

WHEN DOES COORDINATION REQUIRE CENTRALIZATION?*

Ricardo Alonso
Northwestern University

Wouter Dessein
University of Chicago

Niko Matouschek
Northwestern University

First version: September 2005

This version: July 2006

Abstract

This paper compares centralized and decentralized coordination when managers are privately informed and communicate strategically. We consider a multi-divisional organization in which decisions must be responsive to local conditions but also coordinated with each other. Information about local conditions is dispersed and held by self-interested division managers who communicate via cheap talk. The only available formal mechanism is the allocation of decision rights. We show that a higher need for coordination improves horizontal communication but worsens vertical communication. As a result, no matter how important coordination is, decentralization dominates centralization if the division managers are not too biased towards their own divisions and the divisions are not too different from each other (e.g. in terms of division size).

Keywords: coordination, decision rights, cheap talk, incomplete contracts

JEL classifications: D23, D83, L23

*We thank Jim Dana, Mathias Dewatripont, Luis Garicano, Rob Gertner, Bob Gibbons, Andrea Prat, Canice Prendergast, Luis Rayo, Scott Schaefer and Jeroen Swinkels for their comments and suggestions. We would also like to thank seminar participants at Bristol University, Chicago GSB, Columbia University, Duke University, Ecares, Essex University, HKUST, Indiana University, Northwestern Kellogg, Oxford University, the University of British Columbia, the University of Toulouse and conference participants at the 2005 MIT Sloan Summer Workshop in Organizational Economics and the 2006 Utah Winter Business Economics Conference for their comments. A previous version of this paper has been circulated under the title "A Theory of Headquarters."

1 Introduction

Multi-divisional organizations exist primarily to coordinate the activities of their divisions. To do so efficiently, they must resolve a trade-off between coordination and adaptation: the more closely activities are synchronized across divisions, the less they can be adapted to the local conditions of each division. To the extent that division managers are best informed about their divisions' local conditions, efficient coordination can be achieved only if these managers communicate with the decision makers. A central question in organizational economics is whether efficient communication and coordination are more easily achieved in centralized or in decentralized organizations. In other words, are organizations more efficient when division managers communicate horizontally and then make their respective decisions in a decentralized manner or when they communicate vertically with an independent headquarters which then issues its orders?

This question has long been debated among practitioners and academics alike. Chandler (1977), for instance, argues that coordination requires centralization:

*“Thus the existence of a managerial hierarchy is a defining characteristic of the modern business enterprise. A multiunit enterprise without such managers remains little more than a federation of autonomous offices. [...] Such federations were often able to bring small reductions in information and transactions costs but they could not lower costs through increased productivity. They could not provide the administrative coordination that became the central function of modern business enterprise.”*¹

Consistent with this view, many firms respond to an increased need for coordination by abandoning their decentralized structures and moving towards centralization.² There are, however, also numerous managers who argue that efficient coordination can be achieved in decentralized organizations provided that division managers are able to communicate with each other. Alfred Sloan, the long-time President and Chairman of General Motors, for instance, organized GM as a multi-divisional firm and granted vast authority to the division managers.^{3,4} To ensure coordination between them, Sloan set up various committees that gave the division managers an opportunity to exchange ideas.

¹Chandler (1977), pp.7-8.

²See for instance the DaimlerChrysler Commercial Vehicles Division (Hannan, Podolny and Roberts 1999), Procter & Gamble (Bartlett 1989) and Jacobs Suchard (Eccles and Holland 1989).

³Our discussion of General Motors is based on “Co-ordination By Committee,” Chapter 7 in *My Years With General Motors* by Alfred Sloan (1964).

⁴Writing to some of his fellow GM executives in 1923, for instance, Sloan stated that “According to General Motors plan of organization, to which I believe we all heartily subscribe, the activities of any specific Operation are under the absolute control of the General Manager of that Division, subject only to very broad contact with the general officers of the Corporation.” (Sloan 1964, p.106).

These committees did not diminish the authority of division managers to run their own divisions and instead merely provided “a place to bring these men together under amicable circumstances for the exchange of information and the ironing out of differences.”⁵ Reflecting on the first such committee, the General Purchasing Committee that was created to coordinate the purchasing of inputs by the different divisions, Sloan wrote in 1924 that:

*“The General Purchasing Committee has, I believe, shown the way and has demonstrated that those responsible for each functional activity can work together to their own profit and to the profit of the stockholders at the same time and such a plan of coordination is far better from every standpoint than trying to inject it into the operations from some central activity.”*⁶

The apparent success of Alfred Sloan’s GM and other decentralized firms in realizing inter-divisional synergies suggests that in many cases coordination can indeed be achieved without centralization.⁷

The aim of this paper is to reconcile these conflicting views by analyzing when coordination does and does not require centralization. Decentralized organizations have a natural advantage at adapting decisions to local conditions since the decisions are made by the managers with the best information about those conditions. However, such organizations also have a natural disadvantage at coordinating since the manager in charge of one decision is uncertain about the decisions made by others. Moreover, self-interested division managers may not internalize how their decisions affect other divisions. One might therefore reason naively that centralization is optimal whenever coordination is sufficiently important relative to the desire for responsiveness. We argue that this reasoning is flawed and show that decentralization can be optimal even when coordination is very important. Intuitively, when coordination becomes very important, division managers recognize their interdependence and communicate and coordinate very well under decentralization. In contrast, under centralization, an increased need for coordination strains communication, as division managers anticipate that headquarters will enforce a compromise. As a result, decentralization can be optimal even when coordination becomes very important. In particular, this is so if the division

⁵Sloan (1964), p.105.

⁶Sloan (1964), p.104.

⁷Other, and more recent, examples of firms that rely on the managers of largely autonomous divisions to coordinate their activities without central intervention are PepsiCo (Montgomery and Magnani 2001) and AES Corporation (Pfeffer 2004). In the 1980s and early 90s PepsiCo centralized very few of the activities of its three restaurant chains Pizza Hut, Taco Bell and KFC and ran them as what were essentially stand-alone businesses. The management of PepsiCo believed that coordination between the restaurant chains could be achieved by encouraging the division managers to share information and letting them decide themselves on their joint undertakings: “In discussing coordination across restaurant chains, senior corporate executives stressed that joint activity should be initiated by divisions, not headquarters. Division presidents should have the prerogative to decide whether or not a given division would participate in any specific joint activity. As one explained, ‘Let them sort it out. Eventually, they will. It will make sense. They will get to the right decisions.’” (Montgomery and Magnani 2001, p.12).

managers' incentives are sufficiently aligned and the divisions are not too different from each other. Thus, only in organizations in which divisions are quite heterogeneous, for instance in their relative size, or in which implicit and explicit incentives strongly bias division managers towards their own divisions, does coordination require centralization.

We propose a simple model of a multi-divisional organization with three main features: (i.) decision making involves a trade-off between coordination and adaptation. In particular, two decisions have to be made and the decision makers must balance the benefit of setting the decisions close to each other with that of setting each decision close to its idiosyncratic environment or 'state.' Multinational enterprises (MNEs), for instance, may realize scale economies by coordinating the product designs in different regions. These cost savings, however, must be traded off against the revenue losses that arise when products are less tailored to local tastes. (ii.) information about the states is dispersed and held by division managers who are biased towards maximizing the profits of their own divisions rather than those of the overall organization. Moreover, the division managers communicate their information strategically to influence decision making in their favor. In the above MNE example, regional managers are likely to be best informed about the local tastes of consumers and thus about the expected revenue losses due to standardization. In communicating this information they have an incentive to behave strategically to influence the decision making to their advantage. (iii.) the organization lacks commitment and, in particular, can only commit to the ex ante allocation of decision rights (Grossman and Hart 1986 and Hart and Moore 1990). This implies that decision makers are not able to commit to make their decisions dependent on the information they receive in different ways. Communication therefore takes the form of cheap talk (Crawford and Sobel 1982). In this setting we compare the performance of two organizational structures: under Decentralization the division managers communicate with each other horizontally and then make their decisions in a decentralized manner while under Centralization the division managers communicate vertically with a headquarters manager who then makes both decisions.

Underlying our results are differences in how centralized and decentralized organizations aggregate dispersed information. In our model vertical communication is always more informative than horizontal communication. Essentially, since division managers are biased towards the profits of their own divisions while headquarters aims to maximize overall profits, the preferences of a division manager are more closely aligned with those of headquarters than with those of another division manager. As a result, division managers share more information with headquarters than they do with each other. Thus, while division managers are privately informed about their own divisions' local conditions and therefore have an *information advantage*, headquarters has a *communication*

advantage. This communication advantage, however, gradually disappears as coordination becomes more important. In particular, whereas an increased need for coordination leads to worse communication under Centralization, it actually improves communication under Decentralization. Intuitively, when coordination becomes more important, headquarters increasingly ignores the information that it receives from the division managers about their local conditions. This induces each manager to exaggerate his case more which, in turn, leads to less information being communicated. In contrast, under Decentralization an increase in the need for coordination makes the managers more willing to listen to each other to avoid the costly coordination failures. As a result, the managers' incentives to exaggerate are mitigated and more information is communicated.

The fact that as coordination becomes more important the communication advantage of headquarters gradually disappears, while the division managers' information advantage does not, drives our central result: Decentralization can dominate Centralization even when coordination is extremely important. Specifically, we show that no matter how important coordination is, centralizing destroys value if the division managers are not too biased towards their own divisions and the divisions are not too different from each other. Thus, coordination only requires centralization if the divisions are quite heterogenous or division managers care mostly about their own divisions.

In the next section we discuss the related literature. In Section 3 we present our model which we then solve in Sections 4 and 5. The main comparison between Centralization and Decentralization is discussed in Section 6. In Section 7 we extend our model by allowing for asymmetries between the divisions. Finally, we conclude in Section 8.

2 Related Literature

Our paper is related to, and borrows from, a few different literatures.

Coordination in Organizations: A number of recent papers analyze coordination in organizations. Hart and Holmstrom (2002) focus on the trade-off between coordination and the private benefits of doing things 'independently:' under decentralization the division managers do not fully internalize the benefits of coordination while under centralization the decision maker ignores the private benefits that division managers realize if they act independently. In Hart and Moore (2005), some agents specialize in developing ideas about the independent use of one particular asset, while others think about the coordinated use of several assets. They analyze the optimal hierarchical structure and provide conditions under which coordinators should be superior to specialists. Finally, in Dessein, Garicano and Gertner (2005), 'product managers' are privately informed about the benefits of running a particular division independently, whereas a 'functional manager' is privately informed

about the value of a coordinated approach. They endogenize the incentives for effort provision and the communication of this private information and show that functional authority is preferred when effort incentives are less important.

A key difference between these papers and ours is that in their models a trade-off between centralization and decentralization arises because the incentives of the central decision-maker are distorted (and biased towards coordination). In contrast, in our paper, authority is allocated to a benevolent principal under centralization.⁸ Decentralization may nevertheless be strictly preferred because it allows for a better use of dispersed information. Bolton and Farrell (1990) have also emphasized this trade-off between coordination and the use of local information. In a model of entry they show that decentralization is good at selecting a low cost entrant but also results in inefficient delay and duplication of entry. Unlike our paper, Bolton and Farrell (1990) rule out communication. In contrast to our result, they find that centralization is preferred provided that coordination costs are sufficiently large relative to the value of private information.⁹

Following Gertner, Scharstein and Stein (1994) and Stein (1997), there is finally a large literature on internal capital markets which argues that centralization can facilitate the efficient allocation of capital across otherwise independent divisions. Closest in spirit to our paper are Stein (2002) and Friebel and Raith (2006) who assume that the information necessary for an efficient allocation is dispersed and needs to be aggregated.¹⁰ Unlike in our paper, communication is always vertical.¹¹ *Information Processing in Organizations*: The large literature on team theory, starting with Marshak and Radner (1972), constitutes the first attempt by economists to understand decision making within firms. Team theory analyzes decision making in firms in which information is dispersed and physical constraints make it costly to communicate or process this information. In doing so it abstracts from incentive problems and assumes that agents act in the interest of the organization. A team-theoretic model that is closely related to ours in spirit is Aoki (1986) who also compares the efficiency of vertical and horizontal information structures.¹² In contrast to this paper, and to

⁸Otherwise, the trade-off between coordination and adaptation is similar to the trade-offs between coordination and ‘independence’ considered in the above papers. Whereas in the above papers coordination is a binary choice, we allow for decisions to be more or less coordinated.

⁹Also in Gertner (1999) headquarters is unbiased but it only intervenes when bargaining between divisions breaks down. His model further differs from ours in that sharing information is costly. He shows how the presence of an independent arbitrator may foster information sharing.

¹⁰See also Ozbas (2005) and Wulf (2005).

¹¹Our paper is also related to the large political economy literature on fiscal federalism which studies the choice between centralization and decentralization in the organization of states (see Oates 1999 and Lockwood 2005 for surveys). However, the institutions they look at, such as elections and legislatures, are different from the institutions that determine decision making within firms and that we focus on.

¹²In addition to Aoki (1986), we follow Dessein and Santos (2006) in modeling a trade-off between adaptation and coordination. This paper shows how, in the presence of imperfect communication, extensive specialization results in organizations that ignore local knowledge. Also Cremer, Garicano and Prat (2006) study how physical communication

team theory in general, our results do not depend on assumptions about physical communication constraints. Instead, we endogenize communication quality as a function of incentive conflicts. As such our analysis is related to the mechanism design approach to organizational design which also focuses on the incentives of agents to misrepresent their information.¹³ This literature, however, concentrates on settings in which the Revelation Principle holds and in which centralized organizations are therefore always weakly optimal. In contrast, we develop a simple model in which the Revelation Principle does not hold since agents are unable to commit to mechanisms. As a result decentralized organizations can be strictly optimal.

Our no-commitment assumption is in line with a number of recent papers that adopt an incomplete contracting approach to organizational design and model communication as a cheap talk.¹⁴ Dessein (2002) considers a model in which a principal must decide between delegating decision rights to an agent versus keeping control and communicating with that agent.¹⁵ Harris and Raviv (2005) consider a similar set up but allow the principal to have private information while Alonso and Matouschek (2006) endogenize the commitment power of the principal in an infinitely repeated game. These papers, however, do not analyze coordination, nor do they allow for horizontal communication. A paper that does analyze coordination in a setting with horizontal communication is Rantakari (2006). He has independently developed a model that is similar to ours and shows that if division managers care about coordination to different degrees, then it may be optimal to centralize control in the division that cares less about coordination. Since in his model, managers care exclusively about their own division, centralization is always optimal when coordination becomes important.

Cheap Talk: From a methodological perspective our paper is related to the large cheap talk literature that builds on Crawford and Sobel (1982). A technical difference between our model and that in Crawford and Sobel (1982) is that we allow for the preferred decisions of the senders and receivers to coincide. As such our paper is related to Melumad and Shibano (1990) who also allow for this

constraints limit coordination. In their model, organizations face a trade-off between adopting a common technical language, which allows for better coordination among units, or several specialized, distinct languages which are better adapted to each unit. Other related team theory papers include Bolton and Dewatripont (1994), on the firm as a communication network, Garicano (2000), on the organization of knowledge in production, Qian, Roland and Xu (2006), on coordinating change in organizations, and more generally, the large literature on information processing organizations (Radner 1993 and others; see Van Zandt 1999 for an overview).

¹³For a survey of this literature see Mookherjee (2006).

¹⁴Another related literature analyzes how adding a prior cheap talk stage matters in coordination games or games with asymmetric information. Farrell (1987), for example, shows how cheap talk can improve coordination in a battle of the sexes game. More closely related to our set-up, Farrell and Gibbons (1989) analyzes how cheap talk can avoid a break-down in a bargaining game with asymmetric information. A recent paper along these lines is Baliga and Morris (2002). Farrell and Rabin (1996) provide an overview of related cheap talk applications.

¹⁵See also Marino and Matsusaka (2005).

possibility. A key difference between their analysis and ours is that they focus on communication equilibria with a finite number of intervals while we allow for equilibria with an infinite number of intervals. We show that such equilibria maximize the expected joint surplus and that they are computationally straightforward since they avoid the integer problems associated with finite interval equilibria. For this reason we believe that our model is more tractable than the leading example in Crawford and Sobel, the traditional workhorse for cheap talk applications.¹⁶ We further differ from both Crawford and Sobel (1982) and Melumad and Shibano (1990) in that we allow for multiple senders. Our paper is therefore related to Battaglini (2002) and Krishna and Morgan (2002) who consider models in which a principal consults multiple informed “experts” who all observe the same piece of information. The question they investigate is whether, and if so how, the principal can illicit this information from the experts.¹⁷ In contrast, in our model the senders observe different pieces of independent information which makes it impossible to achieve truth-telling.

3 The Model

An organization consists of two operating divisions, Division 1 and Division 2, and potentially one headquarters. Division $j \in \{1, 2\}$ generates profits that depend on its local conditions, described by $\theta_j \in \mathbf{R}$, and on two decisions, $d_1 \in \mathbf{R}$ and $d_2 \in \mathbf{R}$. In particular, the profits of Division 1 are given by

$$\pi_1 = K_1 - (d_1 - \theta_1)^2 - \delta (d_1 - d_2)^2, \quad (1)$$

where $K_1 \in \mathbf{R}_+$ is the maximum profit that the division can realize. The first squared term captures the *adaptation loss* that Division 1 incurs if decision d_1 is not perfectly adapted to its local conditions, that is if $d_1 \neq \theta_1$, and the second squared term captures the *coordination loss* that Division 1 incurs if the two decisions are not perfectly coordinated, that is if $d_1 \neq d_2$. The parameter $\delta \in [0, \infty)$ then measures the importance of coordination relative to adaptation. The profits of Division 2 are similarly given by

$$\pi_2 = K_2 - (d_2 - \theta_2)^2 - \delta (d_1 - d_2)^2, \quad (2)$$

where $K_2 \in \mathbf{R}_+$ is the maximum profit that Division 2 can realize. Without loss of generality we set $K_1 = K_2 = 0$. Headquarters does not generate any profits.

Information: Each division is run by one manager. Manager 1, the manager in charge of Division 1, privately observes his local conditions θ_1 but does not know the realization of θ_2 . Similarly,

¹⁶ A recent paper which shares this feature is Kawamura (2006).

¹⁷ See also Ottaviani and Sorensen (2001) who study the impact of reputational concerns on communication in a multiple experts setting.

Manager 2 observes θ_2 but does not know θ_1 . The HQ Manager, i.e. the manager in charge of headquarters, observes neither θ_1 nor θ_2 . It is common knowledge, however, that θ_1 and θ_2 are uniformly distributed on $[-s_1, s_1]$ and $[-s_2, s_2]$ respectively, with s_1 and $s_2 \in \mathbf{R}_+$. The draws of θ_1 and θ_2 are independent.

Preferences: We assume that Manager 1 maximizes $\lambda\pi_1 + (1 - \lambda)\pi_2$ whereas Manager 2 maximizes $(1 - \lambda)\pi_1 + \lambda\pi_2$, where $\lambda \in [1/2, 1]$. The parameter λ thus captures how biased each division manager is towards his own division's profits. The HQ Manager simply aggregates the preferences of the two division managers and thus maximizes by $\pi_1 + \pi_2$.¹⁸ For simplicity we take the preferences of the managers as given and do not model their origins. Intuitively, factors outside of our model, such as career concerns and subjective performance evaluations, are prone to bias division managers towards maximizing the profits of the division under their direct control as their managerial skills and effort will be mainly reflected in the performance of this division. In contrast, the skills and effort of the HQ Manager are more likely to be reflected in the overall performance of the organization, rather than in that of one particular division. In principle, the organization might attempt to neutralize the division managers' biases towards their own divisions by compensating them more for the performance of the rest of the organization than for that of their own division. As will become clear below, if it were possible to contract over λ , the organization would always set $\lambda = 1/2$ and all organizational structures would perform equally well. However, it will typically be undesirable for the organization to fully align managerial incentives if the division managers have to make division specific effort choices.¹⁹ Furthermore, to the extent that divisions need to make many decisions, the allocation of one particular decision right is likely to have only a negligible impact on endogenously derived incentives. It therefore seems reasonable, as a first step, to assume that the division managers' biases do not differ across organizational structures.²⁰

Contracts and Communication: We follow the property rights literature (Grossman and Hart 1986,

¹⁸For the results presented below it is not important that the HQ Manager is entirely unbiased. Qualitatively similar results would be obtained as long as her utility function is a convex combination of that of Managers 1 and 2.

