РЭШ
Российская
экономическая
школа

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

MASTER THESIS

Тема: Влияние иерархии общения на эффективность двусторонней коммуникации

Title: Communication in organizations: sequential versus simultaneous cheap talk

Студент/ Student:

Starkov Egor Alexeevich
(Ф.И.О. студента, выполнившего работу)

Научные руководители/ Advisors:

NES Associate professor of Economics
Sergei Izmalkov, PhD;
NES Assistant professor of Economics
Efthymios Athanasiou, PhD
(ученая степень, звание, место работы, Ф.И.О.)

Оценка/ Grade:

____

Подпись/ Signature:

____

Москва 2013

Communication in Organizations: Sequential versus Simultaneous Cheap Talk*

Egor Starkov

Abstract

This paper looks into the question of optimal design of communication network within a company. The principal trade-off in managerial decision-making is often identified as adaptation to local environment versus coordination with other divisions within a firm. This trade-off creates a conflict of interests between managers of different departments and prevents them from communicating truthfully with each other. We explore different communication structures in order to optimize the communication process and find out that if the divisions differ sufficiently in size and the smaller division depends heavily on coordination then sequential communication with larger firm as leader is more preferable by the firm, while with divisions of similar sizes simultaneous communication yields better performance.

Целью данной работы является сравнение различных методов иерархического общения менеджеров внутри фирмы. В работе рассматривается модель коммуникации, в которой менеджеры сталкиваются с необходимостью компромисса между адаптацией к локальным условиям своего отдела и координацией действий с другими отделами. В результате возникает конфликт интересов, ведущий к зашумленности общения. В данной работе показывается, что схема последовательного общения в таких условиях оказывается более предпочтительной для фирмы в случае сильной разницы в размерах отделов, в то время как для отделов схожих размеров лучшие результаты показывает схема одновременной коммуникации.

*The author would like to thank Sergei Izmalkov and Efthymious Athanasiou for the attentive advisorship and plenty of rich feedback throughout the process of writing this thesis, as well as Sergei Stepanov for a quantity of valuable comments on the final version of the paper. Acknowledgement also goes to Elena Paltseva and Eyal Winter who, without knowing this, were very helpful during the search for the topic.

Contents

1 Introduction

Cheap talk models in their majority explore settings with one-sided informational asymmetry – in most papers, from seminal work by Crawford and Sobel (1982) to recent contributions, one agent possesses some private knowledge that he wants to some extent share with other agent, whose interests are somewhat, but not absolutely, correlated with his own. No doubt that this topic is indeed of interest and such models have a wide variety of real-life applications, e.g. in the context of expert advice. Moreover, some problems with two-sided asymmetry may be reduced to or separated into few problems with one-sided asymmetry: for example, if today I ask your advice on the wallpapers that would best fit my living room, and tomorrow you ask for my advice concerning monetary policy of Russian Central Bank – these would be two different "uninformed principal & informed expert" problems, and if we neglect dynamic issues like reputational concerns then they can easily be solved separately.

Nonetheless, such decomposition of problem is not always possible. Leading example that motivates such setting (of two-sided communication) and, thus, our paper, comes from managerial economics and concerns the interaction of managers who are in charge of different divisions of the firm. The issue is that every manager running some division of a firm faces two goals that need to be fulfilled. Primary goal is adaptation: "The essence of management is coping with change" (Chakravarthy, 1982, p.35). Any firm exists in a vastly changing conditions, and in order to survive it needs to keep up to the changes, which may include but are not limited to fluctuations in demand, competitors' behaviour, shocks to productivity, innovations etc. Same is true for division-level managers – similarly, they have to optimize business processes within their divisions and adapt to everchanging local environment. What is important, the information about such local conditions (or any other dispersed division-specific "local information") is perceived by managerial literature to be dispersed within the firm (Grant, 1996; Cramton, 2001), i.e. it is possessed by managers of divisions to which it is relevant, but not by other managers and/or principal (headquarters).

Another goal that needs to be fulfilled by a division manager is coordinating activities with other divisions (Malone, 1987; Heath and Staudenmayer, 2000). A nice illustration for this fact would be interaction between marketing and production departments. Consider a car manufacturing company that is about to introduce a new model in its model row. Marketing department has an aim of maximizing firm sales and knows that consumers love red cars the most, so offering red color would boost sales significantly. However, Henry from production insists that every car that walks down the assembly line should be painted in a shade of black. The reason is that black paint is cheaper, so that would result in lower costs, which is one of his primary goals. Optimally, representatives of both departments should meet and discuss, which option would yield maximal sales at minimal cost. This is, of course, a somewhat jocular example, but substitute the color with a choice between steel, aluminum and carbon as a main material for car body or with size and power of engine installed on a car – and the example becomes more lively, doesn't it?

In fact, trade-off between adaptation to local conditions and coordination with other divisions is recognized as the principal trade-off in managerial decision-making (Alonso et al., 2008b).

That is exactly where the problem arises: in order to coordinate with someone, you need to know what his decision is likely to be. However, when making his decision, the manager will try to adapt to his local conditions, information about which is unavailable to you, which makes predicting his decision a tough problem. In fact, not only you want to know the other agent's information to predict his decision, but also you would like to tell him something about your decision, so that he would coordinate his decision with yours and thus mitigate the coordination problem (it is in his interest as well). The other manager faces a completely symmetric problem, so in the end there is a mutual desire to communicate. The problem is that because of the need for adaptation to differing local conditions, managers' interests diverge: each of them would like *both* decisions to answer the needs of *his* division, which undermines the incentives for truthful communication, since each manager would like to manipulate the beliefs and decisions of the other one in his advantage. Therefore, communication is noisy. Is there anything we can do to reduce the noise?

In the literature such communication is usually modeled as cheap talk, i.e. an exchange of non-verifiable, non-binding and payoff-irrelevant messages. The issue is that the communication process itself is usually assumed to be simultaneous, which means that all agents (managers) send their messages at the same time. While this may be reasonable in some settings (e.g. if they prepare written reports for each other, roughly at the same time), it seems somewhat silly in others (e.g. in direct verbal communication), where it might make much more sense to introduce sequential communication schemes, in which one agent speaks first, and the other listens to him and responds, keeping in mind the information he has just received. We ask whether and under what conditions such sequential communication would be better for agents and for the firm as compared to the conventional simultaneous communication.

Speaking more formally, we consider a two-agent cheap talk model with symmetric agents[1]. Each agent has a piece of private information and possesses decision rights on some issue. Each agent's preferences on the decision space depend on the information he possesses and on the decision of another agent – note that agent does not exactly care about other's information, only about his decision. Before the round of decision-making agents have an opportunity to communicate with each other, and our question is how does sequential communication scheme perform on this round in comparison to simultaneous communication.

We find out that simultaneous communication scheme strictly dominates the

[1]This is unlike conventional "expert advice" cheap talk model, where agents are distinguished into informed agent (expert) and uninformed principal – in our setting each agent plays both of these roles.

sequential mechanism in case the divisions are more or less similar in size. However, if one division is significantly larger than another and this other division incurs sufficiently large losses from miscoordination then sequential communication may be more preferable, with the larger division as first-mover.

The paper is organized as follows: Section 2 presents the review of relevant literature. Section 3 sets up the model, Section 4 derives and describes the set of equilibria in our game – these two sections mostly replicate the results from earlier papers. Novel results start from Section 5, in which we discuss some properties of equilibria, while Section 6 discusses the differences between simultaneous and sequential communication schemes. Section 7 concludes.

2 Related literature

Speaking about cheap talk, one should certainly start with seminal paper by Crawford and Sobel (1982), who introduced this method of communication into the analysis of an "expert advice" model. They discovered that if interests of principal and agent, while differing, are sufficiently aligned (i.e. bias of an expert is not too strong) then transmission of information is possible, though not perfectly. They identify the existence of multiple equilibria, which differ in precision of transmitted information, and reasonably conclude that stronger expert's bias in preferences lowers the upper bound on the precision of communication. We are mostly interested in extensions of this model that (a) consider two-sided informational asymmetries and (b) explore various non-trivial communication mechanisms (i.e. all but one-round simultaneous signal-sending).

de Barreda (2010) extends Crawford & Sobel's model for principal to have private information as well. The author discovers that this private information reduces the informativeness of messages used by expert, and in wide range of environments presence of this private information cannot make up for the loss in precision of expert's message, so welfare of both agents decreases. Similar results were obtained by Lai (2009). Chen (2009) also finds out in a similar setting that giving principal an opportunity to send a message before the expert sends his does not lead to significant improvement, since under mild conditions principal would be unwilling to truthfully reveal his information on this first step. Therefore, in the end we arrive at the result that two-sided informational asymmetry generally makes things worse, even if it is supposed to make the decision-maker more informed.

