

# Costly Communication and Incentives

Mathias Dewatripont  
ECARES, Université Libre de Bruxelles

March 23, 2006

# Introduction

- Standard theory (mechanism design, cheap talk) has stressed asymmetric information and at times impossibility to communicate but not really *cost* of communication.
- Communication is costly, however.
- On the other hand, one can economize on communication costs, e.g. by exchanging 'summary information'.
- A more recent literature, started by Radner (1993) has stressed communication costs, but in teams, i.e. without incentive problems.
- It has generated many insights on internal organization design, as well as on the optimal size of organizations.

- In this lecture:
  - I first offer a brief summary of selected results from the costly-communication-in-team literature;
  - then I discuss a specific recent model from the literature, due to Dessein and Santos (2003);
  - then I introduce incentives in this setting, following an approach I developed with Jean Tirole.
  
- One crucial result I will stress: A key advantage of the introduction of incentives in a costly-communication setting is to allow for the *coexistence of various forms of communication*, in particular cheap talk versus costly communication. This considerably enriches the analysis of organizations.

# 1. A brief overview of the communication-in-team literature

## 1.1. Reasons for team work

If it is costly for one individual to know what another individual already knows, why incur communication costs by engaging in team work? Several reasons have been invoked:

- Radner (1993) and Van Zandt (various papers): concern for delay and the advantage of parallel processing of information.
- Bolton-Dewatripont (1994) and Dessein-Santos (2003): Smithian returns from specialization.
- Garicano (2000): economize on costly training for 'hard-to-solve', infrequent problems.

## 1.2. Selected results

- Models where total amount of information to be gathered (and communicated) is exogenous (e.g. 'aggregate' information about a given set of  $N$  projects):
  - Radner and Van Zandt analyze whether 'regular hierarchies' emerge as optimum.
  - Endogenous specialization: Bolton and Dewatripont show how 'thick reports go to the top more quickly'.
  - Endogenous specialization: Garicano shows that more highly skilled individuals should not be called upon 'at the start'.
- All this is broadly consistent with insights from management literature (e.g. Mintzberg 1979).

- Models with endogenous amount of information collected and therefore endogenous organization size:
  - Meagher-Orbay-Van Zandt (2001): setting in which comparing info about the environment is time-consuming and the environment itself is changing; yields inverted U-shaped relation between organization size and environment volatility.
  - Bolton-Dewatripont (1995): decentralize when labor time constraint is tight relative to budget constraint (too costly then to 'fine-tune' allocation of funds).
- Focus of this lecture on a model with endogenous probability that communication takes place: Dessein-Santos, to be discussed in more detail.

## 2. Dessein-Santos setting

- Reason for team work: Smithian returns from specialization.
- Reconsiders the idea that benefits from specialization are limited by miscoordination costs (expressed for example by Becker and Murphy 1992).
- Shows the importance in this discussion of taking into account the extent to which the organization needs to be 'adaptive' to local information.
- Endogenizes the degree of centralization, the adaptiveness of the organization as well as the quality of communication and compares with empirical results documenting 'organizational shift' towards more communication and more decentralization (Caroli 2001, Caroli and Van Reenen 2001, Bresnahan et al. 2002)

- Simplify setting here: only two actions to be taken: single 'primary' action  $a$  and single 'complementary' action  $b$ .
- Action  $a$  should ideally be as close as possible to 'local information'  $\theta$ , a random variable with mean  $\theta_0$  and variance  $\sigma_\theta^2$ . One could think of  $\theta_0$  as the 'status quo'.
- Action  $b$  should ideally be as close as possible to action  $a$  (think of  $b$  as an 'input' which has to 'fit' with  $a$ ).
- Expected cost to the organization:

$$E \left[ \phi(a - \theta)^2 + \beta(b - a)^2 \right]$$

where  $\phi$  is the weight given to misadaptation and  $\beta$  the weight given to miscoordination.

