

WHEN DOES COORDINATION REQUIRE CENTRALIZATION?*

Ricardo Alonso
University of Southern California

Wouter Dessein
University of Chicago

Niko Matouschek
Northwestern University

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 adapted 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, decentralization can dominate centralization even when coordination is extremely important relative to adaptation.

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

JEL classifications: D23, D83, L23

*We are grateful to the Editor Jeremy Bulow and to two anonymous referees for their comments and suggestions. We also thank Jim Dana, Mathias Dewatripont, Luis Garicano, Rob Gertner, Bob Gibbons, Tom Hubbard, Paul Milgrom, Andrea Prat, Canice Prendergast, Luis Rayo, Scott Schaefer and Jeroen Swinkels for very helpful discussions. Finally, we thank participants at various seminars and conferences for their comments.

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. Alfred Chandler, 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

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).

exchange ideas. In 1923, for example, he established the General Technical Committee to facilitate coordination between the engineers in the various parts of the corporation. The committee did not diminish the authority of the division managers 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.*”^{5,6} In describing his plan to set up the committee to his fellow executives he stated that

*“I believe that such a plan properly developed gives the necessary balance between each Operation and the Corporation itself and will result in all the advantages of co-ordinated action where such action is of benefit in a broader way without in any sense limiting the initiative of independence of action of any component part of the group.”*⁷

The apparent success of Alfred Sloan’s GM and other decentralized firms in realizing inter-divisional synergies suggest that in many cases coordination can indeed be achieved without centralization.⁸

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 need for adaptation. We argue that this

⁵Sloan (1964), p.105.

⁶Specifying the functions of the General Technical Committee Sloan stated that “The Committee would not, as to principle, deal with the specific problems of any individual Operation. Each function of that Operation would be under the absolute control of the General Manager of that Division.” (Sloan 1964, p.107). Reflecting on the committee later, he stated that it “produced a free exchange of new and progressive ideas and experience among division engineers. In short, it co-ordinated information. [...] the General Technical Committee was the mildest kind of organization. Its most important role was that of a study group. It became known as a seminar. [...] Sometimes the committee’s discussion would conclude with the approval of a new device or method, or a recommendation on engineering policy and procedure, but more often the results were simply that information was transmitted from one to all” (Sloan 1964, p.109).

⁷Sloan (1964), p.106.

⁸Other, 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).

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.

To investigate coordination in organizations 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. In particular, the only formal mechanism the organization can commit to is 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 headquarter 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. The difference in the quality of horizontal and vertical communication, however, diminishes, and eventually vanishes, 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 costly coordination failures. As a result, the managers' incentives to exaggerate are mitigated and more information is communicated.

The fact that the difference in the quality of horizontal and vertical communication diminishes as coordination becomes more important drives our central result: Decentralization can dominate Centralization even when coordination is extremely important. Specifically, in symmetric organizations – in which divisions are of equal size and have the same need for coordination and in which decisions are made simultaneously – decentralization *always* outperforms centralization when the division managers' incentives are sufficiently aligned. The same result also holds in asymmetric organizations in which decisions are made sequentially or in which the divisions differ in their need for coordination. In organizations in which the divisions differ in terms of their size the result also holds as long as the size difference is not too large.

2 Related Literature

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

Coordination in Organizations: A number of recent papers analyze coordination in organizations. Hart and Holmström (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 biased towards coordination. In contrast, in our paper, authority is allocated to a benevolent principal under centralization.⁹ 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, however, Bolton and Farrell (1990) rule out communication.¹⁰

Our rationale for decentralization is also related to the literature on influence activities (Milgrom 1988, Milgrom and Roberts 1990), which argues that centralization induces agents to engage in wasteful activities in an attempt to influence decision making. But whereas in Milgrom (1988) these influence activities are pure waste, in our model they take the form of distorting local information that is useful for decision making. In a related vein, Stein (2002), Ozbas (2005) and Friebel and Raith (2006) argue that the distortion of local information by division managers limits the value of centralization as a way to improve the efficient reallocation of capital or resources. Unlike our paper, however, communication is always vertical.¹¹

Information Processing in Organizations: The large literature on team theory, starting with Marschak 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

⁹Otherwise, 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.

¹⁰Also in Gertner (1999) headquarters is unbiased but it only intervenes when bargaining between divisions breaks down. He shows how the presence of an independent arbitrator may foster information sharing.

¹¹Our paper is further 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).

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

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 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 (2007) 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 recent and independently developed paper that complements ours is Rantakari (2006). He also analyzes coordination in organizations in which information is dispersed but focuses on settings in which divisions differ in their need for coordination. Among other results, he shows that in such asymmetric settings it can be optimal to put in place asymmetric organizational structures in which, for instance, all decision rights are concentrated in one division.

An alternative to our no-commitment assumption is to adhere to the mechanism design approach but impose restrictions on communication, as in Melumad, Mookherjee and Reichelstein (1992). In their model, a principal must decide whether to contract directly with two agents (centralization) or to contract with one of them who then contracts himself with the second agent (decentralization). Because of exogenous communication restrictions decentralized contracting allows for a better use of local information, as in our model, and may therefore be preferred over centralized contracting. Unlike our model, however, decentralization results in a hierarchical relationship between the two

organizations that ignore local knowledge. Also Cremer, Garicano and Prat (2007) study how physical communication 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. In Qian, Roland and Xu (2006), finally, different organizational forms (M-form and U-form) result in different information structures, affecting the organization's ability to coordinate change.

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, Farrell and Gibbons 1989 or more recently Baliga and Morris 2002).

¹⁵See also Marino and Matsusaka (2005).

agents. The control loss associated with decentralized contracting, therefore, does not come in the form of a loss of coordination.

Cheap Talk and Expert Literature: From a methodological perspective our paper contributes to the cheap talk and expert literatures that build on Crawford and Sobel (1982). In these models, a receiver makes a decision after consulting with one or several privately informed but biased senders or ‘experts.’ 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 possibility.¹⁶ 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 (1982), the traditional workhorse for cheap talk and expert applications.¹⁷ We further differ from both Crawford and Sobel (1982) and Melumad and Shibano (1990) in that we allow for multiple senders. Also Battaglini (2002) and Krishna and Morgan (2001) consider models in which a principal consults multiple informed experts, but these experts all observe the same piece of information. The question they investigate is whether, and if so how, the principal can elicit this information from the experts.¹⁸ 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*

¹⁶See also Stein (1989).

¹⁷See also Blume, Board and Kawamura (2007).

¹⁸See also Ottaviani and Sorensen (2001) who study the impact of reputational concerns on communication in a setting with multiple experts.

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, Manager 2 observes θ_2 but does not know θ_1 . The HQ Manager, that is, 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 $\pi_1 + \pi_2$.¹⁹ 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 explicit incentive contracts, 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.^{20,21} 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

¹⁹For 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.

²⁰In the conclusion we sketch how one could endogenize the own-division bias by allowing the organization to design the compensation schemes of the division managers.

²¹In some cases implicit incentives may actually soften the own-division bias of division managers that are created by explicit incentive schemes and other factors. See, for instance, our discussion of BP in Section 7.

if the division managers have to make division specific effort choices.²² 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.²³

Contracts and Communication: We follow the property rights literature (Grossman and Hart 1986, 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.²⁴

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 headquarters. We refer to communication between the division managers as *horizontal communication* and that between the division managers and headquarters as *vertical communication*. It is well known in the literature on cheap talk games that repeated rounds of communication may expand the set

²²The 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. See also Footnote 21.

²³In 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).

²⁴A 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 8.