Also of some relevance may be the papers by Battaglini (2003) and McGee and Yang (2013), which explore settings with uninformed principal and two informed agents. They would apply for an alternative setup of our model – if division managers communicate their local information to the headquarters instead of each other, and the HQ possesses all the decision rights (centralized organizational structure). Battaglini (2003) finds out that in a setting where state of the world can be described by a *multidimensional* variable (which is our case – local conditions of one division would determine one coordinate of this descriptive variable) if all experts possess full information and linearly independent biases then full revelation is possible in the model. He also notes that the result is robust to the choice of signaling order, i.e. it holds in both simultaneous and sequential communication schemes. However, in our story managers' information is *local*, meaning that each manager knows only his own state, but not states of other divisions, so of more relevance is the paper by McGee and Yang (2013), who explore a similar setting, in which experts' information is non-overlapping. They discover that "when senders have type-independent biases, their information transmissions exhibit strategic complementarity: more information transmitted by one sender leads to more information being transmitted by the other sender", while in case of type-dependent biases signaling can exhibit strategic substitutability.

A more general result for two-player cheap talk games with two-sided asymmetry is presented by Amitai (1996), who characterizes the set of equilibria in such games.

He discovers that unlike in one-sided information case, the set of equilibria depends heavily on the set of possible messages. However, his results apply to finite games only (with finite sets of actions and messages), while most cheap talk games are continuous in the sense that they allow for continual sets of actions and messages.

Galeotti et al. (2013) explore a problem similar to ours – cheap talk in the context of payoffs depending on others' decisions, – in a network setting. However, they focus mainly on network effects of such communication, which are not explored in our paper. In particular, they find out that agent's incentives to truthfully communicate with another agent, speaking in social choice terms, do not satisfy IIA – Independence of Irrelevant Agents, – and actually depend on the number of agents already truthfully communicating with the recipient. They also discover that social welfare depends not just on the absolute amount of truth in the society, but on the equability of distribution of truth as well – meaning that all agents should hear similar amounts of truth.

As for non-trivial communication mechanisms, Aumann and Hart (2003) characterize the set of equilibrium payoffs in general two-person cheap talk games with one-sided asymmetry, where unboundedly long talks are allowed. They discover that multi-stage communication opens path for equilibria which Pareto-dominate the ones that are possible under single round of communication.

Ottaviani and Sørensen (2001) investigate a problem different from ours: they try to establish an optimal order of speech of committee members during a debate and argue that "optimizing over the order of speech can improve the extraction of information, but not perfectly so".

Concerning cheap talk applications to organizational economics, it is worth mentioning papers by Dessein (2002); Alonso and Matouschek (2007); Alonso et al. (2008a,b, 2009); Rantakari (2008); Bester (2009), which examine the effect of communication in the presence of coordination-adaptation trade-off on the optimal allocation of decision rights – i.e. whether it is worthwhile to delegate the decision rights to division managers or it is better to retain them at headquarters. Most of them find out that both organizational structures may be optimal depending on the situation and derive conditions that identify areas in the parameter space, where one structure dominates the other.

3 Model

In order to investigate the effect of local information on the performance of different communication schemes we employ the firm model proposed by Rantakari (2008). This model is also quite similar to the one used by Alonso, Dessein and Matouschek (2008b; hereinafter referred to as ADM), so some results from these papers will also be applied to our paper. In particular, in our analysis we employ part of Lemma A2 and Proposition 2 from ADM as well as some pieces of intuition from Rantakari. Also, as was noted in the Introduction, this section (model set-up) as well as the next one (search for equilibria) inevitably repeat the corresponding pieces of these papers, since the model we employ possesses no significant differences from the ones presented there. Most novel results obtained in this paper are concentrated in sections 5 and 6.

The model describes an enterprise consisting of 2 divisions, each of which is controlled by a separate manager. Each division faces a trade-off between adapting to local environment, expressed by $\theta_i \sim i.i.d.\, U[-s;s]$, $\forall i \in \{1;2\}$, and coordinating with other divisions. Manager of each division i should make a decision $d_i \in [-s;s]$ so as to maximize the profit of his division[2], which is given by the following profit function:

$$\pi_i = \Pi - r_i \cdot \left[\delta_i \left(d_i - \theta_i\right)^2 + \left(1 - \delta_i\right)\left(d_i - d_j\right)^2\right] \tag{1}$$

The first term, $\Pi \in \mathbb{R}$, stands for maximal possible profit that a division can

[2]For simplicity we ignore the possibility of stimulating manager by bonuses for performance of the whole firm (which also includes profits of other divisions) and assume that his salary depends on the performance of his division only.

achieve in the absence of any losses from misadaptation or miscoordination represented by the following two terms. This term is purely illustrative and does not bear any functional load, so instead of considering profits we might as well consider the loss functions:

$$L_i = r_i \cdot \left[\delta_i \left(d_i - \theta_i\right)^2 + \left(1 - \delta_i\right)\left(d_i - d_j\right)^2\right] \tag{2}$$

Multiplier $r_i \in \mathbb{R}_+$ stands for relative size of the division (or weight that headquarters puts on its profits). The first quadratic term in equation (2), $(d_i - \theta_i)^2$, represents losses from misadaptation (we will also call them adaptation losses whenever convenient). We assume that division i incurs quadratic losses if the decision d_i of its manager does not coincide with the local conditions θ_i, i.e. the division does not fully adapt to local environment. This setting allows for rather painless light misadaptation, while fully ignoring local conditions and not adapting to them at all is very costly. This cost may arise, for example, from the demand side of the market: if consumers' preferences are not perfectly matched by the supply of corresponding division, some consumers may not buy the good, and, therefore, sales and profits will be lower. Similarly, the second quadratic term, $(d_i - d_j)^2$, represents costs of miscoordination (sometimes referred to as coordination costs): division i suffers losses if it does not match the decision of another division. Such cost is more likely to arise due to some features of firm's technology of production, which, e.g., requires coordination of production decisions across divisions. Coefficient $\delta_i \in (0;1)$[3] stands for relative importance of adaptation for division i, so $(1 - \delta_i)$ would be relative importance of coordination or "dependability".

As for informational setting: parameters $\{r_i;\delta_i\}_{i=1}^{2}$ and the functional form of profit/loss function are common knowledge. On the other hand, information about local environment, θ_i, is available only to the manager of the corresponding divi-

[3] Including 0 and 1 in the domain of δ_i would allow for some artifacts – for example, with $\delta_1 = \delta_2 = 0$ we would have coordination as the only issue, so our game would degenerate into some kind of continuous coordination game, while with $\delta_1 = \delta_2 = 1$ only adaptation matters, so there is actually no sense in communication. Mixed cases with $\delta_i = 1$, $\delta_j = 0$ are also possible, but are of no interest either. Therefore, we restrict our analysis to interior values of δ's.

sion and is unknown to another manager – that is exactly why it is *local* information. However, there is a possibility for communication: before the round of decision-making one round of communication occurs, i.e. each manager makes an announcement $\tilde{m}_i \in [-s; s]$ about his state. This announcement is heard by the other manager, but there is no opportunity for him to verify the truthfulness of the incoming message. Moreover, we make a reasonable assumption that managers are unable or unwilling to ex-ante commit to strategies that depend on the information received. Finally, this announcement has no direct effect on payoffs. These three characteristics of messages (non-verifiable, non-binding and payoff-irrelevant) imply that communication takes form of cheap talk (Crawford and Sobel, 1982).

The timing of the game is as follows:

1. Nature makes a random draw of $\{\theta_i\}_{i=1}^2$, where $\theta_i \sim i.i.d.\, U[-s; s]$, $\forall i \in \{1; 2\}$

2. Each manager i learns his θ_i

3. Managers simultaneously or sequentially make their announcements $\tilde{m}_i$

4. After each manager i learns the announcement made by other agent, decisions d_i are made

5. Payoffs are realized

4 Solving for equilibrium

In this paper we use the concept of Perfect Bayesian equilibrium, which implies that communication rules are optimal given decision rules, the latter are optimal given beliefs, and beliefs are calculated by Bayes rule whenever possible. We solve the model by backward induction.

On the last step, both agents possess some beliefs about another agent's state (expressed by some posterior belief distribution $F_i(\theta_j|\tilde{m}_j)$), and make optimal decisions given these beliefs:

$$\begin{aligned}\mathbb{E}_i L_i &= \mathbb{E}_i \left[r_i \cdot \left(\delta_i \left(d_i - \theta_i \right)^2 + \left(1 - \delta_i \right) \left(d_i - d_j \right)^2 \right) | \tilde{m}_j \right] \to \max_{d_i \in [-s_i; s_i]} \qquad (3) \\ &= r_i \delta_i \left(d_i - \theta_i \right)^2 + r_i \left(1 - \delta_i \right) \mathbb{E}_i \left[\left(d_i - d_j \right)^2 | \tilde{m}_j \right] \to \max_{d_i \in [-s_i; s_i]}\end{aligned}$$

This is a simple maximization problem. Assuming interior solution, first order condition w.r.t. d_i looks as follows:

$$\begin{aligned}2r_i \left[\delta_i \left(d_i - \theta_i \right) + \left(1 - \delta_i \right) \left(d_i - \mathbb{E}_i \left(d_j | \tilde{m}_j \right) \right) \right] &= 0 \\ \Leftrightarrow d_i = \delta_i \theta_i + \left(1 - \delta_i \right) \mathbb{E}_i \left(d_j | \tilde{m}_j \right) & \qquad (4)\end{aligned}$$

We see that optimal decision rule is a convex (since $\delta_i \in (0; 1)$) combination of $\theta_i \in [-s; s]$ and $\mathbb{E}_i \left(d_j | \tilde{m}_j \right)$, where d_j also belongs to $[-s; s]$. Consequently, d_i will belong to $[-s; s]$ as well, so the solution is indeed interior.

