Organizing to Adapt and Compete

By Ricardo Alonso, Wouter Dessein, and Niko , Dessein, and Matouschek*

We examine the relationship between the organization of a multi-divisional firm and its ability to adapt production decisions to changes in the environment. We show that even if lower-level managers have superior information about local conditions, and incentive conflicts are negligible, a centralized organization can be better at adapting to local information than a decentralized one. As a result, and in contrast to what is commonly argued, an increase in product market competition that makes adaptation more important can favor centralization rather than decentralization. (*JEL* D23, D83, L23)

The organization theorist Chester Barnard and the economist Friedrich Hayek shared the view that the "economic problem of society is mainly one of rapid adaptation to changes in the particular circumstances of time and place" (Hayek, 1945, p.524). But whereas Hayek viewed adaptation as an autonomous process, undertaken by individual economic actors, Barnard (1938) stressed the ability of organizations to engage in what Oliver Williamson (1996, 2002) calls "*coordinated adaptation*." Williamson (1996, p.103), referring to Barnard and challenging Hayek, argues that:

"*Some kind of disturbances require coordinated responses, lest the individual parts operate at cross-purposes or otherwise suboptimize. Failures of coordination may arise because autonomous parties read and react to signals differently, even though their purpose is to achieve a timely and compatible combined response.* [...] *The authority relationship (fiat) has adaptive advantages over autonomy for transactions of a bilaterally (or multi-laterally) dependent kind.*"

[0]* Alonso: Marshall School of Business, University of Southern California, HOH 615, Mail Code 0804, Bridge Hall-308, 3670 Trousdale Parkway, Ste. 308, Los Angeles, CA 90089, vralonso@marshall.usc.edu; Dessein: Graduate School of Business, Columbia University, 3022 Broadway, Uris Hall 625 New York, NY 10027, wd2179@columbia.edu; Matouschek: Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60201; n-matouschek@kellogg.northwestern.edu. We thank Maria Guadalupe, Michael Raith, Steve Tadelis, and the participants of various conferences and seminars for their comments. All remaining errors are our own.

In this paper we explore the relationship between a firm's organizational structure and its ability to adapt to a changing environment. We ask when adaptation is best achieved in a decentralized organization–through a process of autonomous adaptation as envisioned by Hayek–and when it is best done in a centralized organization–through a process of administrative control and coordinated adaptation, as envisioned by Barnard and Williamson. We then examine how the degree of product market competition affects how the firm optimally adapts to changes in its environment. A commonly made argument is that competition makes it more important for firms to be responsive to changes in their environment and, therefore, favors decentralization. This argument is summarized, for instance, in Tim Harford's recent best-seller "Adapt" (Harford 2011, pp.75-76):

"*Thanks to globalisation, businesses have ventured into new and varied markets, where they face intense competition. The traditional purpose of centralisation is to make sure every business unit is coordinated and nobody is duplicating anyone else's effort* [...]. *But a centralised organisation doesn't work so well when confronted with a diverse, fast-moving range of markets. The advantage of decentralisation, rapid adaptation to local circumstances, has grown.*"

While intuitively appealing, we argue that the above reasoning is flawed when thinking about the large-scale, multi-divisional organizations that have come to dominate many industries.

To study the issues of adaptation, competition, and organization, we consider a multi-divisional firm whose profits depend only on how the production level of each division is adapted to its local demand conditions. The production decisions can be either strategic complements–if there are increasing returns to aggregate production–or strategic substitutes–if there are decreasing returns. We show two main results. First, even when division managers have superior information about their local demand conditions, and incentive conflicts are negligible, the headquarter manager may be better at adapting production to demand conditions than the division managers. Second, we show that an increase in competition does not only make it more important that firms adapt to changes in their environments. Instead, it also changes how firms should adapt to such changes, making coordinated adaptation more important relative to autonomous adaptation. As a result, an increase in competition can actually favor centralization rather than decentralization.

Organizing to Adapt. The result that a centralized organization can be better at adapting to changes in the environment than a decentralized one follows from three observations. The first is that if decisions are interdependent, adaptation involves both autonomous adaptation–adapting each production decision to its local demand conditions–and coordinated adaptation–adapting each decision to the other production decisions, and through this channel to the demand conditions in the other markets. Coordination is therefore an input into adaptation that allows the decision

makers to adapt to demand shocks more aggressively. Moreover, the more interdependent the production decisions are, the more important coordinated adaptation becomes relative to autonomous adaptation.

The second observation is that while an organization's ability to engage in autonomous adaptation depends on how much decision makers know about the demand conditions in their respective markets, their ability to engage in coordinated adaptation depends on how much they know about each others' markets. And organization's ability to adapt therefore depends on both the depth and the breadth of the decision makers' information. Moreover, the more interdependent the production decisions, the more important breadth becomes relative to depth.

The final observation is that even though each division manager knows more about his market, and therefore has an advantage in terms of depth, the headquarter manager may well know more about the other markets, and thus have an advantage in terms of breadth. This will be the case, for instance, if there are economies to knowledge specialization (Becker and Murphy 1992, Bolton and Dewatripont 1994, and Sims 2002) or if cognitive constraints make it more difficult for division managers to talk to each other than to headquarters (Ferreira and Sah 2012). And even when there are no economies to knowledge specialization or cognitive constraints, such an information structure can arise naturally when managers communicate with each other strategically, as we will show below.

Together these three observations imply our first main result: for sufficiently interdependent decisions, a headquarter manager who knows a little bit about all markets is better at adapting production to demand conditions than the division managers who know a lot about their own markets but very little about each others'.

Organizing to Compete. To explore the informal argument that competition favors decentralization, we examine an increase in competition that makes demand more price sensitive. In line with the standard intuition, an increase in the price sensitivity of demand makes it more important to adapt production to demand conditions. What the standard intuition overlooks, however, is that an increase in the price sensitivity of demand also changes the process through which decisions are adapted to demand shocks. In particular, the more price sensitive demand is, the more interdependent are the production decisions, and thus the more important coordinated adaptation and breadth become relative to autonomous adaptation and depth. In our setting, an increase in competition therefore favors centralization over decentralization.

While this argument may be counter-intuitive at first, it is consistent with the recent experience of multi-divisional firms such as Unilever that have responded to increased competition by central-

izing decision rights. In particular, until 1999, Unilever had a very decentralized organizational structure in which division managers had vast decision making authority. Since then, however, Unilever has gone through a series of reorganizations that have centralized authority and limited the power of division managers. Recently, Unilever Chief Executive Patrick Cescau explained these reorganizations:

"*Historically, Unilever's business had been built up around highly autonomous operating companies, with their own portfolio priorities and all the resources they needed – marketing, development, supply chain – to develop their business in whatever way they saw fit. This was a highly effective way of building a truly multinational business almost 50 years before the term was invented. But it had become less suited to an increasingly globalised, competitive landscape, where battles were being fought and won with global scale and know-how, and top-down, strategically driven allocation of resources. In today's world, a hundred different portfolio strategies run the risk of adding up to no strategy at all. It's not efficient, it doesn't leverage your best assets and it doesn't build strong global positions.*"[1]

In line with at least the spirit of our model, competition therefore favored centralization and it favored centralization because it made it more important to adapt to changes by coordinating strategies, realizing scale economies, and quickly moving resources across divisions.

Strategic Communication. Below we first establish our results in a model without incentive conflicts and in which we take the generalist-specialist information structure as given, that is, we exogenously assume that a headquarter manager knows a little bit about all markets, whereas division managers know a lot about their own markets but very little about each others'. In Section V, we then endogenize the information structure by allowing for incentive conflicts and strategic communication. Drawing upon Alonso, Dessein, and Matouschek (2008), we assume division managers privately observe demand conditions in their own market but must rely on (horizontal) cheap talk communication to learn about the demand condition of the other division. Communication is strategic because each division manager puts more weight on the profits of his own division than on those of the other division. The headquarter manager, in contrast, is unbiased but has no private information. She must rely on (vertical) cheap talk communication to learn about demand in both markets. We characterize the quality of both horizontal and vertical communication and show that division managers typically share more information with an unbiased headquarter manager than with each other. The generalist-specialist information structure that is central to our

[1] This quote is taken from a presentation that Patrick Cescau gave at the Unilever Investor Seminar on 13 March 2007. The transcript is available at www.unilever.com.

model then arises endogenously. Focusing on the case where the own division biases are small, we further show that centralized decision-making is preferred to decentralized decision-making if and only if production decisions are sufficiently complementary. This result stands in sharp contrast with Alonso, Desssein, and Matouschek (2008) who show that decentralization is always optimal when the own division bias is small. If production decisions are strategic substitutes, in contrast, we confirm the qualitative results in Alonso, Dessein, and Matouschek (2008).

I. Literature Review

Our paper builds on a growing theoretical literature in organizational economics that examines the relationship between a firm's organizational structure and its ability to coordinate decisions across divisions and other sub-units. Coordination requires the aggregation of dispersed information but is imperfect because of physical communication constraints (Aoki 1986, Hart and Moore 2005, Dessein and Santos 2006, and Cremer, Garicano, and Prat 2007) or because sub-unit managers are biased and communicate strategically (Alonso, Dessein, and Matouschek 2008, Rantakari 2008, Dessein, Garicano, and Gertner 2010, Friebel and Raith 2010).[2] A key argument in this literature is that a centralized structure is better at coordinating decisions, while a decentralized one is better at adapting decisions to the circumstances and opportunities of different sub-units.

We add to this literature in a number of ways. First, we show that despite division managers being better informed about their own local circumstances, a central structure can be better at adapting to those circumstances than a decentralized one. Our model shares with Alonso, Dessein, and Matouschek (2008) that decision making can be decomposed into autonomous adaptation and coordinated adaptation and that coordinated adaptation can be best achieved in a centralized structure. In Alonso, Dessein, and Matouschek (2008), however, a decentralized organization is always more adaptive to local information than a centralized one (see Lemma 2 in that paper). Moreover, that paper does not mention or analyze the benefits of coordinated adaptation which, in fact, are dominated by other factors that affect adaptation, such as incentive conflicts. By abstracting from incentive conflicts, the present model clarifies and emphasizes that centralization can be the optimal organizational structure for purely adaptive/informational reasons. Second, many of the above papers, and a number of recent papers in other fields, model the need for adaptation and coordination by assuming a particular payoff function.[3] In contrast, we derive the payoff function by explicitly modeling the production interdependencies and the variability in

[2] More generally, the literature on team theory examines decision making in firms in the presence of informational constraints. See Van Zandt (1999) for an overview.

[3] See, for example, Morris and Shin (2002), Dessein and Santos (2006), Alonso et al. (2008), Rantakari (2008), Calvó-Armengol et al. (2011), Bolton et al. (2013), and Hagenbach and Koessler (2010).

demand that create the need for adaptation and coordination. This approach allows us to examine the impact of the intensity of competition and the external environment of the firm on its internal organization.

By linking the intensity of competition to the choice between centralization and decentralization, we also relate to a literature that studies the impact of competition on organizational features of firms. Most of this literature examines how competition affects managers' incentives to reduce costs both directly–for given incentive schemes–and indirectly–by changing the managers' incentive schemes (Schmidt 1997, Raith 2003, Vives 2008). Closest to us, Raith (2003) shows that the effect of competition on managerial incentives depends on how it affects the level and the price sensitivity of demand. In his model, an increase in price sensitivity makes low marginal costs more important and thus favors high powered incentives. Similarly, in our setting, an increase in price sensitivity makes it more important that the firm adapts to demand shocks but also makes production decisions more interdependent. As a result, such an increase in price sensitivity favors centralization. Raith (2003) and our paper therefore both identify the same channel through which competition affects organizational design but examine a different organizational design variable.[4]

II. The Model

A multi-divisional firm consists of two operating divisions and one headquarters. Division $j = 1, 2$ produces a single good and sells it in Market j. Each division is a monopolist in its market and headquarters does not engage in any production.

The costs of Division $j = 1, 2$ are given by

$$cq_j + gq_1q_2, \tag{1}$$

where $c \geq 0$, $g \in (-1, 1)$, and q_1 and q_2 are the production levels of Divisions 1 and 2. If $g > 0$, an increase in production by one division increases the average and marginal costs of the other division. Such negative cost externalities may arise, for instance, because of the utilization of common services, such as personnel and IT departments and managerial supervision and know-how. If, instead, $g < 0$, then an increase in production by one division reduces the average and marginal costs of the other division. Such positive cost externalities may arise, for instance, because of learning and scale effects in the production of inputs used by both divisions.

The inverse demand function in Market $j = 1, 2$ is given by

$$p_j = \mu + \theta_j - \frac{1}{b}q_j, \tag{2}$$

[4] In addition, Raith (2003) identifies a scale effect, where a lower level of demand reduces the benefits of lower marginal costs. The level of demand plays no role in our model.

where $b > 0$ is the price sensitivity of demand $|\mathrm{d}q_j/\mathrm{d}p_j|$. Demand shocks θ_1 and θ_2 are independently drawn from a distribution with zero mean and variance σ^2. To avoid cumbersome corner solutions, we assume that the intercept $\mu + \theta_j$ is always sufficiently large for production to take place in both markets.

Given cost functions (1) and inverse demand functions (2), profits of Divisions $j = 1, 2$ are given by

$$\pi_j(q_1, q_2) = \left(\mu + \theta_j - \frac{1}{b}q_j - c - gq_k\right) q_j, \text{ where } k \neq j \ . \tag{3}$$

We denote overall firm profits by $\pi(q_1, q_2) = \pi_1(q_1, q_2) + \pi_2(q_1, q_2)$ and expected profits by $\Pi(q_1, q_2) = \mathrm{E}[\pi(q_1, q_2)]$. Whenever it does not cause any confusion, we simplify notation by omitting the production levels. For instance, we sometimes denote expected profits simply by Π.

There are three managers: the headquarter manager, who is in charge of headquarters, and Division Managers 1 and 2, who are in charge of Divisions 1 and 2 respectively. The managers are risk neutral and care about overall firm profits. They differ, however, in terms of the information that is available to them. In particular, we assume that the headquarter manager is a generalist with broad but shallow information while the division managers are specialists with deep but narrow information: Division Manager $j = 1, 2$ knows more than any other manager does about the demand conditions in Market j, while the headquarter manager knows more than Division Manager j does about the demand conditions in Market $k \neq j$. Formally, Division Manager $j = 1, 2$ privately observes signal $priv_j$ about the demand conditions in Market j. Both division managers also observe signals div_1 and div_2 about the demand conditions in Markets 1 and 2, where "div" stands for "division managers." Finally, the headquarter manager observes signals $head_j$ about the demand conditions in Market j, where "head" stands for "headquarter manager." This information structure is summarized in Figures 1a and 1b. We endogenize the signals div_j and $head_j$ in Section V.[5]

INSERT FIGURE 1 HERE

We refer to Division Manager j's information about the demand conditions in Market $j = 1, 2$ as his "local information" and we denote it by $loc_j = (priv_j,\ div_j)$. We can then measure the depth of the division managers' information with the residual variance

$$V_{\mathrm{d}} \equiv \mathrm{E}\left[(\theta_j - \mathrm{E}[\theta_j \,|loc_j\,])^2\right],$$

[5] It follows that what Manager 2 knows about Market 2 is a sufficient statistic to knowing what Manager 1 knows about Market 2. This simplifies the analysis and corresponds to the signal structure which arises endogenously under cheap talk, as analyzed in Section 6. If this were not to be the case, Manager 2 would need to worry about Manager 1's knowledge and exact response. We conjecture that this would further complicate coordinated adaptation (see Section 4), and hence strengthen the case for centralized adaptation.

where the subscript d stands for "depth." Similarly, we measure the breadth of their information with the residual variance

$$V_{\mathrm{b}} \equiv \mathrm{E}\left[\left(\theta_j - \mathrm{E}\left[\theta_j \,|div_j\,\right]\right)^2\right],$$

where the subscript b stands for "breadth." If division managers have no information about the demand conditions in each others' markets, for instance, then the residual variance V_{b} is equal to the variance σ^2. And if division managers are perfectly informed about the demand conditions in their own markets, then $V_{\mathrm{d}} = 0$. Since each division manager knows more about his market than he does about the other market, we have that $V_{\mathrm{d}} \leq V_{\mathrm{b}}$.