¹⁹The conflict between motivating efficient effort provision, on the one hand, and efficient decision making and/or communication, on the other, has been analyzed in a number of recent papers (Athey and Roberts 2001; Dessein, Garicano and Gertner 2005; Friebel and Raith 2006). These papers show that it is typically optimal for organizations to bias division managers towards their own divisions to motivate effort provision even if doing so distorts their incentives on other dimensions.

²⁰In reality λ can be very big and in some cases it can even be equal to one. GM provides a historical example of such a case: "Under the incentive system in operation before 1918, a small number of division managers had contracts providing them with a stated share in the profits of their own divisions, irrespective of how much the corporation as a whole earned. Inevitably, this system exaggerated the self-interest of each division at the expense of the interests of the corporation itself. It was even possible for a division manager to act contrary to the interests of the corporation in his effort to maximize his own division's profits." (Sloan 1964, p.409).

Hart and Moore 1990) in assuming that contracts are highly incomplete. In particular, the organization can only commit to an ex ante allocation of decision rights. Agents are unable to contract over the decisions themselves and over the communication protocol that is used to aggregate information. Once the decision rights have been allocated, they cannot be transferred before the decisions are made. We focus on two allocations of decision rights. Under *Decentralization* Manager 1 has the right to make decision d_1 and Manager 2 has the right to make decision d_2 and both decisions are made simultaneously. Under *Centralization* both decision rights are held by the HQ Manager.²¹

The lack of commitment implies that the decision makers are not able to commit to paying transfers that depend on the information they receive or to make their decisions depend on such information in different ways. Communication therefore takes the form of an informal mechanism: cheap talk. For simplicity we assume that this informal communication occurs in one round of communication. In particular, under Decentralization Manager 1 sends message $m_1 \in M_1$ to Manager 2 and, simultaneously, Manager 2 sends message $m_2 \in M_2$ to Manager 1. Under Centralization, Managers 1 and 2 simultaneously send messages $m_1 \in M_1$ and $m_2 \in M_2$ to the HQ Manager. We refer to communication between the division managers as *horizontal communication* and that between the division managers and the HQ Manager as *vertical communication*. It is well known in the literature on cheap talk games that repeated rounds of communication can expand the set of equilibrium outcomes even if only one player is informed.²² However, even for a simple cheap talk game such as the leading example in Crawford and Sobel (1982), it is still an open question as to what is the optimal communication protocol. Since it is our view that communication is an ‘informal’ mechanism which cannot be structured by the mechanism designer, it seems reasonable to focus on the simplest form of informal communication. In this sense, we take a similar approach as the property rights literature which assumes that players engage in ex post bargaining but limits the power of the mechanism designer to structure this bargaining game.

A key feature of our model is its symmetry: the divisions are of equal size, they have the same need for coordination and the two decisions are made simultaneously. We focus on a symmetric organization since it greatly simplifies the analysis. In Section 7, however, we allow for asymmetries between the divisions and discuss how such asymmetries affect our results.

The game is summarized in Figure 1. First, decision rights are allocated to maximize the total

²¹A natural variation of Decentralization is to allow for sequential decision making and a natural variation of Centralization is to centralize both decision rights in one of the divisions. It turns out that the former structure dominates the latter. We discuss the former structure in Section 7.

²²See, for example, Aumann and Hart (2003) and Krishna and Morgan (2004).

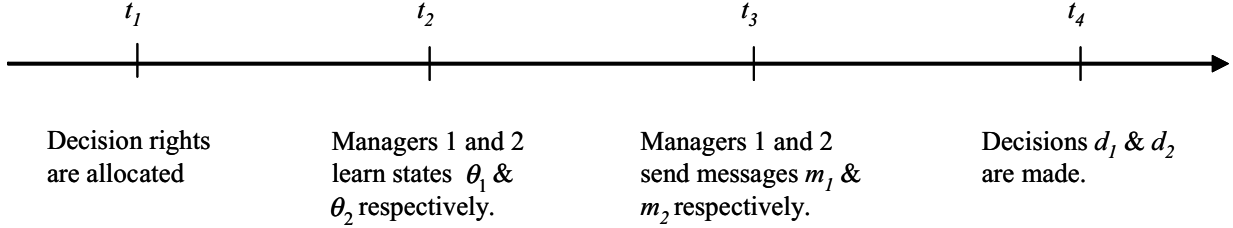


Figure 1: Timeline

expected profits $E[\pi_1 + \pi_2]$. Under Centralization the HQ Manager gets the right to make both decisions and under Decentralization each division manager gets the right to make one decision. Second, the division managers become informed about their local conditions, i.e. they learn θ_1 and θ_2 respectively. Third, the division managers communicate with the decision makers. Under Centralization they engage in vertical communication, sending messages m_1 and m_2 to the HQ Manager, while under Decentralization they engage in horizontal communication, sending messages m_1 and m_2 to each other. Finally, the decisions d_1 and d_2 are made. Each decision maker chooses the decision that maximizes his or her payoff given the information that has been communicated.

4 Decision Making

In this section we characterize decision making under Centralization and Decentralization, taking as given the posterior beliefs of θ_1 and θ_2 . In Section 5 we then characterize the communication subgame and hence endogenize these beliefs. Finally, in Section 6 we draw on our understanding of the decision making and the communication subgame to compare the performance of the two organizational structures. The proofs of all lemmas and propositions are in the appendix.

Under Centralization, the HQ Manager receives messages m_1 and m_2 from the division managers and then chooses the decisions d_1 and d_2 that maximize $E[\pi_1 + \pi_2 | m]$, where π_1 and π_2 are the divisions' profits as given by (1) and (2), respectively, and $m \equiv (m_1, m_2)$. The decisions that solve this problem are convex combinations of the HQ Manager's posterior beliefs of θ_1 and θ_2 given m :

$$d_1^C \equiv \gamma_C E_H[\theta_1 | m] + (1 - \gamma_C) E_H[\theta_2 | m] \tag{3}$$

and

$$d_2^C \equiv (1 - \gamma_C) E_H[\theta_1 | m] + \gamma_C E_H[\theta_2 | m], \tag{4}$$

where

$$\gamma_C \equiv \frac{1 + 2\delta}{1 + 4\delta} \tag{5}$$

and $E_H[\theta_j | m]$, $j = 1, 2$, is the HQ Manager's expected value of θ_j after receiving messages m . Note that γ_C is decreasing in δ and ranges from $1/2$ to 1 . When the decisions are independent, that is when $\delta = 0$, the HQ Manager sets $d_1^C = E[\theta_1 | m]$ and $d_2^C = E[\theta_2 | m]$. As the importance of coordination δ increases, she puts less weight on $E[\theta_1 | m]$ and more weight on $E[\theta_2 | m]$ when making decision d_1 . Eventually, as $\delta \rightarrow \infty$, she puts the same weight on both decisions, i.e. she sets $d_1^C = d_2^C = E[\theta_1 + \theta_2 | m]/2$.

Under Decentralization, the division managers first send each other messages m_1 and m_2 . Once the messages have been exchanged, Manager 1 chooses d_1 to maximize $E[\lambda\pi_1 + (1 - \lambda)\pi_2 | m]$ and, simultaneously, Manager 2 chooses d_2 to maximize $E[(1 - \lambda)\pi_1 + \lambda\pi_2 | m]$. The decision that Manager 1 makes is a convex combination of his local conditions, θ_1 , and of $E_1[d_2 | m]$, i.e. the decision that Manager 1 expects Manager 2 to make:

$$d_1 = \frac{\lambda}{\lambda + \delta}\theta_1 + \frac{\delta}{\lambda + \delta}E_1[d_2 | m]. \quad (6)$$

Similarly, we have that

$$d_2 = \frac{\lambda}{\lambda + \delta}\theta_2 + \frac{\delta}{\lambda + \delta}E_2[d_1 | m], \quad (7)$$

where $E_2[d_1 | m]$ is the decision that Manager 2 expects Manager 1 to make. It is intuitive that the weight that each division manager puts on his state is increasing in the own-division bias λ and decreasing in the importance of coordination δ .

The decisions $E_1[d_2 | m]$ and $E_2[d_1 | m]$ that each division manager expects his counterpart to make can be obtained by taking the expectations of (6) and (7). Doing so and substituting back into (6) and (7) we find that Manager 1's decision is a convex combination of his local conditions θ_1 , his posterior belief about θ_2 and Manager 2's posterior belief about θ_1 :

$$d_1^D \equiv \frac{\lambda}{\lambda + \delta}\theta_1 + \frac{\delta}{\lambda + \delta} \left(\frac{\delta}{\lambda + 2\delta}E_2[\theta_1 | m] + \frac{\lambda + \delta}{\lambda + 2\delta}E_2[\theta_2 | m] \right). \quad (8)$$

Similarly, we have that

$$d_2^D \equiv \frac{\lambda}{\lambda + \delta}\theta_2 + \frac{\delta}{\lambda + \delta} \left(\frac{\lambda + \delta}{\lambda + 2\delta}E_2[\theta_1 | m] + \frac{\delta}{\lambda + 2\delta}E_2[\theta_2 | m] \right). \quad (9)$$

It can be seen that as δ increases, Manager $j = 1, 2$ puts less weight on his private information θ_j and more weight on a weighted average of the posterior beliefs. In the limit, as $\delta \rightarrow \infty$, the division managers only rely on the communicated information and set $d_1^D = d_2^D = (E_2[\theta_1 | m] + E_2[\theta_2 | m])/2$. Note that, for given posteriors, these are exactly the same decisions that the HQ Manager implements when $\delta \rightarrow \infty$.

5 Strategic Communication

In this section we analyze communication under the two organizational structures. A key insight of this section is that while vertical communication is more informative than horizontal communication this difference decreases, and eventually vanishes, as coordination becomes more important. This is so since an increase in the need for coordination improves horizontal communication but worsens vertical communication.

We proceed by investigating the division managers' incentives to misrepresent information in the next sub-section. This then allows us to characterize the communication equilibria in Section 5.2 and to compare the quality of vertical and horizontal communication in Section 5.3. Finally, in Section 5.4 we derive the expected profits and show that they can be expressed as linear functions of the quality of communication.

5.1 Incentives to Misrepresent Information

Regardless of the allocation of control, managers have an incentive to make others believe that their local conditions are more extreme, that is, further from the average than is truly the case. To see this, consider first the incentives of Manager 1 to misrepresent his information under Centralization.²³ For this purpose let $\nu_1 = E_H[\theta_1 | m_1]$ be the HQ Manager's expectation of θ_1 after receiving message m_1 and suppose that Manager 1 can simply choose any ν_1 . In other words, suppose that Manager 1 can credibly misrepresent his information about his state. The question then is why, and by how much, Manager 1 would like to misrepresent his information. Clearly, Manager 1 would like the HQ Manager to have the posterior that maximizes his expected payoff:

$$\nu_1^* = \arg \max_{\nu_1} E \left[-\lambda (d_1 - \theta_1)^2 - (1 - \lambda) (d_2 - \theta_2)^2 - \delta (d_1 - d_2)^2 \mid \theta_1 \right], \quad (10)$$

where $d_1 = d_1^C$ and $d_2 = d_2^C$ as defined in (3) and (4) respectively. We will see below that in equilibrium the expected value of the posterior of θ_2 is equal to the expected value of θ_2 , i.e. that $E_{m_2} [E[\theta_2 | m_2]] = E[\theta_2] = 0$. Assuming that this relationship holds we can use (10) to obtain

$$\nu_1^* - \theta_1 = \frac{(2\lambda - 1)\delta}{\lambda + \delta} \theta_1 \equiv b_C \theta_1. \quad (11)$$

This expression shows that Manager 1 has no incentive to misrepresent his information when $\theta_1 = 0$. The reason for this is that, in the absence of any information about θ_2 , the incentives of Manager 1 and the HQ Manager are perfectly aligned when $\theta_1 = 0$. For $\theta_1 \neq 0$, however, Manager 1 does have an incentive to exaggerate his state, that is to set $|\nu_1^* - \theta_1| > 0$. This is due to his desire to

²³The argument for Manager 2 is analogous.

adapt d_1 more closely to θ_1 . Essentially, since Manager 1 does not fully internalize the need to coordinate the decisions, he always wants d_1 to be set closer to θ_1 than the HQ Manager does. It is therefore in his interest to induce the HQ Manager to choose a larger decision whenever $\theta_1 > 0$ and to choose a smaller decision whenever $\theta_1 < 0$ and he can achieve this by exaggerating his state. Expression (11) also shows Manager 1's desire to exaggerate his state is increasing in $|\theta_1|$. Essentially, the bigger $|\theta_1|$, the bigger Manager 1's marginal benefit of inducing the HQ Manager to implement a more extreme decision and thereby bring d_1 closer to θ_1 .

The division managers' incentives to misrepresent information under Decentralization are similar to those under Centralization. To see this, let $\nu_1 = E_2[\theta_1 | m_1]$ be Manager 2's expectation of θ_1 after receiving message m_1 and suppose again that Manager 1 can simply choose any ν_1 . His optimal choice of ν_1 is given by (10), where $d_1 = d_1^D$ and $d_2 = d_2^D$ as defined in (8) and (9) respectively. If we assume again that $E_{m_2}[E[\theta_2 | m_2]] = E[\theta_2] = 0$, which will be shown to hold in equilibrium, then it follows that for $\delta > 0$

$$\nu_1^* - \theta_1 = \frac{(2\lambda - 1)(\lambda + \delta)}{\lambda(1 - \lambda) + \delta} \theta_1 \equiv b_D \theta_1. \quad (12)$$

Thus, it is again the case that Manager 1 has no incentive to misrepresent his information when $\theta_1 = 0$, that he has an incentive to exaggerate it if $\theta_1 \neq 0$ and that his incentive to exaggerate is increasing in $|\theta_1|$. The incentive to exaggerate is again due to Manager 1's desire to closely adapt decision d_1 to θ_1 . In particular, since Manager 1 puts some weight on coordinating his decision with that of Manager 2, his ability to adapt d_1 to θ_1 is constrained by Manager 2's choice of d_2 . If, for instance, $\theta_1 > 0$, then Manager 1 would like to induce Manager 2 to set a higher d_2 so that he can increase d_1 and thereby bring it closer to θ_1 . He can do so by exaggerating his state, i.e. by setting $\nu_1^* - \theta_1 > 0$.

How do the incentives to misrepresent information depend on the own-division bias and the need for coordination? It is intuitive that under both structures the incentives to misrepresent are increasing in the own-division bias, i.e. $db_l/d\lambda \geq 0$ for $l = C, D$. What might be less intuitive is that an increase in the need for coordination increases the incentives to misrepresent information under Centralization but decreases them under Decentralization, i.e. $db_C/d\delta \geq 0$ and $db_D/d\delta \leq 0$. This suggests that, as coordination becomes more important, the division managers are less willing to share information with the coordinator under Centralization but are more willing to share information with each other under Decentralization. To understand this, recall that under Centralization Manager 1 exaggerates his private information to induce the HQ Manager to make a more extreme decision, i.e. to increase $|d_1|$, while under Decentralization he exaggerates to induce Manager 2 to make a more extreme decision. The key observation is that the more important the

need for coordination, the less sensitive the HQ Manager’s decision making is to the information she receives from Manager 1 and thus the stronger Manager 1’s desire to exaggerate. In contrast, an increase in the importance of coordination makes Manager 2’s decision more sensitive to the information he receives from Manager 1 and thereby reduces Manager 1’s desire to misrepresent his information. Essentially, when coordination becomes more important, division managers are less willing to share information with a headquarters that increasingly ignores their information but they are more willing to share information with each other to avoid costly coordination failures.²⁴

5.2 Communication Equilibria

We now show that, as in Crawford and Sobel (1982), all communication equilibria are interval equilibria in which the state spaces $[-s_1, s_1]$ and $[-s_2, s_2]$ are partitioned into intervals and the division manager only reveal which interval their local conditions θ_1 and θ_2 belong to. In this sense the managers’ communication is noisy and information is lost. Moreover, the size of the intervals – which determines how noisy communication is – depends directly on b_D and b_C as defined in (11) and (12).

A communication equilibrium under each organizational structure is characterized by (i.) communication rules for the division managers, (ii.) decision rules for the decision makers and (iii.) belief functions for the message receivers. The communication rule for Manager $j = 1, 2$ specifies the probability of sending message $m_j \in M_j$ conditional on observing state θ_j and we denote it by $\mu_j(m_j | \theta_j)$. The decision rules map messages $m_1 \in M_1$ and $m_2 \in M_2$ into decisions $d_1 \in \mathbf{R}$ and $d_2 \in \mathbf{R}$ and we denote them by $d_1(m)$ and $d_2(m)$ respectively. Finally, the belief functions are denoted by $g_j(\theta_j | m_j)$ for $j = 1, 2$ and state the probability of state θ_j conditional on observing message m_j .

We focus on Perfect Bayesian Equilibria of the communication subgame which require that (i.) communication rules are optimal for the division managers given the decision rules, (ii.) the decision rules are optimal for the decision makers given the belief functions and (iii.) the belief functions are derived from the communication rules using Bayes’ rule whenever possible.

Since all communication equilibria will be shown to be interval equilibria, we denote by $a_j^{2N} \equiv$

²⁴To see this formally, note that before Manager 1 learns Manager 2’s message, he expects the HQ Manager to make decision $E_{\theta_2}(d_1^C) = \gamma_C \nu_1$, where γ_C is given in (5) and $\nu_1 = E_H(\theta_1 | m_1)$ is the HQ Manager’s posterior belief of θ_1 . Since γ_C is decreasing in δ , the HQ Manager’s decision making becomes less sensitive to the information she receives from Manager 1 when coordination becomes more important. This, in turn, gives Manager 1 an incentive to exaggerate his state more. In contrast, we can use (9) to show that under Decentralization the decision that Manager 1 expects Manager 2 to make before having received Manager 2’s message is given by $E_{\theta_2}(d_2^D) = \nu_1 \delta / (\lambda + 2\delta)$. It can be seen that $E_{\theta_2}(d_2^D)$ is more sensitive to ν_1 when coordination is more important. As a result, an increase in δ actually mitigates Manager 1’s desire to exaggerate θ_1 .

$(a_{j,-N}, \dots, a_{j,-1}, a_{j,0}, a_{j,1}, \dots, a_{j,N})$ and $a_j^{2N-1} \equiv (a_{j,-N}, \dots, a_{j,-1}, a_{j,1}, \dots, a_{j,N})$ the partitioning of $[-s_j, s_j]$ into $2N$ and $2N - 1$ intervals respectively, where $a_{j,-N} = -s_j$, $a_{j,0} = 0$ and $a_{j,N} = s_j$. Thus, a_j^{2N} corresponds to finite interval equilibria with an even number of intervals and a_j^{2N-1} corresponds to those with an odd number of intervals. As will be shown in the next proposition, the end points are symmetrically distributed around zero, i.e. $a_{j,i} = a_{j,-i}$ for all $i \in \{1, \dots, N\}$. The following proposition characterizes the finite communication equilibria when $\delta > 0$.