The right-hand side of expression (4) involves expected decision of another agent. In order to find it let us take similar f.o.c. for agent j and take its expectation:

$$\mathbb{E}_i \left(d_j | \tilde{m}_j \right) = \delta_j \mathbb{E}_i \left(\theta_j | \tilde{m}_j \right) + \left(1 - \delta_j \right) \mathbb{E}_i \left(\mathbb{E}_j \left(d_i | \tilde{m}_i \right) \right) \qquad (5)$$

Note that $\mathbb{E}_i \left(\mathbb{E}_j \left(d_i | \tilde{m}_i \right) \right) = \mathbb{E}_j \left(d_i | \tilde{m}_i \right)$, because agent i can perfectly predict j's reaction to his message $\tilde{m}_i$. Also, denote:

$$m_j := \mathbb{E}_i \left(\theta_j | \tilde{m}_j \right) \qquad (6)$$

as the "effective message" of agent j or belief about θ_j instilled by j in i[4]. Then

[4]Note that due to quadratic loss functions, $\mathbb{E}_i \left(\theta_j | \tilde{m}_j \right)$ is the only information that is used from the whole posterior belief distribution $F_i \left(\theta_j | \tilde{m}_j \right)$, which simplifies our lives a lot.

substituting $\mathbb{E}_j(d_i|m_i)$ from similar f.o.c. for i into the previous expression yields us

$$\begin{aligned}\mathbb{E}_i(d_j|\tilde{m}_j) &= \delta_j m_j + (1-\delta_j)\left[\delta_i m_i + (1-\delta_i)\,\mathbb{E}_i(d_j|\tilde{m}_j)\right] \\ \Leftrightarrow \mathbb{E}_i(d_j|\tilde{m}_j) &= \frac{\delta_j m_j + (1-\delta_j)\,\delta_i m_i}{\delta_i + \delta_j - \delta_i\delta_j} \end{aligned} \tag{7}$$

And so optimal decision rules are given by

$$d_i = \delta_i\theta_i + \frac{1-\delta_i}{\delta_i+\delta_j-\delta_i\delta_j}\left(\delta_j m_j + (1-\delta_j)\,\delta_i m_i\right) \tag{8}$$

Therefore, we see that optimal decision rule is a convex combination[5] of own state, belief about other agent's state and the belief that i has instilled in other agent concerning i's state[6]. These three elements represent three effects as identified by Rantakari. The first is *direct adaptation effect* – decision should be close to own state simply because of losses from misadaptation. As stated by Rantakari, "it measures how much the manager would respond to his information absent any accommodation by the other division". The second term, m_j, represents *coordination effect* – I (as agent i) know that your (j's) decision should by similar logic be close to your state (which I expect to be m_j), and I would like my decision to be close to yours because of the need for coordination, so I adjust my decision towards m_j.

The third term, m_i, stands for *induced adaptation effect.* Intuitively, m_i is the belief that I have instilled in you about my state. While discussing the two previous effects I assumed that adaptation was your only goal, so you would make a decision around your θ_j, and since I want my decision to be between my state θ_i and your decision d_j, I took it as a weighted average of θ_i and m_j. But because of the coordination effect described above, you are also willing to shift your decision d_j towards m_i, your belief about my state, and this action of yours mitigates my coordination problem, thus allowing me to increase the amount of adaptation.

[5]One can easily see that given $\delta_i, \delta_j \in (0;1)$, all three weights (at θ_i, m_i & m_j) also lie in $(0;1)$.

[6]Since all these three elements belong to $[-s;s]$, their convex combination also lies in $[-s;s]$, which is one more evidence in favor of the solution being interior.

Having found the optimal decision rules, let us now turn to communication strategies. We see from equations (8) and (6) that decisions depend only on own states and both expected states given messages. Let us first show that no perfect revelation equilibrium is possible.

Proposition 1. *No perfect revelation equilibrium is possible.*

Proof. Assume there is. Then $\forall \theta_i \; \exists \tilde{m}_i$ s.t. $m_i \left(= \mathbb{E}_j \left(\theta_i | \tilde{m}_i\right)\right) = \theta_i$. Therefore, agent i is able to instill any possible belief in agent j (and vice versa), so we may solve a maximization problem w.r.t. m_i and see what message agent i would like to send in such a situation. Substituting optimal decision rules (8) into the loss function (2) yields the following problem:

$$\begin{aligned} L_i \;\; &= \;\; r_i \left(1-\delta_i\right) \mathbb{E}_i \left[\delta_i \left(1-\delta_i\right) \left(-\theta_i + \frac{\delta_j m_j + \left(1-\delta_j\right) \delta_i m_i}{\delta_i + \delta_j - \delta_i \delta_j} \right)^2 + \right. \\ &\left. + \left(\delta_i \theta_i - \delta_j \theta_j + \frac{m_j \delta_j^2 \left(1-\delta_i\right) - m_i \delta_i^2 \left(1-\delta_j\right)}{\delta_i + \delta_j - \delta_i \delta_j} \right)^2 \right] \to \min_{m_i} \end{aligned} \tag{9}$$

From f.o.c. of this problem (recalling that prior beliefs are $\mathbb{E}_i \theta_j = \mathbb{E}_i m_j = 0$) we obtain the desired belief $m_i = \theta_i \frac{\delta_i + \delta_j - \delta_i \delta_j}{\delta_i - \delta_i \delta_j} \neq \theta_i$, therefore, it is always profitable for agent i to deviate from truth-telling.[7] □

Therefore, we resort to Crawford & Sobel's classic "interval equilibria", in which the whole state/message space is partitioned into intervals, and message $\tilde{m}_i$ indicates that state θ_i belongs to one of these intervals, but does not disclose its exact location. In particular, denote this partition as $\{a_{ik}\}_{k=0}^{N_i}$, where $-s = a_{i0} < a_{i1} < ... < a_{iN_i} = s$. Then message $\tilde{m}_i \in [a_{ik}; a_{i,k+1}]$ is supposed to mean that $\theta_i \in [a_{ik}; a_{i,k+1}]$ and give no further information, so $m_i = \mathbb{E}_j \left(\theta_i | \tilde{m}_i\right) = \frac{a_{ik} + a_{i,k+1}}{2}$. Let us now find this

[7] Crucial here is the assumption that zero and one are not in the domain of delta. In fact, if $\delta_i = 1$ then agent i does not care about coordination and, as a consequence, about communication, so he does not actually have any incentives to misreport his information, and perfect revelation is possible.

partition. It is done from the indifference condition: if $\theta_i = a_{ik}$, then agent i should be indifferent between sending messages $m_i = \frac{a_{ik}+a_{i,k+1}}{2}$ and $m_i = \frac{a_{i,k-1}+a_{ik}}{2}$:

$$\mathbb{E}_i\left(L_i|\theta_i = a_k, m_i = \frac{a_{ik} + a_{i,k+1}}{2}, m_j\right) = \mathbb{E}_i\left(L_i|\theta_i = a_k, m_i = \frac{a_{i,k-1} + a_{ik}}{2}, m_j\right) \tag{10}$$

where m_j refers to i's belief about θ_j *at the moment of choosing* m_i, i.e. $m_j = \mathbb{E}_i\left(\theta_j|\tilde{m}_j\right)$ in case i has already received message $\tilde{m}_j$ (which means that he is a "follower" in a sequential communication game) and $m_j = \mathbb{E}_i\left(\theta_j|\emptyset\right) = \mathbb{E}_i\left(\theta_j\right) = 0$ in case he has not (valid for all other cases – "leader" in a sequential communication game and both agents in simultaneous communication case). Condition (10) after some transformations turns into a second-order recurrent equation:

$$|a_{i,k+1} - a_{i,k}| - |a_{i,k} - a_{i,k-1}| = \frac{4\delta_j}{\left(1-\delta_j\right)\delta_i}\left(a_{i,k} - m_j\right) \tag{11}$$

Solving this equation with initial conditions $a_{i,0} = -s$; $a_{i,N_i} = s$ yields us the sequence $\{a_{ik}\}_{k=0}^{N_i}$, which characterizes the equilibrium communication scheme with N_i communication intervals:

$$a_{i,k} = m_j + \frac{\left(s-m_j\right)\left(x_i^{2k} - x_i^{-2k}\right) - \left(s+m_j\right)\left(x_i^{2(N_i-k)} - x_i^{-2(N_i-k)}\right)}{x_i^{2N_i} - x_i^{-2N_i}} \tag{12}$$

where $x_i = \left(\sqrt{1+\frac{\delta_j}{(1-\delta_j)\delta_i}} + \sqrt{\frac{\delta_j}{(1-\delta_j)\delta_i}}\right)^2 > 1$. Some of our most sagacious readers might infer out of expression (12) that intervals are pretty wide on the edges of the state space (close to s or $-s$, i.e. when $k \to N_i$ or $k \to 0$), while they are much more densely concentrated around m_j. Another interesting point is that unlike in Crawford & Sobel's model, here we do not have an upper bound on N_i, so there would be infinitely many equilibria on communication stage, one for each pair $(N_1; N_2) \in \mathbb{N} \times \mathbb{N}$. Let us then characterize the full set of equilibria[8]:

[8] Characterization for the case of simultaneous communication is completely analogous (except for differences in notation) to the one obtained by ADM and Rantakari. On the other hand, characterization of communication equilibria for the case of sequential communication is a novel

Proposition 2. *For any communication mechanism (sequential or simultaneous) there exists an infinite number of equilibria, each of which can be characterized by a pair of positive integers* $(N_1; N_2) \in \mathbb{N}^2$*, such that* $\forall i = \{1; 2\}$*,* $j = 3 - i$*:*

- *Optimal decision rules* $d_i(\theta_i, m_i, m_j)$ *are given by equation (8);*
- *Communication rules are as follows: any message* $\tilde{m}_i \in [a_{i,k-1}; a_{i,k}]$ *is sent with equal non-zero probability density* $\frac{1}{a_{i,k}-a_{i,k-1}}$ *if* $\theta_i \in [a_{i,k-1}; a_{i,k}]$ *and sent with probability zero otherwise,* $\forall k = \overline{1, N_i}$*;*
- *Beliefs* $F_i(\theta_j|\tilde{m}_j)$ *specify equal non-zero probability density* $\frac{1}{a_{j,k}-a_{j,k-1}}$ *to any state* $\theta_j \in [a_{j,k-1}; a_{j,k}]$ *if message* $\tilde{m}_j \in [a_{j,k-1}; a_{j,k}]$ *and probability zero otherwise, for* $\forall k = \overline{1, N_j}$*;*
- *Partitions* $\{a_{i,k}\}_{k=1}^{N_i}$ *are defined as in equation (12).*

Proof. In the text above. □

In fact, for each pair of numbers of intervals $(N_1; N_2) \in \mathbb{N}^2$ there exist many equilibria with different communication rules (different from the ones described in the second point of Proposition 2). For example, one other equilibrium can be characterized by $\tilde{m}_i = \frac{a_{i-1,k}+a_{i,k}}{2} (= m_i)$ if $\theta_i \in [a_{i,k-1}; a_{i,k}]$ $\forall k = \overline{1, N_i}$. There exists a variety of other equilibria with different message domains, but in case we wanted to use them an issue of out-of-equilibrium beliefs would arise, since not all messages would have been sent in equilibrium. Communication rule from Proposition 2 is the simplest among those that exploit the whole message space (but not unique – messages do not necessarily need to be distributed *uniformly* within one interval), which allows to avoid out-of-equilibrium problems – so we focus on this equilibrium in further analysis, since considering the whole family of communication equilibria would mean nothing in terms of results, but would add plenty of headache in terms of formal analysis.

result.

In fact, Proposition 2 allows us to make one interesting conclusion concerning the relation between equilibrium strategies in sequential and simultaneous games:

Corollary 1. *From Proposition 2 it can be inferred that leader's behaviour in sequential communication game is completely analogous to the one in simultaneous communication game. Behaviour of the follower in sequential game is described by the same decision rule, but different communication rule.*

Therefore, we may actually make a conclusion that the leader (agent who speaks first), while having an opportunity to affect follower's communication strategy, does not exploit it and behaves as if the follower did not base his messages on the one received from the leader. This has an easy explanation behind it: leader's messages affect the partition used by the follower, but do not change follower's expected message: it is easy to see that $\mathbb{E}(m_f) = 0$. And we have already argued that because of quadratic cost assumption, only average values actually matter in our problem – thus leader's messages do not affect follower's communication strategy in such a way which would affect leader's expected payoff, and leader has no incentives to change his behaviour in any way due to this "leadership".

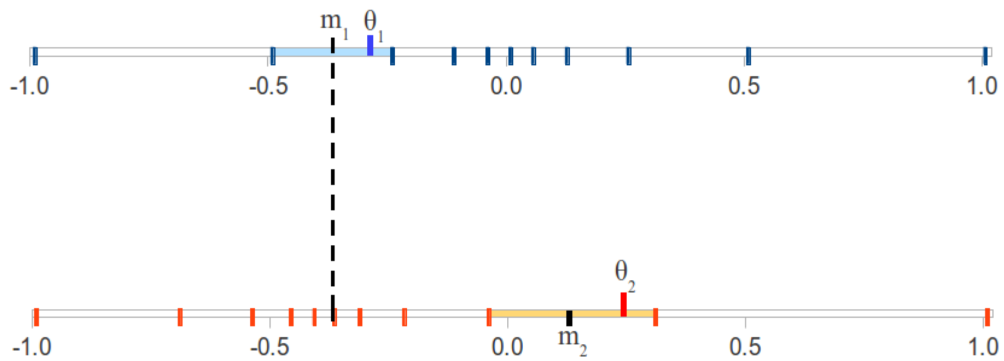


Figure 1: Illustration of sequential communication equilibrium

5 Properties of equilibria

Figure 1 gives an example of equilibrium on communication stage for the case of sequential communication. The upper line represents state/message space of agent 1, who also is the first to send his message. Since he has no non-trivial information about θ_2 (agent 2 has not yet sent his message), agent 1's belief is $m_2 = 0$, so the intervals are most densely located around zero, and the partition is symmetric around zero. In case of simultaneous communication, agent 2 would have a similar partition (meaning concentrated around zero; it would be completely the same only if $\delta_i = \delta_j$ and $N_i = N_j$). However, in our case agent 2 would already have some information to update his beliefs: in particular, he would have already received message $\tilde{m}_1$ and thus obtained knowledge that θ_1 belongs to the blue interval (on the upper line), so his communication intervals would now be concentrated not around zero, but around this new posterior belief m_1 (on the lower line).

Why does this concentration of intervals happen? It originates from the type-dependence of agent's bias which is implicitly present in the model: if an agent expects his local conditions to coincide with those of his counterpart, then there would be no conflict of interests – both agents would be able to fully adapt to their respective conditions and fully coordinate at the same time. Such alignment of interests as perceived by agent j happens if $\theta_j = \mathbb{E}_j\left(\theta_i|\tilde{m}_i\right) = m_i$ for (if j is) the follower in sequential game and if $\theta_j = \mathbb{E}_j\left(\theta_i\right) = 0$ for the leader in sequential and both agents in simultaneous game. We can see that these are points of the highest concentration of intervals – intervals are extremely small around these points, so noise in communication is minimal, agent is willing to communicate truthfully or almost truthfully in these cases. On the other hand, the farther is agent's state θ from this point of no conflict, the more serious is the conflict of interests as perceived by him, so stronger are the incentives for him to misreport the information he possesses. Therefore, he has to introduce noise into his messages, and this noise gets larger – intervals get wider, – the farther agent's state is from the point of no conflict.

In our example we can see that given the same number of intervals, the partition for agent 2 (on the lower line of Figure 1) is more informative about low states (left on the axis) than partition of agent 1, because intervals are concentrated below zero for agent 2; the opposite can be said about high (positive) θ's – agent 1's partition performs much better transmitting information about these. So in the end, which partition of these would be more informative on average? To answer this question we need to introduce some measure of informativeness of equilibrium.

Denote $\sigma_i := \mathbb{V}(m_i) = \mathbb{E}(m_i^2) - \mathbb{E}(m_i)^2$ and call it "message variance". Note that $\mathbb{E}(m_i) = \frac{\int_{-s}^{s} m_i d\theta_i}{\int_{-s}^{s} d\theta_i} = \frac{1}{2s}\sum_{k=1}^{N_i}\int_{a_{i,k-1}}^{a_{i,k}} \frac{a_{i,k}+a_{i,k-1}}{2} d\theta_i = \frac{1}{4s}\sum_{k=1}^{N_i}\left(a_{i,k}^2 - a_{i,k-1}^2\right) =$ $= \frac{1}{4s}\left(-a_{i,0}^2 + a_{i,1}^2 - a_{i,1}^2 + a_{i,2}^2 + ... + a_{i,N_i}^2\right) = \frac{1}{4s}\left(-a_{i,0}^2 + a_{i,N_i}^2\right) = \frac{1}{4s}\left(-s^2 + s^2\right) = 0$. Therefore, $\sigma_i = \mathbb{E}(m_i^2)$. Moreover, referring to Lemma 2 from ADM, we may also say that $\mathbb{E}(m_i^2) = \mathbb{E}(m_i\theta_i)$.

Consider a measure of residual variance: $\mathbb{E}\left[(\theta_i - \mathbb{E}(\theta_i|\tilde{m}_i))^2\right] = \mathbb{E}\left[(\theta_i - m_i)^2\right]$. It characterizes the share of variation of θ_i that is not described by messages m_i. $\mathbb{E}\left[(\theta_i - m_i)^2\right] = \mathbb{E}\left[\theta_i^2 - 2\theta_i m_i + m_i^2\right] = \mathbb{E}(\theta_i^2) - 2\sigma_i + \sigma_i = \frac{s^2}{3} - \sigma_i$. We can see that residual variance is a linear function of σ_i, therefore σ_i may as well be used as a measure of informativeness. Informative equilibrium implies low residual variance (so that smaller share of variance of θ_i remains unexplained), which means higher σ_i. Therefore, σ_i is positively associated with informativeness of equilibria.