In contrast to the division managers, the headquarter manager is equally well informed about the demand conditions in both markets. In particular, after receiving signals $head_1$ and $head_2$, the headquarter manager's residual variance about the demand conditions in Market $j = 1, 2$ equals

$$V \equiv \mathrm{E}\left[\left(\theta_j - \mathrm{E}\left[\theta_j \,|head_j\,\right]\right)^2\right].$$

To capture the assumption that the headquarter is a generalist and the division managers specialists, we assume that

$$V_{\mathrm{d}} < V < V_{\mathrm{b}}.$$

Compared to Division Manager $j = 1, 2$, the headquarter manager therefore knows less about the demand conditions in Market j but more about those in Market $k \neq j$.

To model the firm's organizational choice, we assume that the two goods that the firm produces are so complex that they cannot be fully described in a written contract, neither ex ante nor ex post.[6] Note that this also rules out contracts on the price and the quantity of the goods. The only organizational choice that the firm can make is therefore the allocation of decision rights. We focus on three organizational structures: Centralization, Divisional Centralization, and Decentralization. Under Centralization both production decisions q_1 and q_2 are made by the headquarter manager and under Divisional Centralization they are both made by Division Manager 1. Under Decentralization, in contrast, production decisions q_1 and q_2 are made by Division Managers 1 and 2 respectively.

Finally, the timing is as follows. First, decision rights are allocated to maximize expected profits. Second, the managers learn their information. Third, the decision makers decide on the production levels. Finally, profits are realized and the game ends.

Before we move on to solve this model, it is useful to discuss some of the main assumptions. One key assumption is that the division managers' information is taken as given. Implicitly we

[6] For ex ante non-contractibility, see for instance, Grossman and Hart (1986) and Hart and Moore (1990). For ex post non-contractiblility, see, among others, Aghion et al. (2004).

are therefore ruling out efficient communication between the division managers. We endogenize communication and thus the information structure in Section V.

Another key assumption is that managers only care about overall firm profits. We abstract from incentive conflicts to focus on the role of information in the managers' ability to adapt to demand shocks.

Throughout the paper we assume that the demand shocks in the two markets are independent. In principle we could allow for the demand shocks to be correlated. Doing so, however, would only obscure the difference between depth and breadth that is at the heart of our results. We therefore deliberately focus on independent demand shocks.

Another assumption that we maintain throughout the paper is that the externalities the divisions impose on each other are on the cost side rather than on the demand side. Notice, however, that the profit functions (3) could be generated in a model with constant marginal costs and linear demand functions in which the two goods are imperfect substitutes or complements. A model with demand externalities of this type is therefore identical to ours.

Finally, given our focus on cost externalities, it is natural to assume that the decision makers set quantities and not prices. Notice, however, that since the divisions are monopolists, setting prices is equivalent to setting quantities. Moreover, since we allow for both positive and negative cost externalities, the production decisions can be either strategic complements or substitutes.

III. Adaptation and the Depth and Breadth of Information

To solve the model, we first examine decision making, adaptation, and the firm's performance under Decentralization. In the final sub-section we then derive the expected profits under the two centralized structures and determine the optimal organizational structure.

A. Decision Making

Suppose that the firm is decentralized. After the division managers have observed their signals, they simultaneously decide on the production levels that maximize expected profits. Taking expectations over profits (3) and differentiating, we find that the reaction functions are given by

$$q_1 = \frac{b}{2}\left(\mu - c + \mathrm{E}\left[\theta_1 \left| loc_1 \right.\right]\right) - t\mathrm{E}\left[q_2 \left| div \right.\right] \tag{4}$$

and

$$q_2 = \frac{b}{2}\left(\mu - c + \mathrm{E}\left[\theta_2 \left| loc_2 \right.\right]\right) - t\mathrm{E}\left[q_1 \left| div \right.\right], \tag{5}$$

where $t \equiv gb$ and where $div = (div_1, div_2)$ are the signals that are observed by both division managers. To make the right decision, each division manager therefore has to predict both his

local demand conditions and the decision that the other division manager is going to make. Note that his ability to predict his local demand condition depends on the depth of his information while his ability to predict the other decision depends on its breadth.

In what follows, the parameter t plays an important role. Notice that the sign of t determines whether the production decisions are strategic complements or substitutes and that the absolute value of t measures the degree of interdependence between the production decisions. Notice also that the degree of interdependence $|t|$ is increasing in the price sensitivity of demand b, which will be important for what follows below. To see why this is the case, consider Figures 2a and 2b. In each figure, the flat lines are Division 1's marginal costs for different production levels by Division 2. And the downward sloping lines are Division 1's inverse demand and marginal revenue functions, drawn for less price sensitive demand in Figure 2a and for more price sensitive demand in Figure 2b. The figures show that the more price sensitive demand is, the larger is the impact of a change in Division 1's marginal costs–caused by a change in production by Division 2–on Division 1's optimal production level.

INSERT FIGURE 2 HERE

Next we can solve reaction functions (4) and (5) to find the decision rules

$$q_1 = b\frac{\mu - c}{2(1+t)} + \left(\frac{b}{2}\mathrm{E}\left[\theta_1 \left| loc_1\right.\right] + bt\frac{t\mathrm{E}\left[\theta_1 \left| div_1\right.\right] - \mathrm{E}\left[\theta_2 \left| div_2\right.\right]}{1-t^2}\right) \tag{6}$$

and

$$q_2 = b\frac{\mu - c}{2(1+t)} + \left(\frac{b}{2}\mathrm{E}\left[\theta_2 \left| loc_2\right.\right] + bt\frac{t\mathrm{E}\left[\theta_2 \left| div_2\right.\right] - \mathrm{E}\left[\theta_1 \left| div_1\right.\right]}{1-t^2}\right), \tag{7}$$

where we are abusing notation somewhat by using q_1 and q_2 to denote both the production levels and the decision rules. The first term on the RHS of each expression is the average production level $\overline{q} \equiv \mathrm{E}[q_1] = \mathrm{E}[q_2]$. The second term is then the difference between actual and average production and thus captures the extent to which division managers adapt to demand shocks. Since the division managers are unbiased, they always make the first best decision on average. The firm's performance therefore depends only on how well the division managers adapt to demand shocks.

B. Adaptation

To examine how the division managers adapt to demand shocks, suppose that there is a positive demand shock in Market 1 but not in Market 2. In particular, suppose that $\theta_1 = \Delta$ for some $\Delta > 0$ and $\theta_2 = 0$. Also, and without loss of generality, suppose that there are negative cost externalities, that is, $g > 0$. From (6) and (7) it follows that the expected changes in production are then given

by

$$\mathrm{E}\left[q_1 - \overline{q}\,|\theta_1 = \Delta\right] = \mathrm{E}\left[\frac{b}{2}\mathrm{E}\left[\theta_1\,|loc_1\right] + \frac{b}{2}\frac{t^2}{1-t^2}\mathrm{E}\left[\theta_1\,|div_1\right]|\theta_1 = \Delta\right] \tag{8}$$

and

$$\mathrm{E}\left[q_2 - \overline{q}\,|\theta_1 = \Delta\right] = -\frac{b}{2}\frac{t}{1-t^2}\mathrm{E}\left[\mathrm{E}\left[\theta_1\,|div_1\right]|\theta_1 = \Delta\right]. \tag{9}$$

To interpret these expressions, suppose first that the division managers are perfectly informed about the demand conditions. We can then illustrate the above production changes in Figure 3a, where the solid lines are the reaction functions (4) and (5) if $\theta_1 = \theta_2 = 0$ and the dashed line is Division Manager 1's reaction function if $\theta_1 = \Delta$ and $\theta_2 = 0$. The above production changes then correspond to a move from A to C in the figure.

INSERT FIGURE 3 HERE

Notice that the overall changes in production are the sum of two components. The first component is the change in production by Division Manager 1, holding constant the decision by Division Manager 2. It is given by the first term on the RHS of (8) and it corresponds to a move from A to B in Figure 3a. Since this component does not require any coordination between the division managers, it depends only on how much Division Manager 1 knows about the demand conditions in his own market. As such, it captures the notion of "autonomous adaptation" that we discussed in the Introduction.

The second component, in contrast, is the additional change in production that is due to each division manager adapting production by his division to changes in production by the other, and thus captures the notion of "coordinated adaptation." It is given by the second term on the RHS of (8) and the only term on the RHS of (9) and it corresponds to a move from B to C in Figure 3a. Notice that these terms depend only on how much Division Manager 2 knows about the demand conditions in Market 1. The division managers' ability to engage in coordinated adaptation therefore depends on the breadth of their information while their ability to engage in autonomous adaptation depends on their depth.

The decomposition of the overall change in production into autonomous and coordinated adaptation highlights that coordination is an input into adaptation that allows the division managers to adapt to demand shocks more aggressively. Moreover, the more interdependent the production decisions, the more important coordinated adaptation–and thus breadth–becomes relative to autonomous adaptation–and thus depth. To see this, notice that Division Manager 1's coordinated response–the second term on the RHS of (8)–is the product of his autonomous response and the

term $t^2/(1-t^2)$, which is increasing in $|t|$. The larger $|t|$ is, therefore, the larger Division Manager 1's coordinated response is relative to his autonomous response.

This last observation is also illustrated in Figures 3a and 3b. In Figure 3a the reaction functions were drawn for a smaller g (and thus a smaller $t \equiv gb$) and in Figure 3b they are drawn for a larger g. The figures show that while autonomous adaptation is the same for either value of g, coordinated adaptation is larger, the larger is g. The more interdependent the production decisions are , therefore, the more adaptation depends on the division managers' ability to adapt to demand shocks in a coordinated fashion–and thus the breadth of their information–than their ability to adapt to them autonomously–and thus the depth of their information.

C. Organizational Performance

Since decision rules q_1 and q_2 in (6) and (7) are linear in the conditional expectations we can write expected profits $\Pi(q_1, q_2)$ as

$$\Pi(q_1, q_2) = \Pi(\overline{q}, \overline{q}) + \Pi(q_1 - \overline{q}, q_2 - \overline{q}). \tag{10}$$

The first term on the RHS are the "rigid profits"–the firm's expected profits if it does not adapt production to demand shocks–and the second term is the "gains from adaptation"–the additional profits the firm expects to realize if it does adapt production to demand shocks. If the division managers were perfectly informed about the demand conditions in both markets we would have

$$\Pi(q_1^*, q_2^*) = \frac{b(\mu - c)^2}{2(1+t)} + \frac{b\sigma^2}{2(1-t^2)}, \tag{11}$$

where q_1^* and q_2^* are the first best decision rules. Notice that the gains from adaptation are increasing in the variability of demand as measured by the variance σ^2. To the extent that the firm can adapt production to demand shocks, it therefore benefits from more demand variability. Moreover, the more price sensitive demand is, that is, the larger b is, the more the firm benefits from demand variability. This reflects the standard intuition that the more price sensitive demand is, the more important it becomes for the firm to adapt production to demand shocks. We will return to this intuition when we discuss competition below.

We can now write expected profits as

$$\Pi(q_1, q_2) = \Pi(q_1^*, q_2^*) - \frac{b}{2}\left(V_{\mathrm{d}} + \frac{t^2}{1-t^2}V_{\mathrm{b}}\right). \tag{12}$$

Recall that the division managers make the first best decisions on average and thus realize first best rigid profits. The second term on the RHS of the above expression is therefore the difference in the first best and the actual gains from adaptation. This term reflects our two key insights from the

previous section. First, the division managers' ability to adapt to demand shocks depends on both the depth and the breadth of their information. Second, the more interdependent the production decisions are, the more important is the division managers' ability to adapt in a coordinated fashion–and thus the breadth of their information–relative to their ability to adapt to demand shocks autonomously–and thus the depth of their information.

D. The Optimal Organizational Structure

So far we have focused on Decentralization. Since the organizational structures only differ in terms of the decision makers' information, however, we can use the expressions we derived above to characterize the two centralized organizations. Consider, for instance, Centralization, in which case both decisions are made by the headquarter manager. Since the headquarter manager only receives one signal for each market, we can derive her decision rules by substituting $head_j$ for both loc_j and div_j in the decision rules (6) and (7). Similarly, we can obtain expected profits by substituting V for both V_{d} and V_{b} in (12). The difference in expected profits between Centralization and Decentralization are therefore given by

$$\Pi^D - \Pi^C = \frac{b}{2}\left[(V - V_{\mathrm{d}}) - \frac{t^2}{1-t^2}(V_{\mathrm{b}} - V)\right], \tag{13}$$

where "C" and "D" indicate "Centralization" and "Decentralization." The two terms in the squared brackets represent the costs and benefits of decentralization: on the one hand, the division managers know more about their markets and are therefore better at adapting to demand shocks autonomously; on the other hand, however, the headquarter manager knows more than each division manager does about the other market and is therefore better at coordinating her responses. As anticipated, the costs of decentralization outweigh the benefits if the degree of interdependence is sufficiently large. Indeed, no matter how large the division manager's advantage in terms of depth $V - V_{\mathrm{d}} > 0$, and how small the headquarter manager's advantage in terms of breadth $V_{\mathrm{b}} - V > 0$, there exists a critical level of interdependence above which the centralized structure is better at adapting to demand shocks than the decentralized one.

Consider next Divisional Centralization. We can obtain the decision rules under this organizational structure by substituting loc_1 for div_1 and div_2 for loc_2 in the decision rules (6) and (7). Moreover, it is routine to show that the difference in expected profits between Decentralization and Divisional Centralization is then given by

$$\Pi^D - \Pi^{DC} = \frac{b}{4}(V_{\mathrm{b}} - V_{\mathrm{d}})\left(1 - \frac{t^2}{1-t^2}\right). \tag{14}$$

As in the comparison between Centralization and Decentralization, therefore, the centralized structure, in this case Divisional Centralization, outperforms the decentralized one if the production

decisions are sufficiently interdependent. Since the division managers don't differ in terms of the depth and breadth of their information, however, the difference in the residual variances now does not affect the choice between the two organizational structures.

