^*_{i-1}}^{a^*_i} u^s_{12}(\theta_i(c), a)da >\\
&c\xi_0\Big[g''(P_i)f(\theta_{i+1}(c)) + g''(P_i)f(\theta_i(c)) + g''(P_{i-1})f(\theta_i(c)) + g''(P_{i-1})f(\theta_{i-1}(c))\Big]
\end{aligned} \tag{B.27}$$

which leads to a contradiction. So $\xi_0 = 0$, which means, $\theta_j(c) = \theta'_j(c)$ when c is close to 0. $\square$

References

[1] Aghion, Philippe, and Jean Tirole, "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105 (1997), 1-29.

[2] Alonso, Ricardo, and Niko Matouschek, "Optimal Delegation," *Review of Economic Studies*, 75 (2008), 259-293.

[3] Amador, Manuel, and Kyle Bagwell, "The Theory of Optimal Delegation with an Application to Tariff Caps," *Econometrica*, forthcoming, 2012.

[4] Ambrus, Attila, and Satoru Takahashi, "Multi-sender Cheap Talk with Restricted State S-paces," *Theoretical Economics*, 3 (2008), 1-27.

[5] Amir, Rabah, and Natalia Lazzati, "Endogenous Information Acquisition in Bayesian Games with Strategic Complementarities," Working Paper, 2011.

[6] Argenziano, Rossella, Sergei Severinov, and Francesco Squintani, "Strategic Information Acquisition," Working Paper, 2011.

[7] Athey, Susan, and Jonathan Levin, "The Value of Information in Monotone Decision Problems," Working Paper, 2001.

[8] Austen-Smith, David, "Strategic Transmission of Costly Information," *Econometrica*, 62 (1994), 955-963.

[9] Battaglini, Marco, "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, 70 (2002), 1379-1401.

[10] Che, Yeon-Koo, Wouter Dessein, and Navin Kartik, "Pandering to Persuade," *American Economic Review*, 103 (2013), 47-79.

[11] Che, Yeon-Koo, and Navin Kartik, "Opinions as Incentives," *Journal of Political Economy*, 117 (2009), 815-860.

[12] Crawford, Vincent and Joel Sobel, "Strategic Information Transmission," *Econometrica*, 50 (1982), 1431-1451.

[13] Crémer, Jacques, "Arm's Length Relationship," *Quaterly Journal of Economics*, 110 (1995), 275-295.

[14] Dessein, Wouter, "Authority and Communication in Organizations," *Review of Economic Studies*, 69 (2002), 811-838.

[15] Gentzkow, Matthew, and Emir Kamenica, "Disclosure of Endogenous Information," Working Paper, 2012.

[16] Goltsman, Maria, and Gregory Pavlov, "How to Talk to Multiple Audiences?" *Games and Economic Behavior*, 72 (2011), 100-122.

[17] Hellwig, Christian, and Laura Veldkamp, "Knowing What Others Know: Coordination Motives in Information Acquisition," *Review of Economic Studies*, 76 (2009), 223-251.

[18] Holmström, Bengt, "On Incentives and Control in Organization," PhD Dissertation, Stanford University, 1977.

[19] Holmström, Bengt, "Moral Hazard and Observability," *Bell Journal of Economics*, 10 (1979), 74-91.

[20] Ivanov, Maxim, "Informational Control and Organizational Design," *Journal of Economic Theory*, 145 (2010), 721-751.

[21] Kamenica, Emir, and Matthew Gentzkow, "Bayesian Persuasion," *American Economic Review*, 101 (2011), 2590-2615.

[22] Kartik, Navin, Marco Ottaviani, and Francesco Squintani,"Credulity, Lies, and Costly Talk," *Journal of Economic Theory*, 134 (2007), 93-116.

[23] Krishna, Vijay, and John Morgan, "A Model of Expertise," *Quaterly Journal of Economics*, 116 (2001), 747-775.

[24] Levy, Gilat and Ronny Razin, "On the Limits of Communication in Multidimensional Cheap Talk: A Comment," *Econometrica*, 75 (2007), 885-893.

[25] Matějka, Filip, and Christopher A. Sims, "Discrete Actions in Information Constrainted Tracking Problems," Working Paper, 2010.

[26] Morgan, John, and Phillip C. Stocken, "An Analysis of Stock Recommendations," *RAND Journal of Economics*, 34 (2003), 183-203.

[27] Myatt, David, and Chris Wallace, "Endogenous Information Acquisition in Coordination Games," *Review of Economic Studies*, 79 (2012), 340-374.

[28] Olszewski, Wojciech, "Informal Communication," *Journal of Economic Theory*, 117 (2004), 180-200.

[29] Ottaviani, Marco, "The Economics of Advice," Working Paper, 2000.

[30] Pei, Di, "Uncertainty about Uncertainty in Communication," Working Paper, 2012.

[31] Shapley, Lloyd and Herbert Scarf, "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1 (1974), 23–37.

[32] Sims, Christopher A., "Implications of Rational Inattention," *Journal of Monetary Economics*, 50 (2003), 665-690.

[33] Szalay, Dezsö, "The Economics of Clear Advice and Extreme Options," *Review of Economic Studies*, 72 (2005), 1173-1198.

[34] Watson, Joel, "Information Transmission When the Informed Party is Confused," *Games and Economic Behavior* 12 (1996), 143-161.

[35] Yang, Ming, "Optimality of Debt under Flexible Information Acquisition." Working Paper, 2011.