## 2.1. Decentralized case

- Only one action per individual.
- The individual who chooses  $a$  first observes  $\theta$ , but the individual who chooses  $b$  observes neither  $\theta$  nor  $a$ .
- Call the individual who chooses  $a$  the 'sender' ( $S$ ), since she has to send a message about  $a$  to the other individual, called the 'receiver' ( $R$ ).
- Communication is imperfect:  $S$ 's message is received by  $R$  only with prob  $p$  (with prob  $1 - p$ ,  $R$  learns nothing).
- Timing: *Stage 1*:  $S$  observes  $\theta$ , chooses  $a$  and communicates it to  $R$ ; *Stage 2*:  $R$  receives  $S$ 's message with prob  $p$  and sets  $b$ .

- Since both parties want to minimize

$$E \left[ \phi(a - \theta)^2 + \beta(b - a)^2 \right],$$

$R$  sets  $b = a$  whenever he learns the value of  $a$  and sets  $b = \theta_0$  otherwise.

- Therefore,  $S$  chooses  $a$  to minimize:

$$\phi(a - \theta)^2 + \beta(1 - p)(\theta_0 - a)^2$$

for each realization of  $\theta$ , or:

$$a = \theta_0 + \frac{\phi}{\phi + \beta(1 - p)}(\theta - \theta_0).$$

$\phi / [\phi + \beta(1 - p)]$  is the degree of adaptiveness of the organization, as well as the 'level of discretion' enjoyed by  $S$ . It grows with  $\phi$  and  $p$  and goes down with  $\beta$ .

- Plugging this optimal  $a$  into:

$$E \left[ \phi(a - \theta)^2 + \beta(b - a)^2 \right],$$

one obtains as expected cost of the organization:

$$\frac{\phi\beta(1-p)}{\phi + \beta(1-p)}\sigma_{\theta}^2.$$

This expression grows with  $\sigma_{\theta}^2$ ,  $\phi$  and  $\beta$ , and decreases with  $p$ .

- Next question: What about the choice of  $p$ , in a prior *Stage 0*? For simplicity, assume two possible choices:  $p_L = 0$  and  $p_M > 0$ , where  $p_M$  implies a cost  $F$  for the organization. Under decentralization, the expected cost of the organization is thus:

$$EDC = \min \left\{ \frac{\phi\beta}{\phi + \beta}\sigma_{\theta}^2, \frac{\phi\beta(1-p_M)}{\phi + \beta(1-p_M)}\sigma_{\theta}^2 + F \right\}$$

## 2.2. Centralized case

- One can reduce miscoordination costs by asking a single individual to choose both  $a$  and  $b$ .
- Assume departures from the 'status quo' (i.e.  $a = \theta_0$ ) can still run some risk of miscoordination, i.e. that  $b$  'remains at  $\theta_0$ '. But this risk is limited to  $1 - p_H$ , with  $p_M < p_H \leq 1$ .
- Cost of centralization: loss of Smithian returns, which amount to a cost  $K$  ( $> F$ , to make the problem interesting).
- Under centralization, the expected cost of the organization is thus:

$$ECC = \frac{\phi\beta(1 - p_H)}{\phi + \beta(1 - p_H)}\sigma_\theta^2 + K.$$

## 2.3. Results

- Complementarity between adaptiveness and communication/centralization.
- Higher uncertainty, i.e. a higher  $\sigma_\theta^2$ , favors adaptiveness and communication/centralization.
- Higher costs of misadaptation, i.e. a higher  $\phi$ , does the same as a higher  $\sigma_\theta^2$  but with limits.
- Interestingly, higher miscoordination costs, i.e. a higher  $\beta$ , has ambiguous effects: For  $\beta$  close to 0 or  $\infty$ , choosing  $p = 0$  is optimal, but for intermediate  $\beta$ 's not necessarily:
  - Indeed, a rise in  $\beta$  from a very low level raises the value of communication, in order to keep achieving 'ex-post coordination'.