The only remaining issue then is the choice between Centralization and Divisional Centralization. From (13) and (14), the difference in expected profits between these two organizational structures is given by

$$\Pi^{DC} - \Pi^{C} = \frac{b}{4}\frac{1}{1-t^2}\left[(V - V_{\mathrm{d}}) - (V_{\mathrm{b}} - V)\right].$$

Once again, therefore, Centralization outperforms the alternative, in this case Divisional Centralization, if the headquarter manager's advantage in terms of breadth outweighs the division managers' advantage in terms of depth. The degree of interdependence, however, now does not affect the choice between the two organizational structures. The reason is that both managers are equally good at coordinated adaptation since each now controls both decisions. The headquarter manager, however, is better at adapting q_2 to θ_2 autonomously because of his breadth of information, whereas Manager 1 is better at adapting q_1 to θ_1 autonomously because of his depth of information. Centralization therefore outperforms Divisional Centralization if and only if the headquarter manager's advantage in terms of breadth outweighs the division manager's advantage in terms of depth, that is, if and only if

$$V_{\mathrm{b}} - V > V - V_{\mathrm{d}}. \tag{15}$$

An interpretation of this condition is that there are decreasing returns to knowledge specialization, which is a feature of most commonly used information technologies, such as drawing independent signals to learn about a random variable.[7] For the remainder of this paper, we therefore assume that (15) holds and abstract from Divisional Centralization. Notice, however, that even if (15) does not hold, it is still the case that a centralized structure outperforms the decentralized one if the production decisions are sufficiently interdependent. The only difference is that the centralized structure then refers to Divisional Centralization rather than Centralization and that the critical degree of interdependence does not depend on the depth and the breadth of the decision makers' information. We can now summarize the previous discussion in our main proposition.

PROPOSITION 1. *If* (15) *holds, then Centralization is optimal if*

$$\frac{t^2}{1-t^2} > \frac{V - V_{\mathrm{d}}}{V_{\mathrm{b}} - V}$$

[7] Indeed, one can think of each manager being able to take a fixed number of draws to learn about Market 1 or Market 2. Additional draws then typically reduce the residual variance less than initial draws. This will be the case, for example, if both the signals and the random variables are normally distributed.

and Decentralization is optimal otherwise. If (15) does not hold, then Divisional Centralization is optimal if

$$\frac{t^2}{1-t^2} > 1$$

and Decentralization is optimal otherwise.

As anticipated, a centralized structure is better at adapting to changes in the environment than a decentralized one when decisions are sufficiently interdependent. As such, the model sheds light on the debate about the relative performance of centralized and decentralized organizations in adapting to changes in the environment. It captures both the importance of coordinated adaptation stressed by Barnard and Williamson as well as that of autonomous adaptation stressed by Hayek and others. And it formalizes Williamson's intuition that, overall, "*The authority relationship (fiat) has adaptive advantages over autonomy for transactions of the bilaterally (or multi-laterally) dependent kind.*"

IV. Competition and Adaptation

To explore some of the implications of the result we derived in the previous section, we now examine the effect of an increase in competition on the firm's organizational structure. As discussed in the Introduction, we argue that an increase in competition does not only make adaptation more important. Instead, it also changes how a firm adapts to demand shocks, putting relatively more weight on coordinated adaptation than on autonomous adaptation. As a result, an increase in competition actually favors a centralized structure over a decentralized one.

In general, there are many channels through which competition can affect the behavior of firms. In line with much of the Industrial Organization literature and, in particular, the literature on the effect of competition on the power of managerial incentives and innovation, we focus on the effect of competition on demand (see, for instance, Raith 2003 and Vives 2008). It is well known that an increase in competition can shift firms' inverse demand functions downwards, or it can make demand more price sensitive, or both (Vives 2008). In this section, we therefore explore how the difference in expected profits between Centralization and Decentralization depend on the average intercept of the inverse demand functions μ and on the price sensitivity of demand b. In Appendix A we then provide micro-foundations for this reduced-form approach by modeling competition explicitly and confirming that–depending on its underlying causes–competition either reduces μ, increases b, or both.

From the previous section it is immediate that a reduction in μ alone does not affect the difference in expected profits between Centralization and Decentralization (13). Such a reduction

does diminish the firm's rigid profits. But since rigid profits are equal to first best under both organizational structures it does not affect their relative performance. In our setting, an increase in competition can therefore only affect the firm's organization if it makes demand more price sensitive.

Consider then the effect of an increase in the price sensitivity of demand b. Differentiating the difference in expected profits (13) we have that

$$\frac{\mathrm{d}\left(\Pi^D - \Pi^C\right)}{\mathrm{d}b} = \frac{\partial\left(\Pi^D - \Pi^C\right)}{\partial b} + \frac{\mathrm{d}\left(\Pi^D - \Pi^C\right)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}b},$$

where the first term on the RHS is the "direct effect" and the second is the "indirect effect" that works through t. The direct effect captures the standard intuition that an increase the price sensitivity of demand makes adaptation more important and thus favors the organizational structure that is better at adapting to demand shocks. In particular, it follows from (13) that we can write the direct effect as

$$\frac{\partial\left(\Pi^D - \Pi^C\right)}{\partial b} = \frac{1}{b}\left(\Pi^D - \Pi^C\right).$$

The direct effect therefore magnifies the difference in expected profits between the two organizational structures. Notice, however, that it does not affect the difference in expected profits for a marginal firm that is just indifferent between Centralization and Decentralization. As such, the direct effect does not affect the choice between the two organizational structures. This choice therefore depends only on the indirect effect of an increase in the price sensitivity.

To understand the indirect effect, recall that an increase in the price sensitivity of demand b increases the degree of interdependence $|t| = |gb|$. From our discussion above, it then immediately follows that an increase in the price sensitivity favors coordinated adaptation and breadth over autonomous adaptation and depth. Formally, the indirect effect is given by

$$\frac{\mathrm{d}\left(\Pi^D - \Pi^C\right)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}b} = -\frac{t^2}{\left(1-t^2\right)^2}\left(V_{\mathrm{b}} - V\right) < 0 \quad \text{for all } t \neq 0.$$

For a marginal firm, the overall effect of an increase on the price sensitivity of demand is then to make the centralized structure strictly more profitable than the decentralized one. We therefore have the following proposition.

PROPOSITION 2. *An increase in competition that increases the price sensitivity of demand can result in a shift from Decentralization to Centralization but not the reverse. An increase in competition that does not affect the price sensitivity of demand has no effect on the firm's choice between Centralization and Decentralization.*

A few empirical studies examine the effect of competition on the delegation of decision rights to managers of manufacturing plants and find a positive relationship (see, for instance, Colombo and Delmastro 2004 and Bloom, Sadun, and Van Reenen 2010). The delegation of decision rights to the manager of a manufacturing plant, however, is conceptually different from the delegation of decision rights from headquarters to a division managers, as studied in this paper. For example, more intense competition may make cost reductions more important relative to innovation and brand differentiation. This, in turn, may make it optimal to shift control rights from the marketing and brand managers to plant managers in charge of manufacturing.[8] Delegation may also speed up decision-making, as in Aoki (1986). This aspect of delegation is absent from our framework, but may make delegation more attractive in a more competitive environment.

While not directly touching on the impact of competition on organizational structure, Guadalupe, Li, and Wulf (2013) provide suggestive evidence that large US firms have become more centralized in recent decades. They document a phenomenon termed "functional centralization," where more functional managers (e.g. marketing and R&D specialists) join the executive team surrounding the CEO and where the responsibilities of division managers are reduced (as reflected in lower pay for the divisional managers). As an example, they discuss IBM in the mid 1990's when the new CEO, Lou Gerstner, centralized select functional activities to move away from the "balkanized IBM of the early 1990's."[9] Lou Gerstner, for instance, appointed a Chief Marketing Officer who controlled many of the marketing activities across the various business units. Another example is Procter & Gamble which, in 1989, shifted towards a matrix organization that included functional senior vice presidents to manage functions across business units. While there are many plausible reasons for why IBM and Procter & Gamble may have centralized authority, both organizations were operating in industries that became increasingly competitive.

V. Strategic Communication

In this section we extend our main model by allowing for incentive conflicts and strategic communication. To do so, we draw upon the approach in Alonso, Dessein, and Matouschek (2008) and Rantakari (2008) by assuming that division managers are biased towards their own divisions and communicate their private information using cheap talk. The headquarter manager, in contrast,

[8] In contrast to a plant manager, a division manager has profit and loss responsibility and is typically in charge of coordinating sales, marketing and manufacturing for a product or product range. As documented by Bushman et al. (1995), division managers typically report to a group manager or a COO (Chief Operating Officer), who has the responsibility for coordinating multiple divisions and can be seen as headquarters in our model. In large organizations, several layers of group managers may exist, with higher-level group managers coordinating the decisions of several lower-level group managers (who in turn coordinate the decisions of several division managers) .

[9] Gerstner, Louis, Jr., "Who Says Elephants Can't Dance," Harper Collins, New York, p. 77, 2002, cited by Guadalupe et al. (2012).

remains unbiased but observes no private information. We use this version of the model to show that strategic communication can give rise to the generalist-specialist information structure that is central to our model, even when there are neither economies to knowledge specialization nor cognitive constraints. Focusing on the case where the own divisions biases are small, we further show that centralized decision-making is preferred to decentralized decision-making whenever production decisions are sufficiently complementary. This result stands in sharp contrast with Alonso, Dessein, and Matouschek (2008) who show that decentralization is always optimal when the own division bias is small. If production decisions are strategic substitutes, in contrast, we confirm the qualitative results in Alonso, Dessein, and Matouschek (2008).

A. Allowing for Incentive Conflicts and Strategic Communication

Specifically, we make two changes to the model that we explored above. First, we now assume that the division managers are biased towards their own divisions. In particular, Division Manager $j = 1, 2$ maximizes $\lambda\pi_j + (1 - \lambda)\,\pi_k$, where $k \neq j$ and where $\lambda \in [1/2, 1]$ is Division Manager j's own-division bias. In contrast to the division managers, the headquarter manager continues to care about overall profits. We follow the modelling approach in Alonso, Dessein, and Matouschek (2008) and assume that these preferences are exogenously given.

Second, we now assume that after the firm has decided between Centralization, Divisional Centralization, and Decentralization, Division Manager $j = 1, 2$ observes private signal $priv_j = \theta_j$, where θ_1 and θ_2 are independently drawn from a uniform distribution with support $[-s, s]$, where $s > 0$. The division managers are therefore perfectly informed about the demand conditions in their respective markets and, as a result, the residual variance V_{d} is equal to zero. Next, the division managers engage in one round of cheap talk communication. In particular, under Decentralization, Division Manager $j = 1, 2$ sends message $div_j \in M$ to Division Manager $k \neq j$, where the message space M is the support of their private information $[-s, s]$. Under Divisional Centralization, Division Manager 2 sends message $div_2^{DC} \in M$ to Division Manager 1. Similarly, under Centralization, Division Managers 1 and 2 simultaneously send messages $head_1 \in M$ and $head_2 \in M$ to headquarters. The breadth of the division managers information under Decentralization V_{b}, and under Divisional Centralization V_{b}^{DC}, is now endogenously determined by the quality of horizontal communication. Similarly, the depth and the breadth of the headquarter manager's information V is endogenously determined by the quality of vertical communication. The generalist-specialist information structure then arises if and only if $0 < V < \min\{V_{\mathrm{b}}, V_{\mathrm{b}}^{DC}\}$.

Next we informally discuss the Perfect Bayesian Equilibria of this game. The formal analysis is in Appendix B.

B. The Generalist-Specialist Informatin Structure

Consider first decision making for given posteriors about the demand conditions. The division managers' bias does not affect decision making under Centralization. To obtain the decision rules under Decentralization, we simply need to substitute $\tau \equiv t/(2\lambda)$ for t in the decision rules (6) and (7). In terms of decision making, the only effect of the own-division bias is therefore to reduce the weight that the division managers put on the interdependence between their decisions. Finally, we can obtain the decision rules under Divisional Centralization by first substituting τ for t in reaction function (4) and $\tau\lambda/(1-\lambda)$ for t in (5), and then solving for the equilibrium decision rules.[10]

To derive the posteriors on which the decision rules are based, consider next the communication sub-game. It follows from Crawford and Sobel (1982) that all equilibria of the communication sub-game are partition equilibria in which the state space is partitioned into a finite number of intervals and the division managers only indicate which interval their demand conditions fall into. Formally, in a partition equilibrium, the state space for Market $j = 1, 2$ is partitioned into $(a_0^j, a_1^j, ..., a_{N^j}^j)$, where N^j is the number of intervals and where $a_0^j = \mu - s$ and $a_{N^j}^j = \mu + s$. The length of the intervals is then determined by the difference equation

$$a_{i+1} - a_i = a_i - a_{i-1} + 4b^s(a_i) \text{ for } s = C, D, DC, \tag{16}$$

where $b^s(a_i)$ is the endogenously determined "communication bias." If the absolute size of the communication bias is equal to or larger than 1/4 for all states θ_j for $j = 1, 2$, there can only be a single interval. In this case, the only communication equilibrium is the babbling one. If the absolute size of the communication bias is less than 1/4 for some states, however, there can be multiple equilibria that differ in the number of intervals. We follow the literature by focusing on the most informative communication equilibrium in which the number of intervals is maximized.

To characterize the most informative equilibrium, we first need to characterize the communication biases under the three organizational structures, which we do in the following lemma.

LEMMA 1. (i.) *Under Centralization, the communication bias exists and is given by*

$$b^C = (t - \tau)(\mu - c), \tag{17}$$

with $sign(b^C) = sign(t)$. *If* $\left|b^C\right| \geq 1/4$, *informative communication is not feasible.*

(ii.) *Under Decentralization, if* $(2\lambda - 1)^2 < \left(1 - t^2\right)$, *then the communication bias exists and is given by*

$$b^D(\theta_j) = 2\tau(t - \tau)\frac{(\mu + \theta_j - c) - \tau(\mu - c)}{\tau^2(1 - t^2) - (t - \tau)^2}, \tag{18}$$

[10] We provide details of these calculations in the proof of Lemma 1.

where $b^D(\theta_j)$ *is always positive. If* $(2\lambda-1)^2 \geq \left(1-t^2\right)$*, the communication bias does not exist, and informative communication is not feasible.*

(iii.) *Under Divisional Centralization, the communication bias exists and is given by*

$$b^{DC}(\theta_2) = (t-\tau)\frac{(\mu+\theta_2-c)-\lambda\frac{1-\tau^2}{t-\tau}(\mu-c)}{(1-\lambda)\left(\tau-\lambda\frac{1-\tau^2}{t-\tau}\right)}, \tag{19}$$

with $sign(b^{DC}) = sign(t)$.

Notice that under Centralization the communication bias is constant. The communication equilibria under this organizational structure are therefore as in the well-known constant bias example in Crawford and Sobel (1982). Under both Decentralization and Divisional Centralization, instead, the communication bias depends on the realization of the state θ_j for $j = 1, 2$. Under these organizational structures, the communication equilibria are therefore more similar to those in Alonso, Dessein, and Matouschek (2008) and Rantakari (2008), where the communication bias is also state dependent.[11]

We will see below that differences in the efficiency of communication across organizational structures depend crucially on differences between the communication biases. One key difference is that under the two centralized structures the sign of the bias depends on whether decisions are strategic substitutes or complements while under the decentralized structure the bias always has the same sign. To understand why this is the case, consider first Decentralization. A division manager would then like to induce less production by his counterpart if decisions are strategic substitutes and he would like to induce more production if decisions are strategic complements. In either case, he can do so by convincing his counterpart that–because of very favorable demand conditions in his own market–he will increase production himself. Under Decentralization, division managers therefore always have an incentive to over-report the demand conditions in their markets, which is why $b^D(\theta_j)$ is always positive.