PROPOSITION 1 (Communication Equilibria). *If $\delta \in (0, \infty)$, then for every positive integer N_j , $j = 1, 2$, there exists at least one equilibrium $(\mu_1(\cdot), \mu_2(\cdot), d_1(\cdot), d_2(\cdot), g_1(\cdot), g_2(\cdot))$, where*

i. $\mu_j(m_j | \theta_j)$ is uniform, supported on $[a_{j,i-1}, a_{j,i}]$ if $\theta_j \in (a_{j,i-1}, a_{j,i})$,

ii. $g_j(\theta_j | m_j)$ is uniform supported on $[a_{j,i-1}, a_{j,i}]$ if $m_j \in (a_{j,i-1}, a_{j,i})$,

iii. $a_{j,i+1} - a_{j,i} = a_{j,i} - a_{j,i-1} + 4ba_{j,i}$ for $i = 1, \dots, N_j - 1$

$a_{j,-(i+1)} - a_{j,-i} = a_{j,-i} - a_{j,-(i-1)} + 4ba_{j,-i}$ for $i = 1, \dots, N_j - 1$

with $b = b_C$ under Centralization, where b_C is defined in (11), and

$b = b_D$ under Decentralization, where b_D is defined in (12),

iv. $d_j(m) = d_j^C$, $j = 1, 2$, under Centralization, where d_j^C are given by (3) and (4) and

$d_j(m) = d_j^D$, $j = 1, 2$, under Decentralization, where d_j^D are given by (8) and (9).

Moreover, all other finite equilibria have relationships between θ_1 and θ_2 and the managers' choices of d_1 and d_2 that are the same as those in this class for some value of N_1 and N_2 ; they are therefore economically equivalent.

The communication equilibria are illustrated in Figure 2. In these equilibria each division manager communicates what interval his state lies in. The size of the intervals is determined by the difference equations in Part (iii.) of the proposition. The size of an interval $(a_{j,i+1} - a_{j,i})$ equals the size of the preceding interval $(a_{j,i} - a_{j,i-1})$, plus respectively $4b_C a_{j,i}$ under Centralization and $4b_D a_{j,i}$ under Decentralization, where $a_{j,i}$ is the dividing point between the two intervals. Recall from Section 5.1, that $b_C \theta_j$, $j = 1, 2$, is the difference between the true state of nature θ_j and what Manager j would like the HQ Manager to believe that the state is. Similarly, $b_D \theta_j$, $j = 1, 2$, represents by how much Manager j wants to misrepresent his state when talking to the other division manager under Decentralization. The incentives to distort information thus directly determine how quickly communication deteriorates as θ_j is further away from its mean.²⁵ It is intuitive that since

²⁵This can be related to the leading example in Crawford and Sobel (1982). In that model, there is a fixed difference b between the true state of nature and what the sender would like the receiver to believe is the true state, and

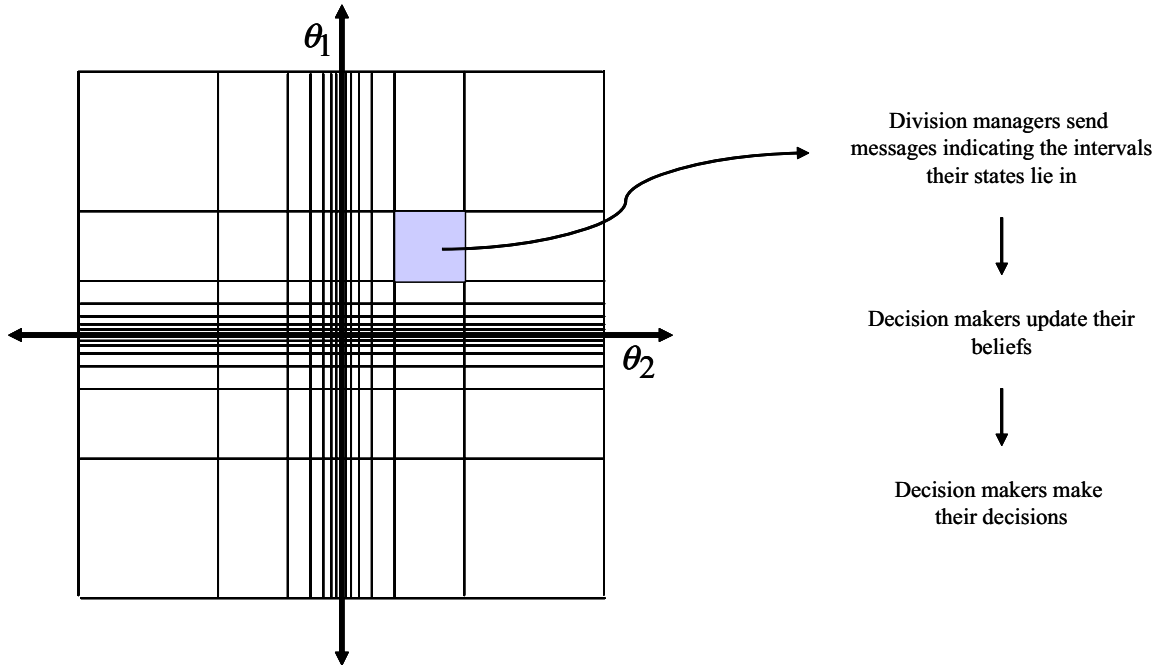


Figure 2: Communication Equilibria

the incentives to misrepresent information are increasing in $|\theta_j|$, not only is it the case that less information is transmitted the larger $|\theta_j|$, but also the rate at which communication becomes noisier is increasing in $|\theta_j|$.

Proposition 1 characterizes communication equilibria for $\delta > 0$. In the absence of any need for coordination the communication equilibria are straightforward. In particular, under Centralization, truth-telling can be sustained if $\delta = 0$ since, in this case, there is no incentive conflict between the division managers. While there are other equilibria, we assume in the remaining analysis that when $\delta = 0$ the managers coordinate on the truth-telling equilibrium. Under Decentralization, communication is irrelevant when $\delta = 0$ since it is optimal for each division manager to set his decision equal to his state which, of course, he observes directly. This implies that for $\delta = 0$ all communication strategies are consistent with a Perfect Bayesian Equilibrium under Decentralization. Merely to facilitate the exposition of one comparative static that we perform later, and without losing generality, we assume that for $\delta = 0$ the communication equilibrium under Decentralization is as those described in Proposition 1.²⁶

Proposition 1 shows that there does not exist an upper limit on the number of intervals that can

equivalently, intervals grow at a fixed rate of $4b$ rather than $4ba_{j,i}$.

²⁶See discussion of equation (17).

be sustained in equilibrium. This is in contrast to Crawford and Sobel (1982) where the maximum number of intervals is always finite. This difference is due to the fact that in our model there exists a state, namely $\theta_j = 0$ for $j = 1, 2$, in which the incentives of the sender and the receiver are perfectly aligned.²⁷ In Crawford and Sobel (1982) this possibility is ruled out. The next proposition shows that in the limit in which the number of intervals goes to infinity, the strategies and beliefs described in Proposition 1 constitute a Perfect Bayesian Equilibrium and that the total expected profits are maximized in this equilibrium.

PROPOSITION 2 (Efficiency). *The limit of strategy profiles and beliefs $(\mu_1(\cdot), \mu_2(\cdot), d_1(\cdot), d_2(\cdot), g_1(\cdot), g_2(\cdot))$ as $N_1, N_2 \rightarrow \infty$ is a Perfect Bayesian Equilibrium of the communication game. In this equilibrium the total expected profits $E[\pi_1 + \pi_2]$ are higher than in any other equilibrium.*

An illustration of an equilibrium in which the number of intervals goes to infinity is provided in Figure 2. In such an equilibrium the size of the intervals is infinitesimally small when θ_j , $j = 1, 2$, is close to zero but grows as $|\theta_j|$ increases. We believe that it is reasonable to assume that the organization is able to coordinate on the equilibrium that maximizes the expected overall profits and we focus on this equilibrium for the rest of the analysis.

5.3 Communication Comparison

We can now compare the quality of communication under the two organizational structures and analyze how it is affected by changes in the need for coordination and the own-division bias. We measure the quality of communication as the residual variance $E[(\theta_j - E[\theta_j|m_j])^2]$ of θ_j , $j = 1, 2$. The next lemma derives the residual variance under vertical and horizontal communication.

LEMMA 1. *In the most efficient equilibrium in which $N_1, N_2 \rightarrow \infty$ the residual variance is given by*

$$E[(\theta_j - E[\theta_j|m_j])^2] = S_l \sigma_j^2 \quad j = 1, 2 \text{ and } l = C, D,$$

where

$$S_l = \frac{b_l}{3 + 4b_l}. \quad (13)$$

The residual variance is therefore directly related to the division managers' incentives to misrepresent information as defined in (11) and (12). In particular, when $b_l = 0$, $l = C, D$, then the division managers perfectly reveal their information and as a result the residual variance is zero. As b_l increases, less information is communicated in equilibrium and the residual variance increases.

²⁷We share this feature with Melumad and Shibano (1991); see the related literature in Section 2.

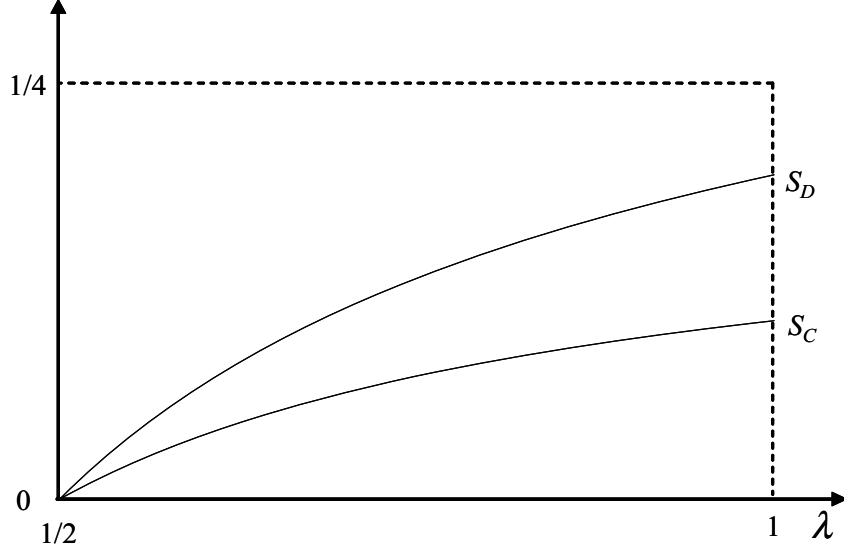


Figure 3: Communication Comparison (λ)

Finally, as $b_l \rightarrow \infty$, the division managers only reveal whether their state is positive or negative and thus $S_l \rightarrow 1/4$. We can now state the following proposition.

PROPOSITION 3 (The Quality of Communication).

- i. $S_C = S_D = 0$ if $\lambda = 1/2$ and $S_D > S_C$ otherwise,
- ii. $\partial S_D / \partial \lambda > \partial S_C / \partial \lambda > 0$,
- iii. $\partial S_C / \partial \delta > 0 > \partial S_D / \partial \delta$ and $\lim_{\delta \rightarrow \infty} S_D = \lim_{\delta \rightarrow \infty} S_C$.

Parts (i.) and (ii.), which are illustrated in Figure 3, show that vertical communication is in general more efficient than horizontal communication. In other words, headquarters has a *communication advantage* vis-a-vis the division managers. To understand this, recall that for $\lambda = 1/2$ communication is perfect under both organizational structures and note, from Part (ii.), that an increase in the own-division bias λ has a more detrimental effect on horizontal than on vertical communication. This is the case since, under Centralization, an increase in λ increases the bias of the senders but does not affect the decision making of the receiver. In contrast, under Decentralization, an increase in λ also leads to more biased decision making by the receiver.

Part (iii.) is illustrated in Figure 4 and shows that the communication advantage diminishes as the need for coordination increases. As discussed in Section 5.1 this key property is due to the fact that a higher δ increases the incentives of division managers to misrepresent their information under Centralization but reduces them under Decentralization. Part (iii.) also shows that as δ increases the communication advantage not only shrinks but actually vanishes. This can be understood by

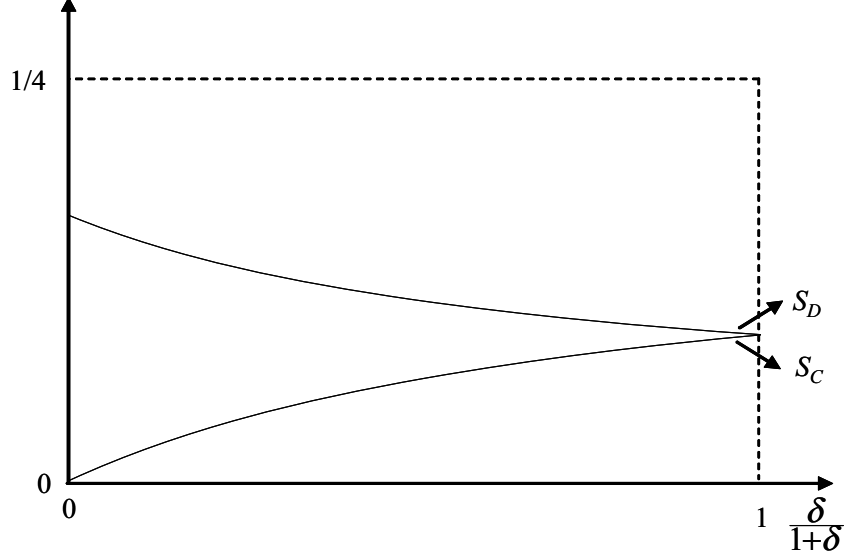


Figure 4: Communication Comparison (δ)

recalling from Section 4 that in the limit in which $\delta \rightarrow \infty$ the decision making under Centralization and under Decentralization converge: under both structures the decisions are set equal to the average posterior. It is then not surprising that the quality of communication also converges.

5.4 Organizational Performance

We can now state the expected profits for each organizational structure.

PROPOSITION 4 (Organizational Performance). *Under Centralization the expected profits are given by*

$$\Pi_C = -(A_C + (1 - A_C) S_C) (\sigma_1^2 + \sigma_2^2), \quad (14)$$

and under Decentralization they are given by

$$\Pi_D = -(A_D + B_D S_D) (\sigma_1^2 + \sigma_2^2), \quad (15)$$

where

$$A_C \equiv \frac{2\delta}{1 + 4\delta}, \quad A_D \equiv \frac{2(\lambda^2 + \delta)\delta}{(\lambda + 2\delta)^2} \quad \text{and} \quad B_D \equiv \delta^2 \frac{4\lambda^3 + 6\lambda^2\delta + 2\delta^2 - \lambda^2}{(\lambda + \delta)^2 (\lambda + 2\delta)^2}. \quad (16)$$

The proposition shows that under both organizational structures the expected profits are a linear function of the underlying uncertainty $(\sigma_1^2 + \sigma_2^2)$ and the sum of the residual variances $S_l (\sigma_1^2 + \sigma_2^2)$, $l = C, D$.

Since the HQ Manager’s decision making is efficient, the first best expected profits would be realized if she were perfectly informed, i.e. if $S_C = 0$. It then follows from (14) that the first best expected profits are given by $-A_C (\sigma_1^2 + \sigma_2^2)$. It is intuitive that these expected profits are decreasing in the need for coordination and are independent of the division managers’ biases.

The expected profits under Centralization differ from the first best benchmark since the HQ Manager is in general not perfectly informed. In particular, the more biased the division managers are towards their own divisions, the less information is communicated and thus the lower the expected profits. The expected profits are also decreasing in δ : not only does an increase in the need for coordination lead to worse communication but it also reduces the expected profits for any given communication quality.

The expected profits under Decentralization differ from the first best both because of imperfect communication and because the division managers’ decision making is biased. Since an increase in the own-division bias leads to worse communication and to more biased decision making it clearly reduces the expected profits. The impact of an increase in δ , in contrast, is ambiguous: it improves communication but it reduces the expected profits for any given communication quality.

6 Centralization versus Decentralization

We can now compare the performance of the two organizational structures. A clear advantage of a decentralized organization is that it puts in control those managers who are closest to the local information. In contrast, in a centralized organization some of this information is lost when it is communicated to the decision maker. Naturally, the lack of local information impairs the ability of a centralized organization to adapt decisions to the local conditions.

However, while Decentralization has an advantage at adapting decisions to local conditions, it has a disadvantage at ensuring that the decisions are coordinated. This is so for two reasons. First, the division managers do not fully internalize the need for coordination and, as such, put excessive weight on adapting their decisions to the local conditions. Second, effective coordination requires that the manager who makes one decision knows what the other decision is. Under Centralization this is naturally the case since both decisions are made by the same manager. In contrast, under Decentralization Manager $j = 1, 2$ is uncertain about what decision Manager $k \neq j$ will make since communication between them is imperfect. In sum, division managers lack both the right incentives and the right information to ensure effective coordination while headquarters lacks the local information to efficiently adapt decisions to the local conditions.

The next lemma shows that these factors translate into a *coordination advantage* of Cen-

tralization and an *adaptation advantage* of Decentralization. To state this lemma, let $AL_l = \mathbb{E} \left[(d_1^l - \theta_1)^2 + (d_2^l - \theta_2)^2 \right]$, $l = C, D$, denote the adaptation losses under the two organizational structures and let $CL_l = \mathbb{E} \left[(d_1^l - d_2^l)^2 \right]$ denote the coordination losses.

LEMMA 2. *For all $\lambda \in [1/2, 1]$ and $\delta \geq 0$, $AL_C \geq AL_D$ and $CL_C \leq CL_D$.*

This lemma supports the basic intuition that centralized structures perform relatively well in terms of coordination and decentralized structures perform relatively well in terms of adaptation. From this one could reason naively that Centralization will prevail if the need for coordination is sufficiently important. A key insight of this section is that this reasoning is flawed. To see this, consider the next proposition which compares the performance of the two organizational structures for $\lambda > 1/2$ and $\delta > 0$. For $\lambda = 1/2$ or $\delta = 0$ both organizational structures achieve the first best expected profits since, in this case, there is no incentive conflict between the managers.

PROPOSITION 5 (Centralization versus Decentralization). *Suppose that $\lambda > 1/2$ and $\delta > 0$:*

- i. δ small:* if $\delta \in (0, \bar{\delta})$, then Decentralization strictly dominates Centralization, where $\bar{\delta} > 0$ is defined in the proof.
- ii. δ large:* if $\delta \in (\bar{\delta}, \infty)$, then Decentralization strictly dominates Centralization for all $\lambda \in (1/2, \bar{\lambda}(\delta))$ and Centralization strictly dominates Decentralization for all $\lambda \in (\bar{\lambda}(\delta), 1]$, where $\bar{\lambda}(\delta) > 17/28$ is defined in the proof.
- iii. $\delta \rightarrow \infty$:* as $\delta \rightarrow \infty$, the expected profits under Decentralization and Centralization converge, i.e. $\lim_{\delta \rightarrow \infty} \Pi_D = \lim_{\delta \rightarrow \infty} \Pi_C$.

The proposition is illustrated in Figure 5, where the downward sloping curve represents $\bar{\lambda}(\delta)$. Part (i.) shows that while organizational structure is irrelevant when $\delta = 0$, Decentralization strictly dominates Centralization if the need for coordination is small but positive and this for any own-division bias $\lambda > 1/2$. Part (ii.) shows that in spite of the coordination advantage of Centralization, Decentralization can strictly outperform Centralization no matter how important coordination is. In particular, whenever $\lambda < 17/28$, Decentralization is optimal for any finite $\delta > 0$. This also implies that the Delegation Principle holds in our model. The Delegation Principle states that decision rights should be delegated to the managers with the most information provided that their incentives are sufficiently aligned and it is satisfied in our model since Decentralization strictly dominates Centralization for sufficiently small values of $\lambda > 1/2$.²⁸ Finally, Part (iii.) shows that in the limit in which coordination becomes all important the expected profits under the two organizational structures converge.