Lemma 1. *Message variance σ_i and, thus, informativeness of equilibrium (ceteris paribus) increases with N_i, and decreases with $|m_j|$.*

Proof. First of all, let us find the closed-form expression for σ_i:

$$\begin{aligned}
\sigma_i(m_j) &= \mathbb{E}(m_i^2|m_j) = \frac{1}{2s}\int_{-s}^{s}\left(\frac{a_{i,k-1}+a_{i,k}}{2}\right)^2 d\theta_i \\
&= \frac{1}{2s}\sum_{k=1}^{N_i}\int_{a_{i,k-1}}^{a_{i,k}}\left(\frac{a_{i,k-1}+a_{i,k}}{2}\right)^2 d\theta_i \\
&= \frac{1}{8s}\sum_{k=1}^{N_i}\left(a_{i,k-1}+a_{i,k}\right)^2\left(a_{i,k}-a_{i,k-1}\right)
\end{aligned}$$

Substituting the expressions for $a_{i,k}$ and taking sums, after some very lengthy calculations, any step of which is too long to be presented here, one can obtain $\sigma_i(m_j) = \frac{s^2}{4}\left(\frac{x_i+1}{x_i^{N_i}-1}\right)^2 \left(\frac{x_i^{2N_i}+x_i^{N_i}+1}{x_i^2+x_i+1} - \frac{x_i^{N_i}}{x_i}\right) - \frac{m_j^2}{4}\left(\frac{x_i-1}{x_i^{N_i}+1}\right)^2 \frac{\left(x_i^{N_i-1}-1\right)\left(x_i^{N_i+1}-1\right)}{x_i^2+x_i+1}$. One can see that $\sigma_i(m_j) = \gamma_1 - \gamma_2 m_j^2$ for some positive γ_2 (its positivity follows directly from the fact that $x_i > 1$). Therefore, σ_i reaches its maximum at $m_j = 0$ and it is diminishing in $|m_j|$ (or, equivalently, in m_j^2).

In order to obtain comparative statics signs for other parameters, let us introduce a function $f(p)$ such that $f\left(x_i^{N_i}\right) = \sigma_i(m_j)$: $f(p) = \frac{s^2}{4}\left(\frac{x_i+1}{p-1}\right)^2\left(\frac{p^2+p+1}{x_i^2+x_i+1} - \frac{p}{x_i}\right) - \frac{m_j^2}{4}\left(\frac{x_i-1}{p+1}\right)^2 \frac{(p-x)\left(p-\frac{1}{x}\right)}{x_i^2+x_i+1}$. One may verify that $f(p)$ is an increasing function of p for $p > 1$: $f'(p) = \frac{\left(x_i^2-1\right)^2}{4x_i\left(x_i^2+x_i+1\right)} \frac{s^2(p+1)^4 - m_j^2(p-1)^4}{(p^2-1)^3}$ is positive if $p > 1$ (since $m_j^2 \leq s^2$). We already know that $x_i > 1$ and $N_i \geq 1$, so in our problem it is always the case that $p > 1$. Therefore, it can be said that σ_i is increasing in $x_i^{N_i}$. Since N_i does not enter the expression for x_i, its only effect on $x_i^{N_i}$ is through the power, and since $x_i > 1$, $x_i^{N_i}$ is increasing in N_i. $\square$

Corollary 2. *Follower's messages are (ceteris paribus) on average less informative than those of the leader or of a player in simultaneous communication game.*

Proof. Recall the definition of m_i: we said that it is the belief about θ_i at the moment of making a communication decision by j. Therefore, $m_i = 0$ for both players in simultaneous communication game, and $m_i = 0$ for the leader (first-mover) in sequential communication case. On the other hand, $m_i \in (-s; s)$ for the follower in sequential communication, so on average $|m_j|$ will be above zero, which by Lemma 1 means that ex ante expected σ_i of the follower would be lower compared to the leader and to the simultaneous communication case. $\square$

One explanation for this (corollary) is geometric: message variance is basically an average of squared distances from each possible point in the state/message space to the center of the corresponding interval – i.e. average squared distance from message $\tilde{m}$ to information m contained in it. We see from Figure 1 that given the same number of intervals, follower's communication strategy will involve very wide intervals on one side of state/message space, so some of mentioned squared distances will be very high on this interval, thus driving an increase in message variance.

More intuitive rephrasing of the previous paragraph involves the argument about follower's incentive compatibility that we had while discussing why the communication intervals are concentrated around a certain point. In particular, we have found out that intervals should be wider the further an agent's state (θ_j) is from his expectation of other agent's state ($\mathbb{E}_j(\theta_i|...)$, which is 0 or m_i depending on j's role in the game). For the leader in sequential game (and both players in simultaneous) this expectation is zero, and both "radical" states, meaning $\theta_j = -s$ and $\theta_j = s$ are equidistant from this expectation (distance is, obviously, equal to s). For the follower maximal distance from this expectation to his state is $s + |m_i|$[9] – this distance is larger than for the leader, so his message in case of such extreme state should be more vague. In fact, this vagueness outweighs higher precision in messages in case of states on the other extreme (on the other edge of the domain) and drives an increase of message variance, meaning that on average follower's messages would be less informative.

Overall, as we have already discussed in Corollary 1, this difference in informativeness is the only difference between the two schemes: leader in sequential communication game behaves exactly as both players do in simultaneous communication game, while the follower has to make adjustments to his communication strategy because of the additional information that he possesses.

But how do agents' payoffs depend on the informativeness of messages in equilibrium (which is measured by sigmas)? If we expand expression (9) and take ex-ante expectation (before θs are realized) of this loss function, we would obtain the fol-

[9] Suppose $m_i < 0$. Then the state which is most distant from this point would be $\theta_j = s$, and the distance would be $s - m_i = s + |m_i|$. Similarly for $m_i > 0$.

lowing:

$$\mathbb{E}L_i = r_i(1-\delta_i)\left(\left(\delta_i+\delta_j^2\right)\frac{s^2}{3} - \sigma_i\frac{\delta_i^2(1-\delta_j)\left[\delta_i(1-\delta_j)+2\delta_j\right]}{\left(\delta_i+\delta_j-\delta_i\delta_j\right)^2} + \right. \\ \left. +\sigma_j\frac{\delta_j^2(1-\delta_i)\left[\delta_i(1-\delta_j)^2-\delta_j^2\right]}{\left(\delta_i+\delta_j-\delta_i\delta_j\right)^2}\right) \tag{13}$$

After taking a look at this expression we can make the following statements:

Lemma 2. *(a)* $\frac{\partial \mathbb{E}L_i}{\partial \sigma_i} < 0$ *for all parameter values, so division profits are positively associated with the informativeness of the message sent.*

(b) $\frac{\partial \mathbb{E}L_i}{\partial \sigma_j} < 0$ *iff* $\delta_i < \left(\frac{\delta_j}{1-\delta_j}\right)^2$ *(or, equivalently, iff* $\delta_j > \frac{\sqrt{\delta_i}}{1+\sqrt{\delta_i}}$*). Therefore, division profits are positively associated with the informativeness of the incoming message if and only if this division is dependent enough (or vice versa – if another division is independent enough).*

A discussion of this Lemma follows up to the end of the current section.

We can see that $\frac{\partial \mathbb{E}L_i}{\partial \sigma_i} = -r_i(1-\delta_i)\frac{\delta_i^2(1-\delta_j)[\delta_i(1-\delta_j)+2\delta_j]}{(\delta_i+\delta_j-\delta_i\delta_j)^2} < 0$, so expected losses decrease in variance of own message. Therefore, it is beneficial to send the most informative message possible as long as it is incentive compatible. The reason for this is that more informative message allows your counterpart to make more precise inferences about your state (and, thus, decision) and accommodate with it better, which on average reduces your coordination losses. However, as we have seen in Proposition 1, you cannot send an absolutely precise message which would also be truthful, so in the end you would like to transmit as much information as possible, while the message is vague enough for you to stay honest.

Since all our equilibria are incentive compatible by definition, agents would prefer the most informative out of them – the most desirable for agent i would be the one with the highest σ_i. Consequently, since each agent is able to choose among equilibria by varying his N_i, we may choose some "focal" equilibrium – this will be

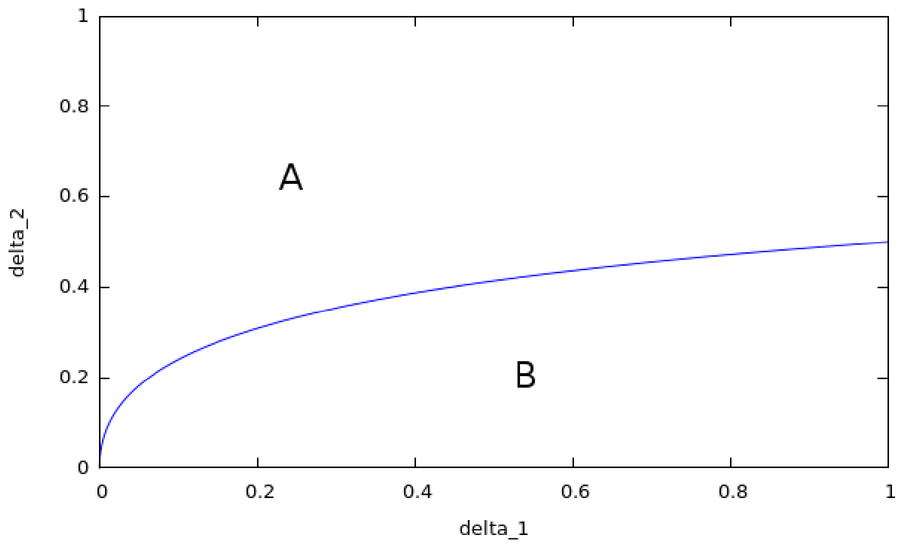


Figure 2: $\frac{\partial \mathbb{E} L_1}{\partial \sigma_2} = 0$ line

the one with $N_i \to \infty \ \forall i \in \{1;2\}$. ADM show in their Proposition 2 that such "limit" equilibrium exists. In fact, if we include N_i in the strategy of player i (thus making it a choice variable), such limit equilibrium would become the unique (up to irrelevant transformations of communication rule, as discussed at the end of Section 4) equilibrium of the game[10].

