Organizational Design and Environmental Volatility

Heikki Rantakari*
University of Southern California
Marshall School of Business

This Version: March 18, 2010

Abstract

I investigate the link between the volatility of a firm's operating environment and its preferred organizational structure, as determined by the allocation of decision rights, the compensation structure of the managers and the degree of operational integration. The results are broadly consistent with the common wisdom regarding the fit between organizational design and the environment. Volatile environments are populated by organizations that have loosely coupled operations to maintain local responsiveness, together with decentralized decision-making and pay that is tied primarily to divisional performance. Stable environments, on the other hand, are populated by organizations that have tightly integrated operations to realize cost savings, together with (typically) centralized decision-making and pay that is closely tied to firm-level performance. The results also refine these relationships. For example, while decentralization and volatility are positively associated in equilibrium, volatility is not directly causing decentralization. Instead, decentralization arises only as a part of the overall adjustment in organizational design and simply decentralizing decision-making as a response to an increase in volatility would actually worsen organizational performance.

JEL Classifications: D2, D8, L2

*Contact: rantakar@marshall.usc.edu. This paper is a heavily revised version of Chapter 2 of my Ph.D. Dissertation. I would like to thank my advisors Robert Gibbons and Bengt Holmström for many helpful discussions and continuous encouragement. The paper has also benefited from comments by Tony Marino, Ricardo Alonso, the participants of ESSET2008, ESAM08, 2007 SITE Summer Workshop, and the seminar participants at various universities. All remaining errors are my own.

"Achieving high performance in a business results from establishing and maintaining a fit among three elements: the strategy of the firm, its organizational design, and the environment in which it operates." (Roberts, 2004:12)

1 Introduction

The need for a fit among a firm's strategy, structure and its operating environment has been extensively discussed by management and strategy scholars at least since Chandler's *Strategy and Structure* (1962). Building on this literature, this paper analyzes from an agency-theoretic perspective the effects of a firm's operating environment on its optimal choice of organizational structure, as captured by the allocation of decisions rights (governance structure), the compensation structure of its managers and its level of operational integration. The framework thus makes the important distinction between *administrative* integration, as captured by the choice of governance structure (centralization vs. decentralization) and *operational* integration (such as the use of shared distribution and marketing channels among the operating divisions), which reflects the level of interdependencies across the operating units, and models them as two distinct choice variables for the firm.

The organization I analyze consists of two operating divisions headed by self-interested division managers and a profit-maximizing headquarters, while the environment is characterized by its volatility and the cost of information. The organizational challenge is one of coordinated adaptation, where the organization needs to respond to changing market conditions while retaining coordination between its activities. The division managers first acquire information about their local conditions, then communicate that information strategically to the decision-maker(s) and finally the decision-maker(s) choose how the divisions will respond, using the information available to them.

This setup is intended to capture the fundamental strategy tradeoff between (local) responsiveness and (global) efficiency.[1] In short, a firm can generate value both through customizing its products and their marketing to meet varying and changing local tastes (increasing customer value) and through large-scale manufacturing and standardization (reducing production costs). An important challenge faced by firms is finding the right balance between the two. A key ingredient of the model is that this tradeoff is not only impacted through the governance and compensation structures employed by the firm, but also through

[1] This particular terminology is most common in the literature on multinational corporations. See, for example, Bartlett (1986), Prahalad and Doz (1987) and Bartlett and Ghoshal (1989). However, the basic tradeoff is also present in the general strategy literature, such as the choice between differentiation and cost leadership (Porter, 1980).

its level of operational integration. Choosing an organizational structure that has a low level of operational integration minimizes the extent of interdependencies across the activities and thus allows for maximal local responsiveness. Increasing the level of operational integration (such as the use of shared components or coordinated product development efforts) can bring efficiency benefits in terms of cost savings and synergies, but such integration increases the level of interdependencies across the firm's activities and thus decreases local responsiveness.

The organizational design problem is then to choose the governance and compensation structures, together with the level of operational integration, to maximize the expected organizational performance given the operating environment of the firm. The level of operational integration, by specifying the balance between efficiency and responsiveness, determines how much coordination is needed between the divisional responses and, as a result, also influences both the nature and severity of the agency conflicts inside the organization. Therefore, each level of operational integration is best managed through a particular combination of governance and compensation structures. And because the governance and compensation structures influence the value actually realized at any given level of operational integration, all three need to be determined together as the optimal response to a given environment.

The results are broadly consistent with the common wisdom regarding the fit between strategy, structure and the environment. Firms that operate in volatile environments pursue a strategy of responsiveness through the combination of loosely integrated operations, decentralized decision-making and a compensation structure that reflects primarily individual divisional performance, while firms that operate in stable environments pursue a strategy of efficiency through the combination of tightly integrated operations, (typically) centralized decision-making and a compensation structure that reflects firm-wide performance.[2] This pattern arises because of three results. First, the equilibrium level of operational integration is decreasing in the volatility of the environment. Intuitively, an increase in volatility increases the value of local responsiveness, which the organization achieves by reducing its level of operational integration. Second, the use of firm-wide incentives is generally increasing in the level of operational integration under both governance structures. This result arises because increasing the level of operational integration reduces the value of information (the acquisition of which is best motivated through division-level incentives) while increasing the agency conflicts in the transmission and use of information (which are best managed through firm-wide incentives). Third, loosely integrated operations are always best managed through a decentralized governance structure, while centralized decision-making can be preferred only

[2] For classic contributions, see, for example, Lawrence and Lorsch (1967), Galbraith (1973,1977) and Minzberg (1979). In the context of MNCs, see, for example, Prahalad and Doz (1987) and Bartlett and Ghoshal (1989). For recent contributions, see, for example, Brickley et al (2003) and Roberts (2004).

when the level of operational integration is sufficiently high.

The results also refine these broad pattens, especially regarding the determinants of decentralization. First, it is commonly argued that a decentralized structure is preferred in more volatile environments because it is able to respond faster and make better use of local information than a centralized structure. Indeed, in the model decentralization and volatility are positively associated in equilibrium. However, volatility is not directly causing decentralization. Instead, volatility reduces the optimal level of operational integration, and it is this reduction in the level of operational integration that raises the relative gains from decentralization. Thus, decentralization arises only as a part of the overall adjustment in organizational design and in fact, exclusively decentralizing decision-making as a response to an increase in volatility would actually decrease organizational performance.[3]

Second, for a given level of operational integration, centralization and the use of firm-wide incentives are substitutes. As a result, decentralization can be the preferred governance structure for *any* level of operational integration, as long as the equilibrium use of firm-wide incentives is sufficiently high. This result follows from the observation that while an increase in the use of firm-wide incentives improves the accuracy of communication and the quality of decision-making under both governance structures, decentralization benefits from these improvements at a relatively faster rate. Therefore, when considering the choice of governance structure, it is not sufficient to only consider the level of operational integration desired, but we also need to consider the optimal compensation structure.

Third, because decentralization can be preferred for any levels of operational integration, there is no simple relationship between the governance and compensation structures. In particular, while centralization is always associated with relatively high levels of operational integration, and thus with relatively heavy use of firm-wide incentives, decentralization can be associated either with low levels of operational integration and limited use of firm-wide incentives, or with high levels of operational integration and even heavier use of firm-wide incentives than a corresponding centralized structure. Overall, the results then suggest that the distinction between administrative and operational integration is a potentially important one, and highlight that when choosing its organizational structure, the firm needs to not only consider the fit with the environment, but also the fit among the design parameters themselves.

[3] The presence of these interdependencies also provides one explanation why many corporate restructurings have yielded disappointing results. For example, the attempt of Brown-Boveri to rationalize its production in the 1970s got thwarted by the failure to simultaneously adjust managerial compensation and authority structures: "Division managers in Germany and France were still measured and evaluated on their own short-term results and had little incentive to help their Italian colleagues. To the contrary, the difficulties in Italy gave the German managers an ability to turn around their business on their own and ignore the joint integration plan." (Prahalad and Doz, 1987/1999:206).

The remainder of the paper is structured as follows. Section 2 reviews the related literature and section 3 outlines the model. Section 4 derives the expected performance of the organization as a function of the environment and the design parameters. Section 5 analyzes the optimal choice of the design parameters and the link between the organization and the environment. Section 6 concludes.

2 Related Literature

The model builds directly on the framework developed in Alonso, Dessein and Matouschek (2008) and Rantakari (2008). However, instead of focusing on the role that the allocation of decision rights plays in managing communication and decision-making in organizations, I focus on the interactions among different organizational design parameters and their joint fit with the environment. Because of this integrative framework, the paper is most closely related to Dessein, Garicano and Gertner (2009), Friebel and Raith (2009) and Athey and Roberts (2001), each of which looks at the simultaneous determination of incentives and decision-making authority from alternative angles. Friebel and Raith (2009) analyze a resource allocation problem, where division managers need to be motivated to exert effort to generate high-quality projects and then to communicate that information (truthfully) to the headquarters. Dessein, Garicano and Gertner (2009) analyze a synergy implementation problem, where the managers need to be motivated to exert productive effort but also need to be induced to make appropriate synergy implementation decisions. Athey and Roberts (2001) combine the problem of inducing productive effort with a project selection problem. In all papers, the basic tradeoff is between providing focused incentives to induce effort and balanced incentives to induce truthful communication and/or appropriate decision-making, with the allocation of decision rights impacting this tradeoff.

The incentive provision problem in my setting also faces the same basic tension: focused incentives to motivate information acquisition and balanced incentives to motivate accurate transmission and use of that information. There are, however, a number of differences. First, the particular tension between motivating information acquisition and then motivating the appropriate use of that information is absent in the other papers, and the analysis highlights some features unique to this tradeoff. Second, while the problem of coordinated decision-making, which is the focus here, is somewhat analogous to the resource allocation and synergy implementation problems, the problem is also qualitatively different. These differences lead to some differences qualitative predictions, contrasted in section 5. Third, I consider the underlying configuration of organizational resources as an additional choice variable, while previous work takes it as exogenously given (or as a binomial choice).

Given the focus on fit, not only between the organization and the environment, but also the organizational design parameters themselves, the analysis is also related to the broader literature on complementarities and fit, such as Milgrom and Roberts (1990,1995) and Holmstrom and Milgrom (1991,1994). The paper also contributes to the growing literature on the role of authority and delegation in managing agency problems. Building on the cheap talk literature that has followed Crawford and Sobel (1982), Dessein (2002), Harris and Raviv (2005) and Alonso (2007), for example, examine how the allocation of decision rights can be used to manage the tradeoff between biased decisions and information losses due to strategic communication.[4] Aghion and Tirole (1997) illustrate how delegation can be used as a motivational tool for information acquisition by allowing the agent to freely use the information he learns. My framework embeds both aspects of the problem and joins them with the possibility of using monetary incentives, which allows us to examine the link between delegation and incentives.[5]

Organizational structures have also been analyzed from various other angles. The paper closest to mine is Dessein and Santos (2006), who analyze, in a team-theoretic model, the limitations that the need for coordinated adaptation imposes on task specialization. Coordination in their model is, however, constrained only because information transmission is exogenously imperfect. Some further perspectives include information processing (for example, Marshak and Radner, 1972, Bolton and Dewatripont, 1994), problem-solving (for example, Garicano, 2000), screening for interdependencies (Harris and Raviv, 2002) and coordination and experimentation (Qian, Roland and Xu, 2006). How the organizational design impacts the competitive position of a firm is investigated in Alonso, Dessein and Matouschek (2009).

Finally, while the economic literature on organizational design is still relatively young, there is a long history of management and strategy scholars that have analyzed the topic of this paper. As a result, this paper owes an intellectual debt to a long string of contributions, including Simon (1947), Chandler (1962,1977), Lawrence and Lorsch (1967), Galbraith (1973,1977), Mintzberg (1979) and Porter (1980), among many others, in particular the later works of Prahalad and Doz (1987), Bartlett and Ghoshal (1989), Brickley et al (2003) and Roberts (2004).

[4] See also Stein (1989) and Melumad and Shibano (1991)

[5] Further, a consequence of the need for coordination is that delegation can actually *decrease* the incentives to acquire information, in contrast to Aghion and Tirole (1997).

3 The Model

The organization I consider consists of two divisions, each managed by a strategic division manager (he), and headquarters (she), who aims to maximize the overall profitability of the organization. This section outlines the payoffs, available actions and the timing of events.

Divisional profits and alternative governance structures: The organization consists of two (symmetric) divisions, i and j. The profitability of each division depends on both how well the activities of the division are aligned with its local conditions and how well the divisions are coordinated with each other. Given the decisions d_i and d_j regarding the operations of divisions i and j, respectively, the realized profits of division i are given by

$$\pi_i(\theta_i, d_i, d_j) = K(\beta) - \beta(d_j - d_i)^2 - \alpha(\theta_i - d_i)^2,$$

where $\theta_i \sim U\left[-\overline{\theta}, \overline{\theta}\right]$ indexes the locally optimal decision for division i, with θ_i and θ_j independently distributed and $\overline{\theta} > 0$ measuring the volatility of the environment. Let $\sigma_\theta^2 = \frac{\overline{\theta}^2}{3}$ denote this ex ante volatility. The alignment of the division with its local conditions is measured by $\alpha(\theta_i - d_i)^2$, while the alignment with the other division is measured by $\beta(d_j - d_i)^2$. The realized profits of the organization are given by $\pi_i + \pi_j$.

The first two design variables for the organization are the allocation of decision rights and the level of operational integration $\beta \in [0, \infty)$. The allocation of decision rights reflects the extent of *administrative* integration in the organization, and is modeled as the choice between *centralization*, under which the headquarters retains control of both divisions, and *decentralization*, under which control over the divisions is delegated to their respective division managers. I will use superscript $g \in \{cent, dec\}$ to denote the two governance structures.

The level of *operational* integration, on the other hand, measures the level of interdependencies across the organization's activities, as determined by the extent to which the organization uses shared manufacturing facilities, distribution networks, coordinated product development efforts and the like. As separate from administrative integration, it thus reflects the operational interdependencies inside the organization. The benefits of such integration are captured by an increasing and continuous function $K(\beta)$, where the benefits are generally viewed as coming from the potential reduction in operating costs that result from the elimination of duplicated assets, the increased scale of the remaining operations and the realization of potential synergies across the activities. The cost of such integration is that as the operations become more tightly coupled, the behavior of the divisions needs to be increasingly coordinated to realize all these benefits. This induced value of coordination is captured by $\beta (d_j - d_i)^2$.

To summarize, the model makes the distinction between administrative integration, as captured by the governance structure of the organization, and operational integration, which reflects the underlying configuration of its productive assets. Further, the types of integration are two separate design parameters for the organization. For example, an organization can choose a tightly integrated operating technology but still choose to manage it in a decentralized fashion.[6]

Division managers: Each division is headed by a risk-neutral division manager (managers i and j, respectively), who are the key strategic actors in the model. Their task consists of information acquisition, communication and, in the case of decentralization, decision-making. Their behavior is controlled through the third choice variable, which is their compensation structure. I assume that manager i is offered a linear incentive contract

$$T_i (\pi_i, \pi_j) = A_i + (1 - s) \pi_i (\theta_i, d_i, d_j) + s\pi_j (\theta_j, d_i, d_j),$$

where the amount of *profit-sharing* $s \in \left[0, \frac{1}{2}\right]$ measures the extent of firm-wide incentives in managerial compensation. If $s = 0$, then the division managers are compensated only based on their divisional performance, while if $s \to 1/2$, then their compensation comes to depend only on firm-wide performance.[7]

Timing of events: The organizational design parameters (g, s, β) are used to manage the unfolding of events summarized in figure 1. First, the division managers invest in acquiring information about their local market conditions. In particular, manager i acquires at personal cost $C(\sigma_\theta^2, q_i)$ a signal t_i that matches the realized state θ_i with probability q_i and is a random draw from the state distribution otherwise. The manager does not learn whether the signal is correct or not, so that his posterior for the local state will be given by

[6] As observed by Porter (1990:17):"A firm faces an array of options in both configuration and coordination for each activity. Configuration options range from concentrated (performing an activity in one location and serving the world from it - e.g., one R&D lab, one large plant) to dispersed (performing each activity in each country). In the latter case, each country would have a complete value chain. Coordination options range from none to very high. For example, if a firm produces its product in three plants, it could, at one extreme, allow each plant to operate with full autonomy – e.g., different product standards and features, different steps in the production process, different raw materials, different part numbers. At the other extreme, the plants could be tightly coordinated by employing the same information systems, the same production processes, the same parts, and so forth."

[7] This profit-sharing rule arises as the equilibrium among the class of linear contracts $T_i (\pi_i, \pi_j) = A_i + s_{ii}\pi_i (\theta_i, d_i, d_j) + s_{ij}\pi_j (\theta_j, d_i, d_j)$, subject to the constraint that $s_{ii} + s_{ji} \leq 1$, so that there is no budget breaker available and the headquarters simply sells the firm to the agents and designs the shares to maximize the net surplus generated subject to the share constraint (teams problem).

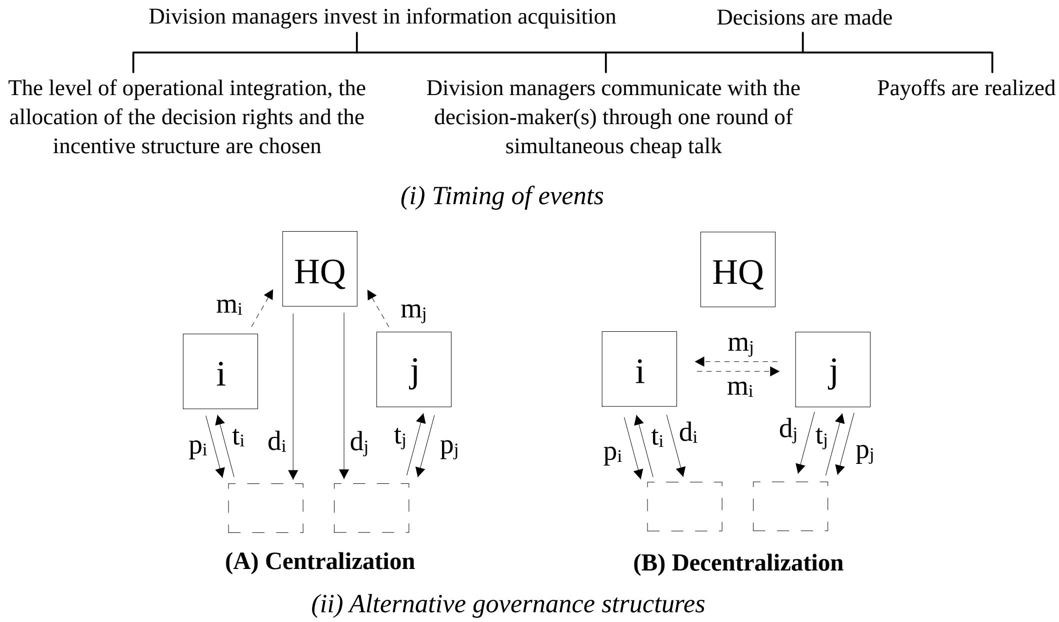


Figure 1: The timing of events and the alternative governance structures.

$E_i(\theta_i|t_i) = q_i t_i$.

Forecasting the results, the value of accuracy in terms of expected profits will be quadratic in q_i. As a result, I define $p_i = q_i^2$ and will refer to p_i as *the quality of primary information.* The cost of information acquisition is assumed to be given by

$$C(\sigma_\theta^2, p_i) = -\mu\sigma_\theta^2 (p_i + \ln(1 - p_i)),$$

where μ parameterizes the (marginal) cost of information. Making the cost of information to be proportional to σ_θ^2 turns out to be a convenient normalization while the particular functional form simplifies the analysis without impacting the qualitative nature of the results. As a final simplification, I assume that the choice of p_i by the managers is observable but not verifiable to the organizational participants.

Having acquired their private information, the division managers strategically communicate their information to the decision-maker(s) through one round of simultaneous cheap talk. In the case of centralization, communication occurs vertically to the headquarters, while in the case of decentralization, communication occurs horizontally between the division managers. Finally, after communication, the decision-maker(s) choose their decisions conditional on the information available to them. In the case of centralization, the headquarters makes decisions to maximize the overall profits of the organization, while in the case of decentralization, the division managers make decisions to maximize their individual payoffs, as determined by their compensation structure.

The organizational design problem: In the beginning of the game, the headquarters chooses the organizational design parameters to maximize the expected net surplus, as determined by the Perfect Bayesian Nash equilibrium of the game described above. That is, she chooses the level of profit-sharing s, the governance structure g and the level of operational integration β to maximize

$$\max_{g,s,\beta} \sum_{k=i,j} \left(E\left(\pi_k^g\left(\beta, s, \lambda\right)\right) - C\left(\sigma_\theta^2, p_k^g\left(\beta, s\right)\right)\right),$$

subject to the equilibrium level of information acquisition, equilibrium accuracy of communication and equilibrium decisions.

Basic assumptions: As described, the model makes a number of simplifying assumptions to keep the analysis tractable. In particular, I do not allow for message-contingent decision rules and transfers (or general non-linear compensation contracts). Commitment to a message-contingent decision rule would make the allocation of authority irrelevant, while maintaining ex post incentive compatibility of decisions but allowing for non-linear compensation schemes would improve the performance of both governance structures, with the relative improvement ex ante ambiguous. However, within the class of linear contracts and the assumed functional form for the profit function, the qualitative results are robust to various alternatives, such as general cost functions. A discussion of these extensions can be found in Appendix C.

4 Expected Profitability and Organizational Design

Before analyzing the link between organizational design and the environment, we need to first understand how the expected profits of the organization depend on the design parameters (g, β, s). Intuitively, the organizational performance will depend on (i) how accurate information is acquired by the division managers and (ii) the use that the organization makes of the information generated. These two aspects are clearly interdependent. In particular, the incentives to acquire information will depend on its value as perceived by the division managers, and thus on how well the organization uses that information. As a result, section 4.1 will first analyze the use of information, as determined by the equilibrium decisions and the accuracy of communication under a given organizational structure (g, β, s), while section

4.2 analyzes how the incentives to acquire information are influenced by the organizational structure. Section 5 completes the analysis by examining the link between the preferred organizational design and the environment.

4.1 Equilibrium use (and value) of information

The solution to the equilibrium use of information follows through backward induction. First, we analyze how the organizational structure influences the equilibrium decisions (section 4.1.1), after which we can determine how the organizational structure influences the accuracy of communication (section 4.1.2). Having the equilibrium decisions and the accuracy of communication, we can then determine the overall value of information to the organization (section 4.1.3). These steps parallel the analysis in Alonso et al. (2008) and Rantakari (2008), with the introduction of imperfect local information.