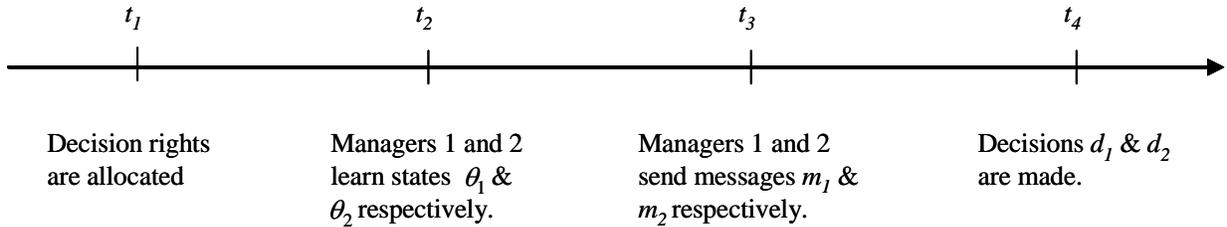


Figure 1: Timeline

of equilibrium outcomes even if only one player is informed.²⁵ 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 8, 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 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, that is, 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 headquarters, while under Decentralization they engage in horizontal communication, exchanging messages m_1 and m_2 with 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.

²⁵See, for example, Aumann and Hart (2003) and Krishna and Morgan (2004).

4 Decision Making

In this section we characterize decision making under Centralization and Decentralization, taking as given the posterior beliefs of the decision makers over 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 Appendix A.

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 $E[\pi_1 + \pi_2 | m]$, that is, her expectation of overall profits given messages $m \equiv (m_1, m_2)$. The decisions that solve this problem are convex combinations of the HQ Manager's posterior beliefs of θ_1 and θ_2 given m :

$$d_1^C \equiv \gamma_C E[\theta_1 | m] + (1 - \gamma_C) E[\theta_2 | m] \quad (3)$$

and

$$d_2^C \equiv (1 - \gamma_C) E[\theta_1 | m] + \gamma_C E[\theta_2 | m], \quad (4)$$

where

$$\gamma_C \equiv \frac{1 + 2\delta}{1 + 4\delta}. \quad (5)$$

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, that is, 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 $E[\lambda\pi_1 + (1 - \lambda)\pi_2 | \theta_1, m]$ and, simultaneously, Manager 2 chooses d_2 to maximize $E[(1 - \lambda)\pi_1 + \lambda\pi_2 | \theta_2, m]$. The decision that Manager 1 makes is a convex combination of his local conditions, θ_1 , and of the decision $E[d_2 | \theta_1, m]$ that he expects Manager 2 to make:

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

Similarly, we have that

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

where $E[d_1 | \theta_2, m]$ 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 $E[d_2 | \theta_1, m]$ and $E[d_1 | \theta_2, m]$ 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 :

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

Similarly, we have that

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

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. As $\delta \rightarrow \infty$, the division managers only rely on the communicated information and set $d_1^D = d_2^D = (E[\theta_1 | \theta_2, m] + E[\theta_2 | \theta_1, m])/2$. Note that, for given posteriors, these are exactly the same decisions that headquarters 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

To understand how communication works in our model, it is useful to start by analyzing the division managers' incentives to misrepresent their information. For this purpose, suppose that

the division managers can credibly misrepresent their information, that is, they can choose the posterior beliefs of the person they are communicating with. Regardless of the organizational structure, the division managers then tend to exaggerate their states. To see this, consider first the incentives of Manager 1 to misrepresent his information under Centralization.²⁶ When making decision d_1 , the HQ Manager puts more weight on coordinating it with d_2 , and less on adapting it to θ_1 , than Manager 1 would like her to. Since $E[\theta_2] = 0$ this implies that if Manager 1 truthfully communicated the state, the decision $E[d_1 | \theta_1]$ that he would expect headquarters to make would not be sufficiently extreme from his perspective. In other words, if $\theta_1 > 0$, he would expect headquarters to make a decision $d_1 > 0$ that is too small from his perspective and if $\theta_1 < 0$ he would expect headquarters to make a decision $d_1 < 0$ that is not small enough. To induce headquarters to choose a more extreme decision, Manager 1 therefore exaggerates his state by reporting $m_1 > \theta_1$ if $\theta_1 > 0$ and $m_1 < \theta_1$ if $\theta_1 < 0$.

Consider next the incentives of Manager 1 to misrepresent his information when he communicates with Manager 2 under Decentralization. When Manager 2 makes decision d_2 , he puts less weight on coordinating it with d_1 , and more on adapting it to θ_2 , than Manager 1 would like him to. Since $E[\theta_2] = 0$, this implies that if Manager 1 truthfully communicated his state, the decision $E[d_2 | \theta_1]$ that he would expect Manager 2 to make would not be sufficiently extreme from his perspective. In particular, if $\theta_1 > 0$, he would expect Manager 2 to make a decision that is too small. To induce Manager 2 to make a larger decision, Manager 1 would like him to believe that he will choose a more extreme decision d_1 . For this reason, Manager 1 again exaggerates his state, that is, he reports $m_1 > \theta_1 > 0$. Similarly, if $\theta_1 < 0$, Manager 1 exaggerates his state by reporting $m_1 < \theta_1$ to induce Manager 2 to make a smaller decision than he otherwise would.

The division managers' incentives to misrepresent their information are therefore qualitatively similar under the two organizational structures. Moreover, it is intuitive that under both structures, the division managers have a stronger incentive to exaggerate their information, the more biased they are towards their own divisions. There is, however, an important difference in how the incentives to exaggerate are affected by changes in the need for coordination. In particular, while under Centralization an increase in the need for coordination exacerbates the division managers' incentives to exaggerate their information, it mitigates them under Decentralization. To understand this, recall that under Centralization the HQ Manager puts *more* weight on coordinating d_1 with d_2 , and less weight on adapting it to the communicated value of θ_1 , than Manager 1 would like her to. An increase in the need for coordination then makes the HQ Manager even less

²⁶The argument for Manager 2 is analogous.

responsive to the communicated information and thus increases the division manager's incentives to exaggerate his information. In contrast, under Decentralization, Manager 2 puts *less* weight on coordinating d_2 with what he expects decision d_1 to be, and more weight on adapting it to θ_2 , than Manager 1 would like him to. An increase in the need for coordination then makes Manager 2 more responsive to the communicated information and thereby reduces the division manager's incentive to exaggerate his information.

To understand the incentives to misrepresent information more formally, consider first Centralization. Let $\nu_1 = E[\theta_1 | m_1]$ 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. Ideally, 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). 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 , that is, that $E_{m_2} [E[\theta_2 | m_2]] = E[\theta_2] = 0$. Assuming that this relationship holds we can use (10) to obtain

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

Since $b_C \geq 0$ this confirms the above intuition that Manager 1 exaggerates his state whenever $\theta_1 \neq 0$. Only when $\theta_1 = 0$ does he have an incentive to communicate truthfully. Moreover, it can be seen that his incentives to exaggerate are increasing in $|\theta_1|$. It is also straightforward to verify that b_C is increasing in λ and δ . Thus, as explained above, Manager 1's incentive to exaggerate is increasing in the own-division bias and in the need for coordination.

Consider next the division managers' incentives to misrepresent their information under Decentralization. For this purpose, let $\nu_1 = E[\theta_1 | \theta_2, m_1]$ 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). If we assume again that $E_{m_2} [E[\theta_2 | \theta_1, m_2]] = E[\theta_2] = 0$, 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 $|\theta_1|$. Moreover, it can be verified that b_D is increasing in λ and decreasing in δ .

Thus, as anticipated above, Manager 1's incentive to exaggerate is increasing in his own-division bias and decreasing in the need for coordination.

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 $\mu_j(m_j | \theta_j)$. Under Centralization, 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 $d_1^C(m)$ and $d_2^C(m)$. Under Decentralization, the decision rule for Manager 1 maps the state θ_1 and messages $m_1 \in M_1$ and $m_2 \in M_2$ into decision $d_1 \in \mathbf{R}$ while the decision rule for Manager 2 maps the state θ_2 and messages $m_1 \in M_1$ and $m_2 \in M_2$ into decision $d_2 \in \mathbf{R}$, and we denote them by $d_1^D(m, \theta_1)$ and $d_2^D(m, \theta_2)$. 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. Formally, whenever $\mu_1(m_1 | \theta_1) > 0$ then $m_1 \in \arg \max_{m \in M_1} E[\lambda \pi_1^l + (1 - \lambda) \pi_2^l | \theta_1]$ for $l = C, D$, where π_1^l and π_2^l are the profits of Divisions 1 and 2 given that decisions are made according to $d_1^l(\cdot)$ and $d_2^l(\cdot)$. The requirement for $\mu_2(m_2 | \theta_2)$ is analogous. Under Centralization the decision rules $d_1^C(\cdot)$ and $d_2^C(\cdot)$ solve $\max_{(d_1, d_2)} E[\pi_1 + \pi_2 | m]$ and under Decentralization the decision rules $d_1^D(\cdot)$ and $d_2^D(\cdot)$ solve $\max_{d_1} E[\lambda \pi_1 + (1 - \lambda) \pi_2 | m, \theta_1]$ and $\max_{d_2} E[(1 - \lambda) \pi_1 + \lambda \pi_2 | m, \theta_2]$ respectively. Finally, the belief functions satisfy $g(\theta_j | m) = \mu_j(m_j | \theta_j) / \int_P \mu_j(m_j | \theta_j) d\theta_j$, where $P = \{\theta_j : \mu_j(m_j | \theta_j) > 0\}$ and $j = 1, 2$.