As for dependence on other agent's σ, it is ambiguous and may go both ways: $\frac{\partial \mathbb{E} L_i}{\partial \sigma_j} = r_i (1-\delta_i) \frac{\delta_j^2 (1-\delta_i)\left[\delta_i (1-\delta_j)^2 - \delta_j^2\right]}{(\delta_i + \delta_j - \delta_i \delta_j)^2} \gtrless 0$. Specifically, $\frac{\partial \mathbb{E} L_i}{\partial \sigma_j} > 0 \Leftrightarrow \delta_i > \left(\frac{\delta_j}{1-\delta_j}\right)^2$. Graphically, this looks as in Figure 2: if parameters $(\delta_1; \delta_2)$ fall into the upper region (denoted as A), then $\frac{\partial \mathbb{E} L_1}{\partial \sigma_2} < 0$, i.e. profits of agent 1 increase in σ_2. Vice versa, if vector of deltas falls into region B, then $\mathbb{E} L_1$ depends positively on σ_2.

Obviously, an explanation of such non-monotonicity is required – it is not really straightforward why in some cases (for some values of parameters) I want to receive as much information from you as possible, and in others – vice versa, I do not really want you to transmit any information. An educated guess would be that (at least) two effects are present here, and each of them dominates for some parameter values – and this guess would be correct. For the sake of clarity let us name our two

[10]However, the issue is that players would have to precommit to their respective Ns *before* the game itself and before they learn their θs – it appears to us that this might be troublesome.

managers as Alice and Bob (and denote them using subscripts A and B respectively; subscripts i and j are applicable to any of them, but within one formula we always imply $i \neq j$).

So suppose that we want to examine the effect that the informativeness of Bob's message has on Alice's profits. From Section 4 we know that Bob's message affects decisions of both players: it enters Alice's decision $d_A(\theta_A, m_B, m_A)$ in coordination term and Bob's decision $d_B(\theta_B, m_A, m_B)$ in induced adaptation term (consult Section 4 in case you need to refresh the knowledge of which effects does agent's optimal decision rule accommodate). Also, recall that Alice's loss (or profit) function is composed of adaptation loss term $(d_A - \theta_A)^2$ and coordination loss term $(d_A - d_B)^2$. Expanding loss function and collecting terms containing σ_B yields us the following expression:

$$\mathbb{E}L_A = ... + \sigma_B \cdot \left(\underbrace{\frac{\delta_B^2 \delta_A (1 - \delta_A)}{(\delta_A + \delta_B - \delta_A \delta_B)^2}}_{\text{adaptation}} \underbrace{- \frac{(1 - \delta_A) \delta_B^3 (2\delta_A + \delta_B - \delta_A \delta_B)}{(\delta_A + \delta_B - \delta_A \delta_B)^2}}_{\text{coordination}} \right) \tag{14}$$

It tells us that informativeness of Bob's messages is positively associated with Alice's adaptation losses and negatively – with her coordination losses.

The only element in adaptation term of the loss function that depends on m_B is d_A. Therefore, Bob's message affects Alice's adaptation losses only through the coordination term in Alice's decision. We have discussed earlier that better precision of incoming message allows for better coordination – in this case Alice would be able to coordinate better with Bob's decision because she has more precise information concerning this decision. However, this coordination comes at cost of adaptation. In order to understand this cost, note that a more informative communication rule $m_B(\theta_B)$ is in fact a mean-preserving spread of a less informative communication rule. Let us denote the situation of less informative communication using superscript

L, and the one with more informative communication with superscript M – then $m_B^M \geqslant_{MPS} m_B^L$ (meaning m_B^M is a mean-preserving spread of m_B^L). Also, since Alice's decision d_A is a linear combination of θ_A, m_A and m_B, and the first two elements stay the same during such transition (from less to more informative communication on Bob's behalf), then we can also say that $d_A^M \geqslant_{MPS} d_A^L$, i.e. more information applies a mean-preserving spread on Alice's decision meaning that this decision becomes more volatile (for fixed θ_A). A straightforward conclusion from this fact is given by Lemma 3.

Lemma 3. *If $d_A^M \geqslant_{MPS} d_A^L$ for any given θ_A then $\mathbb{E}\left(d_A^M - \theta_A\right)^2 > \mathbb{E}\left(d_A^L - \theta_A\right)^2$.*

Proof. Expected adaptation losses are computed in the following way: $\mathbb{E}\left(d_A - \theta_A\right)^2 = \int_{-s}^{s}\left[\int_{-s}^{s}\left(d_A\left(\theta_A, m_B\left(\tilde{m}_B\right), m_A\left(\theta_A\right)\right) - \theta_A\right)^2 d\tilde{m}_B\right] d\theta_A$ where we can write the inner integral like that because, as a matter of fact, we defined communication equilibria in such a way that any message $\tilde{m}_B$ is sent with equal probability in any of them, and what communication rule actually defines is other player's belief rule $m_B\left(\tilde{m}_B\right)$. Then we can rewrite the integral as follows:

$$\begin{aligned}\mathbb{E}\left(d_A - \theta_A\right)^2 &= \int_{-s}^{s}\left[\int_{-s}^{s}\left(d_A^2\left(\theta_A, \tilde{m}_B\right) - 2\theta_A d_A\left(\theta_A, \tilde{m}_B\right) + \theta_A^2\right) d\tilde{m}_B\right] d\theta_A = \\ &= \int_{-s}^{s}\left[2s\theta_A^2 - 2\theta_A \mathbb{E}_{\tilde{m}_B}\left(d_A\right) + \mathbb{E}_{\tilde{m}_B}\left(d_A^2\right)\right] d\theta_A \end{aligned} \tag{15}$$

and given that $d_A^M \geqslant_{MPS} d_A^L$ (for given θ_A), we can say that in the previous expression $\mathbb{E}_{\tilde{m}_B}\left(d_A^M\right) = \mathbb{E}_{\tilde{m}_B}\left(d_A^L\right)$, while $\mathbb{E}_{\tilde{m}_B}\left(d_A^{M^2}\right) > \mathbb{E}_{\tilde{m}_B}\left(d_A^{L^2}\right)$ (because $\mathbb{E}\left(d_A^M\right) = \mathbb{E}\left(d_A^L\right)$ and $\mathbb{V}\left(d_A^M\right) > \mathbb{V}\left(d_A^L\right)$ follow trivially from the definition of a mean-preserving spread). Therefore, the integrand in equation (15) will be point-wise larger for d_A^M than for d_A^L, which means that the whole integral in equation (15) will also be larger, so $\mathbb{E}\left(d_A^M - \theta_A\right)^2 > \mathbb{E}\left(d_A^L - \theta_A\right)^2$. □

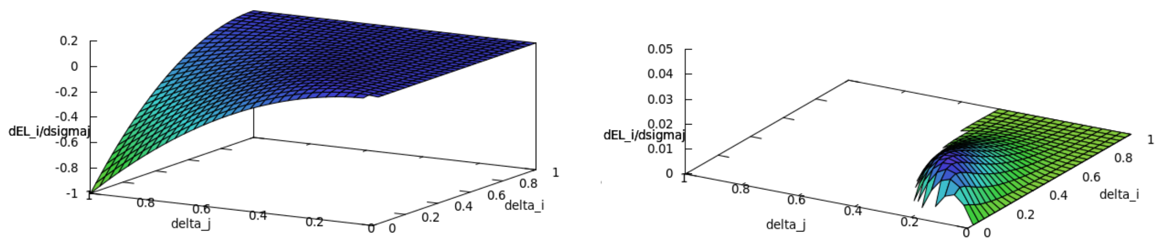
The conclusion is that more precise incoming message increases adaptation losses.

This conclusion is confirmed by equation (14) – one can see that the coefficient at σ_B coming from adaptation loss term is positive.

Now let us turn to the coordination loss term in the loss function. Bob's message m_B enters twice here: as a coordination term in Alice's decision d_A and induced adaptation term in Bob's decision d_B. It seems straightforward that if only the first of these channels was active, then Alice's coordination losses would decrease, since better message would allow her to coordinate better with Bob. However, Bob is aware of this improvement in coordination on Alice's behalf, and is thus able to exercise more induced adaptation – so his decision will now on average be closer to his state. What is more important, his decision will become more volatile and more difficult to coordinate with, which partly negates Alice's coordination effort. In other words, Alice is still able to improve on coordination, but because of Bob's shift in behaviour this improvement is not as large as it could be. In particular, we can see from expression 14 that expected coordination losses are indeed decreasing with σ_B (informativeness of Bob's message).