4.1.1 Equilibrium decisions

In the decision-making stage, the decision-maker(s) use the information available to them to maximize their individual payoffs. Let t_i and t_j denote the signals acquired by the division managers and m_i and m_j denote the messages sent to the decision-maker(s) in the communication stage. Then, in the case of centralization, the decisions are made by the headquarters and she solves $\max_{d_i,d_j} E\left(\pi_i + \pi_j | m_i, m_j\right)$, while in the case of decentralization, the decisions are made by the division managers, where manager i solves $\max_{d_i} E\left((1-s)\pi_i + s\pi_j | t_i, m_i, m_j\right)$. The equilibrium decisions are summarized in the following proposition:

Proposition 1 ***Equilibrium decisions:***

(i) $d_i^{cent} = \frac{(\alpha+2\beta)E_{HQ}(\theta_i|m_i)+2\beta E_{HQ}(\theta_j|m_j)}{\alpha+4\beta}$

(ii) $d_i^{dec} = \frac{(1-s)\alpha}{(1-s)\alpha+\beta} E_i\left(\theta_i|s_i\right) + \frac{\beta}{(1-s)\alpha+2\beta} E_i\left(E_j\left(\theta_j|s_j\right)|m_j\right) + \frac{\beta^2}{((1-s)\alpha+\beta)((1-s)\alpha+2\beta)} E_j\left(E_i\left(\theta_i|s_i\right)|m_i\right)$

(iii) Conditional on the quality of information, the bias in the equilibrium decisions under decentralization is decreasing in the amount of profit-sharing s and it is non-monotone in the level of operational integration β. For any s, the bias is maximized at an intermediate level of operational integration $\beta = \frac{\alpha}{4}\sqrt{2(1-s)}$.

Proof. See Appendix A.1 ∎

As the level of operational integration β is increased, the decisions under both governance structures naturally become less responsive to information about the local conditions of that division and more responsive to information about the local conditions of the other division, reflecting the increasing need for coordination between the divisions. Under centralization, the decisions are (by assumption) profit-maximizing conditional on the information available to the headquarters. Under decentralization, on the other hand, the fact that the division managers place relatively more weight on the performance of their own division whenever $(1-s) > s$ leads to biased decision-making. This bias is naturally decreasing in the level of profit-sharing s. More importantly, and what will play an important role later in the analysis, this bias is non-monotone in the importance of coordination. When $\beta = 0$, the two divisions are fully independent and maximizing joint profits is equivalent to maximizing the individual performance of each division. Thus, the decisions of the division managers will be profit-maximizing even when $s = 0$. Similarly, when $\beta \to \infty$, so that divisional performance becomes extremely sensitive to the amount of coordination between the divisions, the division managers are willing to coordinate even if they care only about the performance of their own divisions. It is only for intermediate β, or when the tension between coordination and local adaptation is the largest, when the decisions under decentralization are particularly biased. Finally, it is worth noting that I have only referred to the size of the bias, not its actual payoff consequences. The payoff consequences are analyzed in section 4.1.3, after having analyzed the equilibrium communication and derived the equilibrium expected profits.

4.1.2 Equilibrium communication

The communication stage is modeled as a cheap-talk game between the privately informed division managers and the decision-maker(s). Knowing how the equilibrium decisions depend on the beliefs of the decision-maker(s), the division managers send simultaneously non-verifiable messages m_i and m_j regarding the signals s_i and s_j they received in the information acquisition stage. A preference conflict between the sender and the receiver over the preferred response to the information held by the sender will lead the sender to try to strategically manipulate his messages to induce more favorable decisions. Of course, in equilibrium, such attempts to mislead the decision-maker(s) are futile and only lead to loss of information. Following Crawford and Sobel (1982) and the subseqent literature, I focus on the most informative partition equilibrium of the cheap talk game, where the division managers reveal only that the signals they received belong to a particular interval of the state space.

The organizational structure (g, s, β) influences both how sensitive the organizational performance is to strategic communication and how accurate that communication is. With respect to the accuracy of communication, the quadratic structure of the payoffs allows us to capture the loss of information (accuracy of communication) through the expected variance of the recipients beliefs around the true information held by the sender, $E\left(t_i - E\left(t_i|m_i\left(t_i\right)\right)\right)^2$. In the present setting, we can further write this as $\left(1 - V_i^g\left(s, \beta\right)\right)\sigma_\theta^2$, where σ_θ^2 is the ex ante variance of the local conditions and $V_i^g\left(s, \beta\right) \in [\frac{3}{4}, 1]$ measures the accuracy of communication, as determined by the equilibrium to the cheap talk game. The solution and its properties are summarized in the following proposition:

Proposition 2 ***Equilibrium communication:***

(i) $V_i^{cent}(s, \beta) = \frac{3(1-s)(2\beta+\alpha)}{\beta(1-2s)+3(1-s)(2\beta+\alpha)}$ *and* $V_i^{dec}\left(s, \beta\right) = \frac{3(1-s)((1-s)\alpha+2\beta)}{((1-s)\alpha+\beta)(1-2s)+3(1-s)((1-s)\alpha+2\beta)}$

(ii) $\frac{\partial V_i^g(s,\alpha,\beta)}{\partial s} \geq 0$, $\frac{\partial V_i^{cent}(s,\alpha,\beta)}{\partial \beta} \leq 0$ *and* $\frac{\partial V_i^{dec}(s,\alpha,\beta)}{\partial \beta} \geq 0$.

(iii) $V_i^{cent}(s, \beta) \geq V_i^{dec}\left(s, \beta\right)\ \forall s, \beta$.

Proof. See Appendix A.2 ■

As expected, the accuracy of communication is increasing in the level of profit-sharing under both governance structures, as profit-sharing aligns the division managers' interests both with each other and with the goal of overall profit-maximization. As $s \to \frac{1}{2}$, $V_i^g\left(s, \beta\right) \to 1$ and communication becomes perfect under both governance structures. The impact of operational integration, on the other hand, depends on the governance structure. Increasing the level of operational integration reduces the accuracy of communication under centralization while it increases the accuracy of communication under decentralization. This difference results from the differential purpose of communication under the two governance structures. Under centralization, the division managers communicate to the headquarters to induce local adaptation. When the level of operational integration is increased, the headquarters becomes increasingly insensitive to those local needs and the vertical conflict between the division managers and the headquarters increases, reducing the accuracy of communication. Under decentralization, the division managers communicate with each other to achieve better coordination. When the level of operational integration is increased, the division managers become increasingly responsive to each others' needs, reducing the level of horizontal conflict and improving the accuracy of communication.

Note, however, that while the accuracy of communication is influenced by the organizational structure, so is the *value* of that communication. Under centralization the value of communication is decreasing in β, as the headquarters becomes increasingly non-responsive to the information acquired. Under decentralization, on the other hand, the value of communication is increasing in β, reflecting the increasing value of coordination between the division managers. Because in both cases the value and accuracy of communication move in the same direction, we cannot yet say anything about the profit consequences of the equilibrium inaccuracies in the information transmission. Further, while communication is more accurate under centralization, it is intuitively also more valuable because of the separation of decision-making authority from direct access to local information. Thus, to make more precise comparisons in terms of the payoff consequences of the communication equilibrium, we need to solve how the expected profits depend on the accuracy of communication.

4.1.3 Expected profits and the value of information

Substituting the equilibrium decisions and the quality of communication to the profit functions and taking expectations allows us to write the expected profitability of the organization as

$$E\left(\sum_{k=i,j}\pi_k^g(s,\beta,\lambda)\right)=\sum_{k=i,j}\left(K\left(\beta\right)-\alpha\sigma_\theta^2+p_k^g(\beta,s,\mu)\Psi_k^g\left(s,\beta\right)\sigma_\theta^2\right),$$

where $p_k^g(\beta,s)$ is the accuracy of the signals acquired by the division managers (*quality of primary information*), and $\Psi_k^g\left(s,\beta\right)$ measures how well the organization uses that information (*value of information*), as a function of its organizational structure (g,s,β).

To better understand the determinants behind the value of information $\Psi_k^g\left(s,\beta\right)$, we can view it to be composed of four parts. First, if the two divisions were completely independent, then the value of information would simply be α, the value of adaptation. When $\beta>0$, then the value of information is reduced below α for three reasons. First, the need for coordination makes even the profit-maximizing decisions less adaptive to local conditions. Second, biased decision-making further reduces the ex ante value of information. Third, some of the information is lost due to strategic communication. Indeed, the quadratic structure of the payoffs makes the four components separable, so that the value of information takes the form

$$\Psi_k^g\left(s,\beta\right)=\underbrace{\alpha-\Lambda_k^{FB}\left(\beta\right)}_{\text{value absent agency conflicts}}-\underbrace{\Delta\Lambda_k^g(s,\beta)}_{\text{loss due to biased decisions}}-\underbrace{\Gamma_k^g(s,\beta)\left(1-V_k^g(s,\beta)\right)}_{\text{loss due to strategic communication}}.$$

The basic comparative statics are summarized in the following proposition:

Proposition 3 ***Properties of the value of information,*** $\Psi_i^g(s,\beta)$***:***

(i) The value of information is decreasing in the level of operational integration and increasing in the amount of profit-sharing under both governance structures: $\frac{\partial \Psi_i^g(s,\beta)}{\partial \beta} \leq 0$ *and* $\frac{\partial \Psi_i^g(s,\beta)}{\partial s} \geq 0$*.*

(ii) The value of information is higher under centralization whenever β *is sufficiently high and* s *is sufficiently low:* $\Psi_i^{cent}(s,\beta) \geq \Psi_i^{dec}(s,\beta)$ *if and only if* $\beta \geq \widehat{\beta}(s)$*, with* $\frac{\partial \widehat{\beta}(s)}{\partial s} > 0$*.*

Proof. See appendix A.3 ■

The basic impact of β and s on the value of information is thus similar under the two governance structures. For any given β, increasing the amount of profit-sharing s reduces the losses due to strategic communication and biased decision-making, thus increasing the value of information. Similarly, for any given level of s, increasing the level of operational integration β reduces the sensitivity of the equilibrium decisions to the information generated and thus reduces the value of information.[8]

While the two governance structures share the same qualitative features on how $\Psi_i^g(s,\beta)$ depends on β and s, they differ systematically in the quantitative impact that the rest of the organizational structure has on the value of information. To build intuition for this result, note that any differences in the value of information arise from differences in the loss due to biased decisions, $\Delta\Lambda_k^g(s,\beta)$, and in the loss due to strategic communication, $\Gamma_k^g(s,\beta)\left(1 - V_k^g(s,\beta)\right)$. As discussed in section 4.1.1, the bias in the equilibrium decisions under decentralization is largest for intermediate levels of operational integration and decreasing in the level of profit-sharing. The payoff consequences of this bias are illustrated in panel (i) of figure 2. Paralleling the size of the bias, the losses are largest for intermediate β and converging to zero whenever $\beta \to 0, \beta \to \infty$ or $s \to 1/2$. As discussed in section 4.1.2, while the accuracy of communication is always higher under centralization, so is the value of accurate communication. The net effect is illustrated in panel (ii), which shows that the loss due to strategic communication is generally higher under centralization. This result reflects the observation that while communication is more accurate under centralization, that

[8] Under centralization, this effect is reinforced by the fact that the loss due to strategic communication is also monotonically increasing in β, thus resolving the ambiguity left in section 4.1.2. Under decentralization, the loss due to strategic communication is also increasing in β almost everywhere. While the loss due to biased decisions is non-monotone, this non-monotonicity is outweighed by the other forces.

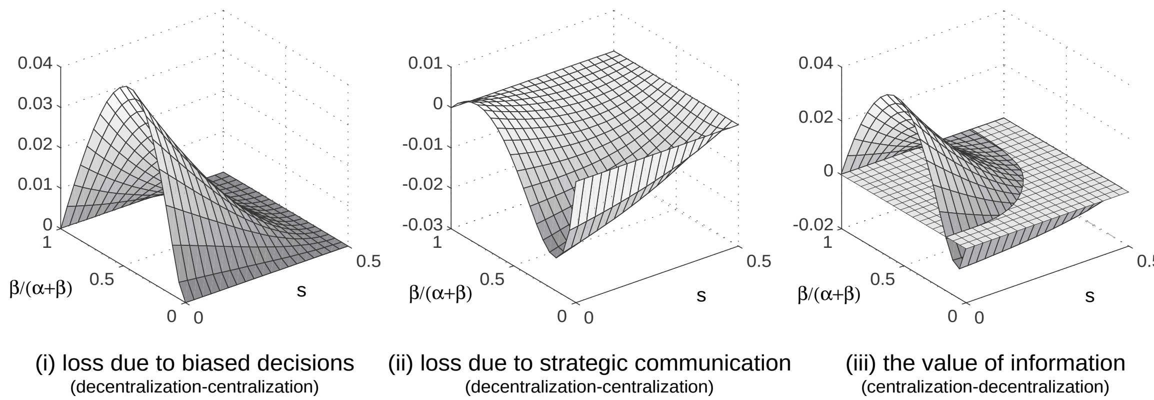


Figure 2: The relationship between the governance structure and $\Psi_i^g(s, \beta)$

accuracy is even more valuable, so that the expected loss due to strategic communication is actually smaller under decentralization. Centralization performs particularly poorly for low β, in which case the information held by the division managers is particularly valuable to headquarters, while a significant portion of it is lost when communicated to the headquarters.

The difference in the value of information, which is then the net effect of the losses in panels (i) and (ii) is given in panel (iii). For sufficiently low β, the bias in the equilibrium decisions under decentralization is not that large and, as a result, the benefit of removing this bias through centralized decision-making is never big enough to compensate the large cost in terms of strategic communication. Relatedly, while increasing the amount of profit-sharing naturally improves both the quality of decision-making under decentralization and the accuracy of communication under both governance structures, it improves the quality of decision-making at a relatively faster rate. As a result, when we increase the level of profit-sharing, the remaining agency losses are increasingly due to strategic communication, making decentralization increasingly attractive. It is thus only for sufficiently high β and sufficiently low s that centralization makes better use of the available information, with the advantage of centralization being largest for intermediate β, or when decision-making under decentralization would be at its worst.

4.2 Managerial information acquisition

Having analyzed the value of information, we can now complete the picture on how organizational performance depends on the organizational structure by analyzing the information acquisition choice of the managers. For this result, it is important to note that the information acquired will have an asymmetric impact on the profitability of the two divisions, benefiting the division whose manager is doing the acquisition while imposing a negative externality on the other division. Analogous to the decomposition in terms of the sources of losses, we can write the value of t_i as

$$\Psi_i^g(s,\beta) = \underbrace{\alpha - \phi_i^g(s,\beta)}_{\text{benefit to division } i} - \underbrace{\phi_j^g(s,\beta)}_{\text{negative externality on division } j}.$$

In other words, the value of t_i to division i is $\alpha - \phi_i^g(s,\beta)$, where $\phi_i^g(s,\beta)$ is the total reduction in the value of information for division i (as determined by (g,s,β)), while $\phi_j^g(s,\beta)$ is the negative externality that the information imposes on the other division. The reason for the presence of this negative externality is the fact that whenever $\beta > 0$, an improvement in the precision of t_i causes d_i to be more locally responsive, which in turn increases division j's coordination costs.

Now, because the profit-sharing level s induces the manager to place a weight $(1-s)$ on the performance of his division while placing a weight s on the performance of the other division, we can write the value of information to the division managers (*the perceived value of information*) as

$$\widetilde{\Psi}_i^g(s,\beta) = (1-s)(\alpha - \phi_i^g(s,\beta)) - s\phi_j^g(s,\beta) = (1-s)\Psi_i^g(s,\beta) + (1-2s)\phi_j^g(s,\beta).$$

The value of information to the division managers is thus proportional to the true value of information, *plus* the non-internalized part of the negative externality on the other division. With this notation, manager i's optimization problem becomes simply

$$\max_{p_i} p_i \widetilde{\Psi}_i^g(s,\beta)\sigma_\theta^2 - C(\sigma_\theta^2, p_i),$$

which gives us the equilibrium quality of primary information as $p_i^g(s,\beta,\mu) = \frac{\widetilde{\Psi}_i^g(s,\beta)}{\widetilde{\Psi}_i^g(s,\beta)+\mu}$, the properties of which are summarized in the following proposition:

Proposition 4 ***Properties of the quality of primary information,*** $p_i^g(\beta,s,\mu)$*:*

(i) The quality of primary information is decreasing in the level of operational integration (β)*, the cost of information* (μ) *and the amount of profit-sharing* (s) *under both governance structures:* $\frac{\partial p_i^g(\beta,s,\mu)}{\partial\beta}$, $\frac{\partial p_i^g(\beta,s,\mu)}{\partial\mu}$ *and* $\frac{\partial p_i^g(\beta,s,\mu)}{\partial s} \leq 0$.

(ii) For given (s,β)*, the quality of primary information acquired under decentralization is higher unless* β *is sufficiently high and* s *is sufficiently low:* $p_i^{cent}(\beta,s,\mu) \geq p_i^{dec}(\beta,s,\mu)$ *if and only if* $\beta \geq \overline{\beta}(s)$*, with* $\frac{\partial\overline{\beta}(s)}{\partial s} > 0$

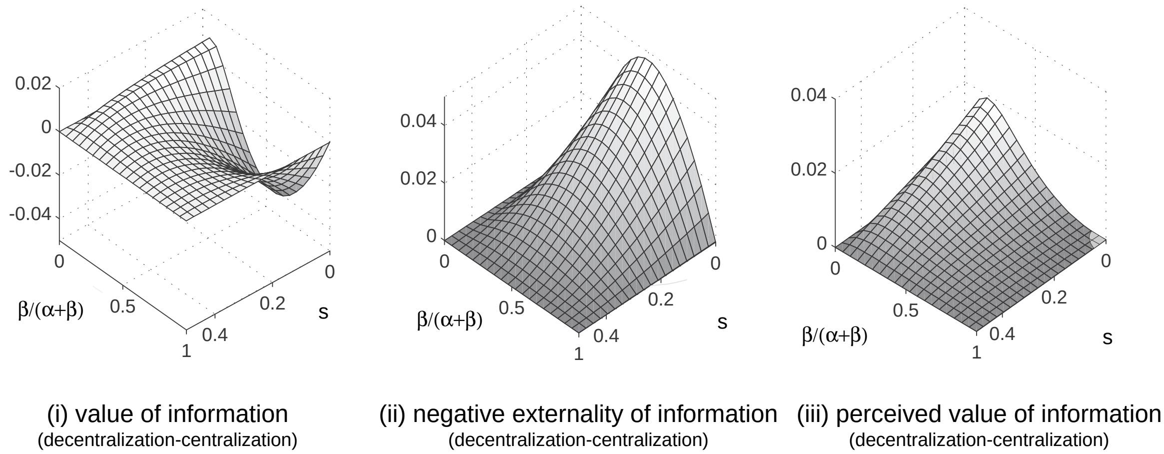


Figure 3: The relationship between the governance structure and $p_i^g(s, \beta, \mu)$

Proof. See appendix A.4 ■

As with the value of information, the accuracy of the signals acquired under the two governance structures share the qualitative comparative statics with respect to the rest of the variables. Not surprisingly, the quality of information acquired is decreasing in the cost of that information. Similarly, it is decreasing in the level of operational integration, as an increase in β reduces the value of information. More importantly, the quality of information is decreasing in the level of profit-sharing, despite the fact that the value of information is increasing. The reason for this result is that as we increase the level of profit-sharing, the more weight the manager places on the negative externality that the information acquired imposes on the other division and the less weight he places on the positive value realized by his division. As a result, while the value of information is increasing to the organization, the (perceived) value of that information to the division managers is decreasing. Finally, because I have assumed that the cost of information is proportional to σ_θ^2, p_i^g is independent of σ_θ^2.

While the two governance structures thus share the basic qualitative features of $p_i^g(\beta, s, \mu)$, they again differ in a systematic way in the quantitative impact that the the rest of the design parameters have on the quality of primary information. To analyze these differences, note that the difference in the quality of information acquired is directly related to the difference in the perceived value of information, and the difference in the perceived value of information is, in turn, determined by differences in the true value of information and in the size of the negative externality. As discussed in section 4.1.3, the true value of information is higher under decentralization whenever β is sufficiently low or s is sufficiently high. Second, the negative externality is always larger under decentralization. This result follows because $\phi_j^g(s,\beta)$ reflects the coordination costs that d_i imposes on the other division, and the equilibrium decisions are always more adaptive under decentralization. These differences are illustrated in figure 3.

The net effect of these two forces is that the *perceived* value of information (and thus the quality of information acquired by the division managers) is almost always higher under decentralization for a given (s,β), with the difference largest for low β and s. For intermediate β, the large negative externality is countered by the significantly lower true value of information, while the negative externality gets eliminated when $s \to 1/2$ or $\beta \to \infty$. In short, for a given (s,β), decentralization almost always generates more information but a centralized structure is able to make better use of that information for β sufficiently high and s sufficiently low. With these results, we can then analyze how (g,s,β) are determined in equilibrium.

5 Choice of Organizational Design

Given the analysis of section 4, we can write our optimization problem as

$$\max_{g,s,\beta} \sum_{k=i,j} \left(K(\beta) - \tfrac{\alpha\bar{\theta}^2}{3} + \left(p_k^g(\beta,s,\mu)\Psi_k^g(s,\beta) - \mu C(p_k^g(\beta,s,\mu))\right)\sigma_\theta^2 \right)$$

$$\text{s.t.} \qquad p_k^g(\beta,s,\mu) = \tfrac{\widetilde{\Psi}_i^g(s,\beta)}{\widetilde{\Psi}_i^g(s,\beta)+\mu}.$$

We can decompose the design problem into three steps. First, we can solve for the level of profit-sharing for a given governance structure. Second, we can consider the choice of governance structure itself. Since this part is independent of σ_θ^2, the compensation and governance structures are solely determined by (β,μ).[9] Knowing the equilibrium g and s, we can then determine the optimal level of operational integration and examine how the overall organizational design varies with the environment. The basic relationships are summarized in the following proposition:

[9] Any potential dependency on σ_θ^2 can be understood through changes in μ through $\frac{d\mu}{d\sigma_\theta^2}$

Proposition 5 ***Choice of organizational structure:***

(i) The equilibrium level of profit-sharing is unique given (g,β,μ) and it is decreasing in the cost of information and typically increasing in the level of operational integration under both governance structures: $\frac{\partial s^g}{\partial \mu} \leq 0$ *and* $\frac{\partial s^g}{\partial \beta} \gtrsim 0$.

(ii) Centralization is preferred over decentralization if and only if the level of operational integration (β) *and the cost of information* (μ) *are sufficiently high:* $\beta \geq \underline{\beta}(\mu)$, *with* $\frac{\partial \underline{\beta}(\mu)}{\partial \mu} \leq 0$. *Further, whenever centralization is preferred,* $s^{cent}(\beta, \mu) < s^{dec}(\beta, \mu)$.