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

$(a_{j,-N}, \dots, a_{j,-1}, a_{j,0}, a_{j,1}, \dots, a_{j,N})$ 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, that is, $a_{j,i} = a_{j,-i}$ 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 $i = 1, \dots, N_j - 1$

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

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

$$b = b_D \text{ under Decentralization, where } b_D \text{ is defined in (12),}$$

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

$$d_j(m, \theta_j) = d_j^D$$
, $j = 1, 2$, 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 $4b_C a_{j,i}$ 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.²⁷ It is intuitive that since

²⁷This 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

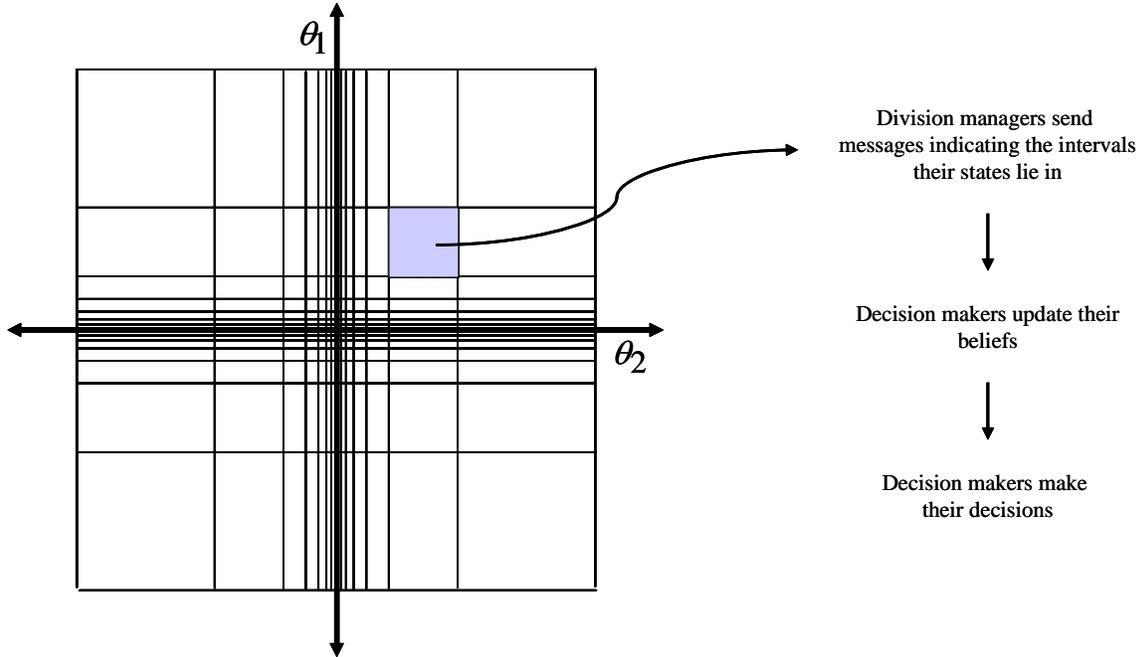


Figure 2: Communication Equilibria

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 1.²⁸

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 $4ba_{j,i}$.

²⁸See our discussion of equation (18).

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.²⁹ 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.

²⁹We share this feature with Melumad and Shibano (1991); see the related literature in Section 2.

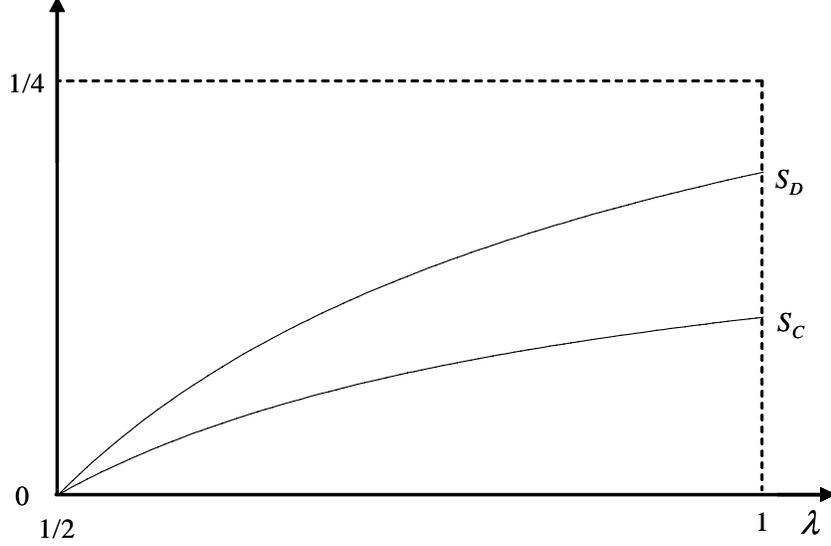


Figure 3: Communication Comparison (λ)

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.) are illustrated in Figure 3 and show that vertical communication is in general more efficient than horizontal communication. 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 difference in the quality of the two modes of communication 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 difference in the quality of the two modes of communication not only shrinks but actually vanishes. This can be understood by recalling from Section 4 that

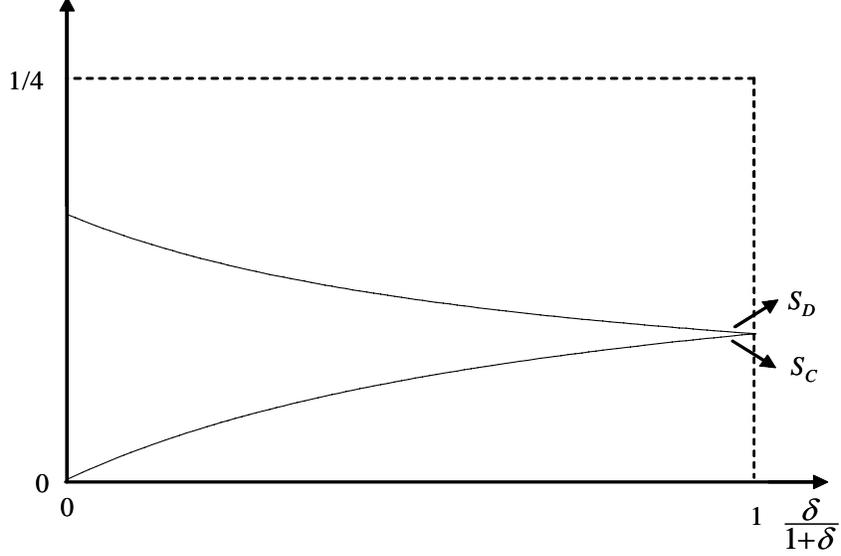


Figure 4: Communication Comparison (δ)

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, that is, 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' bias.

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 information to efficiently adapt decisions to the local conditions.

The next lemma shows that because of these factors coordination losses are always lower un-

der Centralization and adaptation losses are always lower under 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 $\delta \in [0, \infty)$, $AL_C \geq AL_D$ and $CL_C \leq CL_D$.