²⁸For the Delegation Principle see, for instance, Milgrom and Roberts (1992).

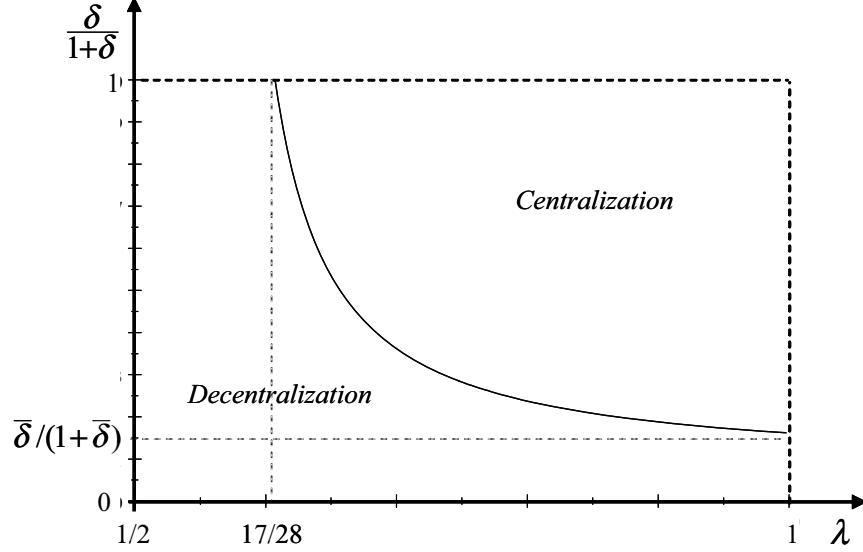


Figure 5: Organizational Performance

To understand these results, it is useful to decompose the effect of a change in δ on the expected profits of the two organizational structures:

$$\frac{d\Pi_l}{d\delta} = - \left(CL_l + \left(\frac{\partial AL_l}{\partial \delta} + 2\delta \frac{\partial CL_l}{\partial \delta} \right) + \left(\frac{\partial AL_l}{\partial S_l} + 2\delta \frac{\partial CL_l}{\partial S_l} \right) \frac{dS_l}{d\delta} \right) \text{ for } l = C, D. \quad (17)$$

The first term denotes the direct effect of an increase in δ on the expected profits, holding constant decision making and communication: an increase in δ puts more weight on the coordination loss and thus reduces the expected profits under both organizational structures. Since, from Lemma 2, $CL_C \leq CL_D$, this effect makes Centralization more attractive relative to Decentralization.

The second term reflects how an increase in δ affects decision making holding constant communication and the weight on the coordination loss in the profit function. Under Centralization this effect is equal to zero while under Decentralization it is negative for small values of δ and positive when δ is sufficiently big.²⁹ Intuitively, whereas for $\delta = 0$ division managers are unbiased, for $\delta > 0$ they put too little weight on coordination. For δ large enough, however, more interdependence aligns the incentives of the decision makers. Indeed, recall from Section 4 that in the limit as coordination becomes all important, decision making under Centralization and Decentralization converge.

²⁹The zero effect under Centralization is due to the Envelope Theorem: an increase in δ induces the HQ Manager to change her decision making so as to reduce the coordination loss CL_C and increase adaptation loss AL_C ; the change in decision making, however, does not have a first order effect on the expected profits.

Finally, the last term represents the impact of an increase in δ on communication, holding constant decision making and the weight on the coordination loss in the profit function. Under Centralization this effect is always negative: an increase in δ leads to worse communication which reduces the expected profits. In contrast, under Decentralization this effect is always positive since an increase in δ actually improves communication and thereby increases the expected profits. In a model with non-strategic communication the first two terms on the RHS of (17) would still be present but the third term would not. We show below that the fact that the quality of communication is endogenous in our model, and therefore the third term is in general not zero, is key for the results in Proposition 5.

δ small Consider first how the different effects play out when δ is small. Intuitively, even when the need for coordination is very small communication is noisy which affects the ability of the HQ Manager to adapt the decisions to the local conditions and the ability of division managers to coordinate them. Since for small δ coordination is much less important than adaptation, the centralized structure suffers more from imperfect communication than the decentralized one. As a result, Decentralization dominates Centralization.

To see this intuition more formally, note that for $\delta = 0$, $d_1 = \theta_1$ and $d_2 = \theta_2$ under both organizational structures. The coordination loss, and thus the first term on the RHS of (17), is then the same under both structures: $CL_C = CL_D = E[(\theta_1 - \theta_2)^2]$. The second term on the RHS of (17), which captures the effect of changes in the decision making on the expected profits, is also the same. This is so since, at $\delta = 0$, decision making is efficient under both organizational structures and as a result small changes in the decision making do not have a first order effect on the expected profits. The differences in the expected profits for small δ are therefore determined by how a change in the need for coordination affects communication, i.e. the third term on the RHS of (17). In particular, for $\delta = 0$

$$\begin{aligned} \frac{d\Pi_D}{d\delta} - \frac{d\Pi_C}{d\delta} &= \left(\frac{\partial AL_C}{\partial S_C} + 2\delta \frac{\partial CL_C}{\partial S_C} \right) \frac{dS_C}{d\delta} - \left(\frac{\partial AL_D}{\partial S_D} + 2\delta \frac{\partial CL_D}{\partial S_D} \right) \frac{dS_D}{d\delta} \\ &= \frac{\partial AL_C}{\partial S_C} \frac{dS_C}{d\delta} > 0. \end{aligned}$$

Under Centralization an increase in δ results in a deterioration in communication which limits the HQ Manager's ability to adapt the decisions to the local conditions. There is no such adaptation loss under Decentralization since division managers do not rely on communication for their adaptation decisions. Changes in the quality of communication also affect the ability of the decision makers to coordinate the decisions. This effect is given by the terms $2\delta \partial CL_l / \partial S_l$, $l = C, D$, in above

expression which is zero at $\delta = 0$. Thus, for small δ Decentralization dominates Centralization since noisy communication has a more detrimental effect on the HQ Manager's ability to adapt than on the division managers' abilities to coordinate.

δ large The quality of communication is also key for understanding the relative performance of the two organizational structures when coordination is very important. To see this, consider first what happens in the limit in which $\delta \rightarrow \infty$. We know from Section 4 that in this case the decision making under the two structures converge. Essentially, when coordination becomes sufficiently important the division managers recognize their interdependence and set both decisions equal to the average state, just like the HQ Manager. Since we know from Section 5 that the quality of communication also converges in the limit, it is intuitive that $\lim_{\delta \rightarrow \infty} \Pi_C = \lim_{\delta \rightarrow \infty} \Pi_D$, as shown in Part (iii.) of the above proposition. If the quality of communication did not converge, then Centralization would always dominate Decentralization in the limit. In particular, fixing the quality of communication at some level S_C and S_D and taking the limit of the expected profits as δ goes to infinity yields:

$$\lim_{\delta \rightarrow \infty} \Pi_D|_{S_D} = -\frac{1}{2} (1 + S_D) (\sigma_1^2 + \sigma_2^2) < -\frac{1}{2} (1 + S_C) (\sigma_1^2 + \sigma_2^2) = \lim_{\delta \rightarrow \infty} \Pi_C|_{S_C}, \quad (18)$$

where the inequality follows from $S_D > S_C$ for $\delta \in (0, \infty)$ and $\lambda \in (1/2, 1]$. Moreover, since the expected profits are continuous in δ , (18) implies that for sufficiently high but finite values of δ Centralization would always strictly dominate Decentralization if the quality of communication were held constant. Thus, both the fact that Decentralization can be optimal for any finite $\delta > 0$ and the fact that the two structures perform equally well when $\delta \rightarrow \infty$ are driven by the endogenous quality of communication. In other words, if communication were non-strategic, so that the third term on the RHS of (17) were zero, coordination would require centralization.

To shed more light on why Decentralization can be optimal for any finite $\delta > 0$, it is useful to decompose the effect of a change in λ on the expected profits of the two organizational structures as

$$\frac{d\Pi_l}{d\lambda} = \frac{\partial \Pi_l}{\partial \lambda} + \frac{\partial \Pi_l}{\partial S_l} \frac{dS_l}{d\lambda} \quad \text{for } l = C, D.$$

The first term relates to the impact of λ on decision making, keeping the communication constant, and the second term reflects the impact of a change in λ on the quality of communication. Recall that at $\lambda = 1/2$ both organizational structures achieve the first best expected profits. As a result, the first term is zero under both organizational structures when $\lambda = 1/2$. For small λ , therefore, the preferred organizational structure will be the structure which suffers least from the deterioration in the quality communication.

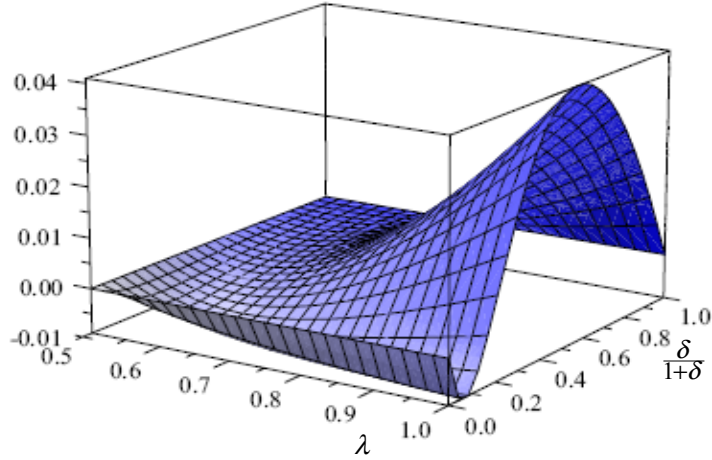


Figure 6: The Relative Value of Decentralization: $\Pi_C - \Pi_D$

Under both organizational structures even a very small own-division bias has a first order impact on communication quality, that is for $\lambda = 1/2$, $dS_l/d\lambda > 0$ for $l = C, D$. A deterioration in communication quality affects the ability of the HQ Manager to adapt to local information and it affects the ability of the division managers to coordinate their decisions. The impact of an information loss on the HQ Manager's ability to adapt, however, is bigger than its impact on the division managers' ability to coordinate. Indeed, one can verify that, for $\lambda = 1/2$,

$$\frac{\partial \Pi_C}{\partial S_C} \frac{dS_C}{d\lambda} < \frac{\partial \Pi_D}{\partial S_D} \frac{dS_D}{d\lambda} < 0.$$

It follows that for any finite $\delta > 0$, Decentralization strictly dominates Centralization when the own-division bias $\lambda > 1/2$ is sufficiently small. The fact that the Delegation Principle is satisfied in our model, therefore, is also due to the endogenous quality of communication.

6.1 The Relative Value of Centralization

So far we have focused on the information cost of centralization and have abstracted from any resource costs. It is evident, however, that setting up a headquarters and hiring a manager to run it involves additional resource costs, at the very least the opportunity cost of the additional manager. In this sub-section we briefly digress from our main analysis to investigate the comparison between Centralization and Decentralization when setting up a headquarters involves such a resource cost.

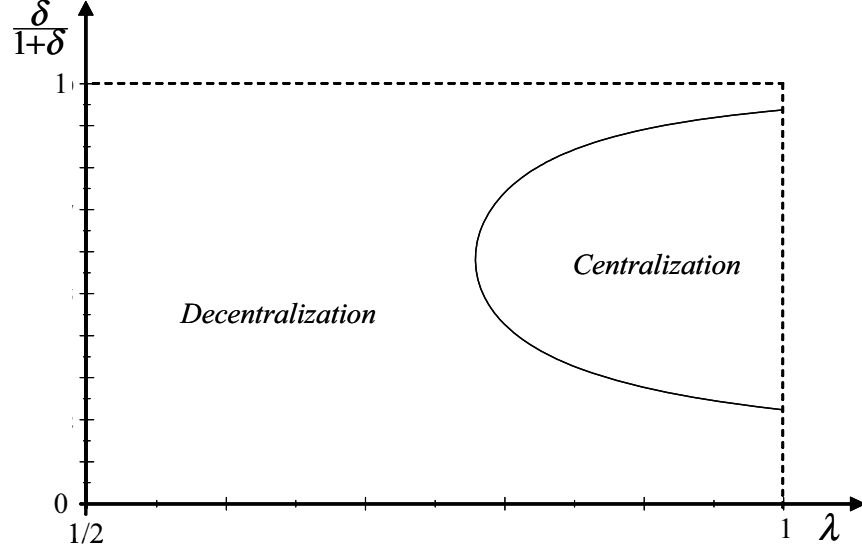


Figure 7: Organizational Performance with Resource Costs

For this purpose, consider first Figure 6 which plots $\Pi_C - \Pi_D$ for $(\sigma_1^2 + \sigma_2^2) = 1$.³⁰ Confirming our previous results, it can be seen that $\Pi_C - \Pi_D < 0$ if either the own-division bias or the need for coordination are small and that $\Pi_C - \Pi_D \geq 0$ otherwise. Suppose now that Centralization involves a resource cost, namely a fixed wage $w > 0$ for the HQ Manager. The expected profits under Centralization are then given by $\bar{\Pi}_C = \Pi_C - w$. It is evident that once we allow for a resource cost Decentralization will still dominate Centralization when either λ or δ are sufficiently small. However, now Decentralization also dominates Centralization for very high values of δ . This is the case since, as δ becomes very large, $\Pi_C - \Pi_D$ converges to zero which, of course, implies that $\bar{\Pi}_C - \Pi_D$ becomes negative. To see this graphically, consider again Figure 6 and note that $\bar{\Pi}_C - \Pi_D$ can be obtained by simply shifting $\Pi_C - \Pi_D$ down by w .

Consider next Figure 7 which compares the performance of the two organizational structures for $w = 0.01$. The figure illustrates two key features of the model with resource costs. First, Decentralization dominates Centralization when the need for coordination is either very small or very big and, provided that the own-division bias is sufficiently big, Centralization dominates for intermediate values of δ . Second, the effect of an increase in δ on the dominant organizational structure is ambiguous: an increase in δ can make it optimal to switch from Decentralization to Centralization or vice versa.

³⁰Since both Π_C and Π_D are proportional to $(\sigma_1^2 + \sigma_2^2)$ setting $(\sigma_1^2 + \sigma_2^2) = 1$ is without loss of generality.

6.2 Information Aggregation

Headquarters are often decried as costly and badly informed bureaucracies that needlessly interfere with the decisions of better informed division managers. This sentiment reflects the natural information advantage of division managers who learn relevant information directly through their involvement in the day-to-day operations of an organization while managers at headquarters rely on the communication of this information by the division managers. However, since operating managers are more willing to share information with headquarters than with each other, managers at headquarters enjoy a counteracting communication advantage. As a result, a headquarters manager who does not observe any information directly may still be a better aggregator of information than a division manager who does observe some information himself but also relies on communication to obtain other information. In this section we investigate whether this can indeed be the case. So far we allowed for differences in the variances of the local conditions σ_1^2 and σ_2^2 . For expositional convenience we now restrict attention to the case in which the two division managers are equally informed, i.e. $\sigma_1^2 = \sigma_2^2$, and define $\sigma^2 \equiv \sigma_1^2 = \sigma_2^2$.

There are many ways in which one could measure how well a manager aggregates information. A natural measure, and the one we focus on, is the average accuracy of a manager's posterior beliefs about each state as given by the average residual variance of θ_1 and θ_2 . We know from the analysis above that under Centralization the average residual variance of the HQ Manager is given by $S_C\sigma^2$ while under Decentralization the average residual variance of each division manager is given by $S_D\sigma^2/2$. The information advantage of the division managers is reflected by the residual variance of their own state being zero while the HQ Manager's communication advantage is reflected by the fact that $S_C \leq S_D$. The next proposition shows under what conditions the communication advantage dominates the information advantage.

PROPOSITION 6 (Information Aggregation). *Suppose that $\sigma_1^2 = \sigma_2^2$. Then $S_C \leq S_D/2$ for all $\delta \in [0, \delta_1]$ and $S_C \geq S_D/2$ for all $\delta \in [\delta_1, \infty)$, where*

$$\delta_1 \equiv \frac{\lambda}{8\lambda - 1} \left(2 - \lambda + \sqrt{20\lambda + \lambda^2 + 1} \right) > 0.$$

Thus, the HQ Manager is a better aggregator if and only if coordination is not very important. In light of our previous discussion the intuition for this result is straightforward: for small δ the HQ Manager's communication advantage is significant and dominates the division managers' information advantage; as δ increases, however, the communication advantage continuously shrinks and, eventually, is dominated by the information advantage.

7 Asymmetric Organizations

A key feature of our model is the symmetry of the organization: the two divisions are of equal size, they have the same need for coordination and the two decisions are made simultaneously. In this section we relax the symmetry assumptions and investigate how asymmetries between the divisions affect the relative performance of centralized and decentralized organizations. We show that such asymmetries tend to favor centralization when coordination is very important. However, provided that the differences between the divisions are not too big, it is still the case that Decentralization can be optimal even if the need for coordination is arbitrarily high. Our key results therefore continue to hold, albeit in a weaker form. To streamline the exposition, and in contrast to our previous analysis, we now assume that both division managers have the same amount of private information, in the sense that $\sigma_1^2 = \sigma_2^2$. Also, we relegate the formal analysis to the appendix.

7.1 Sequential Decision Making

The first asymmetry we consider concerns the timing of decision making under Decentralization. In many settings one of the division managers may have a first mover advantage, that is he may be able to make his decision before the other division manager can do so. To explore this possibility, we now consider a version of our model in which the decision making under Decentralization takes place sequentially. In particular, suppose that Manager 1 is the Leader and Manager 2 the Follower. After both division managers have observed their local conditions, the Follower sends a single message to the Leader. Next the Leader makes his decision, the Follower observes it and then makes his own.

At first it would seem that sequential decision making facilitates coordination since it eliminates the Follower's uncertainty about the Leader's decision. This would suggest that allowing for sequential decision making further strengthens the result that Decentralization can be optimal even when coordination is very important. There are, however, two countervailing effects. First, under sequential decision making the Leader is more uncertain about the Follower's decision than he is under simultaneous decision making. This is the case since the Follower communicates less information to the Leader than he does to a peer. Second, when decision making takes place sequentially the Leader adapts his decision more closely to his local conditions since he knows that the Follower will be forced to adjust his decision accordingly. As a result, decisions can be less coordinated under sequential than under simultaneous decision making. The next proposition investigates how these different effects play out when either coordination is very important or the own-division bias is very small.

PROPOSITION 7 (Sequential Decision Making).

- i.* For any $\lambda \in (1/2, 1]$ Centralization strictly dominates Decentralization with sequential decision making when coordination is sufficiently important.
- ii.* For any $\delta \in (0, \infty)$ Decentralization with sequential decision making strictly dominates Centralization when the own-division bias $\lambda > 1/2$ is sufficiently small.

Part (i.) highlights a key difference between the model with sequential decision making and the symmetric model while Part (ii.) highlights a key similarity. In particular, Part (i.) shows that when decision making takes place sequentially and coordination is very important, then efficient coordination does require centralization. This is the case since even in the limit in which $\delta \rightarrow \infty$ decision making under Decentralization is biased and headquarters' communication advantage does not vanish. In contrast, when decision making takes place simultaneously then, in the limit, decision making and communication are as efficient under Decentralization as under Centralization and the two structures perform equally well.