(iii) The optimal level of operational integration is decreasing in the volatility of the environment and increasing in the cost of information: $\frac{d\beta}{d\sigma_\theta^2} \leq 0$ *and* $\frac{d\beta}{d\mu} \geq 0$. *Whenever there is a change in the governace structure, this change is discrete.*

Proof. See Appendix A.5 ■

To build some intuition for the results, consider first the choice of profit-sharing. The first-order condition is

$$\left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta)\right)\frac{\partial p_i^g}{\partial s} + p_i^g \frac{\partial \Psi_i^g(s,\beta)}{\partial s} = 0,$$

where the first part measures the marginal value of further information acquisition while the second part measures the improvement in the use of that information. Under both governance structures, the basic tradeoff is thus the same: increasing the level of profit-sharing reduces the quality of primary information acquired by the division managers $\left(\frac{\partial p_i^g}{\partial s} < 0\right)$ while improving the use of that information $\left(\frac{\partial \Psi_i^g(s,\beta)}{\partial s} > 0\right)$. Now, as μ increases, p_i^g decreases and thus the value of profit-sharing is reduced. Intuitively, since there is less information, it is less damaging for the organization not to use that information accurately. Instead, s is decreased to restore some of the incentives to acquire information. Similar logic applies with respect to β. As the level of operational integration is increased, both the agency losses and the negative externality $\phi_j^g(s,\beta)$ increase, while the value of information decreases, making profit-sharing both more valuable and less costly to the organization.[10]

The intuition for the choice of governance structure, in turn, follows from our discussion in section 4. The choice of s balances the losses due to suboptimal information acquisition

[10] These effects are countered by the fact that as β increases, information becomes less valuable and so p_i^g also decreases. Further, under decentralization, the agency losses can be non-monotone in β. However, for the equilibrium governance structure, the result holds (for a region under decentralization it is negative but for those parameters, centralization is never the preferred governance structure).

and suboptimal use of that information. The governance structure, in turn, helps to relax this tradeoff. In section 4 we saw that whenever β is sufficiently low or s is sufficiently large, then decentralization is able to both generate more information and make better use of that information. Given the monotonicity of s^g in μ, we can thus conclude that decentralization is a more efficient choice both when μ is sufficiently low and when β is sufficiently low. Centralization is preferred only when the value of information is sufficiently higher under centralization to outweigh the motivational advantage of decentralization in terms of information acquisition. Further, because the value of information must be higher under centralization whenever centralization is preferred, decentralization will correspondingly benefit relatively more from additional profit-sharing and thus $s^{cent}(\beta, \mu) < s^{dec}(\beta, \mu)$. In other words, the advantage of centralization is to make relatively good use of the information generated even at low levels of profit-sharing. This advantage is most valuable when μ is high because then the need to motivate information acquisition is the highest (and s^g will be the lowest).

Finally, the choice of operational integration balances the benefit of improved efficiency $K(\beta)$ with the cost of reduced equilibrium adaptiveness of the organization. While the choice of s had a unique solution, the choice of β can easily have multiple local maxima. However, we can establish that the profit function is submodular in β and σ_θ^2 while being supermodular in β and μ, even after accounting for the equilibrium choice of g and s. Intuitively, an increase in the volatility of the operating environment increases the costs of compromised adaptiveness and so the level of operational integration will be decreasing in σ_θ^2. Similarly, an increase in μ reduces p_i^g, which in turn reduces the cost of operational integration because there is less information that the organization can respond to.

We can thus summarize the main relationships as follows: the level of profit-sharing is increasing in the level of operational integration while decreasing in the cost of information under both governance structures. Centralization arises as the preferred governance structure only when both the level of operational integration and the cost of information are sufficiently large. Finally, the optimal level of operational integration is decreasing in the level of volatility and increasing in the cost of information. Having identified the key relationships, we can now move on to considering how the overall organizational design varies with the environment.

5.1 The relationship between organizational design and the environment

The general pattern of the relationship between organizational design and the environment that arises from proposition 5 matches closely the common wisdom in the management literature. Organizations operating in stable environments are tightly integrated, typically centrally governed and have a significant portion of the division managers' pay tied to firm-level performance. In contrast, firms operating in volatile environments are loosely integrated, with decentralized decision-making and pay tied mainly to divisional performance.

To illustrate the results, figure 4 shows the equilibrium solution for a particular parameterization of $K(\beta)$.[11] The first panel illustrates the comparative statics with respect to σ_θ^2. When the environment is stable, the the organization is tightly integrated, centralized and makes relatively heavy use of firm-wide incentives. As the level of volatility increases, the level of operational integration and the amount of profit-sharing decrease. Eventually, the preferred governance structure switches to decentralization, which, as shown in proposition 5, is associated with a discrete reduction in the level of operational integration and a discrete increase in the level of profit-sharing. If the baseline cost of information was lower, then the comparative statics would be similar but the level of operational integration would be lower for any σ_θ^2 and the switch to decentralization would occur earlier.

The second panel illustrates the comparative statics with respect to μ. Now, when information is cheap, the organization is decentralized with a low level of operational integration and low level of profit-sharing to maintain local responsiveness. As the cost of information increases, the level of operational integration and the amount of profit-sharing increase and eventually the governance structure switches to centralization, with the switch associated with a discrete reduction in the amount of profit-sharing and a discrete increase in the level of operational integration. If the baseline volatility would be higher, then the level of operational integration would be lower at any cost of information and the switch to centralization would occur later.

Beyond confirming this much-discussed pattern, there are three results that deserve separate discussion.

The relationship between decentralization and volatility: It is commonly argued that when firms face an increase in the volatility of their market conditions, they should decentralize their decision-making authority to improve their local responsiveness. Indeed,

[11] $K(\beta) = K\left(1 - e^{-\mu_1\left(\frac{r^{\mu_2}}{\overline{\mu} - r^{\mu_3}}\right)}\right)$, where $K = 4, \mu_1 = 2.5, \mu_2 = 0.5, = \overline{\mu} = 1.2$ and $\mu_3 = 0.5$

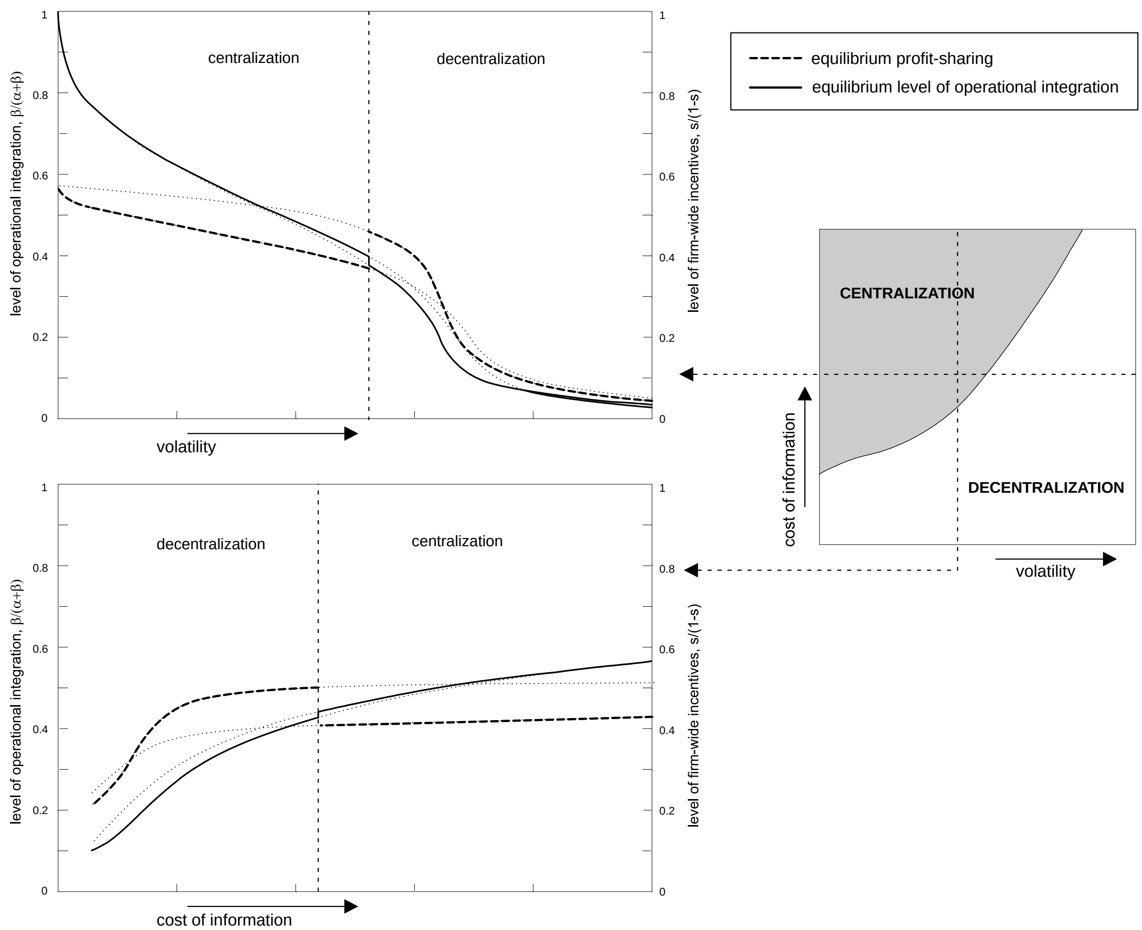


Figure 4: An illustration of the equilibrium comparative statics

there is emerging empirical evidence supporting a positive relationship between decentralization and volatility (e.g. Nagar 2002 and Wulf 2006) and the model also generates such an equilibrium relationship. However, and what is the key here, volatility is not directly causing decentralization in the model. Instead, an increase in volatility warrants a reduction in the level of operational integration, and it is that reduction in the level of operational integration that converts the organization into one that is better managed through a decentralized decision-making structure. Indeed, because the optimal governance structure is directly independent of the level of volatility, simply decentralizing decision-making authority would *worsen* organizational performance.

While this comparative static is made stark by the assumption that the cost of information is proportional to the volatility of the environment, having understood how μ and σ_θ^2 independently impact the optimal organizational design, we can now consider any co-movements between the two. Indeed, from proposition 5 we can directly infer that, absent adjustments in β, decentralization becomes more attractive if and only if the relative cost of information is decreasing in $\sigma_\theta^2 : \frac{\partial \mu}{\partial \sigma_\theta^2} < 0$. We also have a clear intuition for why this would be the case: a reduction in the relative cost of information relaxes the tradeoff between motivating information acquisition and then using that information appropriately. The relaxation of this constraint allows the organization to increase the equilibrium level of profit-sharing, which in turn benefits decentralization relatively more and makes it a more attractive governance structure.

The substitutability of centralization and firm-wide performance pay: It is often suggested that as it becomes increasingly possible to align division managers with the goals of the firm, the more attractive decentralization becomes because the managers now take better into account the firm-wide consequences of their actions.[12] This verbal account, however, ignores the fact that increasing the degree of incentive alignment also makes the managers more cooperative with the headquarters. In short, increasing the alignment of the division managers with the goals of the firm benefits both centralized and decentralized governance structures and it becomes an issue of quantitative differences in the benefits.

In the present model, the suggestion holds because the benefits of incentive alignment are accruing at a faster rate under decentralization. This result also leads to another conclusion, which is that when determining whether to centralize or decentralize decision-making, it is not only enough to look at the level of operational integration that is desired, but also to what extent the organization can use profit-sharing to align the incentives of the division managers. If profit-sharing is sufficiently easy (cost of information is low), then decentral-

[12] Also referred to as the alignment principle (see, for instance, Milgrom and Roberts, 1992).

ization continues to be preferred independent of the level of operational integration. It is only when information is sufficiently costly (and so profit-sharing sufficiently expensive) that centralized decision-making can be preferred.

The relationship between the use of firm-wide incentives and the governance structure: Some recent studies have examined the relationships among the governance structure, the use of firm-wide incentives and other variables. One strand, focusing on delegation and volatility, has found that firms operating in environments that exhibit more volatility and higher informational asymmetries are more decentralized and that decentralization is generally associated with pay that is more tied to firm-level performance (Nagar 2002 and Wulf 2006, for example). The second strand, focusing on the role of operational integration, has found that an increase in interdependencies across operating units is generally associated with a decrease in delegation and an increase in the use of firm-level performance measures (Bushman et al 1995, Christie et al 2003, Colombo and Delmasto 2004 and Abernethy et al 2004).

The results from the model are clearly consistent with these empirical results. In addition, the results from the model help to reconcile the apparently conflicting conclusions that both decentralization and centralization are associated with heavier use of firm-level performance measures. By noting that the optimal compensation structure depends on both the level of operational integration and the governance structure, we can actually identify three different regions: (i) centralized organizations with high levels of operational integration, making heavy use of firm-wide incentives, (ii) decentralized organizations with comparable levels of operational integration, making even heavier use of firm-wide incentives and (iii) decentralized organizations with low levels of operational integration, making little use of firm-wide incentives. Thus, depending on the level of operational integration, decentralization can indeed be associated with either more or less firm-level performance pay than centralization.

The first result is, to my knowledge, new in the literature while the latter two are also qualitatively different from both Dessein, Garicano and Gertner (2009) and Friebel and Raith (2009). The reason is that in both papers, the "integrated" task is always better performed through a "centralized" structure, thus leading to a one-to-one relationship between what I have called administrative and operational integration (of course, their models deal with different organizational tasks but the parallels are there), which then caused a one-to-one relationship between the compensation and governance structures, with centralization always associated with heavier use of firm-wide incentives. The results here highlight that administrative and operational integration are two distinct design parameters, where decentralized decision-making can be preferred even at high levels of operational integration, and so also eliminating the simple relationship between equilibrium compensation and governance structures.

5.2 Some extensions

5.2.1 General cost functions

The results of section 5 were derived under the assumption that the cost of information is given by $C(p_i, \sigma_\theta^2) = -\mu\sigma_\theta^2 (p_i + \ln(1 - p_i))$. The qualitative results extend, however, to any well-behaved cost functions. The reason for this lies in the fact that the primary economic driver for the differences in organizational design lie in the differential impact that the governance structures have on $\Psi_i^g (s, \beta)$ and $\widetilde{\Psi}_i^g (s, \beta)$, not what the common $C(p_i, \sigma_\theta^2)$ looks like. However, because the exact results do make use of both $C''(.)$ and $C'''(.)$, some of the results may not hold locally, even if the global tradeoffs are unchanged. The extent to which the results generalize is discussed in Appendix C.1. For example, it is possible that the relationship between decentralization and both μ and σ_θ^2 is non-monotone, even if it continues to be the case that as the level of operational integration is sufficiently low or the equilibrium level of profit-sharing is sufficiently high, decentralization is always preferred, while for sufficiently high β and low s, centralization is preferred. There is, however, a more economically relevant reason why the relationship between volatility and decentralization can be non-monotone, discussed next.

5.2.2 Endogenous strength of incentives

A key simplification with respect to the compensation structure of the managers was that the analysis only focused on a profit-sharing agreements where the total profits $\pi_i + \pi_j$ were divided between the two division managers. Another dimension in which the compensation of managers generally varies is in the strength of incentives - how sensitive is pay to performance in the first place. The need to maintain tractability precludes fully general analysis of the cost of incentives, but we can introduce such costs in two ways. The first is to assume that there is some exogenous cost of incentives function, $G(\lambda)$, where λ is the overall strength of incentives. The second is to assume that no ex ante transfers are possible, so that the headquarters maximizes $(1 - \lambda) E(\pi_i + \pi_j)$ instead of the total surplus.

In both cases, the results parallel those of section 5, but with one important exception. Now, for intermediate costs of information, the relationship between decentralization and volatility is inherently non-monotone. In particular, decentralization arises now as the preferred governance structure both when the level of operational integration is very high and when the level of operational integration is very low, with centralization remaining the preferred governance structure only for intermediate levels of operational integration. The economic intuition behind this non-monotonicity is as follows. As discussed in section 4, the main advantage of centralization is its ability to make relatively good use of the information generated even at low levels of profit-sharing, with this advantage being largest at intermediate levels of operational integration. Above, this advantage manifested itself in centralization choosing a lower level of profit-sharing and arising as the preferred governance structure for high enough β and μ. When the strength of incentives is endogeous, this advantage affects also the margin on the strength of incentives: because of its ability to tolerate conflict between the divisions (low s), centralization is able to use conflict instead of the strength of incentives to motivate information acquisition and thus to economize on the strength of incentives relative to decentralization. Because this ability to maintain high levels of information acquisition is more valuable for intermediate than high β, it is for intermediate levels of operational integration where centralization retains its advantage the longest. An illustration of these results is provided in Appendix C.2.

5.2.3 Cost versus complexity

Complexity of the environment is sometimes offered, together with volatility, as one of the key determinants when analyzing the link between organizational design and the environment. Cost of information can be seen as one aspect of environmental complexity but I have not discussed it as such because there are other aspects to complexity that are not captured by the cost of information and which can have a different impact on the organizational design. As an example, another dimension of complexity is the difficulty of transmitting information. The recipient of the messages can misunderstand the content of the messages, or information can be "hardened" but only at a cost, with both the likelihood of misunderstandings and costs of hardening the information increasing in the complexity of the environment. An increase in either decreases the attractiveness of centralization, implying that the impact on the optimal governance structure is exactly opposite to that of cost of information.[13] For empirical purposes, it is then crucial whether a given measure reflects the difficulty of generating information or transmitting that information.

[13] Results are available from the author on request.

6 Conclusion

I have examined the fit between the overall organizational design (composed of the level of operational integration, the allocation of decision rights and the compensation structure of division managers) and the environment (as determined by its volatility and the cost of information). The results were broadly consistent with the acepted wisdom regarding the fit between the organizational design and the environment. Organizations operating in volatile environments were loosely coupled, with decentralized decision-making and a compensation structure primarily based on divisional performance. Conversely, organizations operating in stable environments were tightly integrated, with generally centralized decision-making and a compensation structure based of firm-wide performance.

While the results were broadly consistent with the accepted wisdom, they also refined the relationship between organizational design and the environment. First, while decentralization and volatility were positively associated in equilibrium, volatility was not directly causing decentralization. Instead, an increase in volatility led to a decrease in operational integration, and it was that decrease in operational integration that made decentralization to be the preferred governance structure. Indeed, only decentralizing decision-making would have worsened organizational performance.

Second, centralized decision-making and incentive alignment were shown to be substitutes: while both governance structures benefited from increased profit-sharing in terms of more accurate communication and improved decision-making, these benefits were accruing at a faster rate under decentralization. Thus, when determining the optimal governance structure, an organization should not only consider its level of operational integration but also the optimal level of profit-sharing. If the optimal level of profit-sharing is sufficiently high, then decentralization will be preferred even if the level of operational integration is high.

Third, because of this substitutability, the results pointed out that there should be no clear empirical relationship between the governance structure and the compensation structure of the division managers. Instead, the compensation structure is determined by both the governance structure and the level of operational integration. Depending on the level of operational integration, a decentralized governance structure can then make more or less heavy use of firm-level incentives than a corresponding centralized governance structure. Overall the results suggest that the distinction between administrative and operational integration is potentially an important one, and illustrate how the optimal organizational structure has to account not only for the interactions between governance and compensation structures, but how they are both further impacted by the level of operational integration.

References

[1] Abernethy, M.A., J. Bouwens and L. van Lent (2004), "Determinants of Control System Design in Divisionalized Firms," *The Accounting Review*, 79(3), pp. 545-570

[2] Aghion, P. and J. Tirole (1997), "Formal and Real Authority in Organizations," Journal of Political Economy, 105, pp. 1-29

[3] Alonso, R. (2007), "Shared Control and Strategic Communication," working paper, University of Southern California

[4] Alonso, R., W. Dessein and N. Matouschek (2008), "When Does Coordination Require Centralization?," *American Economic Review*, 98(1), pp. 145-79

[5] Alonso, R., W. Dessein and N. Matouschek (2009), "Organize to Compete," working paper, USC, Northwestern university and Columbia university

[6] Athey, S. and J. Roberts (2001), "Organizational Design: Decision Rights and Incentive Contracts," The American Economic Review, 91(2), pp. 200-205

[7] Bartlett, C.A. (1986), "Building and managing the transnational: the new organizational challenge." In: Porter, M.E. (Ed.), Competition in Global Industries. Harvard Business School Press, Boston, MA.