In what follows, we refer to $\Delta AL \equiv AL_C - AL_D \geq 0$ as the *adaptation advantage* of the decentralized structure and to $\Delta CL \equiv CL_C - CL_D \geq 0$ as the *coordination advantage* of the centralized one. The relative performance of the two structures can then be stated as

$$\Pi_C - \Pi_D = -\Delta AL + 2\delta\Delta CL. \quad (17)$$

From the result that Centralization has a coordination advantage, one could reason naively that Centralization will prevail if the need for coordination δ is sufficiently important. A key insight of this section is that this reasoning is flawed. To see this, consider first Figure 5 which plots the relative performance $\Pi_C - \Pi_D$ for $(\sigma_1^2 + \sigma_2^2) = 1$.³⁰ The figure shows that when $\lambda > 1/2$ is sufficiently small, Decentralization strictly outperforms Centralization for any finite δ . In other words, the Delegation Principle – which states that decision rights should be delegated to the best informed managers provided that their incentives are sufficiently aligned – always holds, independent of the need for coordination.³¹ Thus, as claimed above, Decentralization can dominate Centralization even when coordination is very important. Essentially, when the own-division bias is very small and coordination is very important, the division managers have both the incentives and the information to coordinate their decisions with each other. In contrast, the ability of headquarters to adapt the decisions to their states is limited by the fact that vertical communication becomes less efficient as coordination becomes more important. As a result, the adaptation advantage of the decentralized structure can dominate the coordination advantage of the centralized one even when coordination is very important.

Figure 5 also shows that Decentralization strictly dominates Centralization for any own-division bias $\lambda > 1/2$ when the need for coordination $\delta > 0$ is sufficiently small. Centralization therefore only dominates Decentralization when both the need for coordination and the own-division bias are large enough. Finally, the figure shows that the two organizational structures perform equally well when either $\lambda = 1/2$, $\delta = 0$ or $\delta \rightarrow \infty$. When $\lambda = 1/2$ or $\delta = 0$, there is no incentive conflict between the division managers or between them and headquarters and, as a result, first best expected profits are realized under both structures. The result that the two structures

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

³¹For the Delegation Principle see, for instance, Milgrom and Roberts (1992).

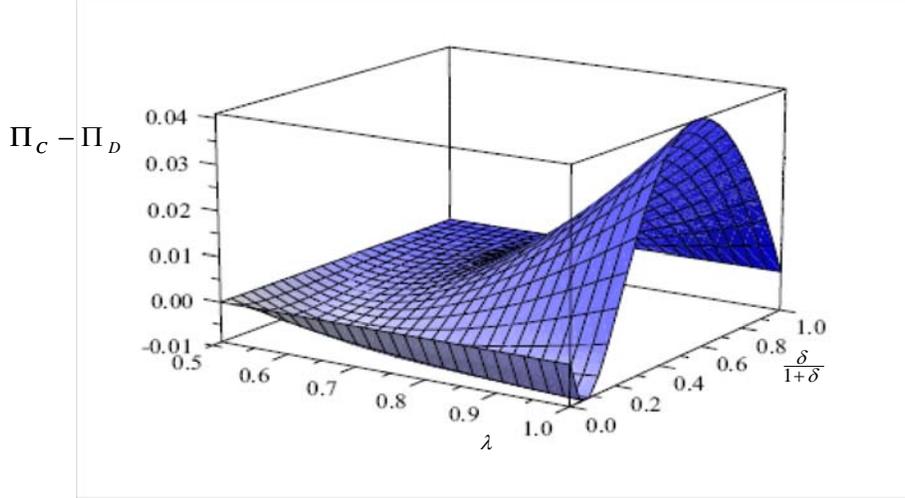


Figure 5: Organizational Performance in Three Dimensions

perform equally well when $\delta \rightarrow \infty$ is an immediate consequence of the fact that decision making and communication under the two structures both converge in the limit in which coordination becomes all important, as was shown previously. The next proposition compares the relative performance of the two organizational structures more formally.

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

- i. λ small:* *if $\lambda \in (1/2, 17/28)$, then Decentralization strictly dominates Centralization.*
- ii. δ small:* *if $\delta \in (0, \bar{\delta})$, then Decentralization strictly dominates Centralization, where $\bar{\delta} > 0$ is defined in the proof.*
- iii. δ and λ large:* *if $\delta \in (\bar{\delta}, \infty)$, then Decentralization strictly dominates Centralization for all $\lambda \in [17/28, \bar{\lambda}(\delta))$ and Centralization strictly dominates Decentralization for all $\lambda \in (\bar{\lambda}(\delta), 1]$, where $\bar{\lambda}(\delta) > 17/28$ is defined in the proof.*

The proposition is illustrated in Figure 6 and summarizes the main insights that we discussed above. We now turn to the intuition for the main results reported in the proposition.

Small λ The first part of Proposition 5 shows that when the own-division bias $\lambda > 1/2$ is sufficiently small, the adaptation advantage of the decentralized structure outweighs the coordination advantage of the centralized one. In particular, it shows that this is the case even when the importance of coordination is arbitrarily high. To understand this, note that when the own-division bias is small, the inefficiency of decentralized decision making is very limited and, as a result, the

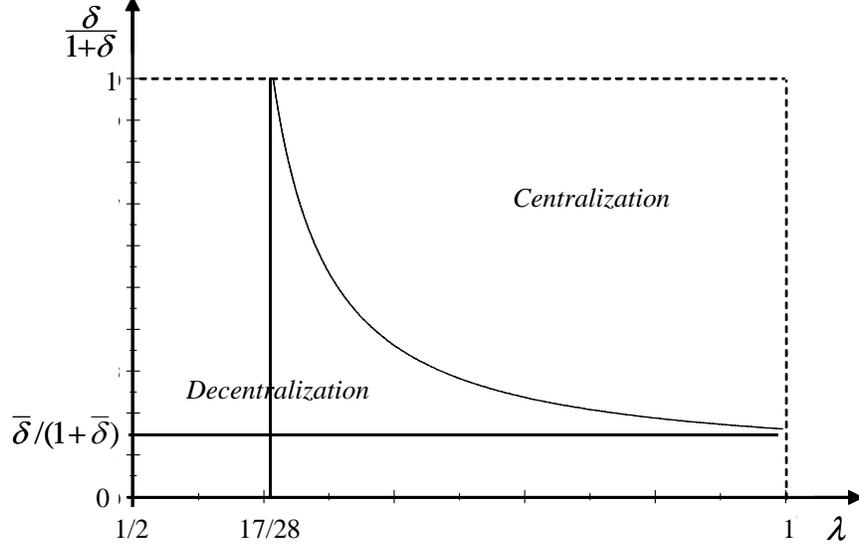


Figure 6: Organizational Performance in Two Dimensions

relative organizational performance of the two structures is determined entirely by the differences in the quality of horizontal and vertical communication. Note next that the more important coordination, the better horizontal communication and thus the easier it is for the division managers to coordinate their decisions. In contrast, an increase in the need for coordination worsens vertical communication which makes it more difficult for headquarters to adapt the decisions to the local information. As coordination becomes more important, the coordination advantage of the centralized structure therefore becomes smaller and smaller relative to the adaptation advantage of the decentralized structure. As a result, Decentralization outperforms Centralization even when coordination is very important.

To see this more formally, recall that the difference in the expected profits between the two organizational structures is given by (17), where ΔAL is the adaptation advantage of the decentralized structure and ΔCL is the coordination advantage of the centralized structure. At $\lambda = 1/2$ there is no incentive conflict between the division managers and between them and headquarters. As a result, the two structures perform equally well in terms of coordination and adaptation, that is, $\Delta CL = \Delta AL = 0$. Consider now the effect of a marginal increase in the own-division bias:

$$\frac{d(\Pi_C - \Pi_D)}{d\lambda} = \frac{\partial \Pi_C}{\partial S_C} \frac{\partial S_C}{\partial \lambda} - \frac{\partial \Pi_D}{\partial S_D} \frac{\partial S_D}{\partial \lambda} = -\frac{d\Delta AL}{d\lambda} + 2\delta \frac{d\Delta CL}{d\lambda} \quad \text{for } \lambda = 1/2.$$

The first equality shows that, at $\lambda = 1/2$, an increase in the own-division bias only affects the relative performance of the two organizational structures through its effect on the quality of communication.