However, it is still the case that, no matter how important coordination, Decentralization strictly dominates Centralization if the own-division bias $\lambda > 1/2$ is sufficiently small, as shown in Part (ii.) of the proposition. The reason is the same as when decision making takes place simultaneously: for λ close to $1/2$ the information advantage of the decentralized structure dominates the communication advantage of the centralized one. The central insight of the symmetric model therefore continues to hold in this extension, albeit in a weaker form. In other words, coordination does not necessarily require centralization, even when decisions are made sequentially.

7.2 Different Needs for Coordination

In many organizations divisions differ in how much their profits depend on coordination. To allow for this possibility we now extend the model considered in Section 3 by assuming that the weight that Division 1 puts on coordination is given by $\delta_1 \in [0, \infty)$ while that of Division 2 is given by $\delta_2 \in [0, \infty)$. As in the previous sub-section, we again focus on the relative performance of the two organizational structures when either coordination is very important or the own-division bias is very small:³¹

PROPOSITION 8 (Different Needs for Coordination).

- i.* For any $\lambda \in (1/2, 1]$ and $\delta_j \in [0, \infty)$, $j = 1, 2$, Centralization strictly dominates Decen-

³¹For a full analysis of this model when $\lambda = 1$ see Rantakari (2006).

tralization when coordination is sufficiently important for Division $k \neq j$.

- ii. For any $\delta_1, \delta_2 \in (0, \infty)$ Decentralization strictly dominates Centralization when the own-division bias $\lambda > 1/2$ is sufficiently small.

Part (i.) again highlights a key difference with the symmetric model while Part (ii.) highlights a key similarity. In particular, Part (i.) shows that when coordination becomes very important for one of the divisions but not the other, then efficient coordination does require centralization. In spite of this fact, however, the central result of the symmetric model continues to hold, as shown in Part (ii.): no matter how important coordination is for each division, Decentralization dominates Centralization when the own-division bias $\lambda > 1/2$ is sufficiently small. Thus, even when the divisions have different needs for coordination, coordination does not necessarily require centralization.

7.3 Different Division Sizes

Finally, we allow for differences in the size of the two divisions. In particular, we now extend the model considered in Section 3 by assuming that the profits of Divisions 1 and 2 are given by $2\alpha\pi_1$ and $2(1 - \alpha)\pi_2$ respectively, where π_1 and π_2 are defined in (1) and (2) and $\alpha \in [1/2, 1)$ is a parameter that measures the relative size of the two divisions. Once again we focus on the performance of the two organizational structures when either the need for coordination is very big or the own-division bias is very small:

PROPOSITION 9 (Different Division Sizes).

- i. For any $\lambda > 1/2$ and $\alpha > 1/2$ Centralization strictly dominates Decentralization when coordination is sufficiently important.
- ii. For any $\delta \in (0, \infty)$ Decentralization strictly dominates Centralization when the own-division bias $\lambda > 1/2$ and the difference in the division sizes $\alpha > 1/2$ are sufficiently small.

Thus, as in the two previous extensions, Centralization strictly dominates Decentralization for any $\lambda > 1/2$ when coordination is sufficiently important. Moreover, provided that the difference in the division sizes is sufficiently small, it is again the case that for any finite δ , Decentralization strictly dominates Centralization if the own-division bias $\lambda > 1/2$ is sufficiently small. If, however, the difference in the division sizes is ‘too big,’ then Centralization can be optimal even if the incentives of the division managers are extremely closely aligned, i.e. if $\lambda > 1/2$ is arbitrarily close to $1/2$. The extent to which coordination requires centralization therefore depends on the relative size of the divisions. Essentially, the bigger the size difference between the divisions, the more likely it is that coordination requires centralization.

8 Conclusions

When does coordination require centralization? In this paper we addressed this question in a model in which information about the costs of coordination is dispersed among division managers who communicate strategically to promote their own divisions at the expense of the overall organization. We showed that vertical communication is more efficient than horizontal communication: division managers share more information with an unbiased headquarters than they do with each other. However, headquarters' communication advantage diminishes, and eventually vanishes, as coordination becomes more important. As a result, decentralization can be optimal even if coordination is very important. In particular, this is so if the incentives of the division managers are sufficiently aligned and the divisions are not too different from each other. Thus, only in organizations in which divisions are quite heterogeneous, for instance in their relative size, or in which implicit and explicit incentives strongly bias division managers towards their own divisions, does coordination require centralization.

The analysis of our model is surprisingly simple and, arguably, more tractable than the leading example in Crawford and Sobel (1982), the traditional workhorse for cheap talk applications. This is so since in our setting the efficient communication equilibrium is one in which the number of intervals goes to infinity. As a result, one can avoid the integer problem that is associated with the finite interval equilibria in the standard model. We hope that our framework will prove useful in approaching various applied problems ranging from organizational design, the theory of the firm to fiscal federalism.

Our central result – that decentralization can be optimal even if coordination is very important – is reminiscent of Hayek (1945) and the many economists since who have invoked the presence of 'local knowledge' as a reason to decentralize decision making in organizations. In the management literature, this view has become known as the Delegation Principle.³² Our insight differs from the standard Delegation Principle rationale on two important dimensions.

First, the standard argument posits that efficient decision making requires the delegation of decision rights to those managers who possess the relevant information. The alternative – communicating the relevant information to those who possess the decision rights – is discarded on the basis of physical communication constraints (Jensen and Meckling 1992). Is the local knowledge argument still relevant in a world in which technological advances have slashed communication costs? Our analysis suggests that it is: even in the absence of any physical communication costs, local knowledge remains a powerful force for decentralization. The same factors that make decentraliza-

³²See, for instance, Milgrom and Roberts (1992).

tion of decision rights unattractive – biased incentives of the local managers – are even more harmful for the transfer of information. As long as division managers are not too biased, the distortion of information in a centralized organization outweighs the loss of control under decentralization.

Second, the standard rationale for the Delegation Principle presumes that decentralization eliminates the need for communication. If, however, decision relevant information is dispersed among multiple managers – as is likely to be the case when decisions need to be coordinated – then efficient decision making always requires communication. Our analysis has highlighted that while local knowledge gives division managers an information advantage, headquarters has a communication advantage. As a result, headquarters can actually be a more efficient aggregator of dispersed information. Despite the need for information aggregation, we show that decentralization is often preferred, even if coordination is very important. The reason is that as coordination and information aggregation become more important, division managers communicate more efficiently with each other and headquarters' communication advantage diminishes and eventually disappears.

Beyond its implication for organizational design, it is tempting to interpret our theory as one of horizontal firm boundaries. In order to have a compelling theory of horizontal integration, however, it would be necessary to combine our approach with the possibility of some bargaining between division managers or between division managers and headquarters. It would further be important to endogenize managerial incentives, as managers in independent firms may have different objectives than those belonging to the same organization. These issues, and others, await future research.

9 Appendix A

We first define the random variable \bar{m}_i as the posterior expectation of the state θ_i by the receiver of message m_i . The following lemma will be used throughout the appendix.

LEMMA A1. For any communication equilibrium considered in Proposition 1 $E_{\theta_2}[\bar{m}_1\bar{m}_2] = E_{\theta_2}[\theta_1\bar{m}_2] = E_{\theta_2}[\theta_2] = E_{\theta_2}[\bar{m}_2] = E_{\theta_2}[\theta_1\theta_2] = 0$ and $E_{\theta_1}[\bar{m}_1\bar{m}_2] = E_{\theta_1}[\theta_2\bar{m}_1] = E_{\theta_1}[\theta_1] = E_{\theta_1}[\bar{m}_1] = E_{\theta_2}[\theta_1\theta_2] = 0$.

Proof: All equalities follow from independence of θ_1 and θ_2 and that in equilibrium $E_{\theta_i}[\bar{m}_i] = E_{\theta_i}[\theta_i] = 0$. ■

Proof of Proposition 1: We first note that in any Perfect Bayesian Equilibrium of the communication game, optimal decisions given beliefs satisfy (3) and (4) for the case of Centralization and (8) and (9) for the case of Decentralization. Now we will establish that communication rules in equilibrium are interval equilibria.

For the case of Centralization let $\mu_2(\cdot)$ be any communication rule for Manager 2. The expected utility of Manager 1 if the Headquarter Manager holds a posterior expectation ν_1 of θ_1 is given by

$$E_{\theta_2}[U_1 | \theta_1, \nu] = -E_{\theta_2} \left[\lambda \left(\hat{d}_1^C - \theta_1 \right)^2 + (1 - \lambda) \left(\hat{d}_2^C - \theta_2 \right)^2 + \delta \left(\hat{d}_1^C - \hat{d}_2^C \right)^2 \right], \quad (19)$$

with

$$\hat{d}_1^C \equiv \gamma_C \nu_1 + (1 - \gamma_C) E[\theta_2 | \mu_2(\cdot)] \quad (20)$$

$$\hat{d}_2^C \equiv (1 - \gamma_C) \nu_1 + \gamma_C E[\theta_2 | \mu_2(\cdot)]. \quad (21)$$

It is readily seen that $\frac{\partial^2}{\partial \theta_1 \partial \nu_1} E_{\theta_2}[U_1 | \theta_1, \nu_1] > 0$ and $\frac{\partial^2}{\partial \theta_1^2} E_{\theta_2}[U_1 | \theta_1, \nu_1] < 0$. This implies that for any two different posterior expectations of the Headquarter Manager, say $\underline{\nu}_1 < \bar{\nu}_1$, there is at most one type of Manager 1 that is indifferent between both. Now suppose that contrary to the assertion of interval equilibria there are two states $\theta_1^1 < \theta_1^2$ such that $E_{\theta_2}[U_1 | \theta_1^1, \bar{\nu}_1] \geq E_{\theta_2}[U_1 | \theta_1^1, \underline{\nu}_1]$ and $E_{\theta_2}[U_1 | \theta_1^2, \underline{\nu}_1] > E_{\theta_2}[U_1 | \theta_1^2, \bar{\nu}_1]$. But then $E_{\theta_2}[U_1 | \theta_1^2, \bar{\nu}_1] - E_{\theta_2}[U_1 | \theta_1^2, \underline{\nu}_1] < E_{\theta_2}[U_1 | \theta_1^1, \bar{\nu}_1] - E_{\theta_2}[U_1 | \theta_1^1, \underline{\nu}_1]$ which violates $\frac{\partial^2}{\partial \theta_1 \partial \nu_1} E_{\theta_2}[U_1 | \theta_1, \nu_1] > 0$. The same argument can be applied to Manager 2 for any reporting strategy $\mu_1(\cdot)$ of Manager 1. Therefore all equilibria of the communication game under Centralization must be interval equilibria.

For the case of Decentralization let $\mu_1(\cdot)$ and $\mu_2(\cdot)$ be communication rules of Manager 1 and Manager 2, respectively. Sequential rationality implies that in equilibrium decision rules must conform to (8) and (9). If ν_1 denotes the expectation of the posterior of θ_1 that Manager 2 holds then $\frac{\partial^2}{\partial \theta_1 \partial \nu_1} E_{\theta_2}[U_1 | \theta_1, \nu_1] > 0$ and the proof follows as in the preceding paragraph.

We now characterize all finite equilibria of the communication game, i.e. equilibria that induce a finite number of different decisions. For this purpose for Manager j let a_j be a partition of $[-s_j, s_j]$, any message $m_j \in (a_{j,i-1}, a_{j,i})$ be denoted by $m_{j,i}$ and $\bar{m}_{j,i}$ be the receiver's posterior belief of the expected value of θ_j after receiving message $m_{j,i}$.

a. Centralization: In state $a_{1,i}$ Manager 1 must be indifferent between sending a message that induces a posterior $\bar{m}_{1,i}$ and a posterior $\bar{m}_{1,i+1}$ so that $E_{\theta_2} [U_1 | a_{1,i}, \bar{m}_{1,i}] - E_{\theta_2} [U_1 | a_{1,i}, \bar{m}_{1,i+1}] = 0$. Using Lemma A1 on (19) we have that

$$\begin{aligned} & E_{\theta_2} [U_1 | a_{1,i}, \bar{m}_{1,i}] - E_{\theta_2} [U_1 | a_{1,i}, \bar{m}_{1,i+1}] \\ &= \lambda (\gamma_C \bar{m}_{1,i+1} - a_{1,i})^2 + (1 - \lambda) (1 - \gamma_C)^2 \bar{m}_{1,i+1}^2 + \delta (2\gamma_C - 1)^2 \bar{m}_{1,i+1}^2 \\ & \quad - \left(\lambda (\gamma_C \bar{m}_{1,i} - a_{1,i})^2 + (1 - \lambda) (1 - \gamma_C)^2 \bar{m}_{1,i}^2 + \delta (2\gamma_C - 1)^2 \bar{m}_{1,i}^2 \right). \end{aligned}$$

Substituting $\bar{m}_{1,i} = (a_{1,i-1} + a_{1,i})/2$ we have that $E_{\theta_2} [U_1 | a_{1,i}, \bar{m}_{1,i}] - E_{\theta_2} [U_1 | a_{1,i}, \bar{m}_{1,i+1}] = 0$ if and only if $a_{1,i} = (a_{1,i-1} + a_{1,i+1}) / (2 + 4b_C)$, where b_C is defined in (11). Rearranging this expression, we get

$$a_{1,i+1} - a_{1,i} = a_{1,i} - a_{1,i-1} + 4b_C a_{1,i}. \quad (22)$$

Let the total number of elements of a_1 be N . Equilibrium with an even number of elements would correspond to the case $N = 2N_1$, and $N = 2N_1 + 1$ when the equilibrium has an odd number of elements. Using the boundary conditions $a_{1,0} = -s_1$ and $a_{1,N} = s_1$ to solve the difference equation (22) gives

$$a_{1,i} = \frac{s_1}{(x_C^N - y_C^N)} (x_C^i (1 + y_C^N) - y_C^i (1 + x_C^N)) \text{ for } 0 \leq i \leq N, \quad (23)$$

where the roots x_C and y_C are given by $x_C = (1 + 2b_C) + \sqrt{(1 + 2b_C)^2 - 1}$ and $y_C = (1 + 2b_C) - \sqrt{(1 + 2b_C)^2 - 1}$ and satisfy $x_C y_C = 1$, with $x_C > 1$. It is readily seen that for each $k \leq N$, $a_{1,k} = a_{1,N-k}$, i.e. the intervals are symmetrically distributed around zero. When $N = 2N_1$, i.e. the partition has an even number of elements, then $a_{1, \frac{N}{2}} = 0$. In this case we can compactly write (23) as

$$a_{1,i} = \frac{s_1}{(x_C^{N_1} - y_C^{N_1})} (x_C^i - y_C^i) \text{ for } 0 \leq i \leq N_1.$$

For an equilibrium with an odd number of elements $N = 2N_1 + 1$ there is a symmetric interval around zero where the Headquarter Manager's expected posterior of θ_1 is zero. In this case we can compactly write (23) as

$$a_{1,i} = \frac{s_1}{(x_C^{2N_1+1} - y_C^{2N_1+1})} (x_C^i (x_C^{N_1} + y_C^{N_1+1}) - y_C^i (x_C^{N_1+1} + y_C^{N_1})) \text{ for } 1 \leq i \leq N_1.$$

The analysis for Manager 2 is analogous.

b. Decentralization: If Manager 1 observes state θ_1 and sends message $m_{1,i}$ that induces a posterior belief $\bar{m}_{1,i}$ in Manager 2 his expected utility is given by

$$\mathbb{E}_{\theta_2} [U_1 | \theta_1, \bar{m}_{1,i}] = -\mathbb{E}_{\theta_2} \left[\lambda (d_1^D - \theta_1)^2 + (1 - \lambda) (d_2^D - \theta_2)^2 + \delta (d_1^D - d_2^D)^2 \right], \quad (24)$$

where d_1^D and d_2^D are given by (8) and (9). It must again be the case that $\mathbb{E}_{\theta_2} [U_1 | a_{1,i}, \bar{m}_{1,i}] - \mathbb{E}_{\theta_2} [U_1 | a_{1,i}, \bar{m}_{1,i+1}] = 0$. Making use of Lemma A1 on (24) we have that

$$\begin{aligned} & \mathbb{E}_{\theta_2} [U_1 | a_{1,i}, m_{1,i}] - \mathbb{E}_{\theta_2} [U_1 | a_{1,i}, m_{1,i+1}] \\ &= \left(\lambda \delta^2 \frac{(\bar{m}_{1,i+1} \delta - (\lambda + 2\delta) a_{1,i})^2}{(\lambda + \delta)^2 (\lambda + 2\delta)^2} + (1 - \lambda) \frac{\delta^2 \bar{m}_{1,i+1}^2}{(\lambda + 2\delta)^2} + \delta \frac{\lambda^2 (\bar{m}_{1,i+1} \delta - (\lambda + 2\delta) a_{1,i})^2}{(\lambda + \delta)^2 (\lambda + 2\delta)^2} \right) \\ & \quad - \left(\lambda \delta^2 \frac{(\bar{m}_{1,i} \delta - (\lambda + 2\delta) a_{1,i})^2}{(\lambda + \delta)^2 (\lambda + 2\delta)^2} + (1 - \lambda) \frac{\delta^2 \bar{m}_{1,i}^2}{(\lambda + 2\delta)^2} + \delta \frac{\lambda^2 (\bar{m}_{1,i} \delta - (\lambda + 2\delta) a_{1,i})^2}{(\lambda + \delta)^2 (\lambda + 2\delta)^2} \right). \end{aligned}$$

Substituting $\bar{m}_{1,i} = (a_{1,i-1} + a_{1,i}) / 2$ and $\bar{m}_{1,i+1} = (a_{1,i} + a_{1,i+1}) / 2$ we have that $\mathbb{E}_{\theta_2} [U_1 | a_{1,i}, m_{1,i}] - \mathbb{E}_{\theta_2} [U_1 | a_{1,i}, m_{1,i+1}] = 0$ if and only if $a_{1,i} = (a_{1,i-1} + a_{1,i+1}) / (2 + 4b_D)$, where b_D is defined in (12). Rearranging this expression, we get

$$a_{1,i+1} - a_{1,i} = a_{1,i} - a_{1,i-1} + 4b_D a_{1,i}. \quad (25)$$

Using the boundary conditions $a_{1,0} = -s_1$ and $a_{1,N} = s_1$ to solve the difference equation (25) gives

$$a_{1,i} = \frac{s_1}{(x_D^N - y_D^N)} (x_D^i (1 + y_D^N) - y_D^i (1 + x_D^N)) \text{ for } j = 1, 2 \text{ and } 0 \leq i \leq N, \quad (26)$$

where the roots x_D and y_D are given by $x_D = (1 + 2b_D) + \sqrt{(1 + 2b_D)^2 - 1}$ and $y_D = (1 + 2b_D) - \sqrt{(1 + 2b_D)^2 - 1}$ and satisfy $x_D y_D = 1$, with $x_D > 1$. It is readily seen that for each $k \leq N$, $a_{1,k} = a_{1,N-k}$, i.e. the intervals are symmetrically distributed around zero. When $N = 2N_1$, i.e. the partition has an even number of elements, then $a_{1, \frac{N}{2}} = 0$. In this case we can compactly write (26) as

$$a_{1,i} = \frac{s_1}{(x_D^{N_1} - y_D^{N_1})} (x_D^i - y_D^i) \text{ for } 0 \leq i \leq N_1.$$

For an equilibrium with an odd number of elements $N = 2N_1 + 1$ there is a symmetric interval around zero where the Headquarter Manager's expected posterior is zero. In this case we can compactly write (26) as

$$a_{1,i} = \frac{s_1}{(x_D^{2N_1+1} - y_D^{2N_1+1})} (x_D^i (x_D^{N_1} + y_D^{N_1+1}) - y_D^i (x_D^{N_1+1} + y_D^{N_1})) \text{ for } 1 \leq i \leq N_1.$$