Summing up, higher precision of incoming message has two effects in terms of agents' actions, which combine to produce two effects on total expected losses (each of which may dominate depending on the parameter values). On the one hand, an improvement in precision of Bob's message allows Alice to coordinate with Bob better. However, this improvement in coordination comes at cost of adaptation, so Alice's average adaptation losses increase. On the other hand, Bob, being aware of Alice's improved coordination, may now give more weight to adaptation, which to some degree negates Alice's coordination effort, and thus Alice's coordination costs decrease, but not as much as desired.

Finally, one more claim that needs to be made is that even when profits are negatively associated with the precision of incoming message (i.e. $\frac{\partial \mathbb{E} L_i}{\partial \sigma_j} > 0$, so additional adaptation losses dominate benefits gained in terms of coordination), the strength of this association is not that high relative to its magnitudes of opposite



(a) $\frac{\partial \mathbb{E}L_i}{\partial \sigma_j}$ (b) Positive part of $\frac{\partial \mathbb{E}L_i}{\partial \sigma_j}$

Figure 3: Sensitivity of expected losses to precision of incoming message

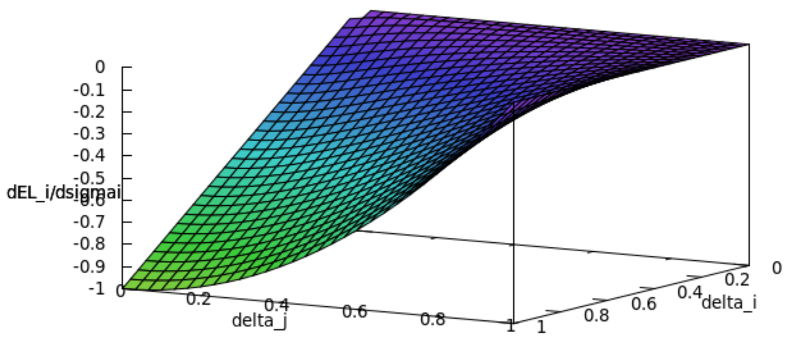


Figure 4: Sensitivity of expected losses to precision of own message – $\frac{\partial \mathbb{E}L_i}{\partial \sigma_i}$

sign (for other pairs of deltas; can be seen on Figure 3) and relative to magnitudes of sensitivity to precision of own message ($\frac{\partial \mathbb{E}L_i}{\partial \sigma_i}$; Figure 4).

6 Simultaneous vs. sequential communication

General result of this section is formalized by Proposition 3:

Proposition 3. *Sequential communication scheme with division 1 as leader is preferred to simultaneous communication scheme:*

(a) by the follower – never; he strictly prefers simultaneous communication;

(b) by the leader – iff $\delta_1 > \left(\frac{\delta_2}{1-\delta_2}\right)^2$*;*

(c) by the firm – iff $\frac{r_1}{r_2}(1-\delta_1)\left(\delta_1\left(1-\delta_2^2\right)-\delta_2^2\right) > (1-\delta_2)\left(\delta_2+2\delta_1-\delta_1\delta_2\right)$.

Proof. In the text below. □

The considerations given in previous subsection allow us to finally compare the two types of communication procedures. Denote the first-mover (leader) in sequential communication game as agent 1; then follower will be agent 2. Then the following list summarizes the necessary facts obtained earlier:

- If we move from simultaneous communication to sequential then σ_1 remains unchanged, while σ_2 decreases;
- $\frac{\partial \mathbb{E} L_i}{\partial \sigma_i} < 0$;
- $\frac{\partial \mathbb{E} L_i}{\partial \sigma_j} > 0$ iff $\delta_i > \left(\frac{\delta_j}{1-\delta_j}\right)^2$.

Combining these three facts, we can make an inference that profits of the follower are lower (losses are higher) in sequential communication case ($\frac{\partial \mathbb{E} L_2}{\partial \sigma_2} < 0$ and σ_2 decreases). As for the leader, his profits increase if $\delta_1 > \left(\frac{\delta_2}{1-\delta_2}\right)^2$ and fall otherwise. Therefore, we see that sequential communication may only be beneficial to the leader, but always harms the follower. In other words, in some cases I may want you to hear my message before you send yours, but you strictly prefer not doing it, since on average that would make your message less precise and harm your profits.

This is a surprising result: we would expect follower to increase his profits because of the informational advantage that he obtains in the sequential communication game, especially in the absence of any changes in behavior of the leader (unlike, for example, in Stackelberg duopoly model, where the leader exploits follower's knowledge in order to manipulate his actions in some sense and obtain higher profit at the cost of the follower). But nonetheless, this informational advantage seems to be playing against its possessor: adjustment of his communication strategy to this new piece of information makes his messages less precise on average, while he wants them to be as precise as possible (as long as it satisfies the incentive compatibility constraint).

But what about total firm profits? When does benefit of one agent (division) outweigh the losses incurred by another? To do this we need to compute the sign of

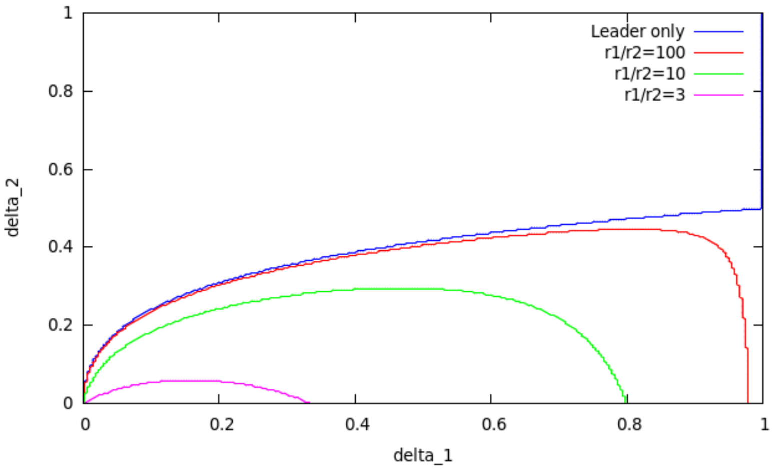


Figure 5: Parameter values and firm profitability

sensitivity of total losses to variance of follower's message.

$$\frac{\partial \mathbb{E} L_\Sigma}{\partial \sigma_2} := \frac{\partial \mathbb{E}(L_1 + L_2)}{\partial \sigma_2} = \frac{\delta_2^2 (1-\delta_1)}{(\delta_1 + \delta_2 - \delta_1\delta_2)^2} \left[r_1 (1-\delta_1)\left(\delta_1\left(1-\delta_2^2\right) - \delta_2^2\right) + \right. \tag{16}$$
$$\left. - r_2 (1-\delta_2)(\delta_2 + 2\delta_1 - \delta_1\delta_2) \right]$$

From equation (16) we can see that $\frac{\partial \mathbb{E} L_\Sigma}{\partial \sigma_2} > 0$ is equivalent to the following condition:

$$\frac{r_1}{r_2}(1-\delta_1)\left(\delta_1\left(1-\delta_2^2\right) - \delta_2^2\right) > (1-\delta_2)(\delta_2 + 2\delta_1 - \delta_1\delta_2) \tag{17}$$

The graphic representation of this condition may be found on Figure 5 for various values of $\frac{r_1}{r_2}$ (relative division sizes). Areas below the corresponding curves represent parameter combinations, for which sequential communication is more profitable on the firm level than simultaneous. It is pretty straightforward that if $\frac{r_1}{r_2} \to \infty$, i.e. the leader's division secures all firm's profits, while follower's division is negligibly small, then condition for sequential communication to be profitable on the firm level coincides with similar condition for division 1 (firm prefers sequential communication whenever the leader does). Then, as $\frac{r_1}{r_2}$ decreases, i.e. follower's department gets relatively larger, the losses it incurs from switching to sequential communication start to matter – we can see that the corresponding sets of parameters (for which sequential communication is more preferable for the firm than simultaneous) get smaller. This set stops intersecting the $[0;1]^2$ parameter space when $\frac{r_1}{r_2}$ falls below 2 – speaking differently, if $r_1 < 2r_2$ then simultaneous communication scheme is always more preferable than sequential communication with division 1 as leader[11].

It would be also of interest to describe the optimal communication order depending on deltas (divisions' dependencies). For the sake of narrative convenience we will look at transition from simultaneous to sequential communication and explore when such transition may deem profitable. Recall that during such transition the only thing that changes is the informativeness of the follower's message – it decreases, to be exact. From the discussion of Lemma 2 we know that two effects arise during such transition:

1. Leader's adaptation losses decrease because his decisions become less volatile.
2. Coordination losses (shared by both players) increase due to worse coordination on leader's behalf (because leader has less knowledge on what to coordinate with), but this effect is mitigated by the follower, who shifts the accents in his decisions from adaptation towards coordination, and whose decisions become less volatile and thus easier to coordinate with.

In order for the transition from simultaneous to sequential communication to be profitable we need the sum of these losses (weighted by division sizes) to decrease in total. When does it happen?