Small changes in decision making do not have a first order effect on the expected profits since decision making is efficient at $\lambda = 1/2$. The second equality shows that the effect of an increase in the own-division bias can be decomposed into its effect on the adaptation advantage of the decentralized structure and the coordination advantage of the centralized one. In particular, an increase in the own-division bias worsens vertical communication which hinders the ability of headquarters to adapt the decisions to the local circumstances. As a result, the adaptation advantage of the decentralized structure increases, that is, $d\Delta AL/d\lambda > 0$ for all $\delta \in (0, \infty)$. An increase in the own-division bias also worsens horizontal communication which hinders the ability of the division managers to coordinate and, as a result, the coordination advantage of the centralized structure increases, that is, $d\Delta CL/d\lambda > 0$ for all $\delta \in (0, \infty)$. Overall, the adverse effect on the ability of headquarters to adapt outweighs the adverse effect on the division managers' ability to coordinate. In other words, when $\lambda = 1/2$, $d\Delta AL/d\lambda > 2\delta d\Delta CL/d\lambda$ for all $\delta \in (0, \infty)$ and thus $d(\Pi_C - \Pi_D)/d\lambda < 0$. Since $\Pi_C - \Pi_D = 0$ for $\lambda = 1/2$, this implies that $\Pi_C - \Pi_D < 0$ when $\lambda > 1/2$ is sufficiently small.

Small δ The second part of Proposition 5 shows that when the need for coordination is sufficiently small, the decentralized structure outperforms the centralized one, independent of the division managers' bias. Essentially, even when the need for coordination is very small, communication is noisy under both structures. Under Centralization this affects the ability of headquarters to adapt the decisions to the local conditions while under Decentralization it affects the ability of the of division managers to coordinate their actions. Since for small δ coordination is much less important than adaptation, the centralized structure suffers more from imperfect communication than the decentralized one does. As a result, Decentralization dominates Centralization.

To see this more formally, consider the relative performance of the two structures in the neighborhood of $\delta = 0$. Clearly, when $\delta = 0$ there is no incentive conflict between the division managers or between them and headquarters and, as a result, $\Pi_C - \Pi_D = 0$. Consider now the effect of a marginal increase in the need for coordination on the relative performance of the two structures which can be obtained by differentiating (17):

$$\frac{d(\Pi_C - \Pi_D)}{d\delta} = -\frac{d\Delta AL}{d\delta} \quad \text{for } \delta = 0. \quad (18)$$

This expression shows that at $\delta = 0$ the effect of a marginal increase in δ on $\Pi_C - \Pi_D$ depends only on how such an increase affects the adaptation advantage of the decentralized structure. This is so since the coordination advantage ΔCL and the weight 2δ on this advantage are both equal to zero. To understand the effect of an increase in the need for coordination on the adaptation advantage,

recall that such an increase worsens vertical communication which limits the ability of headquarters to adapt the decisions to the states. The ability of the division managers to adapt their decisions, in contrast, is unaffected by any changes in the quality of horizontal communication since these managers observe their own states directly themselves. It is therefore intuitive that at $\delta = 0$ an increase in the need for coordination increases the adaptation advantage, that is, $d\Delta AL/d\delta > 0$. Formally, for $\delta = 0$ and $\lambda > 1/2$,

$$\frac{d(\Pi_C - \Pi_D)}{d\delta} = -\frac{d\Delta AL}{d\delta} = -\left[\frac{\partial AL_C}{\partial S_C} \frac{\partial S_C}{\partial \delta} - \frac{\partial AL_D}{\partial S_D} \frac{\partial S_D}{\partial \delta}\right] = -\frac{\partial AL_C}{\partial S_C} \frac{dS_C}{d\delta} < 0.$$

Since, for $\delta = 0$ and $\lambda > 1/2$, $\Pi_C - \Pi_D = 0$ and $d(\Pi_C - \Pi_D)/d\delta < 0$, the decentralized structure outperforms the centralized one in the neighborhood of $\delta = 0$.

Large δ and λ The first two parts of Proposition 5 show that the decentralized structure outperforms the centralized one when either the division managers are not very biased or coordination is not very important. The third part, in contrast, shows that the centralized structure tends to outperform the decentralized one when the division managers are sufficiently biased and coordination is sufficiently important. In particular, for any $\delta \in (\bar{\delta}, \infty)$, Centralization strictly outperforms Decentralization when the own-division bias is sufficiently large. There are two reasons for this. First, when division managers are very biased, vertical communication is a lot more efficient than horizontal communication, as can be seen in Figure 3. Second, since increases in the own-division bias distort decision making by the division managers but not by headquarters, decision making under Centralization is significantly more efficient than decision making under Decentralization when the own-division bias is large. It is because of these two factors that Centralization tends to outperform Decentralization when the division managers are very biased and coordination is important.

7 Empirical Implications and Evidence

In principle, the main predictions of the model could be tested using cross-sectional data. For this purpose one would need information about how balanced the incentives of division managers are, that is, how much their compensation depends on divisional rather than firm wide profits. One would also require information about the need for coordination between divisions and the organizational structure of firms. With such data, one could test the predictions that are summarized in Proposition 5 and illustrated in Figure 6. First, for a given need for coordination, there should, in general, be a positive relationship between the own-division bias and the degree to which firms are

centralized. Second, for a given own-division bias, there should, in general, be a positive relationship between the need for coordination and the degree of centralization. Third, the first correlation – between the own-division bias and the degree of centralization – should only hold if the need for coordination is sufficiently high. Similarly, the second correlation – between the need for coordination and the degree of centralization – should only hold if division managers are sufficiently biased towards their own divisions. Firms in which coordination between divisions is not very important, or in which the incentives of division managers are very balanced, should be mostly decentralized.

While the model can be potentially tested using cross-sectional data, a more fruitful approach might be to exploit exogenous variation in the need for coordination in a group of firms. To illustrate the panel data implications of the model, suppose that a technological innovation increases the scope for inter-divisional coordination in an industry in which a group of firms is active.³² Suppose also that, due to exogenous differences in the needs to motivate division-specific efforts, the incentive contracts of division managers differ across firms in how much weight they put on divisional and firm wide profits.³³ The model then predicts that, in response to the technological innovation, the probability of a reorganization, in particular a move towards centralization, should be higher in firms in which division managers are biased towards their own divisions than in firms in which they are less biased.

We are not aware of any existing econometric studies that investigate these types of predictions. There is, however, some informal evidence that is consistent with a basic implication of the model. In particular, it appears that successful decentralization tends to require the balancing of the division managers' incentives. Before 1918, for instance, GM was essentially a holding company of independent firms in which “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.”³⁴ As discussed in the Introduction, GM dealt with these problems of decentralization by establishing various inter-divisional committees that facilitated horizontal communication. Equally important, however, was the introduction of a new incentive scheme that balanced the division managers' incentives. In

³²Technological innovations in the 1980s and early 90s, for example, lead to the rise of the ‘platform design process’ which allowed multinational car manufacturers to “reap global economies of scale on the parts that can be produced in common for multiple regions of the world and to be locally responsive on those features where such responsiveness is valued” (Hannan, Podolny and Roberts 1999, p.2).

³³In the Conclusions we discuss how a version of our model in which division managers need to make division-specific effort decisions can be used to endogenize the own-division bias.

³⁴Sloan (1964), p.409.

particular, GM adopted the General Motors Bonus Plan which asserted that “each individual should be rewarded in proportion to his contribution to the profit of his own division and of the corporation as a whole.”³⁵ According to Alfred Sloan, “the Bonus Plan played almost as big a role as our system of co-ordination in making decentralization work effectively.”³⁶

British Petroleum provides another example of the importance of balancing incentives in decentralized firms. In the early 1990s the oil and gas exploration division, known as BPX, was headed by John Browne, who later became BP’s chief executive. He decentralized BPX, creating almost fifty semi-autonomous business units. Initially, since “business unit leaders were personally accountable for their units’ performance, they focused primarily on the success of their own businesses rather than on the success of BPX as a whole.”³⁷ To encourage coordination between the business units, BPX established peer groups which were similar to the committees in GM and were intended as forums in which the heads of the different business units could exchange information and ideas. Also as in the case of GM, BPX complemented the introduction of these peer groups by changes in the implicit and explicit incentives of business unit leaders to reward and promote them not just based on the success of their own division but also for contributing to the successes of other business units. As a result, “‘Lone stars’ – those who deliver outstanding business unit performance but engage in little cross-unit collaboration – can survive within BP, but their careers typically plateau.”³⁸

8 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 while such asymmetries tend to favor Centralization, our central result – that Decentralization can be optimal even when coordination is very important – continues to hold, albeit in a weaker form. Since the analyses of the different extensions are very similar, we focus on just one asymmetry, namely sequential decision making. Differences in the importance of coordination and in the division sizes are briefly discussed at the end of this section. The formal analyses of all three extensions are contained in Appendices B, C and D.