The analysis for Manager 2 is analogous. ■

For the proof of Proposition 2 and Lemma 1 we will make use of the following lemma:

LEMMA A2. i. $E[\bar{m}_j \theta_j] = E[\bar{m}_j^2]$ for $j = 1, 2$. ii. $E[\bar{m}_j^2]$ is strictly increasing in the number of intervals N_j for $j = 1, 2$. and

$$E[\bar{m}_j^2] = \frac{s_j^2}{4} \left[\frac{(x_C^{3N_j} - 1)(x_C - 1)^2}{(x_C^{N_j} - 1)^3 (x_C^2 + x_C + 1)} - \frac{x_C^{N_j} (x_C + 1)^2}{x_C (x_C^{N_j} - 1)^2} \right].$$

Proof: i. Given the equilibrium reporting strategies we have that $\bar{m}_j = E[\theta_j | \theta_j \in (a_{j,i-1}, a_{j,i})]$ for $j = 1, 2$, which implies

$$E[\bar{m}_j \theta_j] = E[E[\bar{m}_j \theta_j | \theta_j \in (a_{j,i-1}, a_{j,i})]] = E[\bar{m}_j E[\theta_j | \theta_j \in (a_{j,i-1}, a_{j,i})]] = E[\bar{m}_j^2].$$

ii. Using (23) we can compute $E(\bar{m}_j^2)$ as follows

$$\begin{aligned} E[\bar{m}_j^2] &= \frac{1}{2s_j} \sum_{N_j} \int_{a_{j,i-1}}^{a_{j,i}} \left(\frac{a_{j,i} + a_{j,i-1}}{2} \right)^2 d\theta_j = \frac{1}{8s_1} \sum_{N_j} (a_{j,i} - a_{j,i-1}) (a_{j,i} + a_{j,i-1})^2 \\ &= \frac{s_j^2}{8 (x_C^{N_j} - y_C^{N_j})^3} \sum_{N_j} \left\{ (1 + y_C^{N_j})^3 x_C^{3(i-1)} (x_C + 1)^2 (x_C - 1) \right. \\ &\quad + (1 + x_C^{N_j})^3 y_C^{3(i-1)} (y_C + 1)^2 (1 - y_C) + (1 + x_C^{N_j})^3 y_C^{3(i-1)} (y_C + 1)^2 (1 - y_C) \\ &\quad \left. + (1 + y_C^{N_j}) (1 + x_C^{N_j})^2 y_C^{i-1} (y_C^2 - 1) (x_C + 1) - (1 + y_C^{N_j})^2 (1 + x_C^{N_j}) x_C^{i-1} (x_C^2 - 1) (y_C + 1) \right\}. \end{aligned}$$

Performing the summation in the above expression and using the fact that $x_C y_C = 1$ we get after some lengthy calculations that

$$\begin{aligned} E[\bar{m}_j^2] &= \frac{s_j^2}{4} \left[\frac{(x_C^{3N_j} - 1)(x_C - 1)^2}{(x_C^{N_j} - 1)^3 (x_C^2 + x_C + 1)} - \frac{(x_C^{N_j} + 1)^2 (y_C + 1)(x_C + 1)}{x_C^{N_j} (x_C^{N_j} - y_C^{N_j})^2} \right] = \quad (27) \\ &= \frac{s_j^2}{4} \left[\frac{(x_C^{3N_j} - 1)(x_C - 1)^2}{(x_C^{N_j} - 1)^3 (x_C^2 + x_C + 1)} - \frac{x_C^{N_j} (x_C + 1)^2}{x_C (x_C^{N_j} - 1)^2} \right]. \end{aligned}$$

To see that $E[\bar{m}_j^2]$ is strictly increasing in N_j first define

$$f(p) = \frac{p^3 - 1}{(p - 1)^3} \frac{(x_C - 1)^2}{(x_C^2 + x_C + 1)} - \frac{p}{(p - 1)^2} \frac{(x_C + 1)^2}{x_C}$$

and note that

$$\begin{aligned} f'(p) &= -3 \frac{p+1}{(p-1)^3} \frac{(x_C - 1)^2}{(x_C^2 + x_C + 1)} + \frac{p+1}{(p-1)^3} \frac{(x_C + 1)^2}{x_C} = \\ &= \frac{p+1}{(p-1)^3} \frac{10x_C^2 + x_C^4 + 1}{x_C(x_C^2 + x_C + 1)}. \end{aligned}$$

Therefore $f'(p) > 0$ for $p > 1$. Since $E[\bar{m}_j^2] = \frac{s_j^2}{4} f(x_C^{N_j})$ this establishes that $E[\bar{m}_j^2]$ is strictly increasing in N_j . ■

Proof of Proposition 2: First we establish that the limit of strategy profiles and beliefs $(\mu_1(\cdot), \mu_2(\cdot), d_1(\cdot), d_2(\cdot), g_1(\cdot), g_2(\cdot))$ as $N_1, N_2 \rightarrow \infty$ denoted by $(\mu_1^\infty(\cdot), \mu_2^\infty(\cdot), d_1^\infty(\cdot), d_2^\infty(\cdot), g_1^\infty(\cdot), g_2^\infty(\cdot))$ is indeed a Perfect Bayesian Equilibrium of the communication game, both under Centralization and under Decentralization. To this end we need only verify that the reporting strategies of Manager 1 and Manager 2 are a best response to $(d_1^\infty(\cdot), d_2^\infty(\cdot))$. Now suppose there is a θ_j that induces an expected posterior \bar{m}_1 in the decision maker and has a profitable deviation by inducing a different expected posterior \bar{m}_2 associated with state θ'_j . But then there exists a finite N_j such that θ_j has a profitable deviation by inducing the same posterior that θ'_j contradicting the fact that strategy profiles and beliefs $(\mu_1(\cdot), \mu_2(\cdot), d_1(\cdot), d_2(\cdot), g_1(\cdot), g_2(\cdot))$ constitute an equilibrium for all finite N_j . Thus θ_j cannot profitably deviate and therefore the reporting strategies $(\mu_1^\infty(\cdot), \mu_2^\infty(\cdot))$ are a best response to $(d_1^\infty(\cdot), d_2^\infty(\cdot))$. The fact that the expected profit under the equilibrium with an infinite number of elements for Manager 1 and Manager 2 coincides with the limit of the expected profit obtains by applying the Lebesgue Dominated Convergence Theorem to the expression of Π_l with $l = \{C, D\}$.

Finally, we show that the equilibrium $(\mu_1^\infty(\cdot), \mu_2^\infty(\cdot), d_1^\infty(\cdot), d_2^\infty(\cdot), g_1^\infty(\cdot), g_2^\infty(\cdot))$ yields higher total expected profits than all finite equilibria.

a. Centralization: The expected profits under Centralization are given by

$$\Pi_C = -E \left[(d_1^C - \theta_1)^2 + (d_2^C - \theta_2)^2 + 2\delta (d_1^C - d_2^C)^2 \right]. \quad (28)$$

Using (3), (4) and Lemmas A1 and A2-i. we have that for given N_1, N_2

$$\begin{aligned} E \left[(d_1^C - \theta_1)^2 \right] &= \sigma_1^2 - \gamma_C (2 - \gamma_C) E_{\theta_1}(\bar{m}_1^2) + (1 - \gamma_C)^2 E_{\theta_2}(\bar{m}_2^2) \\ E \left[(d_2^C - \theta_2)^2 \right] &= \sigma_2^2 + (1 - \gamma_C)^2 E_{\theta_1}(\bar{m}_1^2) - \gamma_C (2 - \gamma_C) E_{\theta_2}(\bar{m}_2^2) \\ E \left[(d_1^C - d_2^C)^2 \right] &= (2\gamma_C - 1)^2 (E_{\theta_1}(\bar{m}_1^2) + E_{\theta_2}(\bar{m}_2^2)). \end{aligned} \quad (29)$$

The rate at which total profits change with the variance of the messages $E_{\theta_j} \left(\overline{m}_j^2 \right)$ is given by

$$\frac{\partial \Pi_C}{\partial E_{\theta_j} \left[\overline{m}_j^2 \right]} = (1 + 2\delta) (2\gamma_C - 1).$$

Since $\gamma_C > \frac{1}{2}$ we have that $\frac{\partial \Pi_C}{\partial E_{\theta_j} \left[\overline{m}_j^2 \right]} > 0$. From Lemma A.2-ii we have that $E_{\theta_j} \left[\overline{m}_j^2 \right]$ increases with N_j and therefore expected profits Π_C increase as the number of elements N_j in the partition of Manager j increases.

b. Decentralization: The expected profits under Decentralization are given by

$$\Pi_D = -E \left[(d_1^D - \theta_1)^2 + (d_2^D - \theta_2)^2 + 2\delta (d_1^D - d_2^D)^2 \right]. \quad (30)$$

Using (8), (9) and Lemmas A1 and A2-i. we have that for given N_1, N_2

$$\begin{aligned} E \left[(d_1^D - \theta_1)^2 \right] &= \frac{\delta^2}{(\lambda + \delta)^2} \sigma_1^2 - \delta^3 \frac{2\lambda + 3\delta}{(\lambda + \delta)^2 (\lambda + 2\delta)^2} E_{\theta_1} \left[\overline{m}_1^2 \right] + \frac{\delta^2}{(\lambda + 2\delta)^2} E_{\theta_2} \left[\overline{m}_2^2 \right] \\ E \left[(d_2^D - \theta_2)^2 \right] &= \frac{\delta^2}{(\lambda + \delta)^2} \sigma_2^2 + \frac{\delta^2}{(\lambda + 2\delta)^2} E_{\theta_1} \left[\overline{m}_1^2 \right] - \delta^3 \frac{2\lambda + 3\delta}{(\lambda + \delta)^2 (\lambda + 2\delta)^2} E_{\theta_2} \left[\overline{m}_2^2 \right] \\ E \left[(d_1^D - d_2^D)^2 \right] &= \frac{\lambda^2}{(\lambda + \delta)^2} (\sigma_1^2 + \sigma_2^2) - \lambda^2 \delta \frac{2\lambda + 3\delta}{(\lambda + \delta)^2 (\lambda + 2\delta)^2} (E_{\theta_1} \left[\overline{m}_1^2 \right] + E_{\theta_2} \left[\overline{m}_2^2 \right]). \end{aligned} \quad (31)$$

The rate at which total profits change with the variance of the messages $E_{\theta_j} \left(\overline{m}_j^2 \right)$ is

$$\frac{\partial \Pi_D}{\partial E_{\theta_j} \left[\overline{m}_j^2 \right]} = \frac{\lambda^2 \delta^2}{(\lambda + \delta)^2 (\lambda + 2\delta)^2} \left(6\delta - 1 + 4\lambda + 2 \left(\frac{\delta}{\lambda} \right)^2 \right).$$

Since $\lambda > \frac{1}{2}$ we have that $\frac{\partial \Pi_D}{\partial E_{\theta_j} \left[\overline{m}_j^2 \right]} > 0$. From Lemma A.2-ii we have that $E_{\theta_j} \left(\overline{m}_j^2 \right)$ increases with N_j and therefore Π_D increases as the number of elements N_j in the partition of Manager j increases. ■

Proof of Lemma 1: Taking limits to each term on the RHS of (27) we obtain

$$\begin{aligned} \lim_{N_j \rightarrow \infty} \frac{\left(x_C^{3N_j} - 1 \right) (x_C - 1)^2}{\left(x_C^{N_j} - 1 \right)^3 (x_C^2 + x_C + 1)} &= \frac{x_C - 1}{x_C^2 + x_C + 1} \\ \lim_{N_j \rightarrow \infty} \frac{x_C^{N_j} (x_C + 1)^2}{x_C \left(x_C^{N_j} - 1 \right)^2} &= 0. \end{aligned}$$

Therefore

$$\lim_{N_j \rightarrow \infty} E \left[\overline{m}_j^2 \right] = \frac{s_1^2}{4} \frac{(x_C + 1)^2}{(x_C^2 + x_C + 1)} = s_1^2 \frac{1 + b_C}{3 + 4b_C} = (1 - S_C) \sigma_j^2$$

where $S_C = b_C/(3 + 4b_C)$. Finally, using Lemma A2-i we have that

$$\mathbb{E} \left[(\theta_j - \mathbb{E}[\theta_j | m_j])^2 \right] = \sigma_j^2 - \mathbb{E}[\overline{m}_j^2] = S_C \sigma_j^2.$$

The analysis for Decentralization is analogous (one just needs to replace C with D in the above expressions). ■

Proof of Proposition 3: i. From (11) and (12) it follows immediately that for $\lambda = 1/2$ $b_D = b_C = 0$ and $S_C = S_D = 0$. Furthermore, since the function $b/(3 + 4b)$ is strictly increasing in b it suffices to show that, whenever $\lambda > 1/2$, $b_D > b_C$. Since $1/2 \leq \lambda \leq 1$ then $\lambda + \delta > \lambda + \delta - \lambda^2 > 0$ and taking the ratio of b_D and b_C we have

$$\frac{b_D}{b_C} = \frac{(\lambda + \delta)(\lambda + \delta)}{\delta(\lambda + \delta - \lambda^2)} > 1.$$

ii. By directly differentiating S_C and S_D we find that

$$\begin{aligned} \frac{\partial S_C}{\partial \lambda} &= \frac{3(2\delta + 1)\delta}{(8\lambda\delta - \delta + 3\lambda)^2} \geq 0 \\ \frac{\partial S_D}{\partial \lambda} &= \frac{3(2\lambda\delta + \lambda^2 + 2\delta^2 + 2\lambda^2\delta)}{(8\lambda\delta - \delta - \lambda + 5\lambda^2)^2} > 0. \end{aligned}$$

Computing the difference $\frac{\partial S_D}{\partial \lambda} - \frac{\partial S_C}{\partial \lambda}$ and noting that $\lambda \geq 1/2$ we have

$$\frac{\partial S_D}{\partial \lambda} - \frac{\partial S_C}{\partial \lambda} = \frac{g(\lambda, \delta)}{(8\lambda\delta - \delta - \lambda + 5\lambda^2)^2 (3\lambda - \delta + 8\lambda\delta)^2} > 0,$$

where

$$\begin{aligned} g(\lambda, \delta) &\equiv 3\delta^3(2\lambda - 1) + 3\lambda\delta^2(8\lambda^2 - 2) + 3\lambda^2\delta(22\lambda - 1) + 192\lambda^3\delta^3(2\lambda - 1) \\ &\quad + 3\lambda^4(9 + 41\delta + 110\delta^2) + 3\lambda^2\delta^2(31 + 54\delta). \end{aligned}$$

iii. Since $\partial b_C / \partial \delta = \lambda(2\lambda - 1) / (\lambda + \delta)^2 > 0$ and $\partial b_D / \partial \delta = -\lambda^2(2\lambda - 1) / (\lambda + \delta - \lambda^2)^2 < 0$ it follows that $\partial S_C / \partial \delta = \partial S_C / \partial b_C \cdot \partial b_C / \partial \delta > 0$ and $\partial S_D / \partial \delta = \partial S_D / \partial b_D \cdot \partial b_D / \partial \delta < 0$. Finally, by taking limits we have

$$\begin{aligned} \lim_{\delta \rightarrow \infty} b_C &= (2\lambda - 1) \\ \lim_{\delta \rightarrow \infty} b_D &= (2\lambda - 1) \end{aligned}$$

which implies that $\lim_{\delta \rightarrow \infty} S_C = \lim_{\delta \rightarrow \infty} S_D$. ■

Proof of Proposition 4: *a. Centralization:* From (28) and (29) and Lemma 1 we have that

$$\Pi_C = - (A_C (\sigma_1^2 + \sigma_2^2) + (1 - A_C) S_C (\sigma_1^2 + \sigma_2^2)).$$

b. Decentralization: From (30) and (31) and Lemma 1 we have that

$$\Pi_D = - (A_D (\sigma_1^2 + \sigma_2^2) + B_D S_D (\sigma_1^2 + \sigma_2^2)). \quad \blacksquare$$

Proof of Lemma 2: Substituting (3) and (4) into the definitions of AL_C and CL_C and making use of Lemma A2 gives

$$AL_C = \left(1 - \frac{1 + 8\delta(1 + \delta)}{(4\delta + 1)^2} (1 - S_C) \right) (\sigma_1^2 + \sigma_2^2) \quad (32)$$

and

$$CL_C = \frac{2\delta}{(1 + 4\delta)^2} (1 - S_C) (\sigma_1^2 + \sigma_2^2). \quad (33)$$

Similarly, substituting (6) and (7) into the definitions of AL_D and CL_D and making use of Lemma A2 we obtain

$$AL_D = \delta^2 \left(\frac{1}{(\lambda + \delta)^2} - \frac{(2\delta^2 - \lambda^2)}{(\lambda + \delta)^2 (\lambda + 2\delta)^2} (1 - S_D) \right) (\sigma_1^2 + \sigma_2^2) \quad (34)$$

and

$$CL_D = \frac{2\lambda^2\delta}{(\lambda + \delta)^2} \left(1 - \frac{\delta(2\lambda + 3\delta)}{(\lambda + 2\delta)^2} (1 - S_D) \right) (\sigma_1^2 + \sigma_2^2). \quad (35)$$

Subtracting (34) from (32) shows that $AL_C - AL_D$ is equal to

$$\frac{\lambda\delta(2\lambda - 1)(\omega_0 + \delta\omega_1 + \delta^2\omega_2 + \delta^3\omega_3 + \delta^4\omega_4)(\sigma_1^2 + \sigma_2^2)}{(4\delta + 1)^2(3\lambda + \delta(8\lambda - 1))(\lambda(5\lambda - 1) + \delta(8\lambda - 1))(\lambda + \delta)(\lambda + 2\delta)}, \quad (36)$$

where $\omega_0 \equiv \lambda^2(5\lambda - 1)$, $\omega_1 \equiv \lambda(33\lambda + 100\lambda^2 - 7)$, $\omega_2 \equiv 2(-9\lambda + 240\lambda^2 + 100\lambda^3 - 5)$, $\omega_3 \equiv 2(174\lambda + 460\lambda^2 - 43)$ and $\omega_4 \equiv 16(53\lambda - 10)$. It is straightforward to verify that $\omega_i > 0$, $i = 1, \dots, 4$, for all $\lambda \in [1/2, 1]$. Inspection of (36) then shows that $AL_C - AL_D = 0$ if either $\delta = 0$ or $\lambda = 1/2$ and that $AL_C - AL_D > 0$ otherwise.