- But when  $\beta$  is very large, best to have  $a \rightarrow \theta_0$ , so as to avoid miscoordination costs: 'ex-ante coordination', with very little communication and responsiveness to local information. Decentralization is then optimal, to reap the Smithian returns.
- Lower communication costs, i.e. a lower  $F$ , may result in a move away from centralization (with a move from  $p_H$  to  $p_M$ ) and therefore in *lower adaptiveness*, or in sticking with decentralization (with a move from  $p_L = 0$  to  $p_M$ ) while becoming *more adaptive*.
- In a symmetric multi-task setting, Dessein and Santos also show that lower communication costs, by raising the attractiveness of adaptiveness, may at times imply more (partial) centralization, i.e. more task bundling.

- Conclusion:
  - the Dessein-Santos model offers a rich setting that sheds light on recent moves towards higher adaptiveness in organizations, higher communication and higher task bundling.
  - it is also a good benchmark model to analyze the effects of the introduction of incentives.

### 3. Introducing incentives

- With incentive problems, efficient communication is hindered by:
  - *lack of congruence* between the sender (persuader, speaker, source) and the receiver (addressee, listener).
  - *moral hazard in team*: (i) the sender must expand time, attention and other resources to communicate effectively her knowledge; (ii) the receiver must pay attention, decode, understand, and rehearse the acquired knowledge. In a nutshell, “it takes two to tango”.
- Economic theory has ignored this moral-hazard-in-team because it has focused on the two polar cases of *soft information* (i.e. cheap talk) and *hard information* (which can be verified by the receiver, but at exogenous cost).

- In practice, information often neither soft nor hard but *in between*, and degree of softness *endogenous*.
  
- Sociology and psychology literatures:
  - general emphasis on communication costs in theories of "persuasion".
  
  - various theories of "dual mode" perspective, with distinction between (often costlier) "issue-relevant route" and (often less costly) "issue-irrelevant cues" (expertise or attractiveness of the sender, credibility of her position given her personal stakes, ...). Choice of route depends on communication abilities and stakes of the parties.

- Below, I introduce a simplified version of Dewatripont-Tirole (*JPE* 2005 forthcoming) into the simplified Dessein-Santos model, and I:
  - consider imperfect congruence as well as moral hazard in team.
  - distinguish between *executive (or implementative) decision-making* and *supervisory (or approval) decision-making*: need "to know the details" or not.
  - investigate the *endogenous reliance on various communication modes*: issue-relevant messages vs cues vs cheap talk.
- Focus is on decentralized Dessein-Santos case, where achieving 'issue-relevant' communication, with prob  $p_M$ , requires *each party* to expend one unit of effort, at cost  $F/2$ , i.e. any single effort is wasted if the other party does not expend effort.

### 3.1. Executive decision-making

- Building block for next case.
  
- Since  $R$  'needs to know the details', learning  $a$  requires successful issue-relevant communication (single comm. mode).
  
- Timing:
  - *Stage 0*: Each party privately decides whether or not to expend one unit of effort, at individual cost  $F/2$ .
  
  - *Stage 1*:  $S$  observes  $\theta$ , chooses  $a$  and, if she has expended effort, can at no further cost try and communicate  $a$  to  $R$ ;
  
  - *Stage 2*: If both parties have expended effort,  $R$  learns the value of  $a$  with prob  $p_M$ .

– *Stage 3: R sets b.*

- Payoffs, gross of effort costs:

$$U_S \equiv -s E \left[ \phi(a - \theta)^2 + \beta(b - a)^2 \right]$$

and

$$U_R \equiv -r E \left[ \phi(a - \theta)^2 + \beta(b - a)^2 \right].$$

where  $r$  and  $s > 0$ . Here, I thus assume congruent objectives, except for the privately incurred effort costs.

- Results:
  - Ceteris paribus, an equilibrium with prob  $p_M$  of communication exists for  $F$  low enough, and for  $\min \{s, r\}$  high enough (both parties' stakes should be sufficient).
  - Similarly, a rise in  $p_M$  (better communication technology) also helps generating a positive prob of communication in equilibrium.