³⁵Sloan (1964), p.407.

³⁶Sloan (1964), p.409.

³⁷Hansen and von Oetinger (2001), p.3.

³⁸Hansen and von Oetinger (2001), p.8.

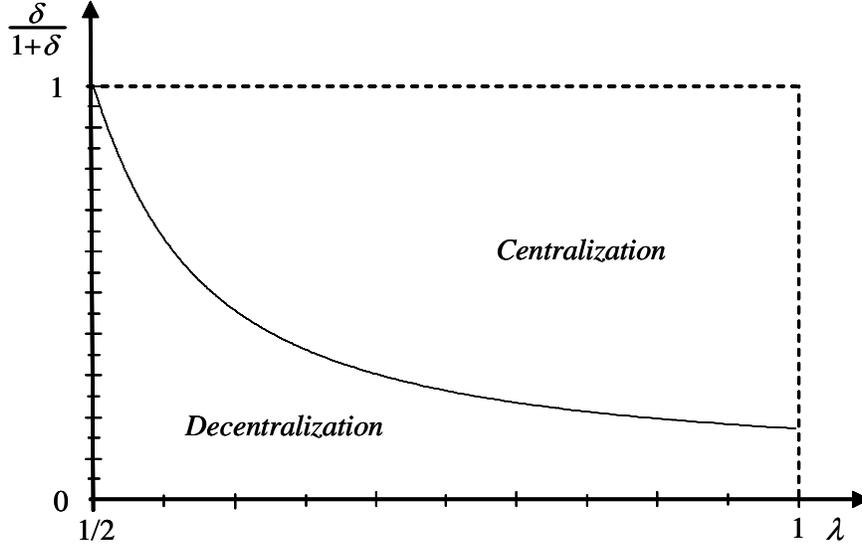


Figure 7: Organizational Performance with Sequential Decision Making when $\sigma_1^2 = \sigma_2^2$

8.1 Sequential Decision Making

In some settings one division manager may have a first mover advantage, that is, he may be able to make his decision before the other division managers can do so. To explore this possibility, we now consider a version of our model in which 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.

Sequential decision making eliminates the Follower's uncertainty about the Leader's decision. As such, it might seem that sequential decision making facilitates coordination and thus further strengthens the result that Decentralization can be optimal even when coordination is very important. This, however, is not the case. To see this, consider Figure 7 which illustrates the relative performance of Centralization and Decentralization with sequential decision making when $\sigma_1^2 = \sigma_2^2$. It can be seen that when coordination is sufficiently important, Centralization strictly dominates Decentralization with sequential decision making for any $\lambda > 1/2$. There are two reasons for the result that sequential decision making tends to favor Centralization. First, 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

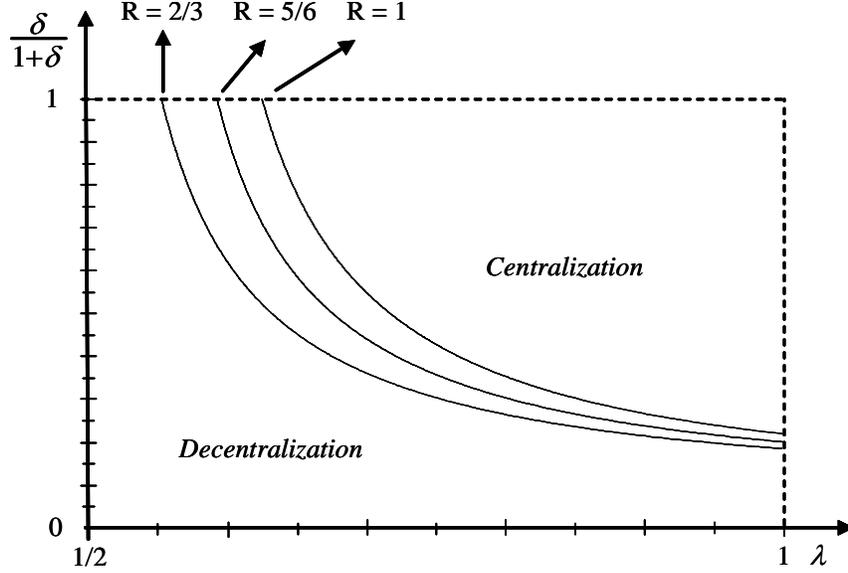


Figure 8: Organizational Performance with Sequential Decision Making when $R \equiv \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} > 1/2$

be less coordinated under sequential than under simultaneous decision making. Second, because the Leader adapts his decision more closely to his state, the Follower has a stronger incentive to misrepresent his information so as to influence the Leader's decision making. As a result, communication under sequential decision making is worse than under simultaneous decision making. The Leader is therefore less well informed about what decision the Follower is going to make than he would be under simultaneous decision making. Other things equal, this tends to make coordination more difficult.

While sequential decision making weakens our result that Decentralization can be optimal even when coordination is very important it does not eliminate it altogether. To see this, consider first Figure 7. While this figure shows that for any $\lambda > 1/2$, Centralization dominates Decentralization when the need for coordination δ is large enough, it also shows that, for any $\delta > 0$, Decentralization dominates Centralization when the own-division bias $\lambda > 1/2$ is small enough. Thus, it is still the case that Decentralization dominates Centralization provided that the division managers are not too biased towards their own divisions. The case for Decentralization is even stronger if the Leader has more private information than the Follower, in the sense that $\sigma_1^2 > \sigma_2^2$. This can be seen in Figure 8 which illustrates the relative performance of Centralization and Decentralization for different values of $R \equiv \sigma_1^2 / (\sigma_1^2 + \sigma_2^2)$. Essentially, the bigger σ_1^2 relative to σ_2^2 , the more the organization benefits from the perfect communication of θ_1 under Decentralization with sequential

decision making, and thus the better Decentralization performs relative to Centralization.

8.2 Other Asymmetries

In some organizations divisions differ in how much their profits depend on coordination. The optimal design of such organizations has recently been analyzed in Rantakari (2006). In Appendix C, we consider a model that follows Rantakari (2006) in allowing the divisions to have different needs for coordination. We show that for any $\lambda > 1/2$, Centralization dominates Decentralization when coordination becomes sufficiently important for one of the divisions. Importantly, however, it is still the case that, for any need for coordination, Decentralization dominates Centralization when the own-division bias $\lambda > 1/2$ is sufficiently small. Our central result therefore continues to hold in this extension.

A similar result holds if divisions have the same need for coordination but differ in their sizes. In Appendix D we consider a model in which the profits of Divisions 1 and 2 are given by $2\alpha\pi_1$ and $2(1 - \alpha)\pi_2$, where $\alpha \in (1/2, 1)$ is a parameter that measures the relative size of the two divisions. Once again it can be shown that for any $\lambda > 1/2$, Centralization dominates Decentralization if coordination is sufficiently important. However, it is still the case that Decentralization always dominates Centralization in the neighborhood of $\lambda = 1/2$, provided that α is not ‘too big.’ Thus, our central result continues to hold in asymmetric organizations, albeit in a weaker form.

9 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, the difference in the quality of the two modes of communication diminishes, and eventually vanishes, as coordination becomes more important. As a result, decentralization can be optimal even if coordination is very important. In particular, in symmetric organizations – in which divisions are of equal size and have the same need for coordination and in which decisions are made simultaneously – decentralization *always* outperforms centralization when the division managers’ incentives are sufficiently aligned. The same result also holds in asymmetric organizations in which decisions are made sequentially or in which the divisions differ in their need for coordination. In organizations in which the divisions differ in terms of their size the result also holds as long as the

size difference is not too large.