Next, subtracting (33) from (35) shows that $CL_D - CL_C$ is equal to

$$\frac{2\lambda\delta^2(2\lambda - 1)(\gamma_0 + \delta\gamma_1 + \delta^2\gamma_2 + \delta^3\gamma_3)(\sigma_1^2 + \sigma_2^2)}{(4\delta + 1)^2(\lambda(5\lambda - 1) + \delta(8\lambda - 1))(3\lambda + \delta(8\lambda - 1))(\lambda + \delta)(\lambda + 2\delta)}, \quad (37)$$

where $\gamma_0 \equiv \lambda^2(65\lambda - 7)$, $\gamma_1 \equiv \lambda(107\lambda + 280\lambda^2 - 17)$, $\gamma_2 \equiv 2(-5\lambda + 224\lambda^2 + 160\lambda^3 - 3)$ and $\gamma_3 \equiv 4(14\lambda + 3)(8\lambda - 1)$. It is straightforward to verify that $\gamma_i > 0$, $i = 1, 2, 3$, for all $\lambda \in$

$[1/2, 1]$. Inspection of (37) then shows that $CL_D - CL_C = 0$ if either $\delta = 0$ or $\lambda = 1/2$ and that $CL_D - CL_C > 0$ otherwise. ■

Proof of Proposition 5: Using the expressions in Proposition 4 we have that $\Pi_C - \Pi_D$ is given by

$$\frac{(\sigma_1^2 + \sigma_2^2) \lambda \delta (2\lambda - 1) f}{(4\delta + 1)(3\lambda + \delta(8\lambda - 1))(\lambda(5\lambda - 1) + \delta(8\lambda - 1))(\lambda + \delta)(\lambda + 2\delta)},$$

where

$$\begin{aligned} f \equiv & 2\delta^3(4\lambda - 1)(28\lambda - 17) + 2\delta^2(5\lambda - 1)(-3\lambda + 16\lambda^2 - 5) \\ & + \lambda\delta(-51\lambda + 50\lambda^2 + 7) - \lambda^2(5\lambda - 1). \end{aligned}$$

Note that the denominator is strictly positive for all $\delta \in [0, \infty)$ and $\lambda \in [1/2, 1]$. Thus, $\Pi_C - \Pi_D$ is continuous in $\delta \in [0, \infty)$ and $\lambda \in [1/2, 1]$. Also, $\Pi_C - \Pi_D = 0$ if either $\delta = 0$, $\lambda = 1/2$ or $f = 0$. Let $\bar{\lambda}(\delta)$ be the values of $\lambda \in [1/2, 1]$ which solve $f = 0$ for $\delta \in [0, \infty)$; $\bar{\lambda}(\delta)$ is plotted in Figure 5. Note that $\lim_{\delta \rightarrow \infty} \bar{\lambda}(\delta) = 17/28$ and that $\bar{\lambda}(\bar{\delta}) = 1$, where $\bar{\delta} \simeq 0.19257$ is implicitly defined by $(3\bar{\delta} + 32\bar{\delta}^2 + 33\bar{\delta}^3 - 2) = 0$. Note also that $\bar{\lambda}(\delta)$ is decreasing in $\delta \in [\bar{\delta}, \infty)$.

Part (i.): Since $\bar{\lambda}(\delta)$ is decreasing in δ it is sufficient to show that for $\lambda = 1$, $\Pi_C - \Pi_D < 0$ if $\delta \in (0, \bar{\delta})$. To see this note that $\text{sign}(\Pi_C - \Pi_D) = \text{sign} f$ and that for $\lambda = 1$, $f = 2(3\delta + 32\delta^2 + 33\delta^3 - 2)$. It can be seen that $f < 0$ if $\delta \in (0, \bar{\delta})$.

Part (ii.): It is sufficient to show that for $\lambda = 1$, $\Pi_C - \Pi_D > 0$ if $\delta \in (\bar{\delta}, \infty)$. That this is so follows immediately from the definition of $\bar{\delta}$ and the facts that $\text{sign}(\Pi_C - \Pi_D) = \text{sign} f$ and $f = 2(3\delta + 32\delta^2 + 33\delta^3 - 2)$ for $\lambda = 1$.

Part (iii.): Taking the limits of (14) and (15) gives:

$$\lim_{\delta \rightarrow \infty} \Pi_C = \lim_{\delta \rightarrow \infty} \Pi_D = -\frac{5\lambda - 1}{8\lambda - 1} (\sigma_1^2 + \sigma_2^2). \quad \blacksquare$$

Proof of Proposition 6: Using the definitions of S_C and S_D we have that $S_C = S_D/2$ if and only if $\delta = 0$ or

$$\frac{(2\lambda - 1)(-3\lambda^2 - \delta^2 - 4\lambda\delta + 8\lambda\delta^2 + 2\lambda^2\delta)}{2(3\lambda - \delta + 8\lambda\delta)(-\lambda - \delta + 5\lambda^2 + 8\lambda\delta)} = 0.$$

Note that the terms in the denominator are not equal to zero for any $\lambda \in [1/2, 1]$ and $\delta \in [0, \infty)$. Thus, $S_C = S_D/2$ if either $\lambda = 1/2$ or $(-3\lambda^2 - \delta^2 - 4\lambda\delta + 8\lambda\delta^2 + 2\lambda^2\delta) = 0$. One solution to this quadratic equation is given by δ_1 as defined in the Proposition and the other solution is negative. Finally, it is straightforward to verify that for $\lambda = 1$, $S_C < S_D/2$ if $\delta \in (0, \delta_1)$. Thus, $S_C \leq S_D/2$ for all $\delta \in [0, \delta_1]$ and $S_C \geq S_D/2$ for all $\delta \in [\delta_1, \infty)$. ■

10 Appendix B - Sequential Decision Making

Decision Making: In the last stage of the game Manager 2 chooses d_2 to maximize his expected utility $E_2 [(1 - \lambda)\pi_1 + \lambda\pi_2 | d_1]$. The optimal decision that solves this problem is given by

$$d_2^S \equiv \frac{\lambda}{\lambda + \delta} \theta_2 + \frac{\delta}{\lambda + \delta} d_1^S. \quad (38)$$

At the previous stage Manager 1 chooses d_1 to maximize his expected profits $E_1 [\lambda\pi_1 + (1 - \lambda)\pi_2 | m]$. The optimal decision is given by

$$d_1^S \equiv \frac{\lambda(\lambda + \delta)^2}{\lambda^3 + 3\lambda^2\delta + \delta^2} \theta_1 + \delta \frac{\lambda^2 + \delta(1 - \lambda)}{\lambda^3 + 3\lambda^2\delta + \delta^2} E_1 [\theta_2 | m]. \quad (39)$$

Communication: Let $\mu_2(m_2 | \theta_2)$ be the probability with which Manager 2 sends message m_2 , let $d_1(m_2)$ and $d_2(m_2)$ be the decision rules that map messages into decisions and let $g_1(\theta_2 | m_2)$ be the belief function which gives the probability of θ_2 conditional on observing m_2 . We can now state the following proposition which characterizes the finite communication equilibria when $\delta > 0$.

PROPOSITION B1 (Communication Equilibria). *If $\delta \in (0, \infty)$, then for every positive integer N_2 there exists at least one equilibrium $(\mu_2(\cdot), d_1(\cdot), d_2(\cdot), g_1(\cdot))$, where*

- i. $\mu_2(m_2 | \theta_2)$ is uniform, supported on $[a_{2,i-1}, a_{2,i}]$ if $\theta_2 \in (a_{2,i-1}, a_{2,i})$,
- ii. $g_1(\theta_2 | m_2)$ is uniform supported on $[a_{2,i-1}, a_{2,i}]$ if $m_2 \in (a_{2,i-1}, a_{2,i})$,
- iii. $a_{2,i+1} - a_{2,i} = a_{2,i} - a_{2,i-1} + 4b_S a_{2,i}$ for $i = 1, \dots, N_2 - 1$,

$$a_{2,-(i+1)} - a_{2,-i} = a_{2,-i} - a_{2,-(i-1)} + 4b_S a_{2,-i} \text{ for } i = 1, \dots, N_j - 1,$$

where $b_S \equiv ((2\lambda - 1)(\lambda + \delta)(\lambda^2 + \delta)) / ((\lambda(1 - \lambda) + \delta)(\lambda^2 + \delta(1 - \lambda)))$ and

- iv. $d_j(m) = d_j^S$, $j = 1, 2$, where d_j^S is given by (38) and (39).

Moreover, all other finite equilibria have relationships between θ_1 and θ_2 and the managers' choices of d_1 and d_2 that are the same as those in this class for some value of N_2 ; they are therefore economically equivalent.

Proof: The proof is analogous to the proof of Proposition 1. Details are available from the authors upon request. ■

PROPOSITION B2 (Efficiency). *The limit of strategy profiles and beliefs $(\mu_2(\cdot), d_1(\cdot), d_2(\cdot), g_1(\cdot))$ as $N_2 \rightarrow \infty$ is a Perfect Bayesian Equilibrium of the communication game. In this equilibrium the total expected profits $E[\pi_1 + \pi_2]$ are higher than in any other equilibrium.*

Proof: The proof is analogous to the proof of Proposition 2. Details are available from the authors upon request. ■

In the remaining analysis we focus on the efficient equilibrium.

LEMMA B1. *In the most efficient equilibrium in which $N_2 \rightarrow \infty$ the residual variance is given by*

$$\mathbb{E} \left[(\theta_2 - \mathbb{E}[\theta_2 | m_2])^2 \right] = S_S \sigma_2^2,$$

where $S_S = b_S / (3 + 4b_S)$.

Proof: The proof is analogous to the proof of Lemma 1. Details are available from the authors upon request. ■

PROPOSITION B3 (Organizational Performance). *Under Decentralization with sequential decision making the expected profits are given by*

$$\Pi_S = - ((A_D + X) (\sigma_1^2 + \sigma_2^2) + (B_D - X) S_S \sigma_2^2), \quad (40)$$

where A_D and B_D are defined in (16) and

$$X \equiv \delta^3 (2\lambda - 1)^2 \frac{2\lambda^4 + \lambda^2 (6\lambda + 1) \delta + 2\lambda (2 + \lambda) \delta^2 + 2\delta^3}{(\lambda + 2\delta)^2 (\lambda^3 + \delta^2 + 3\lambda^2 \delta)^2}.$$

Proof: The proof is analogous to the proof of Proposition 4. Details are available from the authors upon request. ■

We can now prove Proposition 7 which is stated in the main text.

Proof of Proposition 7: i. Applying l'Hopital's Rule to (14) and (40) and using the assumption that $\sigma_1^2 = \sigma_2^2 = \sigma^2$ we obtain

$$\lim_{\delta \rightarrow \infty} \Pi_C - \lim_{\delta \rightarrow \infty} \Pi_S = \frac{8\lambda (4\lambda - 1) (2\lambda - 1)^2}{(8\lambda - 1) (5\lambda - 1)} \sigma^2$$

which is strictly positive for any $\lambda > 1/2$.

ii. Taking the derivative of (14) and (40) and using the assumption that $\sigma_1^2 = \sigma_2^2 = \sigma^2$ we obtain

$$\frac{d(\Pi_S - \Pi_C)}{d\lambda} = \frac{8\delta}{3(1 + 2\delta)(1 + 4\delta)} \sigma^2$$

which is strictly positive for all finite $\delta > 0$. ■

11 Appendix C - Different Needs for Coordination

Since allowing for differences in the needs for coordination only requires adding a parameter in the main model, we do not replicate the full analysis here. Instead we merely state the key expressions and use them to prove Proposition 8. The derivation of these expressions and their interpretation are exactly as in the main model.

11.1 Centralization

The decisions are now given by

$$\begin{aligned} d_1^C &\equiv \left(\frac{1}{1+2(\delta_1+\delta_2)} \left((1+\delta_1+\delta_2) E_H[\theta_1 | m] + (\delta_1+\delta_2) E_H[\theta_2 | m] \right) \right) \\ d_2^C &\equiv \left(\frac{1}{1+2(\delta_1+\delta_2)} \left((\delta_1+\delta_2) E_H[\theta_1 | m] + n(1+\delta_1+\delta_2) E_H[\theta_2 | m] \right) \right). \end{aligned}$$

The residual variance of θ_1 is given by $S_{C,1}\sigma_1^2$ and that of θ_2 is given by $S_{C,2}\sigma_2^2$, where $S_{C,j} \equiv b_{C,j}/(3+4b_{C,j})$, $j = 1, 2$, and

$$\begin{aligned} b_{C,1} &= \frac{(2\lambda-1)(\delta_2+(\delta_1+\delta_2)^2)}{\delta_2+(\delta_1+\delta_2)^2+\lambda(1+3\delta_1+\delta_2)} \\ b_{C,2} &= \frac{(2\lambda-1)(\delta_1+(\delta_1+\delta_2)^2)}{\delta_1+(\delta_1+\delta_2)^2+\lambda(1+\delta_1+3\delta_2)}. \end{aligned}$$

The expected profits are given by

$$\Pi_C = -\sigma^2 \left(2 \frac{\delta_1+\delta_2}{1+2(\delta_1+\delta_2)} + (S_1+S_2) \frac{1+\delta_1+\delta_2}{1+2(\delta_1+\delta_2)} \right). \quad (41)$$

Applying l'Hopital's Rule gives

$$\lim_{\delta_1 \rightarrow \infty} \Pi_C = -2 \frac{5\lambda-1}{8\lambda-1} \sigma^2. \quad (42)$$

We can also use (41) to evaluate $d\Pi_C/d\lambda$ for $\lambda = 1/2$:

$$\frac{d\Pi_C}{d\lambda} = -\frac{4}{3} \frac{(\delta_1+\delta_2)}{1+2(\delta_1+\delta_2)} \sigma^2. \quad (43)$$

11.2 Decentralization

The decisions under Decentralization are now given by

$$\begin{aligned}
 d_1^D &= \frac{\lambda\theta_1}{\lambda + \lambda\delta_1 + (1-\lambda)\delta_2} + \frac{((1-\lambda)\delta_1 + \lambda\delta_2)(\lambda\delta_1 + (1-\lambda)\delta_2)}{(\lambda + \delta_1 + \delta_2)(\lambda + \lambda\delta_1 + (1-\lambda)\delta_2)} \mathbb{E}_2[\theta_1 | m] \\
 &\quad + \frac{\lambda\delta_1 + (1-\lambda)\delta_2}{\lambda + \delta_1 + \delta_2} \mathbb{E}_1[\theta_2 | m] \\
 d_2^D &= \frac{\lambda\theta_2}{\lambda + (1-\lambda)\delta_1 + \lambda\delta_2} + \frac{(1-\lambda)\delta_1 + \lambda\delta_2}{\lambda + \delta_1 + \delta_2} \mathbb{E}_2[\theta_1 | m] \\
 &\quad + \frac{((1-\lambda)\delta_1 + \lambda\delta_2)(\lambda\delta_1 + (1-\lambda)\delta_2)}{(\lambda + \delta_1 + \delta_2)(\lambda + (1-\lambda)\delta_1 + \lambda\delta_2)} \mathbb{E}_1[\theta_2 | m].
 \end{aligned}$$

The residual variance of θ_1 is given by $S_{D,1}\sigma_1^2$ and that of θ_2 is given by $S_{D,2}\sigma_2^2$, where $S_{D,j} \equiv b_{C,j}/(3 + 4b_{C,j})$, $j = 1, 2$, and

$$\begin{aligned}
 b_1 &= \frac{(2\lambda - 1)\delta_1(\lambda + \lambda\delta_1 + (1-\lambda)\delta_2)}{(\lambda(1-\lambda) + \lambda\delta_1 + (1-\lambda)\delta_2)((1-\lambda)\delta_1 + \lambda\delta_2)} \\
 b_2 &= \frac{(2\lambda - 1)\delta_2(\lambda + \lambda\delta_2 + (1-\lambda)\delta_1)}{(\lambda(1-\lambda) + \lambda\delta_2 + (1-\lambda)\delta_1)((1-\lambda)\delta_2 + \lambda\delta_1)}.
 \end{aligned}$$

The expected profits are given by

$$\Pi_D = -\mathbb{E} \left[(d_1^D - \theta_1)^2 + (d_2^D - \theta_2)^2 + (\delta_1 + \delta_2)(d_1^D - d_2^D)^2 \right], \quad (44)$$

where

$$\begin{aligned}
 &\mathbb{E} \left[(d_1^D - \theta_1)^2 \right] \\
 &= \sigma^2 \left(2 \frac{(\delta_2 + \lambda\delta_1 - \lambda\delta_2)^2}{(\lambda + \delta_1 + \delta_2)^2} - \frac{(\lambda\delta_1 + (1-\lambda)\delta_2)^2}{(\lambda + \delta_1 + \delta_2)^2} S_2 \right. \\
 &\quad \left. + \frac{((1-\lambda)\delta_1 + \lambda\delta_2)(\lambda\delta_1 + (1-\lambda)\delta_2)^2 (2\lambda + (1+\lambda)\delta_1 + (2-\lambda)\delta_2)}{(\lambda + \delta_1 + \delta_2)^2 (\lambda + \lambda\delta_1 + (1-\lambda)\delta_2)^2} S_1 \right) \\
 &\mathbb{E} \left[(d_2 - \theta_2)^2 \right] \\
 &= \sigma^2 \left(2 \frac{(-\delta_1 + \lambda\delta_1 - \lambda\delta_2)^2}{(\lambda + \delta_1 + \delta_2)^2} - \frac{((1-\lambda)\delta_1 + \lambda\delta_2)^2}{(\lambda + \delta_1 + \delta_2)^2} S_1 \right. \\
 &\quad \left. + \frac{((1-\lambda)\delta_1 + \lambda\delta_2)^2 (\lambda\delta_1 + (1-\lambda)\delta_2) (2\lambda + (2-\lambda)\delta_1 + (1+\lambda)\delta_2)}{(\lambda + \delta_1 + \delta_2)^2 (\lambda + (1-\lambda)\delta_1 + \lambda\delta_2)^2} S_2 \right)
 \end{aligned}$$

$$\begin{aligned}
& \mathbb{E} \left[(d_1^D - d_2^D)^2 \right] \\
&= \sigma^2 \left(2 \frac{\lambda^2}{(\lambda + \delta_1 + \delta_2)^2} + \frac{\lambda^2 (2\lambda + (1 + \lambda) \delta_1 + (2 - \lambda) \delta_2) ((1 - \lambda) \delta_1 + \lambda \delta_2)}{(\lambda + \delta_1 + \delta_2)^2 (\lambda + \lambda \delta_1 + (1 - \lambda) \delta_2)^2} S_1 \right. \\
&\quad \left. + \frac{\lambda^2 (2\lambda + (2 - \lambda) \delta_1 + (1 + \lambda) \delta_2) (\lambda \delta_1 + (1 - \lambda) \delta_2)}{(\lambda + \delta_1 + \delta_2)^2 (\lambda + (1 - \lambda) \delta_1 + \lambda \delta_2)^2} S_2 \right).
\end{aligned}$$

Applying l'Hopital's Rule gives

$$\lim_{\delta_1 \rightarrow \infty} \Pi_D = -2 \frac{8\lambda^3 - 9\lambda^2 + 6\lambda - 1}{5\lambda - 1} \sigma^2. \quad (45)$$

We can also use (44) to evaluate $d\Pi_C/d\lambda$ for $\lambda = 1/2$:

$$\frac{d\Pi_D}{d\lambda} = -\frac{8}{3} \frac{(\delta_1 + \delta_2)^2}{(1 + 2(\delta_1 + \delta_2))^2} \sigma^2. \quad (46)$$

We can now prove Proposition 8 which is stated in the main text.

Proof of Proposition 8: i. Using (42) and (45) we obtain

$$\lim_{\delta \rightarrow \infty} \Pi_C - \lim_{\delta \rightarrow \infty} \Pi_S = 8\lambda(4\lambda - 1) \frac{(2\lambda - 1)^2}{(8\lambda - 1)(5\lambda - 1)} \sigma^2$$

which is strictly positive for any $\lambda > 1/2$.

ii. Using (43) and (46) we find that the difference in the derivatives at $\lambda = 1/2$ is given by

$$\frac{d(\Pi_C - \Pi_S)}{d\lambda} = \frac{4}{3} \frac{(\delta_1 + \delta_2)}{(1 + 2(\delta_1 + \delta_2))^2} \sigma^2 \quad \text{for } \lambda = 1/2$$

which is strictly positive for all finite $\delta > 0$. ■

12 Appendix D - Different Division Sizes

Since allowing for different division sizes only requires adding a parameter in the main model, we do not replicate the full analysis here. Instead we merely state the key expressions and use them to prove Proposition 9. The derivation of these expressions and their interpretation are exactly as in the main model.