[8] Bartlett, C.A. and S. Ghoshal (1989), *Managing Across Borders: The Transnational Solution*, Harvard Business School Press, Boston, MA

[9] Bolton, P. and M. Dewatripont (1994), "The Firm as a Communication Network," Quarterly Journal of Economics, 109, pp. 809-839

[10] Brickley, J.A., C.W. Smith and J.L. Zimmerman (2003), Designing Organizations to Create Value: From Structure to Strategy, McGraw-Hill, New York, NY

[11] Bushman, R., R. Indjejikian and A. Smith (1995), "Aggregate Performance Measures in Business Unit Manager Compensation: The Role of Intrafirm Interdependencies," *Journal of Accounting Research*, 33 (Supplement), pp.101-128

[12] Chandler, A. (1962), *Strategy and Structure: Chapters in the History of the American Enterprise*, The MIT Press, Cambridge, MA

[13] Chandler, A. (1977), *The Visible Hand: The Managerial Revolution in American Business*, The Belknap Press of Harvard University Press, Cambridge, MA and London, England

[14] Christie, A.A., M.P. Joye and R.L. Watts (2003), "Decentralization of the Firm: Theory and Evidence," *Journal of Corporate Finance*, 9, 3-36

[15] Colombo, M.G. and M. DelMastro (2004), "Delegation of Authority in Business Organizations: An Empirical Test," *The Journal of Industrial Economics*, 52(1), pp.53-80

[16] Colombo, M.G. and M. DelMastro (2008), *The Economics of Organizational Design*, Palgrave Macmillan, New York

[17] Crawford, V. and J. Sobel (1982), "Strategic Information Transmission," Econometrica, 50, pp. 1431-1452

[18] Dessein, W. (2002), "Authority and Communication in Organizations," Review of Economic Studies, 69, pp. 811-838

[19] Dessein, W., L. Garicano and R. Gertner (2008), "Organizing for Synergies," working paper, Columbia University, London School of Economics and University of Chicago

[20] Dessein, W. and T. Santos (2006), "Adaptive Organizations," *Journal of Political Economy,* 114, pp. 956-995

[21] Friebel, G. and M. Raith (2008), "Resource Allocation and Organizational Form," forthcoming, American Economic Journal: Microeconomics

[22] Galbraith, J.R. (1973), *Designing Complex Organizations*, Addison-Wesley Publishing Company, Reading, MA

[23] Galbraith, J.R. (1977), *Organization Design*, Addison-Wesley Publishing Company, Reading, MA

[24] Garicano, L. (2000), "Hierarchies and the Organization of Knowledge in Production," Journal of Political Economy, 108(5), pp. 874-904

[25] Guadalupe, M. and J. Wulf (2008), "The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization," working paper, Columbia University and Harvard University

[26] Harris, M. and A. Raviv (2002), "Organization Design," *Management Science*, 48(7), pp. 852-865

[27] Harris, M. and A. Raviv (2005), "Allocation of Decision-Making Authority," Review of Finance, 9(3), pp. 353-383

[28] Harris, M. and A. Raviv (2007),"A Theory of Board Control and Size," forthcoming, Review of Financial Studies

[29] Hart, O. and J. Moore (2005), "On the Design of Hierarchies: Coordination versus Specialization," Journal of Political Economy, 113(4), pp. 675-702

[30] Harzing, A-W. (2000), "An Empirical Analysis and Extension of the Bartlett and Ghoshal Typology of Multinational Companies," Journal of International Business Studies, 31(1), pp.101-120

[31] Holmstrom, B. (1999), "The Firm as a Subeconomy," Journal of Law, Economics and Organization, 15(1), pp. 74-102

[32] Holmstrom, B. and P. Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," Journal of Law, Economics and Organization, 7, pp. 24-51

[33] Holmstrom, B. and P. Milgrom (1994), "The Firm as an Incentive System," The American Economic Review, 84(4), pp.972-991

[34] Milgrom, P. and J. Roberts (1990), "The Economics of Modern Manufacturing: Technology, Strategy and Organization," The American Economic Review, 80(3), pp.511-528

[35] Lawrence, P. and J. Lorsch (1967), *Organization and the Environment*, Harvard Business School Press, Boston, MA

[36] Marshak J. and R. Radner (1972), *Economic Theory of Teams*, Yale University Press, New Haven and London

[37] Melumad, N. and T. Shibano (1991), "Communication in Settings with No Transfers," The RAND Journal of Economics, 22, pp.173-198

[38] Milgrom, P. and J. Roberts (1992), *Economics, Organization and Management*, Englewood Cliffs, N.J.: Prentice-Hall

[39] Milgrom, P. and J. Roberts (1995), "Complementarities and fit: Strategy, structure, and organizational change in manufacturing," Journal of Accounting and Economics, 19, pp.179-208

[40] Mintzberg, H. (1979), *The Structuring of Organizations*, Prentice Hall, Englewood Cliffs, NJ

[41] Nadler, D.A. and M.L. Tushman (1997), *Competing By Design: The Power of Organizational Architecture*, Oxford University Press, New York, NY

[42] Nagar, V. (2002), "Delegation and Incentive Compensation," *The Accounting Review*, 77(2), pp.379-395

[43] Porter, M.E. (1980), *Competitive Strategy*, The Free Press, New York, NY

[44] Porter, M.E. (1986), "Changing Patterns of International Competition," California Management Review, 28(2), pp.9-40

[45] Porter, M.E. (1996), "What is Strategy?" Harvard Business Review, 74(6), pp. 61-78

[46] Prahalad, C. and Y. Doz (1987), *The Multinational Mission: Balancing Local Demands and Global Vision*, Free Press, New York, NY

[47] Qian, Y., G. Roland and C. Xu (2006), "Coordination and Experimentation in M-Form and U-Form Organizations," Journal of Political Economy, 114(2), pp.366-402

[48] Rantakari, H (2008), "Governing Adaptation," *Review of Economic Studies*, Vol. 75(4), pp. 1257-1285

[49] Roberts, J. (2004), *The Modern Firm: Organizational Design for Performance and Growth*, Oxford University Press, New York, NY

[50] Simon, H.A. (1947), *Administrative Behavior*, The Free Press, New York, NY

[51] Stein, J. (1989), "Cheap Talk and the Fed: A Theory of Imprecise Policy Announcements," The American Economic Review, 79(1), pp. 32-42

[52] Thompson, J.D. (1967), *Organizations in Action*, McGraw-Hill, New York, NY

[53] Woodward, J. (1965), *Industrial Organization: Theory and Practice*, Oxford University Press, London, UK

[54] Wulf, J. (2006), "Authority, Risk, and Performance Incentives: Evidence from Division Manager Positions Inside Firms," forthcoming, *Journal of Industrial Economics*

[55] Zabojnik, J. (2002), "Centralized and Decentralized Decision Making in Organizations," Journal of Labor Economics, University of Chicago Press, 20(1), pp. 1-22

A Proofs

A.1 Proposition 1: Equilibrium decisions

We can write the profits of division i as $\pi_i = K(\beta_i) - \alpha_i(\theta_i - d_i)^2 - \beta_i(d_i - d_j)^2 = K(\beta_i) - k_i\left((1-r_i)(\theta_i - d_i)^2 + r_i(d_j - d_i)^2\right)$, where β_i is the extent to which division i is integrated with division j, $k_i = \alpha_i + \beta_i$ and $r_i = \frac{\beta_i}{\alpha_i+\beta_i}$. In the derivations that follow, I will allow (for completeness) for both asymmetric interdependencies between the two divisions and asymmetric compensation contracts. Finally, let m_i and m_j denote the equilibrium messages exchanged in the communication stage. Given the profit structure, we have then that under centralization, the headquarters solves

$$\max_{d_i,d_j} E_{HQ}(\pi_i + \pi_j).$$

The first-order conditions are

$$\begin{aligned} d_i &= \tfrac{k_i(1-r_i)E(\theta_i|m_i)+(k_ir_i+k_jr_j)d_j}{k_i+k_jr_j} = a_1E(\theta_i|m_i) + a_2d_j \\ d_j &= \tfrac{k_j(1-r_j)E(\theta_j|m_j)+(k_ir_i+k_jr_j)d_i}{k_j+k_ir_i} = b_1E(\theta_j|m_j) + b_2d_i, \end{aligned}$$

and solving the the equilibrium decisions gives then

$$d_i^{cent} = \tfrac{a_1E(\theta_i|m_i)+a_2b_1E(\theta_j|m_j)}{(1-a_2b_2)} \qquad and \qquad d_j^{cent} = \tfrac{b_1E(\theta_j|m_j)+b_2a_1E(\theta_i|m_i)}{(1-a_2b_2)},$$

where $a_1 = \frac{k_i(1-r_i)}{k_i+k_jr_j}, a_2 = \frac{k_ir_i+k_jr_j}{k_i+k_jr_j}, b_1 = \frac{k_j(1-r_j)}{k_j+k_ir_i}$ and $b_2 = \frac{(k_jr_j+k_ir_i)}{k_j+k_ir_i}$.

Similarly, under decentralization we can write manager i's optimization problem as

$$\max_{d_i} \lambda_i E_{HQ}((1-s_i)\pi_i + s_i\pi_j),$$

where $s_i \in [0, 1/2]$ is the relative weight that manager i places on the profits of division j (amount of profit-sharing) and λ_i is the overall strength of incentives. In the case of pure profit-sharing, $\lambda_i = 1$. Then, the first-order conditions become

$$d_i = \tfrac{(1-s_i)k_i(1-r_i)E(\theta_i|t_i)+((1-s_i)k_ir_i+s_ik_jr_j)E_i(d_j|m_j)}{(1-s_i)k_i+s_ik_jr_j} = a_1E(\theta_i|t_i) + a_2E_i(d_j|m_j)$$

and symmetrically for manager j. Solving for the intersection of the reaction functions, we get

$$d_i^{dec} = \tfrac{a_1(1-b_2a_2)E(\theta_i|t_i)+a_2b_1E_i(E(\theta_j|t_j)|m_j)+a_2b_2a_1E_j(E(\theta_i|t_i)|m_i)}{1-b_2a_2}$$

$$d_j^{dec} = \frac{b_1(1-b_2a_2)E(\theta_j|t_j)+b_2a_1E_j(E(\theta_i|t_i)|m_i)+a_2b_2b_1E_i(E(\theta_j|t_j)|m_j)}{1-b_2a_2},$$

where $a_1 = \frac{(1-s_i)k_i(1-r_i)}{(1-s_i)k_i+s_ik_jr_j}, a_2 = \frac{(1-s_i)k_ir_i+s_ik_jr_j}{(1-s_i)k_i+s_ik_jr_j}, b_1 = \frac{(1-s_j)k_j(1-r_j)}{(1-s_j)k_j+s_jk_ir_i}$ and $b_2 = \frac{(1-s_j)k_jr_j+s_jk_ir_i}{(1-s_j)k_j+s_jk_ir_i}$.

Note that since under decentralization, the two decisions are made by different parties, the decision of manager i does not only depend on his beliefs about his state $E(\theta_i|s_i)$ and his beliefs about the other division's state, $E_i(E(\theta_j|t_j)|m_j)$, but also on his beliefs about what manager j thinks θ_i is, because that influences the equilibrium decision d_j and thus the optimal d_i. Note also that the weights under decentralization would match those under centralization if $s_i = s_j = 1/2$.

Under symmetric solution, the equations then simplify to

$$d_i^{cent} = \frac{(\alpha+2\beta)E(\theta_i|m_i)+2\beta E(\theta_j|m_j)}{\alpha+4\beta}$$

and

$$d_i^{dec} = \frac{(1-s)\alpha}{(1-s)\alpha+\beta}E(\theta_i|t_i)+\frac{\beta}{(1-s)\alpha+2\beta}E_i(E(\theta_j|t_j)|m_j)+\frac{\beta^2}{((1-s)\alpha+\beta)((1-s)\alpha+2\beta)}E_j(E(\theta_i|t_i)|m_i).$$

To evaluate the severity of the bias, note that since the coefficients add up to one, any over-responsiveness to own information implies an equivalent under-responsiveness to information regarding the other division. This under-reponsiveness is then given by

$$\frac{2\beta}{\alpha+4\beta} - \frac{\beta}{(1-s)\alpha+2\beta},$$

so that the difference is maximized when

$$\beta = \frac{\alpha}{4}\sqrt{2(1-s)}.$$

A.2 Proposition 2: Equilibrium communication

In the communication stage, the division managers send non-verifiable messages about their local conditions to the decision-maker(s). Knowing how his message influences the decisions, as given by proposition 1, manager i chooses his message from the set of equilibrium messages $\{m_i\}$ to

$$\min_{m_i} E_i\left[\begin{array}{c}(1-s_i)k_i(1-r_i)(\theta_i - d_i(.,m_i))^2 + ((1-s_i)k_ir_i + s_ik_jr_j)(d_i(.,m_i) - d_j(.,m_i))^2 \\ +s_ik_j(1-r_j)(\theta_j - d_j(.,m_i))^2\end{array}\right].$$

The solution requires us to find cutoffs $t_1, t_2, ...$ such that given messages $m_k \rightarrow t_i \in [t_{k-1}, t_k)$

and $m_{k+1} \to t_i \in [t_k, t_{k+1})$, the manager is indifferent between sending messages m_k and m_{k+1} whenever $t_i = t_k$. Then, the fact that the expected decision is increasing in the messages implies that $\frac{\partial^2 U_i}{\partial m_i \partial \theta_i} \geq 0$ and so the agent strictly prefers telling the truth for the other messages. The indifference condition for the agent becomes

$$\begin{aligned}
&E\{((1-s_i)k_i(1-r_i))\left(\left(d_i^H(.,m_i)+d_i^L(.,m_i)\right)-2\theta_i\right)\left(d_i^H(.,m_i)-d_i^L(.,m_i)\right)\\
&+(s_ik_j(1-r_j))\left(d_j^H(.,m_i)+d_j^L(.,m_i)\right)\left(d_j^H(.,m_i)-d_j^L(.,m_i)\right)\\
&+((1-s_i)k_ir_i+s_ik_jr_j)\left(\left(d_i\left(.,m_i^H\right)-d_j\left(.,m_i^H\right)\right)^2-\left(d_i\left(.,m_i^L\right)-d_j\left(.,m_i^L\right)\right)^2\right)\}=0.
\end{aligned}$$

Now, under centralization,

$$\begin{aligned}
&d_i(.,m_i)-d_j(.,m_i)=\tfrac{(a_1-b_2a_1)}{(1-a_2b_2)}E(\theta_i|m_i)-\tfrac{(b_1-a_2b_1)}{(1-a_2b_2)}E(\theta_j|m_j)\\
&d_i^H(.,m_i)-d_i^L(.,m_i)=\tfrac{a_1}{(1-a_2b_2)}\left(E\left(\theta_i|m_i^H\right)-E\left(\theta_i|m_i^L\right)\right)\\
&d_j^H(.,m_i)-d_j^L(.,m_i)=\tfrac{b_2a_1}{(1-a_2b_2)}\left(E\left(\theta_i|m_i^H\right)-E\left(\theta_i|m_i^L\right)\right)\\
&E\left(d_i^H(.,m_i)+d_i^L(.,m_i)\right)=\tfrac{a_1}{(1-a_2b_2)}\left(E\left(\theta_i|m_i^H\right)+E\left(\theta_i|m_i^L\right)\right)\\
&E\left(d_j^H(.,m_i)+d_j^L(.,m_i)\right)=\tfrac{b_2a_1}{(1-a_2b_2)}\left(E\left(\theta_i|m_i^H\right)+E\left(\theta_i|m_i^L\right)\right)
\end{aligned}$$

so that the above simplifies to, noting that $E\theta_i = q_i t_i$, $E\left(\theta_i|m_i^H\right) = q_i\left(\frac{t_{k+1}+t_k}{2}\right)$ and that $E\left(\theta_i|m_i^L\right) = q_i\left(\frac{t_k+t_{k-1}}{2}\right)$

$$|t_{k+1}-t_k| = |t_k-t_{k-1}| + 4\frac{\left[((1-s_i)k_i(1-r_i))a_2b_1-(s_ik_j(1-r_j))b_2^2a_1-((1-s_i)k_ir_i+s_ik_jr_j)\left(a_1b_1^2\right)\right]}{a_1\left[((1-s_i)k_i(1-r_i))+(s_ik_j(1-r_j))b_2^2+((1-s_i)k_ir_i+s_ik_jr_j)b_1^2\right]}t_k.$$

The same logic applies in the case of decentralization, in which case the indifference condition becomes

$$|t_{k+1}-t_k| = |t_k-t_{k-1}| + 4\frac{\left[(1-s_i)k_i(1-r_i)a_2^2b_1+((1-s_i)k_ir_i+s_ik_jr_j)a_1^2b_1-s_ik_j(1-r_j)b_2a_1\right]}{a_1b_2\left[(1-s_i)k_i(1-r_i)a_2^2+((1-s_i)k_ir_i+s_ik_jr_j)a_1^2+s_ik_j(1-r_j)\right]}t_k.$$

As shown in Rantakari (2008), given that the solution in both cases takes the form $|t_{k+1}-t_k| = |t_k-t_{k-1}| + \frac{4}{\varphi}t_k$, the expected variance under the most informative partition is given by

$$E\left(t_i - E_j\left(t_i|m_i\right)\right)^2 = \frac{1}{12}\left(\frac{(1-\zeta(\varphi))^3}{\left(1-\zeta(\varphi)^3\right)}\right)\overline{\theta}^2,$$

where $\zeta(\varphi) = \frac{\varphi}{\left(1+\sqrt{1+\varphi}\right)^2}$. We can then further simplify $\frac{(1-\zeta(\varphi))^3}{\left(1-\zeta(\varphi)^3\right)}$ by noting that $1-\zeta(\varphi) = 1-\frac{\varphi}{\left(1+\sqrt{1+\varphi}\right)^2} = \frac{2}{\left(1+\sqrt{1+\varphi}\right)}$ and then $1-\zeta(\varphi)^3 = \left(1+\zeta(\varphi)+\zeta(\varphi)^2\right)(1-\zeta(\varphi))$,

we get

$$\frac{(1-\zeta(\varphi))^3}{1-\zeta(\varphi)^3} = \frac{(1-\zeta(\varphi))^2}{1+\zeta(\varphi)+\zeta(\varphi)^2} = \frac{\frac{4}{\left(1+\sqrt{1+\varphi}\right)^2}}{1+\frac{\varphi}{\left(1+\sqrt{1+\varphi}\right)^2}+\frac{\varphi^2}{\left(1+\sqrt{1+\varphi}\right)^4}} =$$
$$\frac{4\left(1+\sqrt{1+\varphi}\right)^2}{8+8\sqrt{1+\varphi}+10\varphi+6\varphi\sqrt{1+\varphi}+3\varphi^2} = \frac{4\left(1+\sqrt{1+\varphi}\right)^2}{(4+3\varphi)\left(1+\sqrt{1+\varphi}\right)^2} = \frac{4}{4+3\varphi}.$$

In other words, we can write

$E\left(t_i - E_j\left(t_i|m_i\right)\right)^2 = \frac{1}{12}\left(\frac{(1-\zeta(\varphi))^3}{\left(1-\zeta(\varphi)^3\right)}\right)\overline{\theta}^2 = \frac{1}{12}\frac{4}{4+3\varphi}\overline{\theta}^2$, or $\frac{1}{4+3\varphi}\frac{\overline{\theta}^2}{3}$. To match our formulation of $\left(1-V_i^g(s,\alpha,\beta)\right)\frac{\overline{\theta}^2}{3}$, we have that $V_i^g(s,\alpha,\beta) = 1-\frac{1}{4+3\varphi} = \frac{3+3\varphi}{4+3\varphi}$,where φ is given by the original indifference condition. If $\varphi \to 0$, communication becomes binomial (the managers say whether the state is above or below the ex ante expectation), while as $\varphi \to \infty$, communication becomes perfect. The key part of the solution is that the actual accuracy of the signal, q_i, has no impact on the accuracy of its transmission. The reason for this result is that because both the sender and the receiver discount the information through it's precision the same way, the precision of the signal drops out when determining the relative incentive conflict between the two parties.

In the case of symmetric divisions, the communication equilibrium simplifies considerably with simple substitution giving us

$$V_i^{cent}(s,\beta) = \frac{3(1-s)(2\beta+\alpha)}{\beta(1-2s)+3(1-s)(2\beta+\alpha)} \text{ and } V_i^{dec}(s,\beta) = \frac{3(1-s)((1-s)\alpha+2\beta)}{((1-s)\alpha+\beta)(1-2s)+3(1-s)((1-s)\alpha+2\beta)}.$$

A.3 Proposition 3: Expected profits and the value of information

In the case of decentralization, after substituting in the equilibrium decisions and taking expectations, we can write the expected profits of the organization as

$$\begin{aligned} E\left(\pi_i\right) = K\left(\beta\right) &- \frac{k_i b_1^2\left((1-r_i)a_2^2+r_i a_1^2\right)}{(1-b_2a_2)^2} E\left(\left(E_jE_i\left(\theta_i\right)\right)^2 + \left(E_iE_j\left(\theta_j\right)\right)^2\right) \\ &- k_i\left((1-r_i)\,a_2^2 + r_i a_1^2\right) E\left(E_i\left(\theta_i\right) - E_jE_i\left(\theta_i\right)\right)^2 - k_i r_i b_1^2 E\left(E_j\left(\theta_j\right) - E_iE_j\left(\theta_j\right)\right)^2 \\ &- k_i\left(1-r_i\right) E(\theta_i - E_i\left(\theta_i\right))^2 \end{aligned}$$

and symmetrically for $E\left(\pi_j\right)$, where (a_1, a_2, b_1, b_2) are the weights placed by managers i and j on adaptation and coordination in their first-order conditions for optimal decisions. Note that the only component that depends directly on the precision of the signals is $E(\theta_i - E_i\left(\theta_i\right))^2$. For this, we have that

$$E\left(\theta_i - E_i\theta_i\right)^2 = q_i E\left(\theta_i - q_i\theta_i\right)^2 + \left(1-q_i\right) E\left(\theta_i - q_i x_i\right)^2,$$

where x_i is a random draw from $\left[-\overline{\theta}, \overline{\theta}\right]$. As a result,

$$E\left(\theta_i - E_i\theta_i\right)^2 = \left(1-q_i^2\right)\frac{\overline{\theta}_i^2}{3} = \left(1-p_i\right)\frac{\overline{\theta}_i^2}{3}.$$

For the rest, letting $\frac{k_i b_1^2\left(\left(1-r_i\right)a_2^2 + r_i a_1^2\right)}{\left(1-b_2 a_2\right)^2} = \Lambda_i$, we can write it as (by adding and subtracting $\Lambda_i E\left(E_i\left(\theta_i\right)^2 + E_j\left(\theta_j\right)^2\right)$)

$$E\left(\pi_i\right) = K\left(\beta_i\right) - \Lambda_i E\left(E_i\left(\theta_i\right)^2 + E_j\left(\theta_j\right)^2\right) - \left(k_i\left(\left(1-r_i\right)a_2^2 + r_i a_1^2\right) - \Lambda_i\right)E\left(E_i\left(\theta_i\right) - E_j E_i\left(\theta_i\right)\right)^2$$
$$- \left(k_i r_i b_1^2 - \Lambda_i\right)E\left(E_j\left(\theta_j\right) - E_i E_j\left(\theta_j\right)\right)^2 - k_i\left(1-r_i\right)\left(1-p_i\right)\frac{\overline{\theta}_i^2}{3}.$$

Then, noting that $E_i\left(\theta_i\right) = q_i t_i$ and $E_j E_i\left(\theta_i\right) = q_i E\left(t_i|m_i\right)$, we get that

$$E\left(\pi_i\right) = K\left(\beta_i\right) - \Lambda_i E\left(q_i^2 t_i^2 + q_j^2 t_j^2\right) - \left(k_i\left(\left(1-r_i\right)a_2^2 + r_i a_1^2\right) - \Lambda_i\right)q_i^2 E\left(t_i - E\left(t_i|m_i\right)\right)^2$$
$$- \left(k_i r_i b_1^2 - \Lambda_i\right)q_j^2 E\left(t_j - E\left(t_j|m_j\right)\right)^2 - k_i\left(1-r_i\right)\left(1-q_i^2\right)\frac{\overline{\theta}_i^2}{3}.$$

Let $k_i\left(\left(1-r_i\right)a_2^2 + r_i a_1^2\right) - \Lambda_i = \Gamma_{ii}$ and $\left(k_i r_i b_1^2 - \Lambda_i\right) = \Gamma_{ji}$, which simplifies the expression to

$$E\left(\pi_i\right) = K\left(\beta_i\right) - \Lambda_i E\left(q_i^2 t_i^2 + q_j^2 t_j^2\right) - \Gamma_{ii} q_i^2 E\left(t_i - E\left(t_i|m_i\right)\right)^2$$
$$-\Gamma_{ji} q_j^2 E\left(t_j - E\left(t_j|m_j\right)\right)^2 - k_i\left(1-r_i\right)\left(1-q_i^2\right)\frac{\overline{\theta}_i^2}{3},$$

and then letting $p_i = q_i^2$, noting that $Et_i^2 = \frac{\overline{\theta}^2}{3}$ and that $E\left(t_i - E\left(t_i|m_i\right)\right)^2 = \left(1-V_i\right)\frac{\overline{\theta}^2}{3}$,the expression becomes

$$E\left(\pi_i\right) = K\left(\beta_i\right) - \alpha_i\frac{\overline{\theta}^2}{3} + \left(p_i\left(\alpha_i - \Lambda_i - \Gamma_{ii}\left(1-V_i\right)\right) - p_j\left(\Lambda_i + \Gamma_{ji}\left(1-V_j\right)\right)\right)\frac{\overline{\theta}^2}{3},$$

while adding over the two divisions we get

$$K\left(\beta_i\right) + K\left(\beta_j\right) - \left(\alpha_i + \alpha_j\right)\frac{\overline{\theta}^2}{3} + \left(p_i\Psi_i^g + p_j\Psi_j^g\right)\frac{\overline{\theta}^2}{3},$$

where

$$\Psi_i^g = \alpha_i - \left(\Lambda_i + \Lambda_j\right) - \left(\Gamma_{ii} + \Gamma_{ij}\right)\left(1-V_i\right) \text{ and}$$
$$\Psi_j^g = \alpha_j - \left(\Lambda_i + \Lambda_j\right) - \left(\Gamma_{jj} + \Gamma_{ji}\right)\left(1-V_j\right),$$

the values of information for the two divisions, where $\Lambda_i + \Gamma_{ii}\left(1-V_i\right) = \phi_i$ and $\Lambda_j + \Gamma_{ij}\left(1-V_i\right) = \phi_j$.