- From a normative point of view, since communication has a positive externality on the other party, only potential problem is underinvestment in communication.
- Beyond built-in strategic complementarity, second reason for  $R$  to be more interested in expending effort when he expects  $S$  to do it: When  $S$  expects to communicate, sets  $a$  in a more adaptive manner, so that comm. stakes become endogenous.
- Note: In a setting with only two actions but with continuous  $p$ , Dewatripont-Tirole show that main results survive in the case of strategic substitutability, and also in the case of sequential rather than simultaneous communication efforts.

## 3.2. Supervisory decision-making

- $R$  need not understand  $S$ 's message to learn the value of  $a$ : (costless) 'cheap talk' is possible (and, for simplicity, conveys with prob. 1  $S$ 's message to  $R$ ).
- To make problem interesting, *need* to have different objectives for the two parties: Team approach *cannot* accommodate cheap talk!
- Specifically,  $S$  keeps the same payoff as above:

$$U_S \equiv -s E \left[ \phi(a - \theta)^2 + \beta(b - a)^2 \right]$$

- Instead,  $R$  has a different payoff, namely:

$$U_R \equiv -r [\alpha E |b - a| + (1 - \alpha) E |b - \theta_0|].$$

- Interpretation: With prob  $\alpha$ , it is in  $R$ 's interest to 'follow  $S$ ' and pick  $b = a$  (e.g. because there is 'enough of a market ' for  $R$ ); instead, with prob  $1 - \alpha$ ,  $R$ 's interest is to 'stay with the status quo', and set  $b = \theta_0$ .  $\alpha$  is a measure of the degree of congruence between the two parties.
- Assume neither  $S$  nor  $R$  know  $R$ 's objective function, but  $\alpha$  is common knowledge (generalizes to the case where  $S$  knows  $R$ 's objective function).

- Costly communication allows  $R$  to learn  $a$  as well as 'the details', i.e. his objective function (i.e. whether the project really has potential for him), while cheap talk only conveys the value of  $a$ .
- Two cases obtain:
  - *Low congruence* ( $\alpha < 1/2$ ): here,  $R$  does not select  $b = a$  unless he becomes convinced that his objective function is  $-rE |b - a|$ . We are therefore in a similar case as the previous one:  $R$  will pick  $b = \theta_0$  unless costly communication is successful, and then only with prob  $\alpha$ .
  - *High congruence* ( $\alpha < 1/2$ ): here,  $S$  has "real authority". She can pick  $a = \theta$  and announce it to  $R$  using cheap talk while not expending effort on issue-relevant communication.  $R$  then chooses  $b = a$ .

- Result: An increase in congruence between  $S$  and  $R$  leads to a breakdown of costly communication when  $S$  can count on  $R$  to "rubberstamp" her (costless) recommendation. An increase in congruence raises  $S$ 's payoff monotonically but not  $R$ 's.
- Remark: Normative analysis less straightforward, since  $a$  can be learnt through cheap talk.
- Remark: When  $S$  knows  $R$ 's objective function, possibility of multiple equilibria (with either issue-relevant communication or real authority for  $S$ ) due to self-fulfilling expectations by  $R$ .

### 3.3. "Cues"

- Cues: information not directly relevant to the decision problem but that influences the credibility of  $S$  in the eyes of  $R$ .
- Start from modified set-up:
  - As above,  $R$  expects his payoff to be  $-rE |b - a|$  with prob.  $\alpha$  and  $-rE |b - \theta_0|$  with prob.  $1 - \alpha$ .
  - $\alpha$  is weighted average between a high prob.  $\alpha_H$  (prob.  $\gamma$ ) and a low prob.  $\alpha_L$  (prob.  $1 - \gamma$ ).
  - $S$  knows whether  $\alpha = \alpha_H$  or  $\alpha = \alpha_L$ .
  - All this is common knowledge.