12.1 Centralization

Let $\beta \equiv (1 - \alpha)$. Then decisions are given by

$$\begin{aligned} d_1^C &\equiv \left(\frac{1}{\alpha\beta + \delta} (\alpha(\beta + \delta) E_H(\theta_1 | m) + \beta\delta E_H(\theta_2 | m)) \right) \\ d_2^C &\equiv \left(\frac{1}{\alpha\beta + \delta} (\alpha\delta E_H(\theta_1 | m) + \beta(\alpha + \delta) E_H(\theta_2 | m)) \right). \end{aligned}$$

The residual variance of θ_1 is given by $S_{C,1}\sigma_1^2$ and that of θ_2 is given by $S_{C,2}\sigma_2^2$, where $S_{C,j} \equiv b_{C,j}/(3 + 4b_{C,j})$, $j = 1, 2$, and

$$\begin{aligned} b_{C,1} &= \frac{\beta\delta(2\lambda - 1)(\beta^2 + \delta)}{\alpha\lambda(\beta^2 + \delta^2) + \beta(\delta + \beta^2)(1 - \lambda)\delta + \alpha\beta\lambda(2 + \beta)\delta} \\ b_{C,2} &= \frac{\alpha\delta(2\lambda - 1)(\alpha^2 + \delta)}{\alpha(\delta + \alpha^2)(1 - \lambda)\delta + \beta\lambda(\alpha^2 + \delta^2) + \alpha\beta\lambda(2 + \alpha)\delta}. \end{aligned}$$

The expected profits are given by

$$\Pi_C = -E \left[(d_1^C - \theta_1)^2 + (d_2^C - \theta_2)^2 + 2\delta(d_1^C - d_2^C)^2 \right],$$

where

$$\begin{aligned} E \left[(d_1 - \theta_1)^2 \right] &= \sigma^2 \left(\frac{2\delta^2\beta^2}{(\alpha\beta + \delta)^2} + \alpha(\beta + \delta) \frac{\alpha\beta + (2 - \alpha)\delta}{(\alpha\beta + \delta)^2} S_1 - \frac{\delta^2\beta^2}{(\alpha\beta + \delta)^2} S_2 \right) \\ E \left[(d_2 - \theta_2)^2 \right] &= \sigma^2 \left(\frac{2\alpha^2\delta^2}{(\alpha\beta + \delta)^2} - \frac{\alpha^2\delta^2}{(\alpha\beta + \delta)^2} S_1 + \beta(\alpha + \delta) \frac{\alpha\beta + (1 + \alpha)\delta}{(\alpha\beta + \delta)^2} S_2 \right) \\ E \left[(d_1 - d_2)^2 \right] &= \sigma^2 \left(\frac{2\alpha^2\beta^2}{(\alpha\beta + \delta)^2} - \frac{\alpha^2\beta^2}{(\alpha\beta + \delta)^2} (S_1 + S_2) \right). \end{aligned}$$

Applying l'Hopital's Rule we find that

$$\lim_{\delta \rightarrow \infty} \Pi_C = \frac{-2\alpha(1 - \alpha)(8\lambda - 1)(5\lambda - 1)}{(5\lambda - 1 - \alpha(2\lambda - 1))(3\lambda + (2\lambda - 1)\alpha)} \sigma^2. \quad (47)$$

Also, differentiating we find that at $\lambda = 1/2$

$$\frac{d\Pi_C}{d\lambda} = -\frac{8}{3} \frac{\alpha(1 - \alpha)\delta}{\alpha(1 - \alpha) + \delta} \sigma^2. \quad (48)$$

12.2 Decentralization

The decisions are now given by

$$\begin{aligned} d_1^D &= \frac{(\alpha\lambda\theta_1 + \delta(\alpha\lambda + \beta(1-\lambda))) E_1(d_2^D | m)}{\alpha\lambda(1+\delta) + \beta\delta(1-\lambda)} \\ d_2^D &= \frac{(\beta\lambda\theta_2 + \delta(\alpha(1-\lambda) + \beta\lambda)) E_2(d_1^D | m)}{\beta\lambda(1+\delta) + \alpha\delta(1-\lambda)} \end{aligned}$$

where $\beta \equiv (1-\alpha)$ and

$$\begin{aligned} E_2[d_1^D | m] &= \frac{(\alpha(\alpha\delta(1-\lambda) + \beta\lambda(1+\delta))) E_2[\theta_1 | m] + \beta\delta(\alpha\lambda + \beta(1-\lambda)) E_1[\theta_2 | m]}{(\alpha^2 + \beta^2)\delta(1-\lambda) + \alpha\beta\lambda(1+2\delta)} \\ E_1[d_2^D | m] &= \frac{(\alpha\delta(\alpha(1-\lambda) + \beta\lambda)) E_2[\theta_1 | m] + \beta(\alpha\lambda(1+\delta) + \beta\delta(1-\lambda)) E_1[\theta_2 | m]}{(\alpha^2 + \beta^2)\delta(1-\lambda) + \alpha\beta\lambda(1+2\delta)}. \end{aligned}$$

The residual variance of θ_1 is given by $S_{D,1}\sigma_1^2$ and that of θ_2 is given by $S_{D,2}\sigma_2^2$, where $S_{D,j} \equiv b_{D,j}/(3+4b_{C,j})$, $j = 1, 2$, and

$$\begin{aligned} b_{D,1} &= \frac{\alpha\beta(2\lambda-1)(\alpha\lambda(1+\delta) + \beta(1-\lambda)\delta)}{(\alpha(1-\lambda) + \beta\lambda)(\beta^2(1-\lambda)^2\delta + \alpha^2\lambda^2\delta + \alpha\beta\lambda(1+2\delta)(1-\lambda))} \\ b_{D,2} &= \frac{\alpha\beta(2\lambda-1)(\alpha(1-\lambda)\delta + \beta\lambda(1+\delta))}{(\beta + \alpha\lambda - \beta\lambda)(\alpha^2(1-\lambda)^2\delta + \beta^2\lambda^2\delta + \alpha\beta\lambda(2\delta+1)(1-\lambda))} \end{aligned}$$

The expected profits are given by

$$\Pi_D = -E\left[(d_1^D - \theta_1)^2 + (d_2^D - \theta_2)^2 + 2\delta(d_1^D - d_2^D)^2\right],$$

where

$$\begin{aligned} &E\left[(d_1^D - \theta_1)^2\right] \\ &= \frac{\sigma^2}{((\alpha^2 + \beta^2)(1-\lambda)\delta + \alpha\beta\lambda(1+2\delta))^2} \left(\frac{2\delta^2\beta^2(\alpha\lambda(1+\delta) + \beta(1-\lambda)\delta)^2(\alpha\lambda + \beta(1-\lambda))^2}{(\alpha\lambda(1+\delta) + \beta(1-\lambda)\delta)^2} \right. \\ &\quad \left. + \frac{\alpha\delta^3(\alpha(1-\lambda) + \beta\lambda)(\alpha\lambda + \beta(1-\lambda))^2((\alpha^2 + 2\beta^2)(1-\lambda)\delta + \alpha\beta\lambda(2+3\delta))S_1}{(\alpha\lambda(1+\delta) + \beta(1-\lambda)\delta)^2} \right. \\ &\quad \left. - \beta^2\delta^2(\alpha\lambda + \beta(1-\lambda))^2S_2 \right) \\ &E\left[(d_2^D - \theta_2)^2\right] \\ &= \frac{\sigma^2}{((\alpha^2 + \beta^2)(1-\lambda)\delta + \alpha\beta\lambda(1+2\delta))^2} \left(2\delta^2\alpha^2(\alpha(1-\lambda) + \beta\lambda)^2 - \alpha^2\delta^2(\alpha(1-\lambda) + \beta\lambda)^2S_1 \right. \\ &\quad \left. + \frac{\beta\delta^3(\alpha(1-\lambda) + \beta\lambda)^2(\alpha\lambda + \beta(1-\lambda))(\delta(2\alpha^2 + \beta^2)(1-\lambda) + \alpha\beta\lambda(2+3\delta))S_2}{(\alpha(1-\lambda)\delta + \beta\lambda(1+\delta))^2} \right) \end{aligned}$$

$$\begin{aligned}
& \mathbb{E} \left[(d_1^D - d_2^D)^2 \right] \\
&= \frac{\sigma^2 \lambda^2}{\left((\alpha^2 + \beta^2) (1 - \lambda) \delta + \alpha \beta \lambda (1 + 2\delta) \right)^2} \left(2\alpha^2 \beta^2 \right. \\
&\quad + \frac{\alpha^3 \delta \left((\alpha^2 + 2\beta^2) (1 - \lambda) \delta + \alpha \beta \lambda (3\delta + 2) \right) (\alpha (1 - \lambda) + \beta \lambda) S_1}{(\alpha \lambda (1 + \delta) + \beta (1 - \lambda) \delta)^2} \\
&\quad \left. + \frac{\beta^3 \delta (\alpha \lambda (1 + \delta) + \beta \delta (1 - \lambda)) (\delta (2\alpha^2 + \beta^2) (1 - \lambda) + \alpha \beta \lambda (2 + 3\delta)) (\alpha \lambda + \beta (1 - \lambda)) S_2}{(\alpha \lambda (1 + \delta) + \beta \delta (1 - \lambda)) (\alpha (1 - \lambda) \delta - \beta \lambda (1 + \delta))^2} \right).
\end{aligned}$$

Applying l'Hopital's Rule we find that

$$\lim_{\delta \rightarrow \infty} \Pi_D = \frac{-2\alpha(1-\alpha) \left(2\alpha(1-\alpha) + 2\lambda^2(2\alpha-1)^2(3\lambda-7) - 26\alpha\lambda(1-\alpha) + 9\lambda-1 \right)}{\left(1 - 2\alpha(1-\alpha) - \lambda(2\alpha-1)^2 \right) (3\lambda(1-\lambda) + \alpha(1-\alpha)(6\lambda+1)(2\lambda-1))} \sigma^2. \quad (49)$$

Also, differentiating we find that at $\lambda = 1/2$

$$\frac{d\Pi_D}{d\lambda} = -\frac{16}{3} \alpha(1-\alpha) \delta^2 \frac{1-2\alpha(1-\alpha)}{(\alpha(1-\alpha)+\delta)^2} \sigma^2. \quad (50)$$

We can now prove Proposition 9 which is stated in the main text.

Proof of Proposition 9: i. Using (47) and (49) gives

$$\begin{aligned}
& \lim_{\delta \rightarrow \infty} \Pi_C - \lim_{\delta \rightarrow \infty} \Pi_D \\
&= \frac{6\alpha\lambda(1-\alpha)(2\alpha-1)^2(2\lambda-1)}{\left(1 - 2\alpha(1-\alpha) - \lambda(2\alpha-1)^2 \right) (5\lambda-1-\alpha(2\lambda-1))} \\
&\quad \times \frac{(\alpha(1-\alpha)(2\lambda-1)(42\lambda^2-11\lambda+1) + \lambda(1-\lambda)(5\lambda-1))}{((2\lambda-1)\alpha+3\lambda)(\alpha(1-\alpha)(6\lambda+1)(2\lambda-1)+3\lambda(1-\lambda))} \sigma^2
\end{aligned}$$

which is strictly positive if $\lambda > 1/2$ and $\alpha > 1/2$.

ii. Using (48) and (50) we find that the difference in the derivatives at $\lambda = 1/2$ is given by

$$\frac{d(\Pi_D - \Pi_C)}{d\lambda} = \frac{8}{3} \alpha(1-\alpha) \delta \frac{\alpha(1-\alpha)(1+4\delta) - \delta}{(\alpha(1-\alpha)+\delta)^2} \sigma^2 \quad \text{for } \lambda = 1/2$$

which is strictly positive if

$$\alpha < \frac{1}{2} \left(1 + \sqrt{\frac{1}{1+4\delta}} \right). \quad \blacksquare$$

References

- [1] AOKI, M. 1986. Horizontal vs. Vertical Information Structure of the Firm. *American Economic Review*, 76: 971-983.
- [2] ALONSO, R. AND N. MATOUSCHEK. 2004. Relational Delegation. Mimeo, Northwestern.
- [3] ATHEY, S. AND J. ROBERTS 2001. Organizational Design: Decision Rights and Incentive Contracts. *American Economic Review Papers and Proceedings*, 91: 200-205
- [4] AUMANN, R. AND S. HART 2003. Long Cheap Talk. *Econometrica*, 71, 1619-1660.
- [5] BALIGA, S. AND S. MORRIS. 2002. Coordination, Spillovers, and Cheap Talk. *Journal of Economic Theory*, 105: 450-68.
- [6] BARTLETT, C.A. 1989. Procter & Gamble Europe: Vizir Launch. Harvard Business School Case 9-384-139.
- [7] BATTAGLINI, M. 2002. Multiple Referrals and Multidimensional Cheap Talk. *Econometrica*, 70(4): 1379-1401.
- [8] BOLTON, P. AND M. DEWATRIPONT 1994. The Firm as a Communication Network. *Quarterly Journal of Economics*, CIX, 809-839.
- [9] BOLTON, P AND J. FARRELL. 1990. Decentralization, Duplication and Delay. *Journal of Political Economy*, 98, 803-26.
- [10] CHANDLER, A. 1977. *The Visible Hand: The Managerial Revolution in American Business*. Cambridge: Belknap Press.
- [11] CREMER, J., L. GARICANO AND A. PRAT 2006. Language and the Theory of the Firm. *Quarterly Journal of Economics*, forthcoming.
- [12] CRAWFORD, V. AND J. SOBEL. 1982. Strategic Information Transmission. *Econometrica*, 50:1431-145.
- [13] DESSEIN, W. 2002. Authority and Communication in Organizations. *Review of Economic Studies*, 69: 811-838.
- [14] DESSEIN, W., L. GARICANO AND R. GERTNER 2005. Organizing for Synergies. Mimeo, Chicago GSB.

- [15] DESSEIN, W. AND T. SANTOS 2006. Adaptive Organizations. *Journal of Political Economy*, forthcoming.
- [16] ECCLES, R.G. AND P. HOLLAND 1989. Jacobs Suchard: Reorganizing for 1992. Harvard Business School Case 489106.
- [17] FARRELL, J. 1987. Cheap Talk, Coordination, and Entry. *The RAND Journal of Economics*. 18(1): 34-9
- [18] FARRELL, J. AND R. GIBBONS 1989. Cheap Talk Can Matter in Bargaining. *Journal of Economic Theory*, 48(1): 221-37.
- [19] FARRELL, J. AND M. RABIN 1996. Cheap Talk. *The Journal of Economic Perspectives*, 10(3):103-118.
- [20] FRIEBEL, G. AND M. RAITH 2006. Resource Allocation and Firm Scope. Mimeo, University of Toulouse.
- [21] GARICANO, L. 2000. Hierarchies and the Organization of Knowledge in Production. *Journal of Political Economy*, October, 874-904.
- [22] GERTNER, R. 1999. Coordination, Dispute Resolution, and the Scope of the Firm, Mimeo, GSB Chicago.
- [23] GERTNER, R., D. SCHARFSTEIN AND J. STEIN. 1994. Internal Versus External Capital Markets, *Quarterly Journal of Economics*, 109(4): 1211-30.
- [24] GROSSMAN, S. AND O. HART 1986. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691-719.
- [25] HANNAN, M., J. PODOLNY AND J. ROBERTS 1999. The Daimler Chrysler Commercial Vehicles Division. Stanford Business School Case IB-27.
- [26] HART, O. AND B. HOLMSTROM 2002. A Theory of Firm Scope. Mimeo, MIT.
- [27] HART, O. AND J. MOORE 1990. Property Rights and the Nature of the Firm. *Journal of Political Economy*, 98, 1119-58.
- [28] HARRIS, M. AND A. RAVIV 2005. Allocation of Decision-Making Authority. *Review of Finance*, September 2005, 9, 353-83.

- [29] HART, O. AND J. MOORE 1990. Property Rights and the Nature of the Firm. *Journal of Political Economy*, 98: 1119-58.
- [30] HART, O. AND J. MOORE 2005. On the Design of Hierarchies: Coordination versus Specialization. *Journal of Political Economy*, 113(4): 675-702.
- [31] HAYEK, F. 1945. The Use of Knowledge in Society. *American Economic Review*, 35(4): 519-530.
- [32] JENSEN, M. AND W. MECKLING 1992. Specific and General Knowledge, and Organizational Structure. In *Contract Economics*, (Lars Werin and Hans Wijkander, eds.), Oxford: Blackwell.
- [33] KAWAMURA, K. 2006. Weighted Cheap Talk: Anonymity, Garbling, and Overconfidence in Communication. Mimeo, University of Oxford.
- [34] KRISHNA, V. AND J. MORGAN 2001. A Model of Expertise. *Quarterly Journal of Economics*, 116: 747-775.
- [35] KRISHNA, V. AND J. MORGAN 2004. The Art of Conversation: Eliciting Information from Experts Through Multi-Stage Communication. *Journal of Economic Theory*, 117: 147-179
- [36] LOCKWOOD, B. 2005. Fiscal Decentralization: A Political Economy Perspective, in *The Handbook of Fiscal Federalism* (E. Ahmad and G. Brosio eds.), Edward Elgar, forthcoming.
- [37] MARSHAK J. AND R. RADNER 1972. *Economic Theory of Teams*. Yale University Press, New Haven and London.
- [38] MARINO, A. AND J. MATSUSAKA 2005. Decision Processes, Agency Problems, and Information: An Economic Analysis of Capital Budgeting Procedures. *Review of Financial Studies*, 18(1): 301-325.
- [39] MELUMAD, N. AND T. SHIBANO 1991. Communication in Settings with No Transfers. *Rand Journal of Economics*, 22: 173-198.
- [40] MILGROM, P. AND J. ROBERTS. 1992. *Economics, Organization, and Management*. Prentice Hall, Englewood Cliffs, New Jersey.
- [41] MONTGOMERY, C.A. AND D. MAGNANI. 2001. PepsiCo's Restaurants. Harvard Business School Case 9-794-078.

- [42] MOOKHERJEE, D. 2006. Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective. *Journal of Economic Literature*, XLIV, 367-390.
- [43] OATES, W. 1999. An Essay on Fiscal Federalism. *Journal of Economic Literature*, 37: 1120-1149.
- [44] OZBAS, O. 2005. Integration, Organizational Processes, and Allocation of Resources. *Journal of Financial Economics*, 75(1): 201-242.
- [45] OTTAVIANI, M. AND P. SORENSEN. 2002. Information Aggregation in Debate: Who Should Speak First. *Journal of Public Economics*, 81: 393-421.
- [46] PFEFFER, J. 2004. Human Resources at the AES Corp.: The Case of the Missing Department, Stanford Business School HR3.
- [47] RADNER, R. 1993. The Organization of Decentralized Information Processing. *Econometrica*, 61: 1109-1146.
- [48] RANTAKARI, H. 2006. Managing Change: Allocation of Decision Rights and Strategic Communication. Mimeo, MIT.
- [49] QIAN, Y., G. ROLAND AND C. XU. 2006. Coordination and Experimentation in M-Form and U-Form Organizations. *Journal of Political Economy*, 114: 366-402.
- [50] SLOAN, A.P., JR. 1964. *My Years With General Motors*. New York: Doubleday.
- [51] STEIN, J. 1997. Internal Capital Markets and the Competition for Corporate Resources. *Journal of Finance*, 52: 111-133.
- [52] STEIN, J. 2002. Information Production and Capital Allocation: Decentralized versus Hierarchical Firms, *Journal of Finance*, 57: 1891-1921.
- [53] VAN ZANDT, T. 1999. Decentralized Information Processing in the Theory of Organizations. In *Contemporary Economic Issues*, Vol. 4: Economic Design and Behavior (Murat Sertel ed.), London: MacMillan.
- [54] WULF, J. 2005. Influence and Inefficiency in the Internal Capital Market. Mimeo, University of Pennsylvania.