This is not the case under the centralized structures. To see why, consider the incentives of a division manager to misrepresent the demand conditions in his market under Centralization (the argument for Divisional Centralization is analogous). Importantly, any such misrepresentation now affects production by both divisions. Holding constant production by the division manager's own division, the incentives to misrepresent the demand conditions in his market are similar to those

[11] In contrast to those models, however, in our setting the communication biases $b^D(\theta_j)$ and $b^{DC}(\theta_2)$ are never equal to zero, that is, there is never a point of congruence between the sender and the receiver under Decentralization or Divisional Centralization. As a result, the maximum number of partitions is always finite, which complicates the derivation of the residual variance.

under Decentralization. Holding constant production by the other division, however, the division manager would like to induce more production when decisions are strategic substitutes and less when they are strategic complements. The reason is that headquarters puts more weight on the externality generated by the division's production. If decisions are strategic substitutes, the division manager therefore has an unambiguous incentive to over-report the demand conditions in his market. If decisions are strategic complements, however, the division manager faces *countervailing incentives,* where the net effect is that he would want to under-report the demand conditions in his market.

Having characterized the communication biases, we can now examine the efficiency of communication under the different organizational structures.

LEMMA 2. (i.) *Suppose that* $t < 0$. *Then for any* $\lambda > 1/2$ *we have* $V \leq V_{\mathrm{b}}$ *with strict inequality whenever* $V < \sigma^2$.

(ii.) *There exists a critical value* $t^* > 0$ *and an* $\varepsilon > 0$ *such that for* $\lambda \in (1/2, 1/2 + \varepsilon)$ *we have:* (*a.*) $V < V_{\mathrm{b}}^{DC}$ *for* $t \in (-1, 1)$ *and* (*b.*) $V < V_{\mathrm{b}}$ *if and only if* $t < t^*$.

The lemma provides conditions under which the generalist-specialist information structure arises endogenously and we illustrate them in Figure 4. The first part of the lemma focuses on Centralization and Decentralization and shows that such a generalist-specialist structure then arises endogenously when decisions are strategic complements. The reason is that, as we just saw, the division managers face countervailing incentives when they communicate with headquarters under Centralization but not when they communicate with each other under Decentralization. As a result, communication under Centralization is more efficient than communication under Decentralization, generating a generalist-specialist information structure for any own-division bias λ.

INSERT FIGURE 4 HERE

The second part of the lemma compares communication under all three organizational structures when incentive conflicts are small. It shows that communication under Centralization is always more efficient than communication under Divisional Centralization. And it shows that communication under Centralization is more efficient than communication under Decentralization even when decisions are strategic substitutes, provided that the interdependency is not too strong. For small incentive conflicts, the generalist-specialist information structure therefore arises endogenously even once we allow for all three organizational structures, provided that strategic substitutability is not too strong.

To understand why the generalist-specialist information structure does not arise when the strategic substitutability is sufficiently strong, recall our discussion above of the communication bias.

There we saw that when decisions are strategic substitutes, division managers have an incentive to over-report their demand conditions, both to increase production by their own division and to reduce production by the other division. Because of these reinforcing incentives, division managers actually share less information with headquarters than with each other. As a result, the division managers' information under Decentralization is then both deeper and broader than that of the headquarter manager under Centralization.

C. The Optimal Organizational Structure

The next question is whether the headquarter manager's unbiased decision making and endogenous advantage in terms of breadth can outweigh the division managers' endogenous advantage in terms of depth. With biased division managers, the difference in expected profits between Decentralization and Centralization is given by

$$\begin{aligned}\Pi^D - \Pi^C &= -\frac{b\left(t-\tau\right)^2}{2\left(1+t\right)\left(1+\tau\right)^2}\left(\mu-c\right)^2 - b\left(t-\tau\right)^2\frac{\left(1-t^2\right)+\left(t+\tau\right)^2}{2\left(1-\tau^2\right)^2\left(1-t^2\right)}\left(\sigma^2-V_{\mathrm{b}}\right) \\ &\quad +\frac{b}{2}\left[\left(V-V_{\mathrm{d}}\right)-\frac{t^2}{1-t^2}\left(V_{\mathrm{b}}-V\right)\right]\end{aligned} \tag{20}$$

and the difference in expected profits between Divisional Centralization and Centralization $\Pi^{DC} - \Pi^C$ is given by

$$\begin{aligned}&-\frac{b}{4}\left(t-\tau\right)^2\frac{\left(k-2k\tau+1\right)\left(-2k^2\tau^2+k^2+1\right)}{\left(k+1\right)\left(1-k\tau^2\right)^2\left(1-t^2\right)}\left(\mu-c\right)^2 \\ &-\frac{b}{4}\frac{\left(t-\tau\right)^2\left[\lambda\left(k\left(1+k+\tau^2\left(1-3k\right)\right)\right)\left(\sigma^2-V_{\mathrm{d}}\right)+\left(1-\lambda\right)\left(1+k+k^2\tau^2\left(k-3\right)\right)\left(\sigma^2-V_{\mathrm{b}}\right)\right]}{\left(1-k\tau^2\right)^2\left(1-t^2\right)} \\ &+\frac{b}{4}\frac{1}{1-t^2}\left(\left(V-V_{\mathrm{d}}\right)-\left(V_{\mathrm{b}}-V\right)\right),\end{aligned} \tag{21}$$

where, $k \equiv \lambda/(1-\lambda)$. The first term of each expression is the difference in rigid profits and the second and third terms are the difference in the gains from adaptation. Specifically, the second term is the difference in the gains from adaptation that is due to differences in decision making for given information while the third term is the difference in the gains from adaptation that is due to differences in information. As such, the first two terms capture the standard loss of control and the third captures the gain in information. Since the division managers are biased, the loss of control is always positive, that is, the first two terms are always negative. In contrast to standard models of delegation, however, the gain in information from decentralization may actually be negative. The reason is that in standard models of delegation, such as Dessein (2002), the agent has more information than the principal while, in our setting, the agents have different, but not necessarily

more, information. As a result, decentralization can lead to a loss of information and centralization can be optimal. The next lemma shows that this can be the case even when the division managers' bias is small, in which case the standard loss of control from decentralization is second order.

PROPOSITION 3. *Suppose that the division managers' own-division bias is small. Then Centralized decision-making outperforms Decentralization if and only if production decisions are sufficiently complementary. Formally, there exists a critical value* $t^{**} < 0$ *such that*
(i.) *for* $t < t^{**}$, *there exists an* $\varepsilon > 0$ *such that* $\min\left\{\Pi^C, \Pi^{DC}\right\} > \Pi^D$ *for* $\lambda \in (1/2, 1/2 + \varepsilon)$
(ii.) *for* $t > t^{**}$, *there exists an* $\varepsilon > 0$ *such that* $\Pi^D \geq \max\left\{\Pi^C, \Pi^{DC}\right\}$ *for* $\lambda \in (1/2, 1/2 + \varepsilon)$.

We illustrate the proposition in Figure 5 which also provides additional information about the level of t^{**}. To understand the proposition, consider first the comparison between Centralization and Decentralization. To this end, fix a particular value of $s/(\mu - c)$ and suppose that $t > t^*$, where t^* is defined in Lemma 2. From that lemma we know that if $t > t^*$ communication under Decentralization is more efficient than communication under Centralization. In this case, each division manager therefore knows more about the demand conditions in both markets than headquarters. We thus obtain the standard result: the gain in information is positive and dominates the loss of control, making Decentralization preferable to Centralization. If, instead, $t < t^*$, communication under Centralization is more efficient than communication under Decentralization. As long as $t > t^{**}$, however, the headquarter manager's advantage in terms of breadth is small relative to the division managers' advantage in terms of depth. Moreover, production decisions are not sufficiently interdependent for the advantage in breadth to matter. As a result, Decentralization is still preferable to Centralization. If $t < t^{**}$, however, communication under Centralization is sufficiently more informative than communication under Decentralization – and decisions are sufficiently interdependent – that the headquarter manager's advantage in terms of breadth dominates the division managers' advantage in terms of depth. Centralization is then preferable to Decentralization, even though the incentive conflict is small.

It is noteworthy that Proposition 3 is qualitatively different from Proposition 5 in Alonso, Dessein, and Matouschek (2008), which states that decentralization is *always* optimal when the own-division bias is sufficiently small. In light of the similarities between the model in this section and that in Alonso, Dessein, and Matouschek (2008), this suggests that whether–in models of multi-divisional firms in which information is aggregated via cheap talk–decentralization is optimal for small incentive conflicts depends on the precise nature of externalities between divisions. Intuitively, there are no countervailing incentives in Alonso, Dessein, and Matouschek (2008) when division managers communicate with headquarters. Hence, strategic communication is similar in

nature to the case of strategic substitutes. While we confirm the results in Alonso, Dessein, and Matouschek (2008) when decisions are strategic substitutes, we obtain qualitatively different results when decisions are complements.[12]

INSERT FIGURE 5 HERE

So far we have focused on the comparison between Centralization and Decentralization. Since Centralization dominates Decentralization when $t < t^{**}$, we know that a centralized structure–either Centralization or Divisional Centralization–is optimal in this case. The only remaining question then is whether, for $t > t^{**}$, Decentralization not only dominates Centralization but also Divisional Centralization. The proposition shows that this is indeed the case. To understand why, recall that when $\lambda = 1/2$ the choice between Centralization and Divisional Centralization only depends on the difference $(V - V_{\mathrm{d}}) - (V_{\mathrm{b}} - V)$. In the proof of Lemma 2 we show that for $\lambda \approx 1/2$ the communication bias under divisional centralization is approximately twice the communication bias under centralization. To a first order, we therefore have $V_{\mathrm{b}} = 2V$ and thus $(V - V_{\mathrm{d}}) - (V_{\mathrm{b}} - V) = 0$. In other words, differences between Centralization and Divisional Centralization are only second order for $\lambda \approx 1/2$.[13] In contrast, when $\lambda \approx 1/2$ and $t > t^{**}$, Decentralization is preferred over Centralization to a first order.[14] It follows that Decentralization is then also preferred over Divisional Centralization.

VI. Conclusions

In this paper we explored the relationship between organizational structure and change. We showed that even in a very standard setting, a centralized structure may actually be better at adaptation than a decentralized one and that this may be the case even when division managers have superior information about their local conditions and incentive conflicts are negligible. The reason is that if decisions are interdependent, an effective response to a local shock requires coordinated change. If production is sufficiently interdependent, a headquarter manager with broad but shallow information about each market is then better able to adapt to local shocks than division managers with deep but narrow information about their respective markets. We then showed that this result reverses some of the standard intuitions about organizational structure, such as the notions that competition favors decentralization (Section IV) and that decentralized decision-making ought to be optimal provided that incentive conflicts are sufficiently small (Section V).

[12] See also Alonso (2007) for differences in cheap talk communication when decisions are either complements or substitutes.

[13] Formally, both $\Pi^{DC} - \Pi^{C} = 0$ and $\partial(\Pi^{DC} - \Pi^{C})/\partial\lambda = 0$ at $\lambda = 1/2$.

[14] Formally, while $\Pi^{D} - \Pi^{C} = 0$ at $\lambda = 1/2$, we then have that $\partial(\Pi^{D} - \Pi^{C})/\partial\lambda > 0$.

One potential application of our model are multi-national enterprises. Such firms are often multi-divisional firms in which each division is dedicated to a particular country. General Motors, for instance, started out as a domestic firm that only served the US and Canadian markets. In the late 1920s, however, it became an early multi-national with different subunits producing and selling cars in England, Germany, and other markets. In recent years the role of multi-nationals in the globalized world economy has been hotly debated among the general public, policy makers, and academics (see, for instance, Navaretti and Venables 2004). While some see multi-nationals as catalysts for local economies that transfer know-how and create jobs, others view them as threats to local wealth and national identities. In spite of this debate and an accompanying surge in research on multi-nationals, little is known about what determines the horizontal size of multi-nationals.[15] When, for instance, does globalization induce multi-nationals to expand into ever more countries? And when, instead, does it induce them to divest of some of their divisions, as in the recent breakup of DaimlerChrysler? While our model is explicitly not about the boundaries of multi-divisional firms, it does suggest channels through which globalization may affect the horizontal size of multi-nationals. If, for instance, competition makes adaptation more important, and adaptation is more easily achieved in centralized multi-divisional firms than in stand-alone firms, then competition may lead to the emergence of new and the expansion of existing multi-nationals. But if, instead, competition merely reduces externalities between competitors in the same industries, then it may limit the growth of multi-nationals and even induce them to divest of some of their divisions. A model of the horizontal boundaries of multi-divisional firms that could confirm or reject these speculations awaits future research.

[15] The literature on organizations and international trade focuses on the vertical size of multi-nationals, that is, the classic make versus buy decision, rather than on their horizontal boundaries (see Antràs and Rossi-Hansberg 2009 for a survey).

1 Appendix A: Competition and Adaptation

In this appendix, we explicitly model competition and its impact on demand and the firm's organization. For this purpose, we make three changes to the model described in Section II.

First, we now assume that each division faces a continuum of potential competitors. Each potential competitor consists of a single division and can only operate in one market. To enter a market, a competitor has to pay entry costs $K \geq 0$. We denote the measure of competitors that enter Market $j = 1, 2$ by n_j. To economize on notation, and without loss of generality, we set the competitors' production costs to zero.

Second, instead of simply assuming the divisions' inverse demand functions, we now derive them from the underlying parameters. For this purpose, suppose that there is a measure m of consumers in each market. The utility function of each consumer in Market $j = 1, 2$ is given by

$$U = (1 + \theta_j)\, q_j + \int_0^{n_j} q_{jk} \mathrm{d}k - \frac{1}{2}\left(q_j^2 + \int_0^{n_j} q_{jk}^2 \mathrm{d}k + 2\gamma q_j \int_0^{n_j} q_{jk} \mathrm{d}k\right),$$

where $\gamma \in [0, 1]$, q_{jk} is the amount demanded in Market j of the product of firm $k \in [0, n_j]$ and q_j the amount demanded of Division j's good. We continue to assume that θ_1 and θ_2 are independently drawn from a distribution with mean 0 and variance σ^2. The utility function implies that in each market the goods produced by the competitors are perfect substitutes for each other. Moreover, in each market the goods produced by the competitors are substitutes for the good produced by the local division, where γ measures the degree of substitutability. If $\gamma = 0$, the goods produced by the competitors are independent of those produced by the local division, and if $\gamma = 1$, they are perfect substitutes.

Finally, we need to change the timing of the game to incorporate the competitors. We continue to assume that the firm first decides on its organization and that the firm's decision makers then learn their information and decide on production levels. All potential competitors observe are the firm's production decisions. We make no assumptions about how much the competitors know about the realization of θ_1 and θ_2, or about the firm's organization. Each competitor then decides whether to enter and, if so, how much to produce. Finally, profits are realized and the game ends. Once again we solve for the Perfect Bayesian Equilibrium of this game.

In this setting, there are three parameters that determine the degree of competition that the divisions face: the competitors' entry costs K, market size m, and the degree of substitutability γ. Below we explore how changes in these parameters affect the firm's organization.