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.³⁹ 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 decentralization 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

³⁹See, for instance, Milgrom and Roberts (1992).

with each other. As a result, the difference in the quality of horizontal and vertical communication diminishes and eventually disappears.

A simplifying assumption in our model is that the own-division bias is exogenously given and independent of the organizational structure. A natural extension would be to endogenize the own-division bias by allowing the firm to design the compensation schemes of the division managers. Perhaps the simplest way to do this would be to suppose that the division managers are risk averse and liquidity constrained and that each one has to make a binary effort choice that affects the profits of his division. In such a setting the firm needs to give the division managers a share of divisional profits to motivate effort. In principle it could then induce truth-telling and efficient decision making by giving each division manager an equal share in the profits of the other division. Since division managers are liquidity constrained, however, the firm may find it optimal to give each division manager a smaller share in the other division's profit than in those of his own division. In the context of such an extension, a key feature of our model is that the benefit of reducing the division managers' own-division bias – in terms of improved communication and, in the case of decentralization, improved decision making – tends to be larger in decentralized organizations than in centralized ones and it tends to be larger in organizations in which coordination is more important. As a result, we conjecture that the result that decentralization can be optimal even when coordination is very important continues to hold in such an extension.

The extension with endogenous compensation also underlines the importance of controlling for the need for coordination when empirically investigating the relationship between organizational structure and the own-division bias of division managers. In the main model analyzed in this paper, there is a positive relationship between the own-division bias and the degree of centralization when the need for coordination is high while decentralization is optimal for any own-division bias when the need for coordination is low. We conjecture that in the model with endogenous incentives it is still the case that there is a positive relationship between the own-division bias and the degree of centralization when the need for coordination is high. When the need for coordination is low, however, it is now optimal to delegate the decision rights to division managers who are mainly compensated for the success of their own divisions. In such a model, therefore, one would expect to see decentralized organizations with very biased division managers when coordination is not important and to see decentralized organizations with fairly unbiased division managers when coordination is important.

Beyond its implications for organizational design, it is tempting to interpret our theory as one of horizontal firm boundaries. To do so, however, one would have to address the difference

between a decentralized multi-divisional firm and several independent single-division firms. One difference, emphasized in Holmström (1999) and, informally, in Roberts (2004), is that the owners of a multi-divisional firm can use a much wider array of tools to align the incentives of their division managers than the owners of independent firms can use to align the incentives of their managers across firm boundaries.⁴⁰ This suggests that multi-divisional firms exist to ensure coordination by either centralizing decision rights in a powerful headquarters or, as in the GM and BP examples, by providing balanced incentives for the managers in charge of decentralized divisions. Independent, single division firms, in contrast, sacrifice coordination for the benefits of the high-powered incentives provided by the market. A model of firm boundaries along these lines awaits future research.

⁴⁰See Roberts (2004), pp.106-115.

10 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. 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, that is, 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

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

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 $k \leq N$, $a_{1,k} = a_{1,N-k}$, that is, 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

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

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

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

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

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

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 $k \leq N$, $a_{1,k} = a_{1,N-k}$, that is, 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

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

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

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

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

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

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

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

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

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

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

$$\frac{(\sigma_1^2 + \sigma_2^2) \lambda \delta (2\lambda - 1) f}{(4\delta + 1)(3\lambda + \delta(8\lambda - 1))(\lambda(5\lambda - 1) + \delta(8\lambda - 1))(\lambda + \delta)(\lambda + 2\delta)}, \quad (38)$$

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 6. 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.): Differentiating (38) gives

$$\frac{d(\Pi_C - \Pi_D)}{d\lambda} - \frac{4\delta(\sigma_1^2 + \sigma_2^2)}{3(1 + 4\delta)^2} \quad \text{for } \lambda = 1/2$$

which is negative for all $\delta > 0$. Since $(\Pi_C - \Pi_D) = 0$ for $\lambda = 1/2$ this implies that $\Pi_C - \Pi_D < 0$ in the neighborhood of $\lambda = 1/2$.

Parts (ii.) and (iii.): Since $\bar{\lambda}(\delta)$ is decreasing in δ it is sufficient to show that for $\lambda = 1$, $\Pi_C - \Pi_D < 0$ if $\delta \in (0, \bar{\delta})$ and $\Pi_C - \Pi_D > 0$ if $\delta \in (\bar{\delta}, \infty)$. 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 $\delta \in (0, \bar{\delta})$ and $f > 0$ if $\delta \in (\bar{\delta}, \infty)$. ■

11 Appendix B - Sequential Decision Making

Decision Making: In the last stage of the game Manager 2 chooses d_2 to maximize his expected utility $E[(1 - \lambda)\pi_1 + \lambda\pi_2 | \theta_2, d_1]$. The optimal decision that solves this problem is given by

$$d_2^S \equiv \frac{\lambda}{\lambda + \delta} \theta_2 + \frac{\delta}{\lambda + \delta} d_1^S. \quad (39)$$

At the previous stage Manager 1 chooses d_1 to maximize his expected profits $E[\lambda\pi_1 + (1 - \lambda)\pi_2 | \theta_1, m]$.

The optimal decision is given by

$$d_1^S \equiv \frac{\lambda(\lambda + \delta)^2}{\lambda^3 + 3\lambda^2\delta + \delta^2} \theta_1 + \delta \frac{\lambda^2 + \delta(1 - \lambda)}{\lambda^3 + 3\lambda^2\delta + \delta^2} E[\theta_2 | \theta_1, m]. \quad (40)$$

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 A1 (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 $i = 1, \dots, N_2 - 1$,*
 $a_{2,-(i+1)} - a_{2,-i} = a_{2,-i} - a_{2,-(i-1)} + 4b_S a_{2,-i}$ 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 (39) and (40).*

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 A2 (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 A1. *In the most efficient equilibrium in which $N_2 \rightarrow \infty$ the residual variance is given by*

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

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 A4 (Organizational Performance). *Under Decentralization with sequential decision making the expected profits are given by*

$$\Pi_S = -((A_D + X)(\sigma_1^2 + \sigma_2^2) + (B_D - X)S_S\sigma_2^2), \quad (41)$$

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 the following proposition.

PROPOSITION A5 (Sequential Decision Making). *Suppose that $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Then,*

- 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.*

Proof: i. Applying l'Hopital's Rule to (14) and (41) 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 (41) and using the assumption that $\sigma_1^2 = \sigma_2^2 = \sigma^2$ we get that

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

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

Finally, Figures 7 and 8 are drawn using Propositions 4 and A4.

12 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 A6 which summarizes the claims in the main text. The derivation of these expressions and their interpretation are exactly as in the main model. Also, to simplify we assume that $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

12.1 Centralization

The decisions are now given by

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

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

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

The expected profits are given by

$$\Pi_C = -\sigma^2 \left(2 \frac{\delta_1 + \delta_2}{1 + 2(\delta_1 + \delta_2)} + (S_1 + S_2) \frac{1 + \delta_1 + \delta_2}{1 + 2(\delta_1 + \delta_2)} \right). \quad (42)$$

Applying l'Hopital's Rule gives

$$\lim_{\delta_1 \rightarrow \infty} \Pi_C = -2 \frac{5\lambda - 1}{8\lambda - 1} \sigma^2. \quad (43)$$

We can also use (42) to evaluate $d\Pi_C/d\lambda$:

$$\frac{d\Pi_C}{d\lambda} = -\frac{4}{3} \frac{(\delta_1 + \delta_2)}{1 + 2(\delta_1 + \delta_2)} \sigma^2 \quad \text{for } \lambda = 1/2. \quad (44)$$

12.2 Decentralization

The decisions under Decentralization are now given by

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

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

$$\Pi_D = -\mathbb{E}\left[(d_1^D - \theta_1)^2 + (d_2^D - \theta_2)^2 + (\delta_1 + \delta_2)(d_1^D - d_2^D)^2\right], \quad (45)$$

where

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

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

Applying l'Hopital's Rule gives

$$\lim_{\delta_1 \rightarrow \infty} \Pi_D = -2 \frac{8\lambda^3 - 9\lambda^2 + 6\lambda - 1}{5\lambda - 1} \sigma^2. \quad (46)$$

We can also use (45) to evaluate $d\Pi_C/d\lambda$ for $\lambda = 1/2$:

$$\frac{d\Pi_D}{d\lambda} = -\frac{8}{3} \frac{(\delta_1 + \delta_2)^2}{(1 + 2(\delta_1 + \delta_2))^2} \sigma^2. \quad (47)$$

We can now prove the following proposition.