Repeating the exercise under centralization, the solution takes the same structure with the expected profits after the first substitution equal to

$\frac{k_i b_1^2\left((1-r_i)a_2^2+r_i a_1^2\right)}{(1-b_2 a_2)^2} E\left((E(\theta_i|m_i))^2+(E(\theta_j|m_j))^2\right)+k_i(1-r_i)E(E(\theta_i|m_i)-\theta_i)^2.$

Now, we can write $E(E(\theta_i|m_i)-\theta_i)^2 = E(\theta_i - q_i E(t_i|m_i))^2 = E(\theta_i - q_i t_i + q_i t_i - q_i E(t_i|m_i))^2$,

so that $E(\theta_i - q_i t_i + q_i t_i - q_i E_m t_i)^2 = E(\theta_i - q_i t_i)^2 + q_i^2 E(t_i - E(t_i|m_i))^2$ because the equilibrium beliefs must be unbiased. As result, the expression simplifies to

$$E(\theta_i - q_i E(t_i|m_i))^2 = (1-p_i)\frac{\overline{\theta}_i^2}{3} + p_i E(t_i - E(t_i|m_i))^2,$$

where the first component gives the loss in profits due to inaccurate primary information while the second captures the additional loss due to the strategic communication part. Then, again adding and substituting the component $\Lambda_i E\left(E_i(\theta_i)^2 + E_j(\theta_j)^2\right)$, we can write the expected profits for the division as

$$\begin{aligned}&\Lambda_i E\left(p_i t_i^2 + p_j t_j^2\right) + ((k_i(1-r_i)) - \Lambda_i)\, p_i E(t_i - E(t_i|m_i))^2 \\ &-\Lambda_i p_j E(t_j - E(t_j|m_i))^2 + (k_i(1-r_i))(1-p_i)\frac{\overline{\theta}_i^2}{3},\end{aligned}$$

and letting $\Gamma_{ii} = ((k_i(1-r_i)) - \Lambda_i)$ and $\Gamma_{ji} = -\Lambda_i$, using the equilibrium quality of communication and adding up across the divisions, allows us again the profits as

$$K(\beta_i) + K(\beta_j) - (\alpha_i + \alpha_j)\frac{\overline{\theta}^2}{3} + \left(p_i \Psi_i^g + p_j \Psi_j^g\right)\frac{\overline{\theta}^2}{3},$$

where again

$$\begin{aligned}&\Psi_i^g = \alpha_i - (\Lambda_i + \Lambda_j) - (\Gamma_{ii} + \Gamma_{ij})(1 - V_i) \text{ and} \\ &\Psi_j^g = \alpha_j - (\Lambda_i + \Lambda_j) - (\Gamma_{jj} + \Gamma_{ji})(1 - V_j).\end{aligned}$$

In the case of symmetric divisions and a symmetric profit-sharing rule, the relevant components simplify under centralization to

$$\Lambda_i^{cent} = \Lambda_j^{cent} = \frac{\alpha\beta}{(\alpha+4\beta)},\ \ \Gamma_{ii}^{cent} = \Gamma_{jj}^{cent} = \alpha - \Lambda_i^{cent} \ \text{ and } \ \Gamma_{ij}^{cent} = \Gamma_{ji}^{cent} = -\Lambda_i^{cent},$$

while under decentralization, the coefficients are:

$$\Lambda_i^{dec} = \Lambda_j^{dec} = \frac{\alpha\beta\left(\beta+s^2\alpha\right)}{(s\alpha+2\beta)^2},\ \Gamma_{ii}^{dec} = \Gamma_{jj}^{dec} = \frac{\alpha\beta\left(\beta+s^2\alpha\right)}{(s\alpha+\beta)^2} - \Lambda_i^{dec} \text{ and } \Gamma_{ij}^{dec} = \Gamma_{ji}^{dec} = \beta\left(\frac{s\alpha}{s\alpha+\beta}\right)^2 - \Lambda_i^{dec}.$$

For interpretation, note that Λ_i^g measures the loss in the value of information due to the equilibrium decisions, so that the loss due to the bias in equilibrium decisions is determined by the difference between centralization (first-best decisions) and decentralization,

$$\Delta\Lambda_i^{dec}(s,\beta) = \frac{2\alpha\beta\left(\beta+s^2\alpha\right)}{(s\alpha+2\beta)^2} - \frac{2\alpha\beta}{(\alpha+4\beta)},$$

while the loss due to strategic communication under the two governance structures are then

$$\left(\Gamma_{ii}^{cent} + \Gamma_{ij}^{cent}\right)\left(1 - V_i^{cent}(s,\beta)\right) = \left(\alpha - \frac{2\alpha\beta}{(\alpha+4\beta)}\right)\left(\frac{\beta(1-2s)}{\beta(1-2s)+3(1-s)(2\beta+\alpha)}\right)$$
$$\left(\Gamma_{ii}^{dec} + \Gamma_{ij}^{dec}\right)\left(1 - V_i^{dec}(s,\beta)\right) =$$
$$\left(\frac{\alpha\beta\left(\beta+s^2\alpha\right)}{(s\alpha+\beta)^2} + \beta\left(\frac{s\alpha}{s\alpha+\beta}\right)^2 - 2\frac{\alpha\beta\left(\beta+s^2\alpha\right)}{(s\alpha+2\beta)^2}\right)\left(\frac{((1-s)\alpha+\beta)(1-2s)}{((1-s)\alpha+\beta)(1-2s)+3(1-s)((1-s)\alpha+2\beta)}\right).$$

The totals for the value of information then simplify to

$$\Psi_i^{cent}(s,\beta) = \frac{3\alpha(1-s)(\alpha+2\beta)^2}{(\alpha+4\beta)(\beta(1-2s)+3(1-s)(2\beta+\alpha))}$$
$$\Psi_i^{dec}(s,\beta) = \frac{\alpha(1-s)\left(5s^3\alpha^2(\alpha-2\beta)+s^2\left(5\beta\alpha^2+14\alpha\left(\beta^2-\alpha^2\right)\right)+s\left(16\beta\alpha^2+2\alpha\beta^2+13\alpha^3\right)-\left(4\alpha^3+6\beta^3+11\beta\alpha^2+12\alpha\beta^2\right)\right)}{((1-s)\alpha+2\beta)((1-s)\alpha+\beta)[((1-s)\alpha+\beta)(1-2s)+3(1-s)((1-s)\alpha+2\beta)]}$$

From these expressions, we can then evaluate to comparative statics in the proposition. Since these equations are exact, the easy proof is to simply plot the curves for all s and $\frac{\beta}{\alpha+\beta}$. It is possible to verify the comparative statics also algebraically, but that task is very tedious. As a result, they can be found in the separate appendix B.

A.4 Proposition 4: Quality of primary information

Recall that the solution to the manager's information acquisition problem is given by

$$p_i = \frac{\widetilde{\Psi}_i^g(s,\beta)}{\widetilde{\Psi}_i^g(s,\beta)+\mu}.$$

From this it is then immediate that $\frac{\partial p_i}{\partial \mu} < 0$. To establish $\frac{\partial p_i}{\partial \beta}$ and $\frac{\partial p_i}{\partial s}$, we need to examine their impact on the perceived value of information, where the signs will match $\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial x}$. Here, we have that the perceived value of information under the two governance structures can be written as

$$\widetilde{\Psi}_i^{cent}(s,\beta) = \frac{3\alpha(1-s)(\alpha+2\beta)((1-s)\alpha+\beta(3-4s))}{(\alpha+4\beta)(\beta(1-2s)+3(1-s)(2\beta+\alpha))}$$

$$\widetilde{\Psi}_i^{dec}(s,\beta) = \frac{\alpha(1-s)\left(9\beta^2+11\alpha\beta+13s^2\alpha\beta+\left(4-13s+14s^2-5s^3\right)\alpha^2-12s\beta(\beta+2\alpha)\right)}{((1-s)\alpha+2\beta)[((1-s)\alpha+\beta)(1-2s)+3(1-s)((1-s)\alpha+2\beta)]},$$

from which we can (through some tedious algebra) verify the comparative statics. Because the figure provided is exact, I have relegated the alebra to appendix B. Similarly, for the negative externality, we would have from above that

$$\phi_j^g(s,\beta) = \Lambda_j^g(s,\beta) + \Gamma_{ij}^g(s,\beta)\left(1 - V_i^g(s,\beta)\right)$$

for each governance structure, where the coefficients follow from proposition 3.

A.5 Proposition 5: Choice of organizational design

Recall that our maximization problem is

$$\max_{g,s,\beta} \sum_{k=i,j} \left(K(\beta) - \tfrac{\alpha\bar{\theta}^2}{3} + \left(p_k^g(\beta,s,\mu)\Psi_k^g(s,\beta) - C(\mu, p_k^g(\beta,s,\mu))\right)\tfrac{\bar{\theta}^2}{3} \right)$$

$$\text{s.t.} \qquad p_k^g(\beta,s,\mu) = \tfrac{\widetilde{\Psi}_i^g(s,\beta)}{\widetilde{\Psi}_i^g(s,\beta)+\mu}.$$

A.5.1 Properties of $s^g(\beta,\mu)$

Let us first consider the choice of s, the first-order condition for which is

$$p_i \tfrac{\partial \Psi_i^g(s,\beta)}{\partial s} = -\left(\Psi_i^g(s,\beta) - \mu C'(p_i)\right)\tfrac{\partial p_i}{\partial s}.$$

Then, note that from the manager's first-order condition we have that $\widetilde{\Psi}_i^g(s,\beta) = \mu C'(p_i)$, so that we can also write this first-order condition as

$$p_i \tfrac{\partial \Psi_i^g(s,\beta)}{\partial s} = -\left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta)\right)\tfrac{\partial p_i}{\partial s}.$$

The solution will thus always have $\Psi_i^g(s,\beta) > \widetilde{\Psi}_i^g(s,\beta)$, so that to balance the tradeoff between the generation and use of information, the level of information acquired is always below its first-best level, conditional on s. Now, with respect to the amount of profit-sharing, the key comparative static is with respect to μ. The first-order condition was

$$p_i \tfrac{\partial \Psi_i^g(s,\beta)}{\partial s} + \left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta)\right)\tfrac{\partial p_i}{\partial s} = 0,$$

so that the cross-partial is given by

$$\frac{\partial p_i}{\partial \mu}\frac{\partial \Psi_i^g(s,\beta)}{\partial s}+\left(\Psi_i^g\left(s,\beta\right)-\widetilde{\Psi}_i^g\left(s,\beta\right)\right)\frac{\partial^2 p_i}{\partial s\partial \mu}.$$

now,

$$\frac{\partial p_i}{\partial \mu}=-\frac{\widetilde{\Psi}_i^g(s,\beta)}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^2}\qquad and \qquad \frac{\partial^2 p_i}{\partial s\partial \mu}=\frac{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\left(\widetilde{\Psi}_i^g(s,\beta)-\mu\right)}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^3},$$

so that the above becomes

$$-\frac{\widetilde{\Psi}_i^g(s,\beta)}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^2}\frac{\partial \Psi_i^g(s,\beta)}{\partial s}+\left(\Psi_i^g\left(s,\beta\right)-\widetilde{\Psi}_i^g\left(s,\beta\right)\right)\frac{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\left(\widetilde{\Psi}_i^g(s,\beta)-\mu\right)}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^3},$$

while from the first-order condition we can write

$$\left(\Psi_i^g\left(s,\beta\right)-\widetilde{\Psi}_i^g\left(s,\beta\right)\right)=-p_i\frac{\partial \Psi_i^g(s,\beta)}{\partial s}\frac{\partial s}{\partial p_i},$$

which reduces the above condition to

$$-\frac{\widetilde{\Psi}_i^g(s,\beta)}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^2}\frac{\partial \Psi_i^g(s,\beta)}{\partial s}-p_i\frac{\partial \Psi_i^g(s,\beta)}{\partial s}\frac{\partial s}{\partial p_i}\frac{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\left(\widetilde{\Psi}_i^g(s,\beta)-\mu\right)}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^3},$$

where $\frac{\partial s}{\partial p_i}=\frac{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^2}{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\mu}$ and $p_i=\frac{\widetilde{\Psi}_i^g(s,\beta)}{\widetilde{\Psi}_i^g(s,\beta)+\mu}$, giving

$$\widetilde{\Psi}_i^g\left(s,\beta\right)\frac{\partial \Psi_i^g(s,\beta)}{\partial s}\left(\frac{\widetilde{\Psi}_i^g(s,\beta)}{\mu}\right)\geq 0,$$

which always holds. Thus, $\frac{\partial s^g}{\partial \mu}\leq 0$. However, this applies only around the equilibrium choice. To verify that the solution is unique (so that this result applies globally), we need to see whether $\frac{\partial^2 \pi_i}{\partial s^2}\leq 0$. We have that

$$\frac{\partial^2 \pi}{\partial s^2}=\frac{\partial p_i}{\partial s}\frac{\partial \Psi_i^g(s,\beta)}{\partial s}+p_i\frac{\partial^2 \Psi_i^g(s,\beta)}{\partial s^2}+\left(\frac{\partial \Psi_i^g(s,\beta)}{\partial s}-\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right)\frac{\partial p_i}{\partial s}+\left(\Psi_i^g\left(s,\beta\right)-\widetilde{\Psi}_i^g\left(s,\beta\right)\right)\frac{\partial^2 p_i}{\partial s^2}.$$

Generally, the solution need not be unique because $\frac{\partial^2 \Psi_i^g(s,\beta)}{\partial s^2}$ can be positive, in particular for larger s. As a result, the organization can have two local maxima - one with high quality of primary information but bad relative use of that information and one with low quality of primary information but accurate use of that information. Any change in the equilibrium will occur in the opposite direction since $\frac{\partial^2 \pi}{\partial p_i \partial \mu}\leq 0$. To establish conditions for the uniqueness given the functional forms assumed, we have that

$$\frac{\partial p}{\partial s}=\frac{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\mu}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^2}\leq 0\qquad and \qquad \frac{\partial^2 p}{\partial s^2}=\mu\frac{\left(\frac{\partial^2 \widetilde{\Psi}_i^g(s,\beta)}{\partial s^2}\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)-2\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right)}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^3}\gtreqless 0.^{14}$$

[14] $\frac{\partial^2 p}{\partial s^2}\leq 0$ for μ sufficiently small.

Now, to show uniqueness we use the following logic (as it is clear that generally the second-order condition can be violated): the payoff function is continuously differentiable in s. As a result, if there exists more than one interior maximum, there needs to be at least one interior minimum (or inflection point) to the function. Thus, if we can establish that the second-order condition is satisfied to any solution to the first-order condition (which provides a sufficient restriction to the parameter space), then by implication that solution needs to be unique. Using the condition that

$$\left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta)\right) = -p_i \frac{\partial \Psi_i^g(s,\beta)}{\partial s} \frac{\partial s}{\partial p_i}$$

in equilibrium, we can write the second-order condition as

$$\left(2\frac{\partial \Psi_i^g(s,\beta)}{\partial s} - \frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right) \frac{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\mu}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^2}$$
$$+\frac{\widetilde{\Psi}_i^g(s,\beta)}{\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right)^2}\left(\left(\frac{\partial^2 \Psi_i^g(s,\beta)}{\partial s^2} - \frac{\frac{\partial \Psi_i^g(s,\beta)}{\partial s}}{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}} \frac{\partial^2 \widetilde{\Psi}_i^g(s,\beta)}{\partial s^2}\right)\left(\widetilde{\Psi}_i^g(s,\beta)+\mu\right) + 2\frac{\partial \Psi_i^g(s,\beta)}{\partial s}\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right) < 0.$$

Now, we can establish that for all parameter values and under both governance structures

$$\left(\left(\frac{\partial^2 \Psi_i^g(s,\beta)}{\partial s^2} - \frac{\frac{\partial \Psi_i^g(s,\beta)}{\partial s}}{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}} \frac{\partial^2 \widetilde{\Psi}_i^g(s,\beta)}{\partial s^2}\right)\left(\widetilde{\Psi}_i^g(s,\beta)\right) + 2\frac{\partial \Psi_i^g(s,\beta)}{\partial s}\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right) < 0,$$

so that the above arranges to

$$\mu > \frac{-\widetilde{\Psi}_i^g(s,\beta)\left(\left(\frac{\partial^2 \Psi_i^g(s,\beta)}{\partial s^2} - \frac{\frac{\partial \Psi_i^g(s,\beta)}{\partial s}}{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}} \frac{\partial^2 \widetilde{\Psi}_i^g(s,\beta)}{\partial s^2}\right)\left(\widetilde{\Psi}_i^g(s,\beta)\right) + 2\frac{\partial \Psi_i^g(s,\beta)}{\partial s}\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right)}{\left(\left(\frac{\partial^2 \Psi_i^g(s,\beta)}{\partial s^2} - \frac{\frac{\partial \Psi_i^g(s,\beta)}{\partial s}}{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}} \frac{\partial^2 \widetilde{\Psi}_i^g(s,\beta)}{\partial s^2}\right)\widetilde{\Psi}_i^g(s,\beta) + \left(2\frac{\partial \Psi_i^g(s,\beta)}{\partial s} - \frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right)\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right)},$$

where the RHS is a negative number. Thus, we have established that the second-order condition is strictly negative for any solution to the first-order condition, which implies through continuity that there can be only one solution. Therefore, the solution to the first-order condition is unique. In other words, our assumed cost of information function provides sufficient convexity to the information acquisition problem to overcome the convexity in the returns to alignment part to maintain uniqueness of the solution.

Next, we need to consider how s^g depends on β. For β, we have that

$$\frac{\partial^2 \pi}{\partial s \partial \beta} = \frac{\partial p_i}{\partial \beta}\frac{\partial \Psi_i^g(s,\beta)}{\partial s} + p_i \frac{\partial^2 \Psi_i^g(s,\beta)}{\partial s \partial \beta} + \left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta)\right)\frac{\partial^2 p_i}{\partial s \partial \beta} + \left(\frac{\partial \Psi_i^g(s,\beta)}{\partial \beta} - \frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial \beta}\right)\frac{\partial p_i}{\partial s} \gtreqless 0.$$

To build intuition for the relationship, the first component, $\frac{\partial p_i}{\partial \beta}\frac{\partial \Psi_i^g(s,\beta)}{\partial s}$, measures how much the value of profit-sharing is reduced because the equilibrium quality of information decreases: $\frac{\partial p_i}{\partial \beta} < 0$. The second component, $p_i \frac{\partial^2 \Psi_i^g(s,\beta)}{\partial s \partial \beta}$, measures how the value of profit-sharing itself is impacted because of the change in β. The third component, $\left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta)\right) \frac{\partial^2 p_i}{\partial s \partial \beta}$, measures how the agent's sensitivity to s in terms of the level of information acquisition changes with β, scaled by the current value of additional information acquisition. Finally, $\left(\frac{\partial \Psi_i^g(s,\beta)}{\partial \beta} - \frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial \beta}\right) \frac{\partial p_i}{\partial s}$ measures the change in the marginal value of further information acquisition. In short, an increase in β reduces the value of information, increases the negative externality and increases the agency conflicts, all supporting an increase in profit-sharing. However, the equilibrium quality of information also goes down, which reduces the value of additional profit-sharing. In the present setting, the relationship is always positive under centralization but can be negative under decentralization (but only for parameter values for which decentralization is never the preferred governance structure). For other cost functions, the relationship can be locally negative for either governance structure.

A.5.2 Choice of g

Having established the equilibrium choice of s for each governance structure, we can then consider the choice of governance structure. Decentralization is preferred whenever

$$\begin{aligned} &p_i^{dec}(\beta, s^{dec}, \mu)\Psi_i^{dec}\left(s^{dec}, \beta\right) - \mu C(p_i^{dec}(\beta, s^{dec}, \mu)) \\ &\quad \geq p_i^{cent}(\beta, s^{cent}, \mu)\Psi_i^{cent}\left(s^{cent}, \beta\right) - \mu C(p_i^{cent}(\beta, s^{cent}, \mu)). \end{aligned}$$

While the endogeneity of s and p_i makes the comparison harder, the logic behind the choice is straightforward. First, note that if $\beta = 0$, then $s = 1$ under both governance structures and the both achieve the conditional first-best. Second, if $\beta \to \infty$, then the two solutions converge again. Third, as $\mu \to 0$, each governance structure sets $s = 1/2$ and the performance differentials are again equated. For the rest of the parameter space, we can build the result, after noting the monotonicity between s and μ, through a replication argument. Suppose that β is sufficiently small. Then, from section 4 we know that for any given s, decentralization is able to both produce more information and then use that information better. As a result, centralization can never be preferred (let the $s = s^{cent}(\mu)$, the optimal choice under centralization. Then, decentralization dominates centralization by choosing $s^{dec} = s^{cent}(\mu)$ and can do even better by re-optimizing its choice of s). Same logic applies when considering higher β. Suppose first that μ is sufficiently low so that $s^{cent}(\mu)$ falls in the region where decentralization is able to both generate more information and make better use of that information. Again, it is clear that decentralization must dominate. It is thus only when both β and μ are sufficiently high (and so $s^{cent}(\mu)$ sufficiently low) that decentralization cannot outperform centralization.