1.1 Organization and the Firm's Performance

We solve the game by backward induction, starting with the competitors' production and entry decisions. Then, given the competitors' behavior, we can derive the divisions' inverse residual demand functions. We summarize this result in the next lemma.

LEMMA A1. *Division $j = 1, 2$'s residual inverse demand function is given by*

$$p_j = \mu + \theta_j - \frac{1}{b} q_j, \tag{22}$$

where

$$\mu \equiv \gamma \left(1 - \sqrt{\frac{K}{m}}\right) \quad \textit{and} \quad b \equiv \frac{m}{1-\gamma^2}. \tag{23}$$

Proof: We start by deriving the competitors' production and entry decisions. Suppose that the firm has made production decisions q_1 and q_2 and that measures n_1 and n_2 of competitors have entered the respective markets. The inverse demand function faced by competitor $k \in [0, n_j]$ in Market $j = 1, 2$ is given by

$$p_{jk} = 1 - \frac{2}{m} q_{jk} - \frac{1}{m} Q_j - \frac{\gamma}{m} q_j, \tag{24}$$

where $Q_j = \int_0^{n_j} q_{jk} \mathrm{d}k$ is the total amount produced by the competitors in Market j.

Each competitor determines their quantity choice by solving

$$\max_{q_{jk}} p_{jk} q_{jk},$$

taking as given production of the firm and the other competitors. The first order condition is given by

$$q_{jk} = m - \gamma q_j - Q_j.$$

Note that this condition does not depend on θ_j. It is therefore irrelevant whether the competitors observe θ_j or not. Next we integrate over all n_j competitors to find that the total amount produced by the competitors in Market j is given by

$$Q_j = \frac{n_j}{n_j + 1} \left[m - \gamma q_j\right]. \tag{25}$$

Any competitor k that enters Market j thus produces

$$q_{jk} = \frac{1}{n_j} Q_j = \frac{1}{n_j + 1} \left[m - \gamma q_j\right]$$

and realizes profits

$$\pi_{jk} = \frac{\left[m - \gamma q_j\right]^2}{m \left(n_j + 1\right)^2} - K.$$

Finally, setting π_{jk} equal to zero and solving for n_j we find that a measure

$$n_j = \frac{1}{\sqrt{mK}}\left[m - \gamma q_j - \sqrt{mK}\right]$$

of competitors enters Market j. The residual inverse demand function of Division $j = 1, 2$ characterizes demand for its good given the competitors' behavior. Substituting the competitors' total production (25) in Division j's inverse demand (24), we find that Division j's residual inverse demand is given by (22) and (23). ■

1.2 The Effect of Competition on Organization

In contrast to our main model, the average intercept of the inverse demand functions μ and the slope parameter b are now functions of the underlying parameters. It follows from (22) and (23) that a reduction in the competitors' entry costs K reduces the average intercept but does not affect the price sensitivity of demand. In contrast, an increase in market size m or in the degree of substitutability γ, leads to an anti-clockwise rotation of the residual inverse demand functions. In other words, such changes in the environment reduce the average intercept and increase the price sensitivity. The effect of changes in these parameters on the optimal organizational structure is summarized in Proposition A1 that translates the findings in Proposition 2 to our specific competition model.

PROPOSITION A1. *a. A reduction in entry costs K has no effect on the choice between Centralization and Decentralization.*

b. There exists a critical value of the degree of substitutability $\gamma^* > 0$ and a critical value of the entry costs $K^* > 0$ such that for any $\gamma \leq \gamma^*$ and $K \leq K^*$, an increase in market size m can result in a shift from Decentralization to Centralization but never the reverse.

c. There exists a critical value of the degree of substitutability $\gamma^* < 1$ such that for any $\gamma \geq \gamma^*$ an increase in the degree of substitutability γ can result in a shift from Decentralization to Centralization but never the reverse.

Proof: Follows from Proposition 1 and our previous discussion of the effect of parameter changes on (22) and (23). ■

2 Appendix B: Strategic Communication

We divide the proofs for Section V into three appendices. Appendix B1 characterizes communication equilibria and provides the proof of Lemma 1 while Appendix B2 derives expressions for the quality of horizontal and vertical communication and provides the proof of Lemma 2. Appendix B3 uses the previous results to study the relative performance of Centralization, Divisional Centralization, and Decentralization when the own-division bias is vanishingly small and presents the proof of Proposition 3.

2.1 Appendix B1 - Communication Equilibria

A communication equilibrium under each organizational structure is characterized by (i.) communication rules for the division managers, (ii.) decision rules for the decision makers and (iii.) belief functions for the message receivers. The communication rule for Manager $j = 1, 2$ specifies the probability of sending message $m_j \in M_j$ conditional on observing state θ_j and we denote it by $\mu_j(m_j \mid \theta_j)$, where the message space is $M_j = [-s, s]$. Under Centralization, the decision rules map messages $m = (m_1, m_2)$ into decisions $q_1 \in \mathbf{R}^+$ and $q_2 \in \mathbf{R}^+$, and we denote them by $q_1^C(m)$ and $q_2^C(m)$. Under Decentralization, the decision rule for Manager 1 maps the state θ_1 and messages $m = (m_1, m_2)$ into decision $q_1 \in \mathbf{R}^+$ while the decision rule for Manager 2 maps the state θ_2 and messages $m = (m_1, m_2)$ into decision $q_1 \in \mathbf{R}^+$, and we denote them by $q_1^D(m, \theta_1)$ and $q_2^D(m, \theta_2)$. Under Divisional Centralization, the decision rules map the state θ_1 and message $m = m_2$ into decisions $q_1 \in \mathbf{R}^+$ and $q_2 \in \mathbf{R}^+$, and we denote them by $q_1^{DC}(m, \theta_1)$ and $q_2^{DC}(m, \theta_1)$. Finally, the belief functions are denoted by $g_j(\theta_j \mid m_j)$ for $j = 1, 2$ and characterize the receiver's posterior probability of state θ_j conditional on receiving message m_j.

We focus on Perfect Bayesian Equilibria of the communication subgame which require that communication rules are optimal for the division managers given decision rules. Formally, whenever $\mu_j(m_j \mid \theta_j) > 0$ then $m_j \in \arg\max_{m \in M_j} \mathrm{E}\left[\lambda \pi_j^l + (1-\lambda)\pi_k^l \mid \theta_j\right]$ for $l \in \{C, D, DC\}$, where π_j^l and π_k^l are, respectively, the profits of Divisions j and k, $j \neq k$, given that decisions are made according to $q_1^l(\cdot)$ and $q_2^l(\cdot)$. Perfect Bayesian Equilibria also require that the decision rules are optimal for the decision makers given the belief functions. Thus, under Centralization $q_1^C(\cdot)$ and $q_2^C(\cdot)$ solve $\max_{(q_1,q_2)} \mathrm{E}[\pi_1 + \pi_2 \mid m]$, under Decentralization $q_j^D(\cdot)$ solves $\max_{q_j} \mathrm{E}[\lambda \pi_j + (1-\lambda)\pi_k \mid m, \theta_j]$, and under Divisional Centralization $q_1^{DC}(\cdot)$ and $q_2^{DC}(\cdot)$ solve $\max_{(q_1,q_2)} \mathrm{E}[\lambda \pi_1 + (1-\lambda)\pi_2 \mid m, \theta_1]$. Finally, Perfect Bayesian Equilibria require that the belief functions are derived from the communication rules using Bayes' rule whenever possible, that is, $g_j(\theta_j \mid m) = \mu_j(m_j \mid \theta_j) / \int_P \mu_j(m_j \mid \theta_j) \mathrm{d}\theta_j$, where $P = \left\{\theta_j : \mu_j(m_j \mid \theta_j) > 0\right\}$, $j = 1, 2$.

PROPOSITION B1. *For* $\lambda \in (1/2, 1]$ and $t \neq 0$ *there exists an integer* $\bar{N}(\lambda, t)$*, such that for all* $N \leq \bar{N}(\lambda, t)$ *there exists at least one equilibrium* $(\mu_1(\cdot)\ \mu_2(\cdot)\ , q_1(\cdot),\ q_2(\cdot),\ g_1(\cdot), g_2(\cdot))$*, where*

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

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

c. $a_{j,i+1} - a_{j,i} = a_{j,i} - a_{j,i-1} + 4b^s(a_{j,i})$ *for* $i = 1, ..., N_j - 1$

with $b^s = b^C$ *under Centralization as given by (17),*
$b^s = b^D$ *under Decentralization as given by (18),*
$b^s = b^{DC}$ *under Divisional Centralization as given by (19).*

d. $q_j(m) = q_j^C$, $j = 1, 2,$ *under Centralization, where* q_j^C *are given by (27) and (28), and*

$q_j(m, \theta_1) = q_j^{DC}$, $j = 1, 2,$*under Divisional Centralization, with* q_j^D *as in (29) and (30), and*
$q_j(m, \theta_j) = q_j^D$, $j = 1, 2,$ *under Decentralization, where* q_j^D *are given by (31) and (32).*

Proof: We first show that communication equilibria are interval equilibria. For the case of Centralization let $\mu_2(\cdot)$ be any communication rule for Manager 2. The expected utility of Manager 1 if the headquarter manager holds a posterior expectation ν_1 over θ_1 is given by

$$\mathrm{E}_{\theta_2}[U_1 \mid \theta_1, \nu_1] = \mathrm{E}_{\theta_2}\left[\lambda \pi_1\left(\widehat{q}_1^C, \widehat{q}_2^C, \theta_1\right) + (1-\lambda)\pi_2\left(\widehat{q}_1^C, \widehat{q}_2^C\right)\right], \tag{26}$$

with

$$\hat{q}_1^C \equiv \frac{b}{2(1-t^2)}(\mu - c + \nu_1 - t\mathrm{E}[\mu - c + \theta_2 \,|\mu_2(\cdot)]), \text{ and} \tag{27}$$

$$\hat{q}_2^C \equiv \frac{b}{2(1-t^2)}(\mathrm{E}[\mu - c + \theta_2 \,|\mu_2(\cdot)] - t(\mu - c + \nu_1)). \tag{28}$$

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

Now consider the case of Divisional Centralization. The expected utility of Manager 2 if Manager 1 holds a posterior expectation ν_2 over θ_2 is given by

$$\mathrm{E}_{\theta_1}[U_2 \mid \theta_2, \nu_2] = \mathrm{E}_{\theta_1}\left[(1-\lambda)\pi_1\left(\widehat{q}_1^{DC}, \widehat{q}_2^{DC}\right) + \lambda\pi_2\left(\widehat{q}_1^{DC}, \widehat{q}_2^{DC}, \theta_2\right)\right],$$

with

$$\hat{q}_1^{DC} \equiv \frac{b}{2(1-\frac{\lambda\tau^2}{1-\lambda})}\left(\mu - c + \theta_1 - \tau\left(\mu - c + \nu_2\right)\right), \text{ and} \tag{29}$$

$$\hat{q}_2^{DC} \equiv \frac{b}{2(1-\frac{\lambda\tau^2}{1-\lambda})}\left(\left(\mu - c + \nu_2\right) - \frac{\lambda}{1-\lambda}\tau\left(\mu - c + \theta_1\right)\right). \tag{30}$$

Again, it can be shown that $\frac{\partial^2}{\partial\theta_1\partial\nu_1}\mathrm{E}_{\theta_2}\left[U_2 \mid \theta_2, \nu_2\right] > 0$ and $\frac{\partial^2}{\partial^2\theta_1}\mathrm{E}_{\theta_1}\left[U_2 \mid \theta_2, \nu_2\right] < 0$. By the same argument used previously for Centralization, we can then conclude that all equilibria in this case are, again, interval equilibria.

For the case of Decentralization let $\mu_1(\cdot)$ and $\mu_2(\cdot)$ be communication rules of Manager 1 and Manager 2, respectively. Sequential rationality implies that, in equilibrium, decision rules must conform to

$$\hat{q}_1^{D} = \frac{b}{2}\left(\mu - c + \theta_1\right) - \tau\left(\mathrm{E}\left[q_2 \,|\nu_1, m_2\right]\right), \text{ and} \tag{31}$$

$$\hat{q}_2^{D} = \frac{b}{2}\left(\mu - c + \theta_2\right) - \tau\left(\mathrm{E}\left[q_1 \,|\nu_1, m_2\right]\right), \tag{32}$$

where ν_1 denotes Manager 2's posterior expectation over θ_1. It can readily be seen that

$$\frac{\partial^2}{\partial\theta_1\partial\nu_1}\mathrm{E}_{\theta_2}\left[U_1 \mid \theta_1, \nu_1\right] > 0$$

and the proof follows as in the preceding paragraph.

We now characterize all equilibria of the communication game. For Manager $j = 1, 2$, let a_j be a partition of $[-s, s]$, any message $m_j \in (a_{j,i-1}, a_{j,i})$ be denoted by $m_{j,i}$, and $\overline{m}_{j,i}$ be the receiver's posterior belief of the expected value of θ_j after receiving message $m_{j,i}$.

a. Centralization: The expected utility of Manager 1 in state $a_{1,i}$ is given by

$$\begin{aligned}\mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i}\right] = {} & \frac{b}{4\left(1-t^2\right)}[\left(\mu - c\right)^2\left(1-\lambda\right) - \lambda\left(\overline{m}_{1,i}\right)^2 + \\ & \overline{m}_{1,i}\left(\mu - c\right)\left(2\lambda\left(\mu - c + a_{1,i}\right) + \left(2\lambda - 1\right)t\right) - 2t\lambda\left(\mu - c\right)\left(\mu - c + a_{1,i}\right)].\end{aligned}$$

In state $a_{1,i}$ Manager 1 must be indifferent between sending a message that induces a posterior $\overline{m}_{1,i}$ and a posterior $\overline{m}_{1,i+1}$ implying that $\mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i}\right] - \mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i+1}\right] = 0$. Given decision rules (27) and (28) , and letting $\overline{m}_{1,i} = \left(a_{1,i-1} + a_{1,i}\right)/2$ we have that $\mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i}\right] - \mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i+1}\right] = 0$ if and only if $a_{1,i} = \left(a_{1,i-1} + a_{1,i+1}\right)/\left(2 + 4b^C\right)$ where b^C is given by (17). That is, communication equilibria under Centralization are equivalent to the constant-bias leading example in Crawford and Sobel (1982).