PROPOSITION A6 (Different Needs for Coordination).

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

Proof: i. Using (43) and (46) 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 (44) and (47) 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 $\delta_1, \delta_2 > 0$. ■

13 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 A7 which summarizes the claims in the main text. The derivation of these expressions and their interpretation are exactly as in the main model. Also, to simplify we assume that $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

13.1 Centralization

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

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

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

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

where

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

Applying l'Hopital's Rule we find that

$$\lim_{\delta \rightarrow \infty} \Pi_C = \frac{-2\alpha(1 - \alpha)(8\lambda - 1)(5\lambda - 1)}{(5\lambda - 1 - \alpha(2\lambda - 1))(3\lambda + (2\lambda - 1)\alpha)} \sigma^2. \quad (48)$$

Also, differentiating we find that

$$\frac{d\Pi_C}{d\lambda} = -\frac{8}{3} \frac{\alpha(1 - \alpha)\delta}{\alpha(1 - \alpha) + \delta} \sigma^2 \quad \text{for } \lambda = 1/2. \quad (49)$$

13.2 Decentralization

The decisions are now given by

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

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

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

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

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

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

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

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

Applying l'Hopital's Rule we find that

$$\lim_{\delta \rightarrow \infty} \Pi_D = \frac{-2\alpha(1-\alpha) \left(2\alpha(1-\alpha) + 2\lambda^2(2\alpha-1)^2(3\lambda-7) - 26\alpha\lambda(1-\alpha) + 9\lambda - 1 \right)}{\left(1 - 2\alpha(1-\alpha) - \lambda(2\alpha-1)^2 \right) (3\lambda(1-\lambda) + \alpha(1-\alpha)(6\lambda+1)(2\lambda-1))} \sigma^2. \quad (50)$$

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

$$\frac{d\Pi_D}{d\lambda} = -\frac{16}{3} \alpha(1-\alpha) \delta^2 \frac{1-2\alpha(1-\alpha)}{(\alpha(1-\alpha)+\delta)^2} \sigma^2. \quad (51)$$

We can now prove the following proposition.

PROPOSITION A7 (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.

Proof: (i.) Using (48) and (50) 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 (49) and (51) 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

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

References

- [1] AOKI, M. 1986. Horizontal versus Vertical Information Structure of the Firm. *American Economic Review*, 76: 971-983.
- [2] ALONSO, R. AND N. MATOUSCHEK. 2007. Relational Delegation. *Rand Journal of Economics*, forthcoming.
- [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] BLUME, A., O. BOARD AND K. KAWAMURA. 2007. Noisy Talk. Mimeo, University of Pittsburgh.
- [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] CRAWFORD, V. AND J. SOBEL. 1982. Strategic Information Transmission. *Econometrica*, 50:1431-145.
- [12] CREMER, J., L. GARICANO AND A. PRAT. 2007. Language and the Theory of the Firm. *Quarterly Journal of Economics*, 122(1): 373-407.
- [13] DESSEIN, W. 2002. Authority and Communication in Organizations. *Review of Economic Studies*, 69: 811-838.

- [14] DESSEIN, W. AND T. SANTOS. 2006. Adaptive Organizations. *Journal of Political Economy*, 114(5): 956-995.
- [15] DESSEIN, W., L. GARICANO AND R. GERTNER. 2005. Organizing for Synergies. Mimeo, Chicago GSB.
- [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] FRIEBEL, G. AND M. RAITH. 2006. Resource Allocation and Firm Scope. Mimeo, University of Toulouse.
- [20] GERTNER, R. 1999. Coordination, Dispute Resolution, and the Scope of the Firm, Mimeo, GSB Chicago.
- [21] 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.
- [22] HANNAN, M., J. PODOLNY AND J. ROBERTS. 1999. The Daimler Chrysler Commercial Vehicles Division. Stanford Business School Case IB-27.
- [23] HARRIS, M. AND A. RAVIV. 2005. Allocation of Decision-Making Authority. *Review of Finance*, September 2005, 9, 353-83.
- [24] HART, O. AND B. HOLMSTRÖM. 2002. A Theory of Firm Scope. Mimeo, MIT.
- [25] HART, O. AND J. MOORE. 1990. Property Rights and the Nature of the Firm. *Journal of Political Economy*, 98, 1119-58.
- [26] HART, O. AND J. MOORE. 2005. On the Design of Hierarchies: Coordination versus Specialization. *Journal of Political Economy*, 113(4): 675-702.
- [27] HAYEK, F. 1945. The Use of Knowledge in Society. *American Economic Review*, 35(4): 519-530.

- [28] HOLMSTRÖM, B. 1999. The Firm as a Subeconomy. *Journal of Law, Economics and Organization*, 15(1): 74-102.
- [29] JENSEN, M. AND W. MECKLING. 1992. Specific and General Knowledge, and Organizational Structure. In *Contract Economics*, (Lars Werin and Hans Wijkander, eds.), Oxford: Blackwell.
- [30] KRISHNA, V. AND J. MORGAN. 2001. A Model of Expertise. *Quarterly Journal of Economics*, 116: 747-775.
- [31] 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
- [32] LOCKWOOD, B. 2005. Fiscal Decentralization: A Political Economy Perspective, in *The Handbook of Fiscal Federalism* (E. Ahmad and G. Brosio eds.), Edward Elgar, forthcoming.
- [33] 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.
- [34] MARSHAK J. AND R. RADNER. 1972. *Economic Theory of Teams*. Yale University Press, New Haven and London.
- [35] MELUMAD, N. AND T. SHIBANO. 1991. Communication in Settings with No Transfers. *Rand Journal of Economics*, 22: 173-198.
- [36] MELUMAD, N., D. MOOKHERJEE AND S. REICHELSTEIN. 1992. A Theory of Responsibility Centers. *Journal of Accounting and Economics*, 15(4): 445-484.
- [37] MILGROM, P. 1988. Employment Contracts, Influence Activities, and Efficient Organization Design. *Journal of Political Economy*, 96(1): 42-60.
- [38] MILGROM, P. AND J. ROBERTS. 1992. *Economics, Organization, and Management*. Prentice Hall, Englewood Cliffs, New Jersey.
- [39] MONTGOMERY, C.A. AND D. MAGNANI. 2001. PepsiCo's Restaurants. Harvard Business School Case 9-794-078.
- [40] MOOKHERJEE, D. 2006. Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective. *Journal of Economic Literature*, XLIV, 367-390.

- [41] OATES, W. 1999. An Essay on Fiscal Federalism. *Journal of Economic Literature*, 37: 1120-1149.
- [42] OTTAVIANI, M. AND P. SORENSEN. 2001. Information Aggregation in Debate: Who Should Speak First. *Journal of Public Economics*, 81: 393-421.
- [43] OZBAS, O. 2005. Integration, Organizational Processes, and Allocation of Resources. *Journal of Financial Economics*, 75(1): 201-242.
- [44] PFEFFER, J. 2004. Human Resources at the AES Corp.: The Case of the Missing Department, Stanford Business School HR3.
- [45] RANTAKARI, H. 2006. Governing Adaptation. Mimeo, MIT.
- [46] ROBERTS, J. 2004. *The Modern Firm: Organizational Design for Performance and Growth*. Oxford: Oxford University Press.
- [47] 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.
- [48] SLOAN, A.P., JR. 1964. *My Years With General Motors*. New York: Doubleday.
- [49] STEIN, J. 1989. Cheap Talk and the Fed: A Theory of Imprecise Policy Announcements, *American Economic Review*, 79(1): 32-42.
- [50] STEIN, J. 2002. Information Production and Capital Allocation: Decentralized versus Hierarchical Firms, *Journal of Finance*, 57: 1891-1921.