To establish the result that whenever centralization is preferred, $s^{cent}(\mu) < s^{dec}(\mu)$, note that the implication of the above is that centralization can be preferred only if $\Psi_i^{cent}(s^{cent}(\mu), \beta) > \Psi_i^{dec}(s^{cent}(\mu), \beta)$, because decentralization will always generate more information. Now, consider the first-order condition

$$p_i \frac{\partial \Psi_i^g(s,\beta)}{\partial s} + \left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta) \right) \frac{\partial p_i}{\partial s} = 0$$

and suppose that it is satisfied under centralization. Then, consider the corresponding choice under decentralization. It will be greater than zero for that s (implying that more profit-sharing is desired) if

$$\left(\widetilde{\Psi}_i^{dec}(s,\beta) \right)^2 \frac{\partial \Psi_i^{dec}(s,\beta)}{\partial s} \geq - \left(\widetilde{\Psi}_i^{dec}(s,\beta) \frac{\partial \Psi_i^{dec}(s,\beta)}{\partial s} + \left(\Psi_i^{dec}(s,\beta) - \widetilde{\Psi}_i^{dec}(s,\beta) \right) \frac{\partial \widetilde{\Psi}_i^{dec}(s,\beta)}{\partial s} \right) \mu.$$

Now, $\left(\widetilde{\Psi}_i^{dec}(s,\beta) \frac{\partial \Psi_i^{dec}(s,\beta)}{\partial s} + \left(\Psi_i^{dec}(s,\beta) - \widetilde{\Psi}_i^{dec}(s,\beta) \right) \frac{\partial \widetilde{\Psi}_i^{dec}(s,\beta)}{\partial s} \right)$ is negative for any equilibrium s, so the condition becomes

$$\frac{\left(\widetilde{\Psi}_i^{dec}(s,\beta) \right)^2 \frac{\partial \Psi_i^{dec}(s,\beta)}{\partial s}}{-\left(\widetilde{\Psi}_i^{dec}(s,\beta) \frac{\partial \Psi_i^{dec}(s,\beta)}{\partial s} + \left(\Psi_i^{dec}(s,\beta) - \widetilde{\Psi}_i^{dec}(s,\beta) \right) \frac{\partial \widetilde{\Psi}_i^{dec}(s,\beta)}{\partial s} \right)} \geq \mu.$$

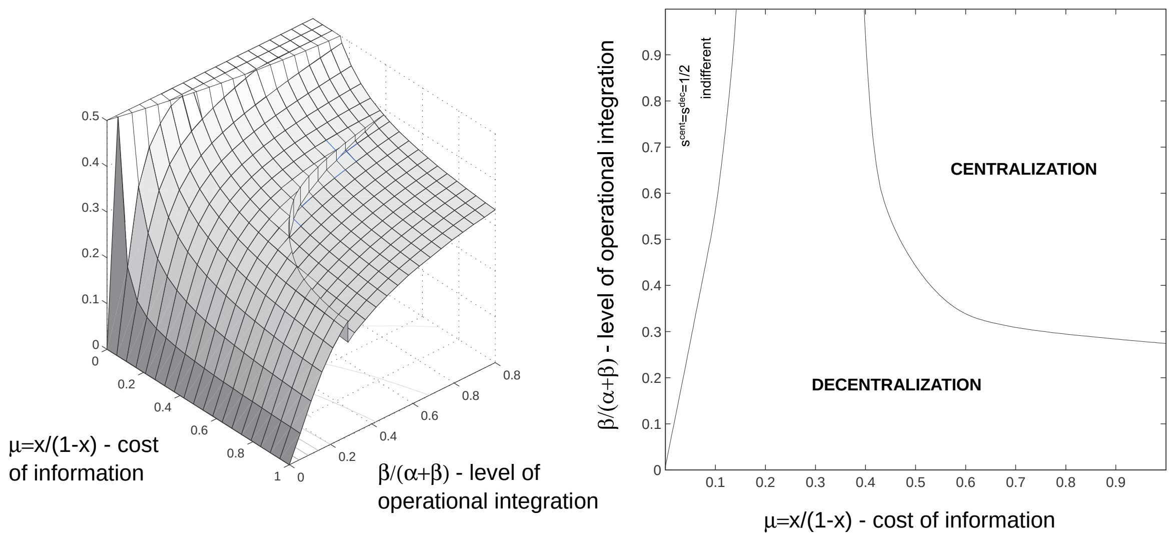
Next, we can back out from the first-order condition under centralization (the s, μ pair is optimal under centralization) that

$$\mu(s^{cent}) = \frac{\left(\widetilde{\Psi}_i^{cent}(s,\beta) \right)^2 \frac{\partial \Psi_i^{cent}(s,\beta)}{\partial s}}{-\left(\widetilde{\Psi}_i^{cent}(s,\beta) \frac{\partial \Psi_i^{cent}(s,\beta)}{\partial s} + \left(\Psi_i^{cent}(s,\beta) - \widetilde{\Psi}_i^{cent}(s,\beta) \right) \frac{\partial \widetilde{\Psi}_i^{cent}(s,\beta)}{\partial s} \right)}.$$

As a result, for a given (β, μ), decentralization benefits from further profit-sharing whenever

$$\frac{-\left(\widetilde{\Psi}_i^{cent}(s,\beta) \frac{\partial \Psi_i^{cent}(s,\beta)}{\partial s} + \left(\Psi_i^{cent}(s,\beta) - \widetilde{\Psi}_i^{cent}(s,\beta) \right) \frac{\partial \widetilde{\Psi}_i^{cent}(s,\beta)}{\partial s} \right)}{\left(\widetilde{\Psi}_i^{cent}(s,\beta) \right)^2 \frac{\partial \Psi_i^{cent}(s,\beta)}{\partial s}} \geq \frac{-\left(\widetilde{\Psi}_i^{dec}(s,\beta) \frac{\partial \Psi_i^{dec}(s,\beta)}{\partial s} + \left(\Psi_i^{dec}(s,\beta) - \widetilde{\Psi}_i^{dec}(s,\beta) \right) \frac{\partial \widetilde{\Psi}_i^{dec}(s,\beta)}{\partial s} \right)}{\left(\widetilde{\Psi}_i^{dec}(s,\beta) \right)^2 \frac{\partial \Psi_i^{dec}(s,\beta)}{\partial s}},$$

which holds for all s such that $\Psi_i^{cent}(s,\beta) \geq \Psi_i^{dec}(s,\beta)$. As a result, at $s^{cent}(\mu)$ for which decentralization cannot immediately outperform centralization by choosing $s = s^{cent}(\mu)$, decentralization chooses to increase its amount of profit-sharing to sacrifice some of the quality of information but increase the accuracy at which that information is used. As a result, we can make the conclusion that whenever centralization is preferred for a given β, $s^{cent}(\mu) < s^{dec}(\mu)$. The logic for the monotonicity in μ is then immediate: as μ decreases, $s^{cent}(\mu)$ decreases. If decentralization simply chooses $s^{cent}(\mu)$, the difference in the value of information (and so in performance) is decreased, and the readjustment of s away from $s^{cent}(\mu)$ becomes more and more likely to be able to beat centralization. The exact solution to the choice of governance structure and profit-sharing rule is illustrated in figure 5. Note



(i) equilibrium level of profit-sharing *(ii) equilibrium choice of governance structure*

Figure 5: Equilibrium choice of s and g

that while graphical, this solution is exact as it solves the model for all μ and β, which are the only variables that impact s and g.

A.5.3 Choice of β

Finally, we can consider the choice of operational integration. For this, we can define

$N(s^g, \beta, g) \;=\; \max_g \left(p_i^g \Psi_i^g\left(s^g, \beta\right) - \mu C(p_i^g)\right)$ as the upper envelope of the information-dependent profits under the two governance structures. Note from above that under each governance structure, $p_i^g \Psi_i^g\left(s^g, \beta\right) - \mu C(p_i^g)$ is monotone decreasing and convex in both β and μ and, given the setting, have one interior intersection. Even if there were multiple intersections, then we must have that at the equal performance boundary $\beta(\mu)$,

$$\tfrac{dN(s^g,\beta,g)}{d\beta}|_{\beta\to\beta(\mu)^-} < \tfrac{dN(s^g,\beta,g)}{d\beta}|_{\beta\to\beta(\mu)^+}$$

because the switch in governance structures occurs only if the loss in profits related to adaptiveness is decreasing at a lower rate under the alternative governance structure. Further, because of this kink, the optimal choice of β will never exist at this point and it is this kink that causes the discrete change in the level of operational integration whenever there

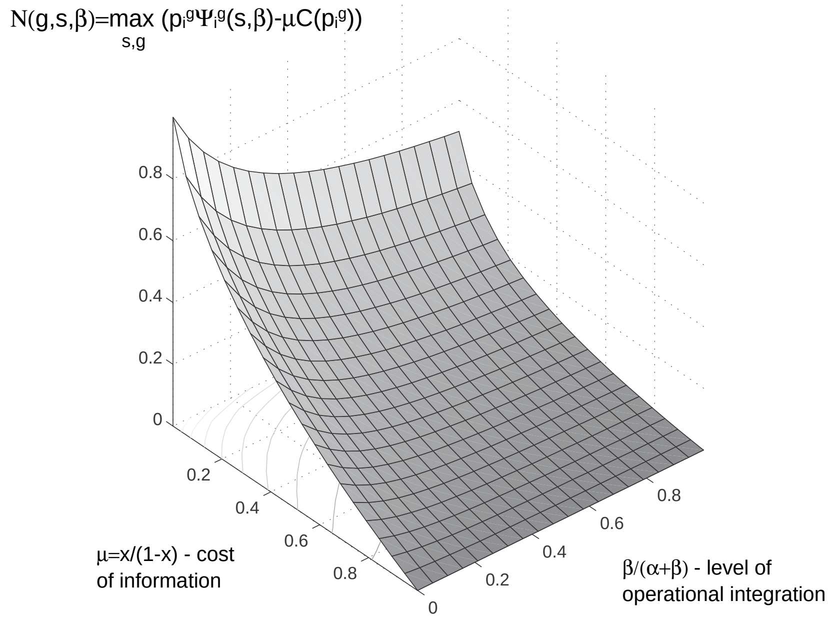


Figure 6: $\max_{g}(p_i^g \Psi_i^g (s^g, \beta) - \mu C(p_i^g))$

is a change in governance structure. Instead, the optimal choice of β is characterized by the first-order condition

$$K'(\beta) + \frac{\partial N(s^g,\beta,g)}{\partial\beta}\sigma_\theta^2 = 0,$$

where the indirect effect through $\frac{\partial s}{\partial \beta}$ is zero due to the envelope theorem. The surface $N(s^g, \beta, g)$ is illustrated in figure 6. Note again that it is the solution for all β, μ given the cost of information function.

To establish the remaining comparative statics, we have that

$$\frac{d\pi}{d\beta} = K'(\beta) + \frac{\partial\left(p_i^g\Psi_i^g(s^g,\beta)-\mu C(p_i^g)\right)}{\partial\beta}\sigma_\theta^2 = K'(\beta) - p_i^g\left(\frac{\partial\Psi_i^g(s,\beta)}{\partial s}\frac{\partial s}{\partial\beta} - \frac{\partial\Psi_i^g(s^g,\beta)}{\partial\beta}\right)\sigma_\theta^2,$$

using the optimality of the agent's choice of effort and the principal's choice of s that follows the choice of β. The submodularity with respect to σ_θ^2 is immediate since

$$\frac{d^2\pi}{d\beta d\sigma_\theta^2} = -p_i^g\left(\frac{\partial\Psi_i^g(s,\beta)}{\partial s}\frac{\partial s}{\partial p_i}\frac{\partial p_i}{\partial\beta} - \frac{\partial\Psi_i^g(s^g,\beta)}{\partial\beta}\right) < 0.$$

Since σ_θ^2 is simply a scaling of the marginal cost surface, the submodularity holds trivially. In the case of multiple equilibria, which are possible due to the convexity of $N(s^g, \beta, g)$, the jump is always towards the lower β. As mentioned above, whenever there is a switch in governance structure due to an increase in σ_θ^2, the kink in the surface causes a discrete drop in the level of operational integration. The comparison with respect to μ is a little more involved because it has a direct impact on s as well which works in the opposite direction ($\frac{\partial s}{\partial \mu} < 0$). However, given the uniqueness of s we can use the implicit function theorem to evaluate $\frac{\partial s}{\partial \mu}$ and establish that $\frac{d^2\pi}{d\beta d\mu} > 0$, with a switch in governance structure then associated with a discrete increase in β.

B Supplemental derivations [not for publication]

B.1 Missing algebra for proposition 3 and 4

(i) Losses due to biased decisions under decentralization

From the expression for profits as derived in the proof for proposition 3, we can write the loss due to biased decisions under decentralization as

$$\Delta\Lambda_i^{dec}(s,\beta) = \left(\frac{\alpha\beta\left(\beta+(1-s)^2\alpha\right)}{((1-s)\alpha+2\beta)^2} - \frac{\alpha\beta}{(\alpha+4\beta)}\right) = \frac{(\alpha\beta)^2(1-2s)^2}{(\alpha+4\beta)((1-s)\alpha+2\beta)^2}.$$

From this, we have that $\frac{\partial\Delta\Lambda_i^{dec}(s,\beta)}{\partial s} = -\frac{2(\alpha\beta)^2(1-2s)}{(\alpha+4\beta)((1-s)\alpha+2\beta)^3} \leq 0$ (so that increased profit-sharing always reduces the loss) and

$$\frac{\partial\Delta\Lambda_i^{dec}(s,\beta)}{\partial\beta} = \frac{2\alpha^2\beta(1-2s)^2\left(2(1-s)\alpha\beta+(1-s)\alpha^2-4\beta^2\right)}{((1-s)\alpha+2\beta)^3(\alpha+4\beta)^2},$$

so that $\frac{\partial\Delta\Lambda_i^{dec}(s,\beta)}{\partial\beta} \geq 0$ for $\beta \leq \overline{\beta}(s,\alpha)$ and $\frac{\partial\Delta\Lambda_i^{dec}(s,\beta)}{\partial\beta} \leq 0$ for $\beta \geq \overline{\beta}(s,\alpha)$, where

$$\overline{\beta}(s,\alpha) = \frac{2(1-s)\alpha+\sqrt{4(1-s)^2\alpha^2+16(1-s)\alpha^2}}{8} = \alpha\left(\frac{(1-s)+\sqrt{(1-s)^2+4(1-s)}}{4}\right).$$

In other words, in terms of the level of operational integration, the loss is maximized at an intermediate level of β for any s. Note that we can also write $\overline{\beta}(s,\alpha)$ as being given by

$$\frac{\overline{\beta}(s,\alpha)}{\alpha+\overline{\beta}(s,\alpha)} = \frac{(1-s)+\sqrt{(1-s)^2+4(1-s)}}{4+(1-s)+\sqrt{(1-s)^2+4(1-s)}} \in \left[\frac{1}{3}, \frac{1+\sqrt{5}}{5+\sqrt{5}}\right].$$

It is also worth noting that the loss due to biased decisions is only second-order whenever $\beta \to 0, \beta \to \infty$ and $s \to 1/2$.

(ii) Losses due to strategic communication under centralization and decentralization

From propositions 2 and 3 we have that the accuracy and the value of accurate communication under centralization is given by

$$V_i^{cent}(s,\alpha,\beta) = \frac{3(1-s)(2\beta+\alpha)}{\beta(1-2s)+3(1-s)(2\beta+\alpha)} \text{ and } \Gamma_i^{cent}(s,\alpha,\beta) = \alpha - \frac{2\alpha\beta}{\alpha+4\beta} = \frac{\alpha(\alpha+2\beta)}{\alpha+4\beta},$$

so we have that :

$$\frac{\partial V_i^{cent}(s,\beta)}{\partial s} = \frac{3\beta(2\beta+\alpha)}{(\beta(2s-1)+3s(2\beta+\alpha))^2} \geq 0 \qquad and \qquad \frac{\partial V_i^{cent}(s,\beta)}{\partial \beta} = -\frac{3(1-s)\alpha(1-2s)}{(\beta(2s-1)+3s(2\beta+\alpha))^2} \leq 0$$
$$\frac{\partial\left(\Gamma_i^{cent}(s,\alpha,\beta)\right)}{\partial s} = 0 \qquad and \qquad \frac{\partial\left(\Gamma_i^{cent}(s,\alpha,\beta)\right)}{\partial \beta} = -\frac{2\alpha^2}{(\alpha+4\beta)^2} < 0,$$

and bringing the two together (the loss due to strategic communication), we get

$$\Gamma_i^{cent}(s,\alpha,\beta)\left(1-V_i^{cent}(s,\alpha,\beta)\right) = \frac{\alpha(\alpha+2\beta)}{\alpha+4\beta}\left(\frac{\beta(1-2s)}{\beta(1-2s)+3(1-s)(2\beta+\alpha)}\right),$$

giving

$$\frac{\partial\left(\Gamma_i^{cent}(s,\alpha,\beta)\left(1-V_i^{cent}(s,\alpha,\beta)\right)\right)}{\partial s} = -\frac{3\alpha\beta(2\beta+\alpha)^2}{(\alpha+4\beta)(\beta(1-2s)+3(1-s)(2\beta+\alpha))^2} \leq 0 \text{ and}$$

$$\frac{\partial\left(\Gamma_i^{cent}(s,\alpha,\beta)\right)\left(1-V_i^{cent}(s,\alpha,\beta)\right)}{\partial \beta} = \frac{\alpha^2(1-2s)\left(2\beta^2(5-4s)+3(1-s)\alpha(4\beta+\alpha)\right)}{(\alpha+4\beta)^2(\beta(1-2s)+3(1-s)(2\beta+\alpha))^2} \geq 0,$$

so that the loss due to strategic communication is monotone decreasing in the degree of profit sharing and monotone increasing in the degree of operational integration. For decentralization, we have from propositions 2 and 3 that

$$V_i^{dec}(s,\beta) = \frac{3(1-s)((1-s)\alpha+2\beta)}{((1-s)\alpha+\beta)(1-2s)+3(1-s)((1-s)\alpha+2\beta)} \text{ and } \Gamma_i^{dec}(s,\alpha,\beta) = \frac{\alpha\beta^2\left(2\beta^2+(1-s)^2\left(3\alpha^2-4\alpha^2 s+6\alpha\beta\right)\right)}{((1-s)\alpha+\beta)^2((1-s)\alpha+2\beta)^2}.$$

Thus, we have that

$$\frac{\partial V_i^{dec}(s,\beta)}{\partial s} = \frac{3(\alpha^2(1-s)^2+2\alpha\beta(1-s)(2-s)+2\beta^2)}{((1-s)\alpha+\beta)(1-2s)+3(1-s)((1-s)\alpha+2\beta)^2} \geq 0$$
$$\frac{\partial V_i^{dec}(s,\beta)}{\partial \beta} = \frac{3(1-s)^2\alpha(1-2s)}{((1-s)\alpha+\beta)(1-2s)+3(1-s)((1-s)\alpha+2\beta)^2} \geq 0$$

$$\frac{\partial \Gamma_i^{dec}(s,\alpha,\beta)}{\partial s} = \frac{2\alpha^2\beta^2(1-2s)\left((1-s)^3\alpha^3-6\beta^2((1-s)\alpha+\beta)\right)}{((1-s)\alpha+\beta)^3((1-s)\alpha+2\beta)^3} \gtreqless 0$$
$$\frac{\partial \Gamma_i^{dec}(s,\alpha,\beta)}{\partial \beta} = \frac{2\alpha^2\beta(1-s)\left((1-s)^3\alpha^2(\alpha(3-4s)+9\beta)+s\beta^2(6(\beta+\alpha)+\alpha(1-s))\right)}{(s\alpha+\beta)^3(s\alpha+2\beta)^3} \geq 0$$

and bringing the two together, we can verify that

$$\frac{\partial\left(\Gamma_i^{dec}(s,\alpha,\beta)\left(1-V_i^{dec}(s,\alpha,\beta)\right)\right)}{\partial s} \leq 0 \text{ and } \frac{\partial\left(\Gamma_i^{dec}(s,\alpha,\beta)\right)\left(1-V_i^{dec}(s,\alpha,\beta)\right)}{\partial \beta} \gtreqless 0,$$

with $\frac{\partial}{\partial \beta} \geq 0 \ \forall \beta < \overline{\beta}(s)$, as given by figure 7. Finally, to compare the two, we have that

$$\Gamma_i^{cent}(s,\alpha,\beta)\left(1-V_i^{cent}(s,\alpha,\beta)\right) - \Gamma_i^{dec}(s,\alpha,\beta)\left(1-V_i^{dec}(s,\alpha,\beta)\right) \geq 0$$

as long as $\beta < \widetilde{\beta}(s)$, as illustrated by figure 7.

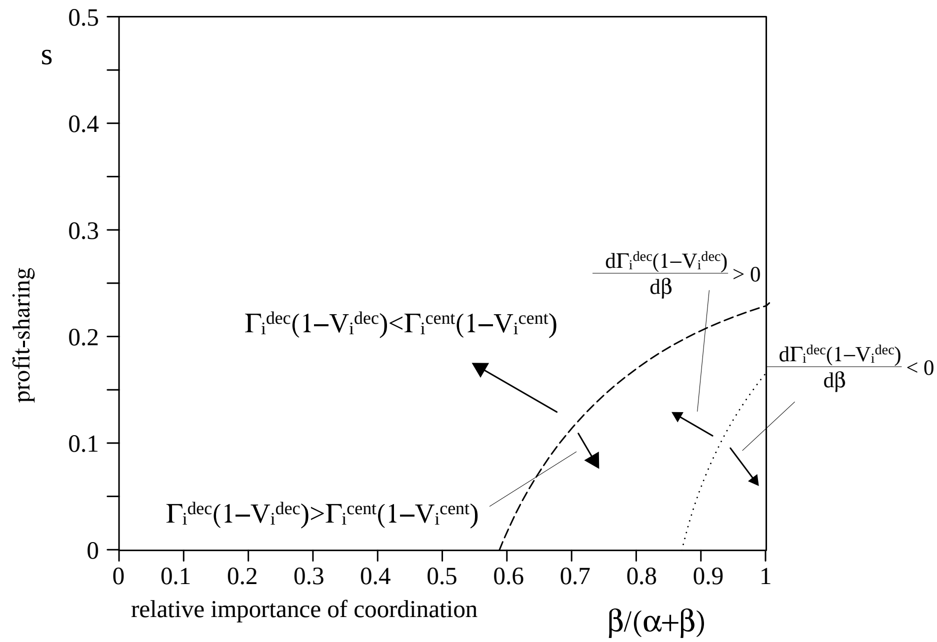


Figure 7: Comparative statics - losses due to strategic communication

(iii) the value of information

From A.3, we have that (at some abuse of notation (but to cut back some) I will use s for $(1-s)$, the weight the division manager places on this own division, and refer to it as interdivisional conflict))

$$\Psi_i^{cent}(s,\beta) = \frac{3\alpha s(\alpha+2\beta)^2}{(\alpha+4\beta)(\beta(2s-1)+3s(2\beta+\alpha))}$$
$$\Psi_i^{dec}(s,\beta) = \frac{\alpha s\left(5s^3\alpha^3-10s^3\alpha^2\beta-s^2\alpha^3+25\beta s^2\alpha^2-14s^2\alpha\beta^2-4s\beta\alpha^2+30\alpha\beta^2 s-4\beta^2\alpha+6\beta^3\right)}{(s\alpha+2\beta)(s\alpha+\beta)(5s^2\alpha-s\alpha+8\beta s-\beta)},$$

from which it is then straightforward but cumbersome to verify that $\frac{\partial\Psi_i^g(s,\beta)}{\partial s} \leq 0$ and $\frac{\partial\Psi_i^g(s,\beta)}{\partial\beta} \lesssim 0$ for both governance structures:

$$\frac{\partial\Psi_i^{cent}(s,\beta)}{\partial s} = -\frac{3\beta\alpha(2\beta+\alpha)^2}{(\alpha+4\beta)(8\beta s-\beta+3s\alpha)^2} \leq 0$$

$$\frac{\partial\Psi_i^{cent}(s,\beta)}{\partial\beta} = -\frac{3s\alpha^2(2\beta+\alpha)(8s\alpha-\alpha+24\beta s-6\beta)}{(\alpha+4\beta)^2(8\beta s-\beta+3s\alpha)^2} \leq 0$$

$$\frac{\partial\Psi_i^{dec}(s,\beta)}{\partial s} = -\frac{\alpha\beta^2\left(\alpha^4 s^3 c_1+\alpha^3\beta s^2 c_2+\alpha^2\beta^2 s c_3+\alpha\beta^3 c_4+12\beta^4\right)}{(5s^2\alpha-s\alpha+8\beta s-\beta)^2(s\alpha+2\beta)^2(s\alpha+\beta)^2} \leq 0,$$

where $c_1 = (160s^3 - 120s^2 + 27s - 2)$, $c_2 = (680s^3 - 516s^2 + 126s + 10)$, $c_3 = (956s^3 - 744s^2 + 204s - 16)$ and $c_4 = (448s^3 - 360s^2 + 120s - 8)$, all greater than zero for $s \in [0.5, 1]$ and

$$\frac{\partial\Psi_i^{dec}(s,\beta)}{\partial\beta} = -\frac{s\alpha^2\left(\alpha^2\beta^2s^2d_1+\beta^4d_2+\alpha\beta^3sd_3+\alpha^3\beta s^3d_4+\alpha^4s^5d_5\right)}{(5s^2\alpha-s\alpha+8\beta s-\beta)^2(s\alpha+2\beta)^2(s\alpha+\beta)^2} \precsim 0,$$

where $d_1 = (10 - 109s + 304s^2 - 18s^3)$, $d_2 = (8 - 94s + 304s^2 - 224s^3)$, $d_3 = (16 - 180s + 564s^2 - 320s^3)$, $d_4 = (2 - 18s + 12s^2 + 140s^3)$ and $d_5 = (2 - 20s + 50s^2)$. All are greater than zero for $s \in [0.5, 1]$, except d_2, which is smaller than zero for $s \lesssim 0.96$. For $s \gtrsim 0.96$, there then exists a β suffciently large so that for all higher levels of operational integration, $\frac{\partial\Psi_i^{dec}(s,\beta)}{\partial\beta} \geq 0$. In this case, the improvement in communication and decision-making that follows an increase in integration outweighs the reduction in the base-value of information. For $s = 1$, this boundary is approximately $\frac{\beta}{\alpha+\beta} \approx 0.94$. For the analysis, this region is irrelevant, however, because in this region, if this would be the optimal structure under decentralization, centralization would always outperform decentralization. Indeed, decentralization could *improve* its performance through further integration so this β cannot be an equilibrium choice. As a result, this exception is irrelevant and not pointed out in the proposition.