From (17), we have that $sign(b^C) = sign(t - \tau) = sign(t)$.

b. Divisional Centralization: Let $k = \lambda/(1-\lambda)$. The expected utility of Manager 2 in state $a_{2,i}$ is given by

$$\begin{aligned}
\mathrm{E}_{\theta_1}\left[U_2 \mid a_{2,i}, \overline{m}_{2,i}\right] &= \frac{b(1-\lambda)}{4\left(1-k\tau^2\right)^2}\left[\left(2k\left(k-1\right)\tau^2+1-k^3\tau^2\right)\mathrm{E}_{\theta_1}\left[\left(\mu-c+\theta_1\right)^2\right]\right. \\
&\quad ... + \left(\overline{m}_{2,i}\right)^2\left(2k\tau^2-k-\tau^2\right)+2k\overline{m}_{2,i}\left(\mu-c+\theta_2\right)\left(1-k\tau^2\right)+ \\
&\quad ... + 2k\left(k-1\right)\tau\left(1-\tau^2\right)\left(\mu-c\right)\overline{m}_{2,i}-2k^2\tau\left(\mu-c\right)\left(\mu-c+\theta_2\right)\left(1-k\tau^2\right).
\end{aligned}$$

In state $a_{2,i}$ Manager 1 must be indifferent between sending a message that induces a posterior $\overline{m}_{2,i}$ and a posterior $\overline{m}_{2,i+1}$ implying that $\mathrm{E}_{\theta_1}\left[U_2 \mid a_{2,i}, \overline{m}_{2,i}\right] - \mathrm{E}_{\theta_1}\left[U_2 \mid a_{2,i}, \overline{m}_{2,i+1}\right] = 0$. Given decision rules (29) and (30) , and letting $\overline{m}_{2,i} = \left(a_{2,i-1}+a_{2,i}\right)/2$ we have that $\mathrm{E}_{\theta_1}\left[U_2 \mid a_{2,i}, \overline{m}_{2,i}\right] - {}_{\theta_1}$ $\left[U_2 \mid a_{2,i}, \overline{m}_{2,i+1}\right] = 0$ if and only if $a_{1,i} = \left(a_{1,i-1}+a_{1,i+1}\right)/\left(2+4b^{DC}\right)$ where b^{DC} is given by (19).

We now show that $sign(b^{DC}) = sign(t)$. Rewrite b^{DC} as $b^{DC}(\theta_2) = b_1^{DC}\theta_2 + b_2^{DC}$, where

$$\begin{aligned}
b_1^{DC} &= \frac{\left(2\lambda-1\right)^2 t^2}{\left(1-\lambda\right)\left(-4\lambda^3+3t^2\lambda-t^2\right)}, \\
b_2^{DC} &= \frac{\left(2\lambda-1\right)\left(-2t+4t\lambda-4\lambda^2+t^2\right)}{2\left(1-\lambda\right)\left(-4\lambda^3+3t^2\lambda-t^2\right)}t(\mu-c).
\end{aligned}$$

For all $\lambda \in [1/2, 1]$ and $t \in [-1, 1]$, we have that $\left(-4\lambda^3+3t^2\lambda-t^2\right) < 0$ and $\left(-2t+4t\lambda-4\lambda^2+t^2\right) < 0$. This implies that $b_1^{DC} < 0$, and $sign(b_2^{DC}) = sign(t)$. That is, the communication bias increases (in absolute value) with the state if and only if $t < 0$. Therefore, to show that $sign(b^{DC}) = sign(t)$ we need only show that

$$b_1^{DC}s + b_2^{DC} > 0 \text{ if } t > 0, \tag{33}$$

$$-b_1^{DC}s + b_2^{DC} < 0 \text{ if } t < 0. \tag{34}$$

The inequality (33) is equivalent to

$$\frac{(\mu-c)}{s} > \frac{2t\left(2\lambda-1\right)}{\left(2t-4t\lambda+4\lambda^2-t^2\right)} > 0.$$

The maximum of the intermediate term above for given $t > 0$ satisfies

$$\max_{\lambda>1/2}\frac{2t\left(2\lambda-1\right)}{\left(2t-4t\lambda+4\lambda^2-t^2\right)} = \frac{t}{(1-t)+\sqrt{1-t^2}} < \frac{1+t}{1-t} < \frac{(\mu-c)}{s},$$

where the last inequality follows from parameter restrictions that ensures positive quantities. This proves that (33) is satisfied. Finally, the inequality (34) is equivalent to

$$\frac{(\mu-c)}{s} > \frac{-2t\left(2\lambda-1\right)}{\left(2t-4t\lambda+4\lambda^2-t^2\right)} > 0.$$

By a similar reasoning as before we have that for given $t < 0$ satisfies

$$\max_{\lambda > 1/2} \frac{-2t\,(2\lambda - 1)}{\left(2t - 4t\lambda + 4\lambda^2 - t^2\right)} = \frac{-t}{(1-t) + \sqrt{1-t^2}} < 1.$$

For $t < 0$ positive quantities is ensured as long as $\mu - c > s$. This proves that (34) is satisfied.

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

$$\mathrm{E}_{\theta_2}\left[U_1 \mid \theta_1, \overline{m}_{1,i}\right] = -\mathrm{E}_{\theta_2}\left[\lambda \pi_1\left(\widehat{q}_1^D, \widehat{q}_2^D, \theta_1\right) + (1-\lambda)\,\pi_1\left(\widehat{q}_1^D, \widehat{q}_2^D\right) \mid \theta_1, \overline{m}_{1,i}\right], \tag{35}$$

where q_1^D and q_2^D are given by (31) and (32). In state $\theta_1 = a_{1,i}$ this can be written as $\mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i}\right]$ being equal to

$$\begin{aligned}
&\frac{b}{4\left(1-t^2\right)}\Big((1-\lambda)\left(1-\tau^2\right)^2 \mathrm{E}_{\theta_2}\left[(\mu - c + \theta_2)^2\right] + \tau^2\left(\lambda - \tau^2 + \lambda\tau^2\right)\mathrm{E}_{\theta_2}\left[m_2^2 | \mu_2(\cdot)\right] \\
&... + \left(\lambda + \lambda\tau^2 - 1\right)\tau^2\left(\overline{m}_{1,i}\right)^2 + \lambda\left(1-\tau^2\right)^2(\mu - c + a_{1,i})^2 \\
&... + 2\lambda\tau^2\left(1-\tau^2\right)\overline{m}_{1,i}\left(\mu - c + a_{1,i}\right) - 2\left(2\lambda - 1\right)\tau^3\overline{m}_{1,i}\mathrm{E}_{\theta_2}\left[m_2 | \mu_2(\cdot)\right] \\
&... - 2\lambda\tau\left(1-\tau^2\right)\mathrm{E}_{\theta_2}\left[(\mu - c + \theta_2)\right](\mu - c + a_{1,i})\Big).
\end{aligned}$$

In state $a_{1,i}$ Manager 1 must be indifferent between sending a message that induces a posterior $\overline{m}_{1,i}$ and a posterior $\overline{m}_{1,i+1}$ implying that $\mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i}\right] - \mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i+1}\right] = 0$. If $(1-\lambda)/\lambda < \tau^2$ the above condition has no solution: essentially Manager 1 would like to induce the highest possible belief in Manager 2. If, however, $(1-\lambda)/\lambda > \tau^2$ then, given sequentially rational decision making (31) and (32) and letting $\overline{m}_{1,i} = \left(a_{1,i-1} + a_{1,i}\right)/2$, we have that $\mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i}\right] - \mathrm{E}_{\theta_2}\left[U_1 \mid a_{1,i}, \overline{m}_{1,i+1}\right] = 0$ if and only if $a_{1,i} = \left(a_{1,i-1} + a_{1,i+1}\right)/\left(2 + 4b^D\left(a_{1,i}\right)\right)$, where $b^D\left(\theta_1\right)$ is given by (18).

We now show that $b^D\left(\theta_1\right)$ is always positive and increasing. The condition $(1-\lambda)/\lambda > \tau^2$ implies that the numerator of (18) is positive and thus $b^D\left(\theta_1\right)$ is increasing. Moreover, if $t < 0$, $(1-\tau)(\mu - c) > (\mu - c) > s$ where the last inequality follows from considering parameter values that induce positive quantities. Therefore $b^D(-s) > 0$. If $t > 0$ then the parameter restriction $(\mu - c)/s > (1+t)/(1-t)$ that guarantees positive quantities, and the fact that for $t > 0$ $(1+t)/(1-t) > 1/(1-t) \geq 1/(1-\tau)$, together imply that $b^D(-s) > 0$. This establishes that $b^D\left(\theta_1\right) > 0$ for all $\theta_1 \in [-s, s]$.

In summary, given the independence of Manager 1 and 2's private information, the multi-sender communication equilibrium decouples into two communication equilibria each of which is equivalent to a sender-receiver game in which the state-dependent bias of the sender satisfies (18).

In particular, since the communication bias $b^D(\theta_1)$ is strictly positive, communication necessarily involves a finite number of intervals as shown in Crawford and Sobel (1982).

Proof of Lemma 1: Proposition B1 derives the expressions and properties for the communication bias under both Centralization, Divisional Centralization and Decentralization. ■

2.2 Appendix B2 - Residual Variance of Communication

In this appendix we first derive closed form expressions for the residual variance under Centralization, Divisional Centralization, and Decentralization. We then prove that these residual variances possess some smoothness properties that enables us to characterize their behavior for λ close to 1/2. We conclude by comparing the informativeness of vertical and horizontal communication.

2.2.1 Vertical Communication

Under Centralization the communication bias is constant, as in the leading example in Crawford and Sobel (1982). Thus, for a given equilibrium with n intervals the residual variance of communication V_n^C satisfies

$$V_n^C = \frac{s^2}{3}(\frac{1}{n^2}) + \frac{1}{3}\left(b^C\right)^2\left(n^2-1\right). \tag{36}$$

The maximum number of intervals $N(b^C)$ satisfies $2N(b^C)(N(b^C)-1)|b^C| \leq 2s$ and is thus given by

$$N(b^C) = int\left(\frac{1}{2}\left(1+\sqrt{1+\frac{4s}{|b^C|}}\right)\right),$$

where $int(z)$ is the largest integer that does not exceed z. Therefore if $\left|b^C\right| < \frac{s}{2}$ the residual variance the communication under Centralization is

$$V^C = V^C_{N(b^C)} = \frac{s^2}{3}(\frac{1}{N(b^C)^2}) + \frac{1}{3}\left(b^C\right)^2\left(N(b^C)^2-1\right),$$

and $V^C = \frac{s^2}{3}$ if $\left|b^C\right| > \frac{s}{2}$.

2.2.2 Horizontal Communication

Under both Divisional Centralization and Decentralization the communication bias takes the form

$$b^s(\theta_i) = b_1^s\theta_i + b_2^s,$$

for $s \in \{DC, D\}$ with

$$b_1^D = (2\lambda-1)/\left(1-\lambda-\lambda\tau^2\right), \qquad b_2^D = (1-\tau)(\mu-c),$$
$$b_1^{DC} = (2\lambda-1)^2 t^2/(1-\lambda)\left(3t^2\lambda-4\lambda^3-t^2\right), \quad b_2^{DC} = \left(1-\lambda\tfrac{1-\tau^2}{t-\tau}\right)(\mu-c).$$

From Proposition B1, $b_1^D, b_2^D > 0$ while $b_1^{DC} < 0$ and $sign(b_2^{DC}) = sign(t)$. We now derive the residual variance for arbitrary $b_1, b_2 > 0$, to obtain (42) below. We will then consider the rate of change of both residual variances as the conflict vanishes.

Consider communication by Manager 1 and let $a_1 = \{a_{1,i}\}_{i=0}^{i=n}$ be a partition of $[-s, s]$ into n intervals. If a_1 characterizes a communication equilibrium by Manager 1 to Manager 2 then it must satisfy the arbitrage condition

$$a_{1,i+1} - a_{1,i} = a_{1,i} - a_{1,i-1} + 4(b_1 a_{1,i} + b_2), \tag{37}$$

with boundary conditions $a_{1,0} = -s$ and $a_{1,n} = s$. Solving this second order linear difference equation we obtain

$$a_{1,i} = Ax^i + B\frac{1}{x^i} - \frac{b_2}{b_1}, \tag{38}$$

where $x = 1 + 2b_1 + 2\sqrt{b_1(b_1+1)} > 1$ is the solution of the characteristic equation associated with (37) that exceeds 1 . Defining r and D as $r = x^n$ and $D = -b_2/sb_1$, the coefficients A and B are

$$\begin{aligned} A &= \left(-\frac{1}{1-r} - \frac{D}{1+r}\right) s, \text{ and} \\ B &= \left(\frac{1}{1-r} - \frac{D}{1+r}\right) sr, \end{aligned}$$

and the size of each interval is given by

$$a_{1,i} - a_{1,i-1} = \frac{(x-1)}{x^i}\left(Ax^{2i-1} - B\right). \tag{39}$$

Maximum Number of Intervals. Let $N(b_1, b_2)$ be the maximum number of intervals in a communication equilibrium. The solution to the second order difference equation characterizes a communication equilibrium as long as the solution (38) is monotonic, i.e. $a_{1,i} - a_{1,i-1} > 0$ is positive. Since $b_1 > 0$ and $b_2 > 0$, we have $A > 0$. Therefore it suffices that $Ax > B$ to guarantee that $Ax^{2i-1} > B$ for all i, $1 \le i \le N(b_1, b_2)$. From the definition of A and B we thus require

$$-r(D-1)(x+1) + r^2(D+1) + x(D+1) > 0.$$

The solution to this quadratic inequality is

$$r = x^n \le \frac{(D-1)}{2(D+1)}(x+1) + \frac{1}{2}\sqrt{\frac{(D-1)^2}{(D+1)^2}(x+1)^2 - 4x}. \tag{40}$$

It follows that $N(b_1, b_2)$ is given by

$$N(b_1, b_2) = int\left(\ln\left(\frac{(D-1)}{2(D+1)}(x+1) + \frac{1}{2}\sqrt{\frac{(D-1)^2}{(D+1)^2}(x+1)^2 - 4x}\right) / \ln x\right).$$

Residual Variance We next compute the residual variance of communication for a communication equilibrium with n intervals, $n \geq 2$. The variance of the message m_1 to Manager 2 is

$$\begin{aligned}
\mathrm{E}_D\left[m_1^2\right] &= \sum_{i=1}^{n} \int_{a_{1,i-1}}^{a_{1,i}} \left(\frac{a_{1,i}+a_{1,i-1}}{2}\right)^2 \frac{1}{2s} d\theta_1 = \frac{1}{8s}\sum_{i=1}^{n}\left(a_{1,i}-a_{1,i-1}\right)\left(a_{1,i}+a_{1,i-1}\right)^2 \\
&= \frac{1}{8s}\sum_{i=1}^{n}\left[\left(a_{1,i}^3-a_{1,i-1}^3\right)+a_{1,i}a_{1,i-1}\left(a_{1,i}-a_{1,i-1}\right)\right] \\
&= \frac{s^2}{4}+\frac{1}{8s}\sum_{i=1}^{n} a_{1,i}a_{1,i-1}\left(a_{1,i}-a_{1,i-1}\right).
\end{aligned}$$

And the residual variance on an $n-$partition equilibrium is

$$V_n^D = \mathrm{E}_D\left[\left(m_1-\theta_1\right)^2\right] = \frac{s^2}{12}-\frac{1}{8s}\sum_{i=1}^{n} a_{1,i}a_{1,i-1}\left(a_{1,i}-a_{1,i-1}\right).$$

We next compute $a_{1,i}a_{1,i-1}\left(a_{1,i}-a_{1,i-1}\right)$. From (38) and the size of each interval (39) we have

$$\begin{aligned}
a_{1,i}a_{1,i-1}\left(a_{1,i}-a_{1,i-1}\right) &= (x-1)\,[\frac{A^3}{x^2}x^{3i}-\frac{xB^3}{x^{3i}}+\frac{A^2\left(x+1\right)Ds}{x^2}x^{2i}-\frac{B^2\left(x+1\right)Ds}{x^{2i}} \\
&\quad +A\frac{\left(Ds\right)^2x+AB+ABx^2-ABx}{x^2}x^i-B\frac{\left(Ds\right)^2x+AB+ABx^2-ABx}{x^ix}].
\end{aligned}$$