Let us start with δ_2 – follower's dependency. As one can see from Figure 5, δ_2 should be low enough for sequential communication to be more preferable, i.e. the follower should be dependent enough. Indeed, if this is the case, then both players' coordination losses should not increase significantly, because the follower will not allow this to happen – if he predicts that the leader will coordinate worse, then the follower will have to offset this by changing his own behavior in favor of coordination,

[11]It is pretty straightforward that if $2r_1 < r_2$ then another sequential communication scheme may be profitable – with division 2 as leader.

since it is valued highly by him. Therefore, if the follower is highly dependent then coordination losses are not likely to be significant, while adaptation gains will still exist – this is why transition from simultaneous to sequential communication is likely to be profitable in general.

As for δ_1, leader's dependency, from Figure 5 one can see that it has a non-monotonic effect on the optimal order of communication: in particular, sequential communication is more likely to be preferred for average values of δ_1. For low values of δ_1 the leader is very dependent and cares little about adaptation, so the gain in adaptation losses is negligible and is insufficient to cover any additional coordination losses (which are unavoidable because of worse communication). On the other hand, if δ_1 is high then the story is similar, but adaptation gains are small due to a different reason: in this case leader *always* adapts well enough, and he actually cares little about coordination, so his actions are affected weakly by incoming messages. At the same time, his decisions are located extremely far from follower's state (most weight in his decisions is given to own state), so the follower generally faces a pretty serious trade-off, incurring heavy costs independently of his decision (these would be either from miscoordination, or from misadaptation). Therefore, any slight change in leader's behaviour would have a strong effect on the follower. The fact that leader will have a less precise message concerning follower's state θ_f means that his decision will on average be further from this θ_f (because follower's message m_f will on average be further from θ_f), which is bad news for the follower: his losses will increase, and because of what we have previously discussed – they will increase significantly. Combined with not-so-high leader's gains on adaptation, this gives us the reason for why under high δ_1 transition from simultaneous to sequential communication is undesirable either, just like in case of low δ_1 – only average values of δ_1 may render sequential communication scheme profitable.

7 Conclusion

In this paper we have examined the issue of choosing a communication mechanism for a situation of cheap talk with two-sided informational asymmetry. In our story both agents possess some private information, which is not of direct relevance to another agent, but at the same time they face a need to coordinate their decisions, while also adapting them to local conditions, which are specific and different for each agent. The existence of this trade-off between adaptation and coordination creates a conflict of interests between agents (since they need to adapt their decisions to different conditions), which prevents them from communicating their intentions truthfully and forces them to introduce noise into their messages. Such setting is particularly relevant in management and organizational economics: trade-off between adaptation and coordination is often outlined as one of most important trade-offs in managerial decision-making, and the phenomena of local information has also shown its relevance in questions of allocation of decision rights and organizational structure of the firm. In particular, we may consider the mentioned agents as managers running their respective divisions within a firm: they need to fulfill their own goals, but at the same time need to somewhat coordinate their actions.

Given this setting and keeping in mind the mentioned imperfections in communication, we ask the following question: is it best to just speak simultaneously or would it be better to first listen to what the other agent has to say? In other words, which communication mechanism is better in such situation: conventional simultaneous message-sending, when both agents send their messages at the same time, or sequential, when one agent receives an incoming message and only after that he sends his own, thus acquiring an informational advantage? The results turned out to be surprising.

The "follower" or "listener", i.e. the agent, who receives the other agent's message before sending his own, turns out to be worse off in this sequential communication game comparing to the symmetric simultaneous communication case. This seems strange given that follower possesses more information in this case and that first-mover's strategy (concerning both messages and decisions) is the same as in simultaneous communication case (i.e. he does not try to exploit follower's new knowledge). However, it turns out that this additional information affects follower's incentives and forces him to send messages, which are less informative on average. This leads to worse coordination on behalf of first-mover (because he has less information on what to coordinate with) and thus increases losses from miscoordination of both agents. Follower then has to sacrifice adaptation to some degree in favor of coordination with the less coordinating counterpart, and his profits decrease.

What is even more surprising in this story is that the leader (agent who sends his message first and is being listened to) may actually benefit from switching from simultaneous communication to the sequential scheme: the fact that the follower will be more willing to coordinate if his message becomes less informative mitigates losses from generally worse coordination of actions due to worse message, and in the end these losses may even not be sufficient to cover the benefit that the leader acquires from adapting better to his own local conditions. Speaking shortly, if the follower has high enough dependency, then the leader is likely to benefit from receiving a less precise message.

As for total firm profits (sum of payoffs of both agents), simultaneous communication strictly dominates the sequential scheme if agents (divisions) are more or less similar in terms of size (or importance to headquarters) or are both sufficiently independent, i.e. both face a rather weak need for coordination. On the other hand, if one division is sufficiently larger than another and possesses some average need for coordination, while this other division is highly dependent, then sequential communication with the larger division as first-mover may be more preferable for the firm in general even despite losses that the lesser department will incur.

It is worth noting that under investigation in this paper is the question of *optimal* communication order, not *equilibrium* one. In fact, there is a variety of reasons of why these two might diverge. One reason is that one can prescribe a manager not to speak before the other one is finished, but one cannot force him to *listen* and comprehend the incoming information: even under the assumption that this action does not require any effort on behalf of the manager (which is arguable), our model predicts that it is simply unprofitable for the manager to listen to his colleague[12], since this in the end reduces profits of his division. Another possibility for deviation involves leadership: if a manager knows that sending a message early may be beneficial for him, then he will try to do so, ignoring all prescriptions. In fact, a situation may arise, in which *both* managers will be willing to send an early message.

Solution of these problems should involve more complicated compensation schemes – we assumed that managers are awarded by the performance of their respective divisions, while in this case overall firm performance might be a better basis for compensation. Of course, deeper investigation of this problem should also take into consideration the moral hazard issues, which have been totally ignored in this paper, since we assumed that all managers' actions are costless.

The main direction for further research is considering other, more complicated communication mechanisms: for example, who said that we should only consider one-round mechanisms? Is it possible that multi-stage communication will yield better results for firm as a whole and each agent personally? Another question has already been outlined in the paragraph above and it asks how we can enforce the optimal communication mechanism, which may be suboptimal for one of participating divisions. Finally, we considered only the decentralized setting, where the headquarters plays no real role, while earlier research shows that sometimes it may be optimal for the HQ to retain the decision rights (instead of delegating them to the managers), and in this case managers will be communicating with HQ instead of each other – optimal mechanism for this case is yet to be found.

[12]To be exact, a manager does not want to know the other's information at the moment of sending a message, but does so at the moment of making a decision.

References

Alonso, R., Dessein, W., and Matouschek, N. (2008a). Centralization Versus Decentralization: an Application to Price Setting by a Multi-Market Firm. *Journal of the European Economic Association*, 6(2-3):457–467.

Alonso, R., Dessein, W., and Matouschek, N. (2008b). When Does Coordination Require Centralization? *American Economic Review*, 98(1):145–179.

Alonso, R., Dessein, W., and Matouschek, N. (2009). Organize to Compete!

Alonso, R. and Matouschek, N. (2007). Relational delegation. *The RAND Journal of Economics*, 38(4):1070–1089.

Amitai, M. (1996). Cheap-talk with incomplete information on both sides. *PhD Thesis, The Hebrew University.*

Aumann, R. J. and Hart, S. (2003). Long cheap talk. *Econometrica*, 71(6):1619–1660.

Battaglini, M. (2003). Multiple referrals and multidimensional cheap talk. *Econometrica*, 70(4):1379–1401.

Bester, H. (2009). Externalities, communication and the allocation of decision rights. *Economic Theory*, 41(2):269–296.

Chakravarthy, B. S. (1982). Adaptation: A promising metaphor for strategic management. *Academy of Management Review*, pages 35–44.

Chen, Y. (2009). Communication with two-sided asymmetric information. *Available at SSRN 1344818.*

Cramton, C. D. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. *Organization science*, 12(3):346–371.

Crawford, V. and Sobel, J. (1982). Strategic Information Transmission. *Econometrica*, 50(6):1431–1451.

de Barreda, I. M. (2010). Cheap talk with two-sided private information. *Job-Market Paper.*

Dessein, W. (2002). Authority and Communication in Organizations. *The Review of Economic Studies*, 69(4):811–838.

Galeotti, A., Ghiglino, C., and Squintani, F. (2013). Strategic information transmission in networks. *Journal of Economic Theory*, forthcoming.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic management journal*, 17:109–122.

Heath, C. and Staudenmayer, N. (2000). Coordination neglect: How lay theories of organizing complicate coordination in organizations. *Research in Organizational Behavior*, 22:153–191.

Lai, K. W. (2009). Essays on strategic information transmission. *PhD Thesis, University of Pittsburgh.*

Malone, T. W. (1987). Modeling coordination in organizations and markets. *Management science*, 33(10):1317–1332.

McGee, A. and Yang, H. (2013). Cheap talk with two senders and complementary information. *Games and Economic Behavior*, 79:181 – 191.

Ottaviani, M. and Sørensen, P. (2001). Information aggregation in debate: who should speak first? *Journal of Public Economics*, 81(3):393–421.

Rantakari, H. (2008). Governing Adaptation. *Review of Economic Studies*, 75(4):1257–1285.