As to the differences between the two governance structures, we can write the difference, $\Psi_i^{cent}(s,\beta) - \Psi_i^{dec}(s,\beta)$ as

$$\frac{s\alpha^2\beta(2s-1)(50s^3\alpha^2\beta-5s^3\alpha^3+160s^3\alpha\beta^2-62s^2\alpha\beta^2+s^2\alpha^3-51\beta s^2\alpha^2+224\beta^3s^2-192\beta^3s+7s\beta\alpha^2-44\alpha\beta^2s+10\beta^2\alpha+34\beta^3)}{(\alpha+4\beta)(8\beta s-\beta+3s\alpha)(5s^2\alpha-s\alpha+8\beta s-\beta)(s\alpha+2\beta)(s\alpha+\beta)}.$$

Now, it is immediate to verify that $\Psi_i^{cent}(s,\beta) - \Psi_i^{dec}(s,\beta) \to 0$ whenever $s \to \frac{1}{2}$, $\beta \to 0$ and $\beta \to \infty$. Further, $\Psi_i^{cent}(s,\beta) \geq \Psi_i^{dec}(s,\beta)$ if and only if

$$50s^3\alpha^2\beta - 5s^3\alpha^3 + 160s^3\alpha\beta^2 - 62s^2\alpha\beta^2 + s^2\alpha^3 - 51\beta s^2\alpha^2 + 224\beta^3s^2 - 192\beta^3s + 7s\beta\alpha^2 - 44\alpha\beta^2s + 10\beta^2\alpha + 34\beta^3 \geq 0$$

Because the other components are weakly positive for all possible parameter values. To examine this condition in more detail, let $r = \frac{\beta}{\alpha+\beta}$, which allows us to write the above as

$$-54rs^2 + 234r^3s^2 + 10r^2 - 141r^3s + 45s^3r^2 + 65s^3r + 43s^2r^2 + s^2 - 58r^2s + 7sr - 5s^3 - 105s^3r^3 + 24r^3 \geq 0.$$

From this, we can establish that $\Psi_i^{cent}(s,\beta) \geq \Psi_i^{dec}(s,\beta)$ if and only if $\beta \geq \underline{\beta}(s)$, where $\frac{d\underline{\beta}(s)}{ds} < 0$ - the higher the degree of interdivisional conflict, the lower the level of operational integration after which the value of information is higher under centralization. Let $\Sigma(r,s)$ denote the above expression. Then, we can make the following observations:

(i) $\Sigma(r=0,s) = s^2(1-5s) < 0$ and $\Sigma\left(r, s=\frac{1}{2}\right) = -\frac{1}{8}(15r + 9r^3 + 21r^2 + 3) < 0$.

(ii) $\Sigma_r(r,s) > 0 \ \forall r \geq \underline{r}(s)$ and $s \gtrsim 0.56$. $\Sigma_r(r,s) < 0 \ \forall r$ if $s \precsim 0.56$.

(iii) $\Sigma_s(r,s) > 0 \ \forall s \geq \underline{s}(r)$ and $r \gtrsim 0.074$. $\Sigma_s(r,s) < 0 \ \forall r \precsim 0.074$.

To establish the second, result, we can write

$$\Sigma_r(r,s) = 9r^2\left(8 - 35s^3 + 78s^2 - 47s\right) + 2r\left(10 + 45s^3 + 43s^2 - 48s\right) + \left(65s^3 - 54s^2 + 7s\right),$$

so that $\Sigma_r(r,s) \geq 0$ if

$$r \geq \underline{r}(s) = \frac{1}{9}\frac{43s^2+10+45s^3-58s-\sqrt{100-1664s+11073s^2-36524s^3+64237s^4-58770s^5+22500s^6}}{(35s^3-78s^2+47s-8)}$$

and $(8 - 35s^3 + 78s^2 - 47s)$ is positive, or $s \gtrsim 0.56$. For $s \lesssim 0.56$, $\Sigma_r(r,s) < 0$. Further, for $s \gtrsim 0.67$, $\underline{r}<0$ so that $\Sigma_r(r,s) > 0 \ \forall r \in [0,1]$.

To establish the third result, we can write

$$\Sigma_s(r,s) = 15s^2\left(9r^2 + 13r - 1 - 21r^3\right) - 2s\left(234r^3 + 43r^2 + 59r + 1\right) - \left(141r^3 + 58r^2 - 7r\right),$$

so that $\Sigma_s(r,s) > 0$ if

$$s \geq \underline{s}(r) = \frac{1}{15}\frac{234r^3-54r+43r^2+1-\sqrt{10341r^6+14107r^4+20889r^5+4074r^3+767r^2-3r+1}}{-9r^2-13r+1+21r^3}$$

and $(9r^2 + 13r - 1 - 21r^3) \geq 0$, or $r \gtrsim 0.074$. For $r \lesssim 0.074$, $\Sigma_s(r,s) \leq 0$, while for $r \gtrsim 0.31$, $\underline{s}(r) < 0.5$ so that $\Sigma_s(r,s) > 0 \ \forall s \in [0.5,1]$.

From results (i)-(iii) we can conclude that decentralization has a higher value of information whenever $s \lesssim 0.56$ or $r \precsim 0.074$ (the limit is negative and marginals are negative, so the function can never become positive). Results (ii) and (iii) establish the converse. First, a necessary requirement for centralization to have a positive value of information, we need to have that either $\Sigma_r(r,s) > 0$ or $\Sigma_s(r,s) > 0$ for a range of parameters to create the possibility that $\Sigma(r,s) > 0$. Further, the results state that if $\Sigma(r,s)$ is ever increasing r, then it is also decreasing for all $\widetilde{r} \geq r$ and similarly for s. Thus, there is at most one interior solution r^* to $\Sigma(r^*,s) = 0$ for any s and vice versa. Further, if $\Sigma(r^*,s) = 0$, then $\Sigma(r,s) > 0 \ \forall r > r^*$ and similarly for s : if $\Sigma(r,s^*) = 0$, then $\Sigma(r,s) > 0 \ \forall s > s^*$. We have thus established the existence of a unique boundary $\widehat{r}(s)$, where centralization is preferred if and only if $r > \widehat{r}(s)$.

To establish that $\frac{d\widehat{r}(s)}{ds} < 0$ (as the degree of interdivisional conflict increases, centralization becomes preferred for a wider range of operational integration) is a little trickier. To begin, note that, using the implicit function theorem, $\frac{d\widehat{r}(s)}{ds} < 0$ if and only if $-\frac{dF/ds}{dF/dr} < 0$, where F is the implicit function $\Sigma\left(\widehat{r}\left(s\right), s\right) = 0$. From above we already know that at the solution, $\Sigma_r\left(r, s\right) > 0$. As a result, $\frac{d\widehat{r}(s)}{ds} < 0$ if $\Sigma_s\left(r, s\right) > 0$. Now, start at $s = 1$. $\Sigma\left(\widehat{r}\left(s\right), s\right) = 0$ simplifies to $18r + 12r^3 + 40r^2 - 4 = 0$ or $r \approx 0.162$. Now, recall that $\Sigma_s\left(r, s\right) > 0$ if $s \geq \underline{s}\left(r\right)$ and note that $\frac{d\underline{s}(r)}{dr} < 0$. In other words, the range of s for which $\Sigma_s\left(r, s\right) > 0$ is positive is increasing in r. For $r = 0.162$, $\underline{s}\left(r\right) \approx 0.676$. Thus, $\frac{d\widehat{r}(s)}{ds} < 0$ at least for $s \in [0.676, 1]$. Now, let $s = 0.7$, which is strictly interior to the range. For this, we obtain that $\widehat{r}\left(s = 0.7\right) \approx 0.36$. And from $\underline{s}\left(r\right)$ we have that for $r \gtrsim 0.31$, $\underline{s}\left(r\right) < 0.5$ so that $\Sigma_s\left(r, s\right) > 0$ $\forall s \in [0.5, 1]$. As result, whenever $\Sigma\left(\widehat{r}\left(s\right), s\right) = 0$, $\Sigma_s\left(r, s\right) > 0$ and so $\frac{d\widehat{r}(s)}{ds} < 0$. At the limit, letting $r = 1$ we find that value of information is higher under centralization whenever $224s^2 + 34 - 192s \geq 0$ or $s \geq 17/28$.

(iv) perceived value of information

The perceived value, in turn, can be written as (continuing to use s for $(1 - s)$), for centralization

$$\frac{3s\alpha(6\alpha\beta s+s\alpha^2+8\beta^2 s-2\beta^2-\alpha\beta)}{(\alpha+4\beta)(8\beta s-\beta+3s\alpha)},$$

so that it is immediate that

$$\frac{\partial\widetilde{\Psi}_i^{cent}(s,\beta)}{\partial s} = \frac{3\alpha(2\beta+\alpha)(3s^2\alpha^2+\beta^2+2s\alpha\beta(10s-1)+8\beta^2 s(4s-1))}{(\alpha+4\beta)(8\beta s-\beta+3s\alpha)^2} > 0$$

$$\frac{\partial\widetilde{\Psi}_i^{cent}(s,\beta)}{\partial\beta} = -\frac{6s\alpha^2(4\beta^2 s(4s-1)+2\alpha\beta s+8\beta s^2\alpha+s^2\alpha^2+s\alpha^2+\beta^2)}{(\alpha+4\beta)^2(8\beta s-\beta+3s\alpha)^2} < 0.$$

Under decentralization, the perceived value of information is given by

$$\frac{\alpha s(5s^3\alpha^2-s^2\alpha^2+13\beta s^2\alpha-2\alpha\beta s+12\beta^2 s-3\beta^2)}{(5s^2\alpha-s\alpha+8\beta s-\beta)(s\alpha+2\beta)},$$

so we get $\frac{\partial\widetilde{\Psi}_i^{dec}(s,\beta)}{\partial s} =$
$\frac{\alpha\left(5\alpha^4 s^5(5s-2)+6\alpha^3\beta s^4(30s-11)+s^4\alpha^4+6\alpha^2\beta^2 s^3(69s-20)+6s^3\alpha^3\beta+4\alpha\beta^3 s^2(104s-23)+9s^2\alpha^2\beta^2+8s\alpha\beta^3+6\beta^4\left(32s^2-8s+1\right)\right)}{(5s^2\alpha-s\alpha+8\beta s-\beta)^2(s\alpha+2\beta)^2} >$
0

$$\frac{\partial\widetilde{\Psi}_i^{dec}(s,\beta)}{\partial\beta} = -\frac{\alpha^2 s^2\left(5\alpha^2 s^3(5s-2)+\alpha\beta s\left(40s^2+2s-2\right)+s^2\alpha^2+\beta^2\left(32s-8s^2-5\right)\right)}{(5s^2\alpha-s\alpha+8\beta s-\beta)^2(s\alpha+2\beta)^2} < 0.$$

For the difference, let us first establish the claim that the negative externality is always larger under decentralization:

$$\phi_j^{dec} - \phi_j^{cent} = \frac{\alpha^2\beta^2 s\left(\alpha^2 s^2\left(\overset{\geq 0}{70s^2-67s+16}\right)+\alpha\beta s\left(\overset{\geq 0}{160s^3-50s^2-61s+23}\right)+\beta^2\left(\overset{\geq 0}{224s^3-200s^2+38s+3}\right)\right)}{(8\beta s-\beta+3s\alpha)(\alpha+4\beta)(5s^2\alpha-s\alpha+8\beta s-\beta)(s\alpha+\beta)(s\alpha+2\beta)} \geq 0.$$

The perceived value, in turn, can be written as

$$\widetilde{\Psi}_i^{dec}(s,\beta) - \widetilde{\Psi}_i^{cent}(s,\beta) = \frac{s\beta\alpha^2(2s-1)(5s^3\alpha^2+20\beta s^3\alpha-s^2\alpha^2-21\beta s^2\alpha-8\beta^2 s^2+10\alpha\beta s+4\beta^2 s+3\beta^2)}{(8\beta s-\beta+3s\alpha)(\alpha+4\beta)(5s^2\alpha-s\alpha+8\beta s-\beta)(s\alpha+2\beta)}.$$

The perceived value of information is thus higher under decentralization whenever

$$5s^3\alpha^2 + 20\beta s^3\alpha - s^2\alpha^2 - 21\beta s^2\alpha - 8\beta^2 s^2 + 10\alpha\beta s + 4\beta^2 s + 3\beta^2 \geq 0,$$

which we can write, using the notation $r = \frac{\beta}{\alpha+\beta}$ as

$$-r^2\left(15s^3 - 12s^2 + 6s - 3\right) + rs\left(10s^2 - 19s + 10\right) + s^2(5s-1) \geq 0.$$

Thus, the perceived value of information is higher under decentralization as long as

$$r \leq \overline{r}(s) = \frac{s}{6}\left(\frac{10s^2-19s+10+\sqrt{400s^4+729s^2-680s^3+112-464s}}{-4s^2+5s^3+2s-1}\right).$$

For $s = 1$, $\overline{r}(s) = \frac{1}{12}\left(1+\sqrt{97}\right) \approx 0.9$ and $\overline{r}(s') = 1$ for $s' = \frac{1}{4}\left(1+\sqrt{7}\right) \approx 0.91$. The perceived value of information is then higher under decentralization for all r when $s \leq s'$ and for $r < \overline{r}(s)$ for $s \in [s', 1]$. Otherwise, the perceived value of information is higher under centralization.

C Extensions [not for publication]

This appendix examines the robustness of the results to various assumptions. Section C.1 considers the solution under a general cost function $C(p_i, \sigma_\theta^2)$. Section C.2 considers the introduction of endogenous strength of incentives, where not only the composition but also the level of incentives can be endogenous. It also considers a natural alternative where, instead of maximizing profits net of the cost of effort, the headquarters actually maximizes the profits net of the share of profits given to the division managers (when the value cannot be extracted from the division managers through fixed transfers). Section C.3 is a collection of shorter observations on various commonly raised questions, such as what is the impact of assuming observable information acquisition, what would occur if the coordination benefits could be separated from the profit functions, what is lost by the assumed symmetry of the solution and so forth.

The bottom line is that not much is lost because of any of the assumptions. The reason is that all the "economic" action in the model comes from the differences in $p_i^g(s,\beta)$ and $\Psi_i^g(s,\beta)$. As long as the profit functions are quadratic and the incentives are linear, so that the equilibrium decisions are linear functions of the beliefs of the decision-maker(s), then any further changes will not qualitatively change the determination of these two components. There is an impact, of course, but it is only quantitative. In other words, where centralization was once marginally preferred, decentralization can now become the preferred alternative and vice versa. But the big picture remains unchanged. It all boils down to how the compensation and governance structure together can limit losses due to strategic communication and biased decision-making, and how they strike the balance between motivating information acquisition and then the appropriate use of that information.

C.1 General cost functions $C(p_i, \sigma_\theta^2)$

The analysis assumed that the cost of information was given by $C(p_i, \sigma_\theta^2) = -\mu\sigma_\theta^2(p_i + \ln(1 - p_i))$. This subsection analyses the solution for $C(p_i, \sigma_\theta^2) = \mu\sigma_\theta^2 C(p_i)$. Note that for section 4, the only change generated by this is that the manager's information acquisition problem is now a solution to the first-order condition

$$\widetilde{\Psi}_i^g(s,\beta) = \mu C'(p_i).$$

Given that $C''(p_i) > 0$, this solution is unique, and it is immediate that the comparative statics of $\frac{\partial p_i}{\partial s}$, $\frac{\partial p_i}{\partial \beta}$ and $\frac{\partial p_i}{\partial \mu}$, as the difference between the two governance structures, continue to hold because with the exception of μ, they are driven by $\widetilde{\Psi}_i^g(s,\beta)$. As to section 5, the general logic of the results naturally continues to hold but some monotonicity conclusions can fail to hold locally and the uniqueness of the equilibrium with respect to s can also fail. The first-order condition for the choice of s continues to be given by

$$p_i \frac{\partial \Psi_i^g(s,\beta)}{\partial s} = -\left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta)\right)\frac{\partial p_i}{\partial s}.$$

The only difference is that now p_i and $\frac{\partial p_i}{\partial s}$ are determined implicitly by the above first-order condition. To revisit some of the results, consider first the impact of μ on s. Here, we have that

$$\frac{d}{d\mu}\left(\frac{\partial \pi}{\partial \mu}\right) = \frac{\partial p_i}{\partial \mu}\frac{\partial \Psi_i^g(s,\beta)}{\partial s} + \left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta)\right)\frac{d}{d\mu}\left(\frac{\partial p_i}{\partial s}\right),$$

while from the optimality of s we get (so we are again establishing the comparative static only around an equilibrium, not globally)

$$-p_i \frac{\partial \Psi_i^g(s,\beta)}{\partial s} \frac{\partial s}{\partial p_i} = \left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta) \right),$$

so that the above becomes

$$\frac{\partial p_i}{\partial \mu} \frac{\partial \Psi_i^g(s,\beta)}{\partial s} - p_i \frac{\partial \Psi_i^g(s,\beta)}{\partial s} \frac{\partial s}{\partial p_i} \frac{d}{d\mu} \left(\frac{\partial p_i}{\partial s} \right).$$

Now, $\frac{\partial p_i}{\partial s} = \frac{\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s}}{\mu C''(p_i)}$, $\frac{\partial p_i}{\partial \mu} = -\frac{C'(p_i)}{\mu C''(p_i)}$ and $\frac{d}{d\mu} \left(\frac{\partial p_i}{\partial s} \right) = -\frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s} \left(\frac{C''(p_i) + \mu C'''(p_i) \frac{\partial p_i}{\partial \mu}}{(\mu C''(p_i))^2} \right)$,

so that the two are local substitutes as long as

$$-C'(p_i) + p_i \left(C''(p_i) - C'''(p_i) \frac{C'(p_i)}{C''(p_i)} \right) < 0.$$

The result thus holds as long as p_i is sufficiently small or always, as long as

$$C''(p_i) < \left(1 + \frac{C'''(p_i)}{C''(p_i)} \right) C'(p_i).$$

The reason for this condition is that if the manager becomes very insensitive to s as μ increases, then the organization might actually want to increase s even more since the problem of losing information due to increased profit-sharing is reduced. In other words, the cost function shouldn't bee "too convex" around the equilibrium for the result to hold generally, but it still holds whenever information is sufficiently expensive independent of the cost function.

For similar reasons, we no longer can establish monotonicity of the equilibrium s in β (of course, even for the particular cost function used, the relationship could be non-monotone under decentralization, but only for parameters under which decentralization was not the equilibrium outcome). If the equilibrium quality of information is sufficiently sensitive, then an increase in β can reduce the equilibrium quality of information sufficiently so that the use of firm-wide incentives will actually decrease in β. Finally, we also lose the uniqueness of s for some cost functions because of the convexity of the returns to s in terms of the improved use of information generated, as discussed in the proof for the uniqueness under the assumed cost function. To consider the conditions for uniqueness, we have that the second-order condition is

$$\frac{\partial p_i}{\partial s} \frac{\partial \Psi_i^g(s,\beta)}{\partial s} + p_i \frac{\partial^2 \Psi_i^g(s,\beta)}{\partial s^2} + \left(\frac{\partial \Psi_i^g(s,\beta)}{\partial s} - \frac{\partial \widetilde{\Psi}_i^g(s,\beta)}{\partial s} \right) \frac{\partial p_i}{\partial s} + \left(\Psi_i^g(s,\beta) - \widetilde{\Psi}_i^g(s,\beta) \right) \frac{d}{ds} \left(\frac{\partial p_i}{\partial s} \right),$$

which, using the conditions for the parameters to constitute a solution, is smaller than zero whenever

$$\left(2\frac{\partial\Psi_i^g(s,\beta)}{\partial s}-\frac{\partial\widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right)\frac{\partial p_i}{\partial s}+p_i\left(\frac{\partial^2\Psi_i^g(s,\beta)}{\partial s^2}-\frac{\partial\Psi_i^g(s,\beta)}{\partial s}\frac{\partial s}{\partial p_i}\frac{d}{ds}\left(\frac{\partial p_i}{\partial s}\right)\right)<0.$$

Again, we have that whenever p_i is sufficiently small, then uniqueness is guaranteed. To obtain bounds on uniqueness, we have that the condition holds always as long as (noting that $\frac{d}{ds}\left(\frac{\partial p_i}{\partial s}\right)=\frac{\frac{\partial^2\widetilde{\Psi}_i^g(s,\beta)}{\partial s^2}\mu C''(p_i)-\frac{\partial\widetilde{\Psi}_i^g(s,\beta)}{\partial s}\left(\mu C'''(p_i)\frac{\partial p_i}{\partial s}\right)}{(\mu C''(p_i))^2}$)

$$\left(\left(2\frac{\partial\Psi_i^g(s,\beta)}{\partial s}-\frac{\partial\widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right)+\frac{\frac{\partial\Psi_i^g(s,\beta)}{\partial s}(C'''(p_i))}{C''(p_i)}\right)\frac{\frac{\partial\widetilde{\Psi}_i^g(s,\beta)}{\partial s}}{\mu}<C''(p_i)\left(\frac{\frac{\partial\Psi_i^g(s,\beta)}{\partial s}}{\frac{\partial\widetilde{\Psi}_i^g(s,\beta)}{\partial s}}\frac{\partial^2\widetilde{\Psi}_i^g(s,\beta)}{\partial s^2}-\frac{\partial^2\Psi_i^g(s,\beta)}{\partial s^2}\right),$$

and noting that $\mu=\frac{\widetilde{\Psi}_i^g(s,\beta)}{C'(p_i)}$, we get

$$\left(\left(2+\frac{C'''(p_i)}{C''(p_i)}\right)\frac{\partial\Psi_i^g(s,\beta)}{\partial s}-\frac{\partial\widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right)\frac{C'(p_i)}{\widetilde{\Psi}_i^g(s,\beta)}\left(\frac{\partial\widetilde{\Psi}_i^g(s,\beta)}{\partial s}\right)^2>C''(p_i)\left(\frac{\partial\Psi_i^g(s,\beta)}{\partial s}\frac{\partial^2\widetilde{\Psi}_i^g(s,\beta)}{\partial s^2}-\frac{\partial\widetilde{\Psi}_i^g(s,\beta)}{\partial s}\frac{\partial^2\Psi_i^g(s,\beta)}{\partial s^2}\right).$$

Now, $\frac{\partial\Psi_i^g(s,\beta)}{\partial s}\frac{\partial^2\widetilde{\Psi}_i^g(s,\beta)}{\partial s^2}-\frac{\partial\widetilde{\Psi}_i^g(s,\beta)}{\partial s}\frac{\partial^2\Psi_i^g(s,\beta)}{\partial s^2}>0$, so again we need that $C''(p_i)$ is not "too large." Finally, in terms of the choice of s, what continues to hold is that $\frac{dp_i}{d\mu}<0$, so that we know in which direction any change in the multiple equilibria will be.