From the sum of a geometric series $\sum_{i=1}^{n} x^{ki} = x^k\frac{1-r^k}{1-x^k}$ we can simplify the summation of the previous terms to obtain

$$\begin{aligned}
\sum_{i=1}^{N} a_ia_{i-1}\left(a_i-a_{i-1}\right) &= \left(r^3-1\right)\left(A^3-\frac{B^3}{r^3}\right)\left(\frac{x}{x^2+x+1}\right)+\left(r^2-1\right)\left(A^2-\frac{B^2}{r^2}\right)\left(Ds\right) \\
&\quad +\left(r-1\right)\left(A-\frac{B}{r}\right)\left(\left(Ds\right)^2+AB\left(\frac{x^2-x+1}{x}\right)\right). \qquad (41)
\end{aligned}$$

To further simplify this expression we first note that

$$\frac{x}{x^2+x+1} = \frac{1}{3+4b_1}$$

and

$$\frac{x^2-x+1}{x} = 1+4b_1,$$

which follows from x being a solution to the characteristic equation of (37). Moreover, from the

definitions of A, B and D we have

$$\begin{aligned}
(r-1)\left(A-\left(\frac{B}{r}\right)\right) &= 2s, \\
\left(r^2-1\right)\left(A^2-\left(\frac{B}{r}\right)^2\right) &= -4Ds^2, \text{ and} \\
\left(r^3-1\right)\left(A^3-\left(\frac{B}{r}\right)^3\right) &= 2s\frac{\left(r^2+r+1\right)}{\left(r^2-1\right)^2}\left(\left(r^2+1\right)\left(3D^2+1\right)-2r\left(3D^2-1\right)\right)s^2,
\end{aligned}$$

which, substituted into (41) yields

$$\begin{aligned}
\sum_{i=1}^{N} a_i a_{i-1}\left(a_i-a_{i-1}\right) &= \frac{\left(r^2+r+1\right)\left(\left(r^2+1\right)\left(3D^2+1\right)-2r\left(3D^2-1\right)\right)}{\left(r^2-1\right)^2}\left(\frac{2}{4b_1+3}\right)s^3 \\
&\quad -2D^2s^3-2r\left(1+4b_1\right)\left(\left(\frac{1}{1-r}\right)^2-\left(\frac{D}{1+r}\right)^2\right)s^3.
\end{aligned}$$

Substituting this expression into V_n^s and after some simplifications we have

$$V_n^s=\frac{s^2}{12}-\left(\frac{1-4D^2b_1}{4\left(4b_1+3\right)}+4\frac{b_1\left(b_1+1\right)}{\left(4b_1+3\right)}\frac{r\left(D^2-\left(\frac{1+r}{1-r}\right)^2\right)}{\left(r+1\right)^2}\right)s^2, \tag{42}$$

where $r = x^n$. Therefore the residual variance of communication is given by $V^s = V^s_{N\left(b_1^s, b_2^s\right)}$, $s \in \{DC, C\}$. ■

2.2.3 Absolute Continuity of Residual Variances

The residual variance $V^s, s = \{C, D, DC\}$, is continuous in the own-division bias λ, although non-differentiable whenever the number of intervals in the most informative communication equilibrium changes value. As the number of intervals tends to infinity as managers become more aligned with each other, the residual variance has an infinite number of points of discontinuity in every neighborhood of $\lambda = 1/2$. Nevertheless, the next lemma shows that V^s retains certain smoothness properties that allows us to characterize its behavior in a neighborhood of $\lambda = 1/2$ through the function $\partial V^s/\partial\lambda$.

LEMMA B1. *The residual variance of communication $V^s, s = \{C, D, DC\}$ is an absolutely continuous function of $\lambda \in [1/2, 1]$ with a well-defined limit $\partial V^s/\partial\lambda$ as λ tends to $1/2$. In particular,*

$$\begin{aligned}
\lim_{\lambda\to 1/2}\frac{\partial V^C}{\partial\lambda} &= \frac{4}{3}s\left(\mu-c\right)|t|, \\
\lim_{\lambda\to 1/2}\frac{\partial V^{DC}}{\partial\lambda} &= \frac{8}{3}s\left(\mu-c\right)|t|, \text{ and} \\
\lim_{\lambda\to 1/2}\frac{\partial V^{D}}{\partial\lambda} &= \frac{2}{9}s^2\left(3\frac{\left(\mu-c\right)\left(1-t\right)}{s}-1\right)\frac{4}{1-t^2}.
\end{aligned}$$

Proof: The function V^s, $s = \{C, D, DC\}$, is continuous and increasing in λ and its derivative is defined except for a countable number of points. To establish absolute continuity of V^s we need to further show that (i.) its derivative is integrable, and (ii.) V^s maps sets of measure zero into sets of measure zero (Luzin N property; see Rudin 1986). Since the set of points of non-differentiability of V^s is countable it follows readily that V^s satisfies the Luzin N property (see Leoni 2009). We will now show that $\partial V^s/\partial\lambda$ is bounded in $[1/2, 1]$, whenever defined, and this will establish integrability.

First, the case of Centralization. Differentiating (36) we obtain

$$|\frac{\partial V_n^C}{\partial b^C}| = \frac{2}{3}|b^C|\left(n^2-1\right) \leq \frac{2}{3}s\frac{N(b^C)+1}{N(b^C)},$$

where the last inequality follows from the definition of $N(b^C)$. Therefore we obtain the uniform bound

$$\sup_{\lambda,n}|\frac{\partial V_n^C}{\partial\lambda}| \leq \sup_{\lambda}\frac{2}{3}s\frac{N(b^C)+1}{N(b^C)}\sup_{\lambda}\frac{\partial b^C}{\partial\lambda} = \frac{2}{3}s\sup_{\lambda}\frac{\partial b^C}{\partial\lambda}.$$

We now show that $\partial V^C/\partial\lambda$ approaches a well-defined limit as $\lambda \longrightarrow 1/2$. From the previous bound we have

$$\lim_{|b^C|\to 0}\frac{\partial V^C}{\partial b^C} = \lim_{|b^C|\to 0}\frac{2}{3}s\frac{N(b^C)+1}{N(b^C)} = \frac{2}{3}s.$$

Given that $b^C = (2\lambda-1)(\mu-c)\tau$ we readily have that

$$\lim_{\lambda\to\frac{1}{2}}\frac{\partial V^C}{\partial\lambda}_{\lambda=\frac{1}{2}} = \frac{2}{3}s\lim_{\lambda\to\frac{1}{2}}\frac{\partial\left|b^C\right|}{\partial\lambda} = \frac{4}{3}s\left(\mu-c\right)|t|. \tag{43}$$

Now we turn to the case of Divisional Centralization and Decentralization. Totally differentiating (42) for $n = N(b_1, b_2)$ we have

$$\frac{\partial V^s}{\partial\lambda} = \left[\frac{\partial V^s}{\partial r}\frac{\partial r}{\partial x}\frac{\partial x}{\partial b_1} + \frac{\partial V^s}{\partial b_1}\right]\frac{\partial b_1}{\partial\lambda} + \frac{\partial V^s}{\partial D}\frac{\partial D}{\partial\lambda}, s = \{D, DC\}$$

where $r = x^{N(b_1,b_2)}$. Computing each element in the previous expression we have

$$\begin{aligned}
\frac{\partial V^s}{\partial r}\frac{\partial r}{\partial x}\frac{\partial x}{\partial b_1} &= N(b_1,b_2)\,r\left(4\frac{\left(D^2-1\right)\left(r^2-2r\frac{D+1}{D-1}+1\right)\left(r^2-2r\frac{D-1}{D+1}+1\right)}{(4b_1+3)}s^2\right)\left(\frac{\sqrt{b_1(b_1+1)}}{(r^2-1)^3}\right),\\
\frac{\partial V^s}{\partial b_1} &= s^2\left(\frac{\left(3D^2+1\right)}{(4b_1+3)^2} - \frac{4\left(6b_1+4b_1^2+3\right)}{(4b_1+3)^2}\frac{r\left(D^2-(\frac{1+r}{1-r})^2\right)}{(r+1)^2}\right), \text{ and}\\
\frac{\partial V^s}{\partial D} &= 2s^2Db_1\left(\frac{(r-1)^2-4rb_1}{(4b_1+3)(r+1)^2}\right).
\end{aligned} \tag{44}$$

To guarantee that $\partial V^s/\partial\lambda$ is bounded we now show that it approaches a finite limit as $\lambda \to 1/2$. From the definition of $r = x^{N(b_1,b_2)}$ in (40) we have

$$\lim_{\lambda\to 1/2} \frac{\frac{(D-1)}{2(D+1)}(x+1) + \frac{1}{2}\sqrt{\frac{(D-1)^2}{(D+1)^2}(x+1)^2 - 4x}}{r} = 1,$$

which implies

$$\lim_{\lambda\to 1/2}\left(r^2 - 2r\frac{D-1}{D+1} + 1\right) = 0$$

and

$$\lim_{\lambda\to 1/2} \frac{r\left(D^2 - (\frac{1+r}{1-r})^2\right)}{(r+1)^2} = \frac{r\left(D^2 - \frac{\frac{4Dr}{D+1}}{\frac{-4r}{D+1}}\right)}{\frac{4Dr}{D+1}} = \frac{1}{4}(D+1)^2.$$

With these limits applied to (44) we have

$$\begin{aligned}
\lim_{\lambda\to\frac{1}{2}} \frac{\partial V^s}{\partial r}\frac{\partial r}{\partial x}\frac{\partial x}{\partial b_1} &= 0 \\
\lim_{\lambda\to\frac{1}{2}} \frac{\partial V^s}{\partial b_1} &= s^2\left(\frac{3D^2+1}{9} - \frac{1}{3}(D+1)^2\right) = \frac{2}{9}s^2(-3D-1) \\
\lim_{\lambda\to\frac{1}{2}} \frac{\partial V^s}{\partial D} &= 0.
\end{aligned}$$

We now consider the cases of Decentralization and Divisional Centralization. For Decentralization, we have that

$$\lim_{\lambda\to\frac{1}{2}} \frac{\partial b_1}{\partial\lambda} = \frac{4}{1-t^2}, \text{ and } \lim_{\lambda\to\frac{1}{2}} \frac{\partial D}{\partial\lambda} = -2t(\mu - c).$$

Therefore the limit of the total derivative as the own-division bias vanishes is

$$\lim_{\lambda\to 1/2} \frac{\partial V^D}{\partial\lambda} = \frac{8}{9}\frac{s^2}{1-t^2}\left(3(1-t)\frac{(\mu-c)}{s} - 1\right), \tag{45}$$

which is bounded for $|t| < 1$. This establishes that $\partial V^D/\partial\lambda$ is bounded and thus integrable.

We now turn to the case of Divisional Centralization. While in this case we have that

$$\lim_{\lambda\to\frac{1}{2}} \frac{\partial b_1}{\partial\lambda} = 0,$$

the derivative $\partial D/\partial\lambda$ becomes unbounded when $\lambda \to \frac{1}{2}$. As $\lim_{\lambda\to\frac{1}{2}} \frac{\partial V^s}{\partial D} = 0$, however, in this case we obtain

$$\lim_{\lambda\to 1/2} \frac{\partial V^{DC}}{\partial D}\frac{\partial D}{\partial\lambda} = \frac{8}{3}s(\mu - c)|t|$$

and

$$\lim_{\lambda\to 1/2} \frac{\partial V^{DC}}{\partial\lambda} = \frac{8}{3}s(\mu - c)|t| = 2\lim_{\lambda\to 1/2} \frac{\partial V^C}{\partial\lambda}. \quad \blacksquare$$

COROLLARY B1. *If* $\lim_{\lambda\to 1/2}\frac{\partial V^s}{\partial\lambda} > \lim_{\lambda\to 1/2}\frac{\partial V^l}{\partial\lambda}$ $s,l\in\{C,DC,D\}$ *then there exists an* $\varepsilon>0$ *such that* $V^s(\lambda)>V^l(\lambda)$ *for all* $\lambda\in(1/2,1/2+\varepsilon)$. *Conversely, if* $\lim_{\lambda\to 1/2}\frac{\partial V^s}{\partial\lambda} < \lim_{\lambda\to 1/2}\frac{\partial V^l}{\partial\lambda}$ *then there exists an* $\varepsilon>0$ *such that* $V^s(\lambda)<V^l(\lambda)$ *for all* $\lambda\in(1/2,1/2+\varepsilon)$.

Proof: As $V^l, l=\{C,DC,D\}$ are absolutely continuous the fundamental theorem of calculus holds (Rudin 1987) and we have that

$$\left[V^s(\lambda)-V^l(\lambda)\right]-\left[V^s(1/2)-V^l(1/2)\right]=\int_{1/2}^{\lambda}\frac{\partial}{\partial\lambda}\left(V^s-V^l\right)d\lambda.$$

As all structures achieve full revelation of information for $\lambda=1/2$, if

$$\lim_{\lambda\to 1/2}\partial\left(V^s-V^l\right)/\partial\lambda>0$$

it follows that there exists an $\varepsilon>0$ such that

$$V^s(\lambda)>V^l(\lambda),\lambda\in(1/2,1/2+\varepsilon).$$

The last claim of the corollary follows from an equivalent argument. ■

2.2.4 Informativeness of Vertical and Horizontal Communication

Proof of Lemma 2: Proposition B1 derives the expressions for b^C, b^{DC}, and b^D. Consider first Part (i.) of the Lemma. We will show that, for any $\theta_i\in[-s,s]$, $\lambda\in(1/2,1]$ and $t<0$ we have $\left|b^C\right|<\left|b^D\right|$. That is, the point-wise communication bias under Decentralization is always larger than under Centralization when $t<0$. Then, it follows from Chen and Gordon (2013) that a smaller point-wise bias leads to more informative communication.

Given that, when $t<0$, b^C is constant and negative and b^D is positive and increasing, it follows that

$$\Delta=\min b^D-|b^C|=\frac{2\lambda-1}{1-\lambda-\lambda\tau^2}\left(\left(1-\lambda\tau-\lambda\tau^3\right)(\mu-c)-s\right).$$

Since $\left(1-\lambda\tau-\lambda\tau^3\right)>0$ for $-1\le t\le 0, 1/2\le\lambda\le 1$ then $\Delta\ge 0$ if and only if

$$\frac{\mu-c}{s}\ge\frac{1}{1-\lambda\tau-\lambda\tau^3}.$$

Since positive quantities requires $(\mu-c)/s>1$ when $t<0$ and

$$\sup_{\substack{-1<t\le 0\\ 1/2\le\lambda\le 1}}\frac{1}{1-\lambda\tau-\lambda\tau^3}=1,$$

then we readily have that $\Delta>0$.

We now turn to Part (ii.) of the Lemma. Part (ii.a) follows from Lemma B1 which states that

$$\lim_{\lambda\to 1/2}\frac{\partial V^{DC}}{\partial\lambda}=2\lim_{\lambda\to 1/2}\frac{\partial V^{C}}{\partial\lambda},$$

and corollary B1.