- Interpret information  $S$  communicates about  $\alpha$  as cues while costly information about  $a$  is an "issue-relevant" message.
- Cue information is hard (only choice for  $S$  is whether to *disclose* it), so only  $\alpha_H$  will ever be disclosed.
- Just as issue-relevant communication, cue communication is a 0–1 choice for both  $S$  and  $R$ , with associated cost  $f/2$  for each party, with  $f \ll F$ .
- Communication costs are additive for each individual ("no-load" condition).
- Timing:  $R$  chooses both communication efforts without having observed  $S$ 's effort levels.
- Focus here on low-congruence case ( $\alpha < 1/2$ ).

- Remark: provided  $f$  is small enough, optimal for  $R$  to look at cues first and then choose whether to expend issue-relevant effort (this allows him to fine-tune effort w.r.t.  $S$ 's credibility).
- Result: The possibility to communicate (relatively cheap) cues: (i) reduces issue-relevant communication when cues are "bad news" about  $S$ 's credibility; (ii) raises issue-relevant communication when cues are "good news" about  $S$ 's credibility but not sufficiently so that  $R$  rubberstamps her recommendation ( $\alpha_H < 1/2$ ); (iii) eliminates issue-relevant communication when cues are sufficiently "good news" about  $S$ 's credibility that  $R$  rubberstamps her recommendation ( $\alpha_H > 1/2$ ).
- In cases (i) and (iii), cues crowd out issue-relevant communication, while the opposite is true in case (ii). Cues are respectively *substitute* or *complement* to issue-relevant communication.

### 3.4. Adaptiveness, cheap talk and costly communication

- At times, better to just rely on cheap talk and be less adaptive rather than going for issue-relevant communication.
- Consider a simple setting where cheap talk may not be fully revealing (as in Crawford-Sobel 1982).
- Specifically,  $\theta \in \{\theta_0 - 2, \theta_0 - 1, \theta_0 + 1, \theta_0 + 2\}$  with each value having ex-ante prob = 1/4. These four  $\theta$ 's and  $\theta_0$  are also the five possible  $a$ 's.
- For simplicity, assume: (i) returns from specialization are high enough to make centralization unattractive; (ii) no possibility of cue communication; (iii)  $S$  has no private information about  $R$ 's objective function; (iv) cheap talk can convey for sure  $S$ 's announcement about  $\theta$ .

- Concentrate first on *cheap talk*.
- Keep  $S$ 's payoff as in subsection 3.2, i.e.:

$$U_S \equiv -s E \left[ \phi(a - \theta)^2 + \beta(b - a)^2 \right]$$

- However,  $R$ 's payoff now becomes:

$$U_R \equiv -r \left[ \alpha E(b - a)^2 + (1 - \alpha) E(b - \theta_0)^2 \right].$$

- In this setting, cheap talk allows  $S$  to convey credible information about  $\theta$  (interests being sufficiently aligned) but not on  $R$ 's preferences. This makes  $R$  reluctant to be very adaptive:
  - $R$  is ready to set  $b = a$  (rather than  $b = \theta_0$ ) when he expects that  $a = \theta_0 \pm 1$ , provided  $\alpha \geq 1/2$ .

- But when he expects  $a = \theta_0 + 2$  (resp.  $\theta_0 - 2$ ), he'd rather set  $b = \theta - 1$  (resp.  $\theta + 1$ ) provided  $1 \leq 4(1 - \alpha)$ , or  $\alpha \leq 3/4$ .
- When  $1/2 \leq \alpha \leq 3/4$ ,  $S$  thus expects  $R$  never to set  $b = \theta_0 \pm 2$  even when this would be the value of  $a$ .
- In this case, under cheap talk,  $R$ 's unwillingness to be very adaptive may imply that  $S$  can do no better than convey the sign of  $\theta - \theta_0$  and then limit its choice of  $a$  to  $\theta_0 \pm 1$  in order to reduce miscoordination; then, perfect coordination is achieved, but with partial adaptiveness.
- Indeed, when  $\theta = \theta_0 \pm 2$ , setting  $a = \theta$  then yields a cost of  $\beta$ . The other possible choice is to have  $a = \theta - 1$  (resp.  $\theta + 1$ ) if  $\theta = \theta_0 + 2$  (resp.  $\theta_0 - 2$ ), which yields a cost of  $\phi$ . This strategy is attractive if  $\phi < \beta$ .