The logic behind the governance structure continues to be the same, with decentralization arising as the preferred governance structure both when β is sufficiently low and when s is sufficiently high. However, what we again cannot establish in full generality is that there is a unique cutoff μ where centralization is preferred for higher and decentralization is preferred for lower costs of information. This result arises because the rates at which p_i^g decrease in μ might make such switch possible. However, what we do know is that when μ is low enough, then s will be large and, as a result, decentralization will be preferred. Similarly, if μ is large enough, $p_i^g\to 0$ and centralization will be preferred.

The possibility of non-monotonicity is also supported by the possible multiplicity of equilibria. For some μ, the preferred governance structure might be centralization with low s, while an increase in μ causes a switch in equilibrium to lower p and higher s, which in turn might warrant a switch to decentralization instead.

Finally, evaluating the impact of σ_θ^2 and μ on β, we have immediately that $\frac{d\beta}{d\sigma_\theta^2}\leq 0$ because σ_θ^2 continues to be a scalar multiplying the profit surface, while for μ we have that

$$\frac{d^2\pi}{d\beta d\mu}=\left(-\frac{dp_i^g}{d\mu}\left(\frac{\partial\Psi_i^g(s,\beta)}{\partial s}\frac{\partial s}{\partial\beta}-\frac{\partial\Psi_i^g(s^g,\beta)}{\partial\beta}\right)-p_i^g\frac{d}{d\mu}\left(\frac{\partial\Psi_i^g(s,\beta)}{\partial s}\frac{\partial s}{\partial\beta}-\frac{\partial\Psi_i^g(s^g,\beta)}{\partial\beta}\right)\right)\sigma_\theta^2.$$

Again, if the cost function is sufficiently convex, the result of the paper on the monotonicity $\frac{d\beta}{d\mu}\geq 0$ can fail to hold. Here, the intuition is that if p_i^g is roughly unchanged, and there is a reduction in s caused by the increase in μ, then there is an increase in the agency costs associated with further integration, leading the organization actually to reduce the level of operational integration to reduce the losses due to strategic decision-making and communication. However, it is again the case that if μ is sufficiently large and so p_i^g is sufficiently low, then the result that $\frac{d\beta}{d\mu} \geq 0$ will hold. In short, while allowing general cost functions can locally make some of the results of section 5 to fail, the global qualitative results continue to hold.

C.2 Endogenous strength of incentives

As mentioned in the text, the key simplification for the analysis was that in addition to focusing on linear compensation contracts, I restricted the attention to pure profit-sharing, where the division managers' compensation was given by

$$T_i(\pi_i, \pi_j) = A_i + (1-s)\pi_i(\theta_i, d_i, d_j) + s\pi_j(\theta_j, d_i, d_j).$$

A simple modification to this is to write the compensation structure as

$$T_i(\pi_i, \pi_j) = A_i + s_{ii}\pi_i(\theta_i, d_i, d_j) + s_{ij}\pi_j(\theta_j, d_i, d_j) = A_i + \lambda\left((1-s)\pi_i(\theta_i, d_i, d_j) + s\pi_j(\theta_j, d_i, d_j)\right),$$

where s continues to measure the composition of incentives but λ measures now the strength of incentives. The only impact that λ has on the solution to the game is that increasing λ increases the level of information acquisition by the division managers. The decision-making and communication stages continue to be solely determined by s. The difference is that the perceived value of information is now $\lambda\widetilde{\Psi}_i^g(s, \beta)$ and so the equilibrium level of information acquisition will solve

$$p_i = \frac{\lambda\widetilde{\Psi}_i^g(s,\beta)}{\lambda\widetilde{\Psi}_i^g(s,\beta)+\mu}.$$

The main difference to the design stage is that there can now be multiple equilibria for the compensation structure even within a governance structure. Intuitively, λ and s can be complements because increasing λ increases the quality of primary information, which in turn increases the value of profit-sharing. In other words, the equilibrium might be one of low strength of incentives and limited profit-sharing or one of high strength of incentives and extensive profit-sharing. Beyond this possible multiplicity, the logic of the results is unchanged. Most importantly, the monotonicity of β in σ_θ^2 and μ remains.

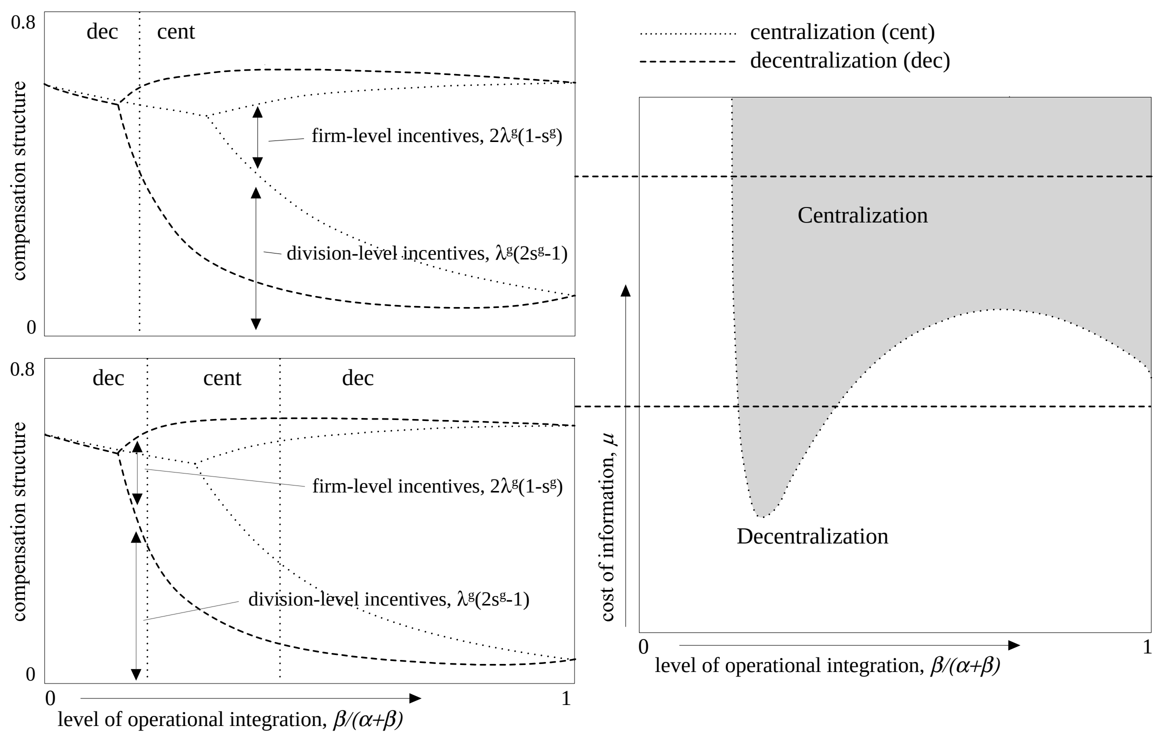


Figure 8: Choice of compensation and governance structures as a function of β and μ

C.2.1 Exogenous cost of incentives

An earlier working paper analyzed the case where the cost of incentives is given by an exogenous function $G(\lambda)$. The detailed analysis is available in the working paper and I will only give the main figures. The key insight is that because centralization is able to use conflict between the divisions to motivate information acquisition, centralization is able to economize on the cost of incentives, which in turn leads to a non-monotone relationship between decentralization and volatility. In the case of the cost of incentives being proportional to volatility, an illustration of the typical solution is given by the figures 8 and 9, first illustrating the link between operational integration and the choice of governance structure, while the second looking at the equilibrium relationship between volatility and the choice of governance structure.

Further, any comparative statics with respect to the cost of incentives parallel those of cost of information: increasing the cost of incentives reduces the equilibrium strength, lowering the quality of primary information acquired, lowering the level of profit-sharing and leading towards a preference for centralization and vice versa.

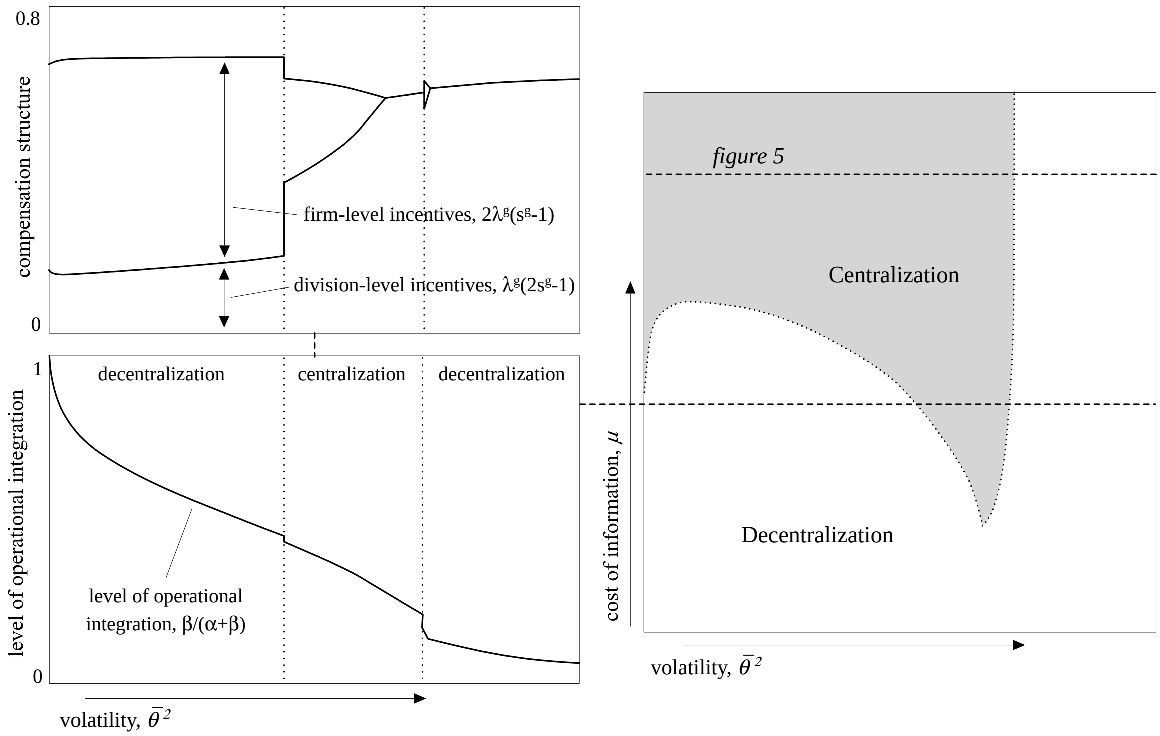


Figure 9: Organizational design and environmental volatility

C.2.2 Without ex ante transfers

The solution in the previous subsection still focuses on maximizing the total surplus, conditional on the cost of incentives $G(\lambda)$. In an alternative formulation, the headquarters focuses on only maximizing the surplus left over after compensating the division managers. As long as the participation constraint is not binding, then the headquarters would be maximizing $(1-\lambda) E(\pi_i + \pi_j)$. The downside of this formulation is that there is no natural benchmark performance level on which to tie $E(\pi_i + \pi_j)$. In particular, other things constant, $E(\pi_i + \pi_j)$ is monotone decreasing in σ_θ^2, so that increasing the strength of incentives becomes inherently cheaper with volatility. Arguably, the organization could choose to hire nobody, set $\beta \to \infty$ and have an expected performance of $K(\infty) - \alpha\sigma_\theta^2$. If we use this as the baseline over which the surplus is divided, then the maximization problem would become

$$(1-\lambda)\left((K(\beta) - K(\infty)) + p_i \Psi_i^g(s,\beta)\,\sigma_\theta^2\right).$$

Then, the first-order condition for the choice of s would be

$$\frac{\partial p_i}{\partial s}\Psi_i^g(s,\beta) + p_i \frac{\partial \Psi_i^g(s,\beta)}{\partial s} = 0.$$

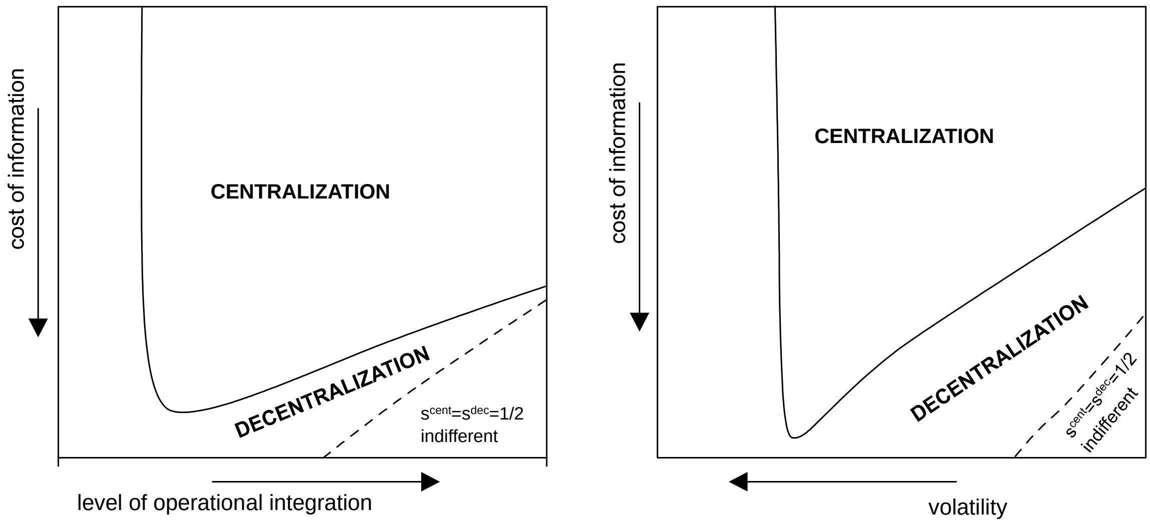


Figure 10: Choice of governance structure as a function of β and σ_θ^2

Note that because the organization now doesn't internalize the cost of information, it has an inherent bias towards motivating additional information acquisition over the use of that information appropriately. This puts pressure on s to be lower than before. Indeed, for λ sufficiently small, we are in the corner solution of $s = 0$. The strength of incentives, in turn, would solve

$$(1-\lambda)\left(\frac{\partial p_i}{\partial \lambda}\Psi_i^g(s,\beta)\,\sigma_\theta^2\right) - \left((K(\beta) - K(\infty)) + p_i\Psi_i^g(s,\beta)\,\sigma_\theta^2\right) = 0,$$

while the choice of governance structure and the choice of β would be unbiased, conditional on (s,λ). However, the organization overall would exhibit a bias towards centralization indirectly through the suboptimally low level of profit-sharing chosen in equilibrium. Beyond that, the basic tradeoffs remain the same, as illustrated in figure 10 mapping the choice of g. The other choices parallel our earlier discussion, and thus not presented. The results are available from the author on request.

C.3 Miscellaenous observations

Finally, let me make few observations on some commonly asked questions regarding the assumptions of the model. Because full analysis would take another paper or two, I will only make intuitive arguments as to their effects. Proofs, where applicable, are available on request.

Are all the stages and components really necessary? In short, yes. If information was freely available to the agents, then it would be optimal to set $s = \frac{1}{2}$, achieving perfect communication and decision-making and making rest of the design problem irrelevant. One could exogenously constraint the degree of alignment, but the results would remain somewhat uninteresting - choose maximal alignment. Similarly, if communication was not strategic, then, in the case of perfect communication, centralization would always be preferred because it would achieve the first-best outcome. Headquarters makes first-best decisions based on the full information communicated while the incentive structure provides first-best incentives to acquire information. Decentralization, on the other hand, would need to balance the incentives to acquire information with incentives to make good decisions. Indeed, this is the result in Athey and Roberts (2001), where truthful communication of the project qualities can be achieved. Exogenously restricting the quality of communication would generate a similar tradeoff between centralization and decentralization. However, it would generate the prediction that under centralization, the division managers are always compensated only based on divisional performance, which appears to run counter to standard intuition. Indeed, one of the main observations from the model is that depending on the level of operational integration, centralization can exhibit more or less profit-sharing than a decentralized structure. As a result, to generate the predictions of the model, we need both strategic information acquisition and strategic communication.

The results would, however, be roughly similar if, instead of an information acquisition task, the managers simply engaged in productive effort for their divisions as in Dessein, Garicano and Gertner (2009). This approach, however, would then ignore the interactions between the quality of information generated and how good use the organization makes of that information, together with the impact that the governance structure alone has on the incentives to acquire information (as standard effort would be independent of the decision-making structure) and the link between organizational design and the difficulty of predicting the environment. Finally, the model also needs continuous (instead of binomial) decisions and information. The reason is that making the information binomial would collapse the communication problem into one between truth-telling and lying, which would fail to appropriately capture the idea that while inducing accurate communication under decentralization is more costly, accurate communication is also less valuable. The key is that the equilibrium here allows for different qualities of communication that can smoothly be controlled through the compensation contract, letting the accuracy track the value.

What happens when information acquisition is not observable? The analysis made the somewhat-unusual assumption that the effort choice of the division managers is observable but not verifiable. While it can be argued that it is possible for the headquarters to be aware of how hard their managers are working without being able to write contracts on that information, it is also natural to ask what happens if the information acquisition is non-observable. The technical difficulty is that when the quality of information is unobservable, then the interpretation of the messages by the recipient comes to be based on beliefs, while messages sent are based on the actual quality of information (as known by the sender) and the expected interpretation by the recipient.

The basic impact, however, is straightforward. When the recipient is less responsive to the information than desired by the sender (which is the case here), then the sender will always acquire less information when the accuracy of information is unobservable. The intuition is follows. Suppose that the quality of information q_i is currently at the level chosen by the manager when q_i is observable. Then, by acquiring a little less accurate information, then the manager is able to reduce the perceived conflict between the him and the recipient, because the recipient will continue to form beliefs based on $E(\theta_i|t_i) = q_i^e t_i$ and so over-respond to the signals when the manager undercuts the expected quality of information (proof for the single sender and recipient available on request). The highest quality of information acquisition that can be sustained will still be above zero because information is valuable.

As to the differences between the two governance structures, we know that the maximum quality of information that can be sustained is the same whenever $\beta \to 0$ and $s = 1/2$, in which cases the quality will be the same as under observability of q (because there are no incentive conflicts between the sender and the receiver) and will converge to the same level when $\beta \to \infty$ (for the same reason that the accuracy of communication converges - the equilibrium decisions are the same under the two governance structures). For the interior, the relative impact is uncertain, but intuitively it would appear that the reduction in the quality will be worse under centralization. Indeed, under centralization $q_i^e = q_i = 0$ is always an equilibrium, while under decentralization there is always a positive value to information because it is the manager who will use it in his decision-making.

What happens if the organization can separately measure the coordination benefits and allocate them to the two divisions optimally? As currently formulated, the benefits of coordination cannot be separated from divisional profits. An alternative formulation allows all the three components, $\alpha(\theta_i - d_i)^2$, $\alpha(\theta_j - d_j)^2$ and $\beta(d_i - d_j)^2$ to be separately contractible. This result could arise, for example, if the headquarters sets up a separate unit that will be serving both of the divisions, where the performance of that unit is given by $\beta(d_i - d_j)^2$.

As long as we focus on a team-production problem, where the payoffs cannot be leveraged, the results continue to be similar. The main reason is that the firm always want to allocate all coordination costs to the two divisions, and do this in a symmetric fashion, so that

$\beta_i = \beta_j = \frac{\beta}{2}$. This symmetry follows because, as discussed in Rantakari (2008), introducing asymmetries between the divisions worsens the average quality of communication under both governance structures and gives a strategic advantage in the decision-making stage to the less integrated division under decentralization.

One difference comes from the compensation structure. Note that the decision-making incentives are based solely on $\alpha\left(\theta_i - d_i\right)^2$ and $\beta\left(d_i - d_j\right)^2$, while the communication incentives rely on all three components. Given that fixing the bias in decision-making is the primary concern under decentralization while improving the accuracy of communication is the primary concern under centralization, we would expect that decentralization will make relatively less use of the other division's performance $\alpha\left(\theta_j - d_j\right)^2$ than centralization, other things constant. This will then allow decentralization to potentially motivate even more information acquisition through $\alpha\left(\theta_i - d_i\right)^2$ and thus improve its relative performance. But again, the difference in quantitative and doesn't alter the fundamental tradeoff between motivating information acquisition and then using that information appropriately.

Is the headquarters really in control of β? The model assumes that the headquarters can control β. This assumption captures the idea that the headquarters can influence what assets the divisions have available to them, even when the divisions are then free to make their own operating decisions conditional on those assets and pre-specified interdependencies in place. We can make two further observations on this assumption. First, headquarters choosing β yields a strictly better outcome than if β was left up to the division managers. The reason is that even if all the costs and benefits of β were borne by the division manager making the choice, they will under-integrate to gain a strategic advantage over the other division (per the benefits and costs of asymmetries between the two divisions, as discussed in Rantakari 2008). Similarly, even when the headquarters is in charge of β, the division managers can have an incentive to try to undo some of this integration through their own choices. Such considerations, however, would only add an additional layer of complexity to the analysis. Further, all the results conditional on β naturally continue to hold. Intuitively, any other set of assumptions will lead to a lower equilibrium β but there is also no a priori reason why, given the difference in level, there should also be further differences in how the equilibrium β responds to changes in σ_θ^2 and μ.