- Could *issue-relevant communication* do better?
- Since cheap talk can convey whether  $\theta$  and  $a$  are below or above  $\theta_0$ , as above, when issue-relevant communication is not successful,  $R$  sets  $b = \theta_0 - 1$  (resp.  $\theta_0 + 1$ ) when he is told that  $a$  is below (resp. above)  $\theta_0$ .
- This means that, if  $\theta = \theta_0 \pm 1$ ,  $S$  will not send hard information to  $R$  *even if she has previously expended effort and conveying information is therefore free at that point in time*. Indeed, if communication is successful and therefore  $R$  learns what her objective function is:
  - with prob  $\alpha$ , it makes no difference for  $S$ , since anyway  $R$  would have set  $b = a$ .
  - with prob  $1 - \alpha$ , it is *bad* for  $S$ , because  $R$  then sets  $b = \theta_0$ .

- Consider now the case where  $\theta = \theta_0 \pm 2$ , where  $S$  has previously expended effort and does send hard information to  $R$ . If communication is successful and therefore  $R$  learns what her objective function is:
  - with prob  $\alpha$ , this is *good* for  $S$ , because  $R$  then sets  $b = a$ .
  - with prob  $1 - \alpha$ , this is *bad* for  $S$ , because  $R$  then sets  $b = \theta_0$ .
- Even in the case where  $\theta = \theta_0 \pm 2$ , the net effect is negative for  $S$  given that she wants to minimize  $E(b - a)^2$ : if she had chosen for example  $a = \theta_0 + 2$ , it would be better for her to have  $R$  choose  $b = \theta_0 + 1$  for sure rather than  $b = \theta_0 + 2$  with prob  $\alpha$  and  $b = \theta_0$  with prob  $1 - \alpha$ , under our assumption that  $1/2 \leq \alpha \leq 3/4$ .
- Consequently, issue-relevant communication would be unattractive for  $S$  even if it were free!

- Conclusion:
  - In this case, in equilibrium, issue-relevant communication has negative value *over what can be communicated through cheap talk*.
  - Through cheap talk, only the sign of  $\theta$  is usefully communicated, since anyway  $R$  and  $S$  restrict themselves to  $a = b = \theta_0 \pm 1$ .
  - This organization therefore exhibits perfect coordination thanks to limiting itself to partial adaptiveness.
  - This example provides an illustration of the rich interplay between cheap talk and costly communication in an incentive model.

### 3.5. Dynamics of communication

- Dewatripont and Tirole look at the role of (two-period) dynamics in *learning* about other party.
- Show that dynamics raises *period-1 efforts* (and can even *enable* communication), because one party may want to "impress" the other about the profitability of their relationship so as to elicit higher period-2 effort from him/her.
- As for *period-2 effort*:
  - It can be lower than static effort when a permanent communication breakdown results from an earlier communication failure
  - On the other hand, optimism following successful communication can under dynamics lead to sustained communication while there would be none in a static context.

## 4. Concluding remarks

- There is now a growing literature on costly communication in teams, generating interesting insights on internal organization design, the boundaries of organizations and even its economy-wide implications (see recent work of Garicano et al.).
- Blending this with incentive perspectives can, as shown hopefully in this lecture, further enrich the approach, in particular with respect to the parties' ability to side-step costly communication and rely on cheap talk.
- Many potential avenues for further research: combination with endogenous control rights (along the lines of Aghion-Tirole 1997 or Dessein 2002), multi-layer organizations, communication in markets, ...