To prove Part (ii.b), first note that, *on average*, the absolute value of the communication bias is larger under Decentralization than under Centralization when $t>0$. As shown in Proposition B1, a necessary condition for informative horizontal communication is that $\frac{\lambda}{1-\lambda}\tau^2<1$. Therefore, whenever b^D is well defined we have

$$\frac{1-\tau}{(1-\lambda)\left(1-\frac{\lambda}{1-\lambda}\tau^2\right)}>\frac{1-\tau}{(1-\lambda)\left(1-\tau^2\right)}>\frac{2}{1+\tau}>1>|\tau|,$$

implying

$$E\left[b^D\right]=(2\lambda-1)(\mu-c)\frac{1-\tau}{(1-\lambda)\left(1-\frac{\lambda}{1-\lambda}\tau^2\right)}>|b^C|.$$

However, unlike the case where $t<0$, we could have cases where the point-wise communication bias under Decentralization is smaller than under Centralization. To see this, note that when $t>0$, we have

$$\Delta=\min b^D(\theta_i)-|b^C|=\frac{2\lambda-1}{1-\lambda-\lambda\tau^2}\left(\left(-2\tau+\lambda\tau+\lambda\tau^3+1\right)(\mu-c)-s\right)$$

Since $\left(-2\tau+\lambda\tau+\lambda\tau^3+1\right)>0$ for $0\le t\le 1, 1/2\le\lambda\le 1$ then $\Delta\ge 0$ if, and only if,

$$(\mu-c)/s\ge 1/\left(-2\tau+\lambda\tau+\lambda\tau^3+1\right).$$

We first note that

$$\max_{1/2\le\lambda\le 1}\frac{1}{\left(-2\tau+\lambda\tau+\lambda\tau^3+1\right)}=\frac{2}{(t+2)(1-t)^2}.$$

Since the restriction to positive quantities requires $(\mu-c)/s>(1+t)/(1-t)$, and $(1+t)/(1-t)>2/(t+2)(1-t)^2$ for $t<\sqrt{2}-1$ it then follows that

$$\Delta\ge 0 \text{ if } 0\le t\le\sqrt{2}-1.$$

If, however, $t>\sqrt{2}-1$, then for λ close to $1/2$ we can have

$$(\mu-c)/s<1/\left(-2\tau+\lambda\tau+\lambda\tau^3+1\right),$$

implying that $b^D(-s)<|b^C|$. Since $b^D(\theta_j)$ is increasing in θ_j the existence of a state $\bar{\theta}_j$ where $b^D(\theta_j)<b^C$ for $\theta_j<\bar{\theta}_j$ follows. We next show that this reversal may translate into horizontal communication being more informative than vertical communication for small own-division bias.

Applying Corollary B1 we conclude that horizontal communication is more informative than vertical in a neighborhood of $\lambda = 1/2$ if, and only if,

$$\frac{1+t}{1-t} < \frac{\mu - c}{s} < \frac{2}{-9t + 3t^3 + 6},$$

where the first inequality follows from the parameter restriction that ensures positive quantities in equilibrium and the second from comparing the limits (43) and (45). For $t < 0.876\,49$, we have $\frac{1+t}{1-t} > \frac{2}{-9t+3t^3+6}$ implying that vertical communication is more informative than horizontal for λ close to 1/2. Define

$$t^* = \max\left\{0.876\,49, \min\left\{t : \frac{\mu - c}{s} < \frac{2}{-9t + 3t^3 + 6}\right\}\right\}.$$

If $t > t^*$ then we have that both $\frac{1+t}{1-t} < \frac{2}{-9t+3t^3+6}$ and $\frac{\mu-c}{s} < \frac{2}{-9t+3t^3+6}$, implying that there exists an $\varepsilon > 0$ such that for $\lambda \in (1/2, 1/2 + \varepsilon)$ horizontal communication is more informative than vertical. ■

2.3 Appendix B3 - Relative Performance

We now turn to comparing the relative performance of Centralization, Divisional Centralization, and Decentralization for a vanishing small own-division bias. We start with a technical lemma that translates the smoothness properties of $V^s, s = \{C, DC, D\}$ derived in Lemma B1 to the comparison of profits under all organizational structures.

LEMMA B2. *The difference in performance $\Pi^s - \Pi^D$, $s = \{C, DC\}$, is an absolutely continuous function of λ. The limit $\partial\left(\Pi^s - \Pi^D\right)/\partial\lambda$, $s = \{C, DC\}$, as λ tends to 1/2 exists and, if its positive, there exists an $\varepsilon > 0$ such that $\Pi^s > \Pi^D$ for $\lambda \in (1/2, 1/2 + \varepsilon)$, while, if it is negative, then there exists an $\varepsilon > 0$ such that $\Pi^s < \Pi^D$ for $\lambda \in (1/2, 1/2 + \varepsilon)$.*

Proof: Lemma B1 establishes that the residual variances of communication V^C, V^{DC}, and V^D are absolutely continuous and the limit of $\partial V^s/\partial\lambda$ $s = \{C, DC, D\}$ exist as $\lambda \to 1/2$. As Π^C, Π^{DC} and Π^D are differentiable functions of V^s it follows that $\Pi^s - \Pi^D$, $s = \{C, DC\}$ is absolutely continuous for $\lambda \in [1/2, 1]$ with a well defined limit as $\lambda \to 1/2$. By application of the fundamental theorem of calculus we have

$$\left[\Pi^s(\lambda) - \Pi^D(\lambda)\right] - \left[\Pi^s(1/2) - \Pi^D(1/2)\right] = \int_{1/2}^{\lambda} \partial\left(\Pi^s - \Pi^D\right)/\partial\lambda d\lambda.$$

Since both Centralization, Divisional Centralization and Decentralization achieve first best performance for $\lambda = 1/2$, if $\lim_{\lambda\to 1/2} \partial\left(\Pi^s - \Pi^D\right)/\partial\lambda > 0$ it follows that there exists an $\varepsilon > 0$ such that

$$\Pi^s(\lambda) > \Pi^D(\lambda), \lambda \in (1/2, 1/2 + \varepsilon).$$

The case $\lim_{\lambda\to 1/2} \partial\left(\Pi^s - \Pi^D\right)/\partial\lambda < 0$ follows similarly from an equivalent argument. ■

The previous analysis showed that vertical communication is more informative than horizontal communication under Decentralization whenever $t < 0$. We now show that this communication advantage may be sufficient for the organization to move to a centralized structure even for a vanishing small conflict.

Proof of Proposition 2: First, we show that

$$\lim_{\lambda\to 1/2} \frac{\partial\left[\Pi^C - \Pi^{DC}\right]}{\partial\lambda} = 0. \tag{46}$$

That is, to a first order, there are no differences between Centralization and Divisional Centralization for a vanishingly small conflict of interest. In particular, this implies that

$$sign\left(\lim_{\lambda\to 1/2} \frac{\partial\left[\Pi^{DC} - \Pi^D\right]}{\partial\lambda}\right) = sign\left(\lim_{\lambda\to 1/2} \frac{\partial\left[\Pi^C - \Pi^D\right]}{\partial\lambda}\right)$$

and, by Lemma B2, it follows that whenever Decentralization dominates Centralization, it also dominates Divisional Centralization, and, conversely, whenever Decentralization is dominated by Centralization, it is also dominated by Divisional Centralization.

To prove (46), we have that the difference $\Pi^{DC} - \Pi^C$ in the performance of Centralization and Divisional Centralization is given by (21) and the limit of the rate of change of this difference is

$$\lim_{\lambda\to 1/2} \frac{\partial\left[\Pi^{DC} - \Pi^D\right]}{\partial\lambda} = \frac{b}{2(1-t^2)}\left(\frac{1}{2}\lim_{\lambda\to 1/2}\frac{\partial V^{DC}}{\partial\lambda} - \lim_{\lambda\to 1/2}\frac{\partial V^C}{\partial\lambda}\right).$$

Then (46) follows since Lemma B1 implies that $\lim_{\lambda\to 1/2} \partial V^{DC}/\partial\lambda = 2\lim_{\lambda\to 1/2} \partial V^C/\partial\lambda$.

We now turn to the comparison between Centralization and Decentralization. The difference $\Pi^C - \Pi^D$ is given by (20) and the limit of the rate of change of this difference is

$$\lim_{\lambda\to 1/2} \frac{\partial\left[\Pi^C - \Pi^D\right]}{\partial\lambda} = \frac{b}{2(1-t^2)}\left(t^2\lim_{\lambda\to 1/2}\frac{\partial V_i^D}{\partial\lambda} - \lim_{\lambda\to 1/2}\frac{\partial V^C}{\partial\lambda}\right).$$

From Lemma B2, Centralization dominates Decentralization for λ close to $1/2$ if $\lim_{\lambda\to 1/2} \partial\left(\Pi^C - \Pi^D\right)/\partial\lambda > 0$, which translates to

$$t^2\lim_{\lambda\to 1/2}\frac{\partial V_i^D}{\partial\lambda} > \lim_{\lambda\to 1/2}\frac{\partial V_i^C}{\partial\lambda}.$$

Using the limits (43) and (45), the previous inequality translates into

$$\frac{8}{9}\frac{s^2t^2}{1-t^2}\left(3(1-t)\frac{(\mu-c)}{s} - 1\right) > \frac{4}{3}s\left(\mu-c\right)|t|,$$

which can be written as

$$\frac{s}{\mu - c} < \frac{3(1-t)\left(2|t| - t - 1\right))}{2|t|}.$$

For $t > 0$ this condition is never satisfied. If $t < 0$ we can solve for t to obtain that Centralization dominates Decentralization when $t < t^{**}$ with

$$t^{**} = \frac{1}{9}\left(\left(3 - \frac{s}{\mu - c}\right) - \sqrt{(\frac{s}{\mu - c})^2 - 6\frac{s}{\mu - c} + 36}\right). \blacksquare$$

References

[1] Aghion, Philippe, Mathias Dewatripont, and Patrick Rey. 2004. "Transferable Control." *Journal of the European Economic Association* 2(1): 115-138.

[2] Aoki, Masahiko. 1986. "Horizontal vs. Vertical Information Structure of the Firm." *American Economic Review* 76: 971-983.

[3] Alonso, Ricardo. 2007. "Shared Control and Strategic Communication. Mimeo.

[4] Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. "When Does Coordination Require Centralization?" *American Economic Review* 98(1): 145-179.

[5] ———, ———, And ———. 2010. "Strategic Communication: Prices versus Quantities." *Journal of the European Economic Association* 2-3: 365-376.

[6] Antras, Pol and Esteban Rossi-Hansberg. 2009. "Organizations and Trade." *Annual Review of Economics* 1: 43-64.

[7] Athey, Susan and Scott Stern. 2002. "The Impact of Information Technology on Emergency Health Care Outcomes." *RAND Journal of Economics* 33(3): 399-432.

[8] Barnard, Chester. 1938. *The Functions of the Executive.* Harvard University Press.

[9] Becker, Gary and Kevin Murphy. 1992. "The Division of Labor, Coordination Costs, and Knowledge." *Quarterly Journal of Economics* 107(4): 1137-1160.

[10] Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2010. "Does Product Market Competition Lead Firms to Decentralize?" *American Economic Review: P&P* 100(2): 434-438

[11] Bolton, Patrick and Mathias Dewatripont. 1994. "The Firm as a Communication Network." *Quarterly Journal of Economics* 109(4): 809-839.

[12] ———, Markus Brunnermeier and Laura Veldkamp. 2013. "Leadership, Coordination and Mission-Driven Management." *Review of Economic Studies* 80(2): 512-537.

[13] Bushman, Robert, Raffi Indjejikian, and Abbie Smith. 1995. "Aggregate Performance Measures in Business Unit Managers Compensation: The Role of Intrafirm Interdependencies." *Journal of Accounting Research* 33: 101-128.

[14] Calvo-Armengol, Antoni, Joan de Marti, and Andrea Prat. 2011. "Communication and Influence." Mimeo.

[15] Chen, Ying and Gordon, Sidartha. 2013. "Information Transmission in Nested Sender-Receiver Games." Mimeo.

[16] Colombo, Massimo and Marco Delmastro. 2004. "Delegation of Authority In Business Organizations: An Empirical Test." *The Journal of Industrial Economics* 52(1): 53-80.

[17] Crawford, Vincent and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(6): 1431-145.

[18] Cremer, Jacques, Luis Garicano, and Andrea Prat. 2007. "Language and the Theory of the Firm." *Quarterly Journal of Economics* 122(1): 373-407.

[19] Dessein, Wouter 2002. "Authority and Communication in Organizations." *Review of Economic Studies* 69: 811-838.

[20] ——— and Tano Santos. 2006. "Adaptive Organizations." *Journal of Political Economy* 114(5): 956-995.

[21] ———, Luis Garicano and Robert Gertner. 2010. "Organizing for Synergies." *American Economic Journal: Microeconomics* 2(4): 77-114.

[22] Ferreira, Daniel and Raaj Sah. 2012. "Who Gets to the Top? Generalists versus Specialists in Managerial Organizations." *Rand Journal of Economics* 43(4): 577-601.

[23] Friebel, Guido and Michael Raith. 2010. "Resource Allocation and Organizational Form." *American Economic Journal: Microeconomics* 2(2): 1-33.

[24] Grossman, Sandford and Oliver Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy* 94(4): 691-719.

[25] Guadalupe, Maria, Hongyi Li and Julie Wulf. 2013. "Who Lives in the C-Suite? Organizational Structure and the Division of Labor in Top Management." *Management Science* 60(4): 824-844.

[26] Hagenbach, Jeanne and Frederic Koessler. 2010. "Strategic Communication Networks." *Review of Economic Studies* 77(3): 1072–1099.

[27] Harford, Tim. 2011. *Adapt: Why Success Always Starts with Failure.* Farrar, Straus, and Giroux. New York, NY.

[28] Hart, Oliver and John Moore. 1990. "Property Rights and the Nature of the Firm." *Journal of Political Economy* 98(6): 1119-1158.

[29] ——— and ———. 2005. "On the Design of Hierarchies: Coordination versus Specialization." *Journal of Political Economy* 113(4): 675-702.

[30] Hayek, Friedrich. 1945. "The Use of Knowledge in Society." *American Economic Review* 35(4): 519-530.

[31] Leoni, Giovanni. 2009. *A First Course in Sobolev Spaces*, Graduate Studies in Mathematics, v.105, American Mathematical Society.

[32] Milgrom, Paul and John Roberts. 1992. *Economics, Organization, and Management.* Prentice-Hall. Englewood Cliffs, NJ.

[33] Morris, Stephen and Hyun Shin. 2002. "Social Value of Public Information." *American Economic Review* 92(5): 1521-1534.

[34] Navaretti, Giorgio and Anthony Venables. 2004. *Multinational Firms in the World Economy.* Princeton University Press. Princeton, NJ.

[35] Raith, Michael. 2003. "Competition, Risk, and Managerial Incentives." *American Economic Review* 93(4): 1425-1436.

[36] Rantakari, Heikki. 2008. "Governing Adaptation." *Review of Economic Studies* 75(4): 1257-1285.

[37] Rudin, Walter. 1986. *Real and Complex Analysis*, 3rd Edition, McGraw-Hill.

[38] Schmidt, Klaus. 1997. "Managerial Incentives and Product Market Competition." *Review of Economic Studies* 64(2): 191-213.

[39] Sims, Christopher. 2002. "Implications of Rational Inattention." *Journal of Monetary Economics* 50: 665-690.

[40] Vives, Xavier. 2008. "Innovation and Competitive Pressure." *The Journal of Industrial Economics* 56(3): 419-469.

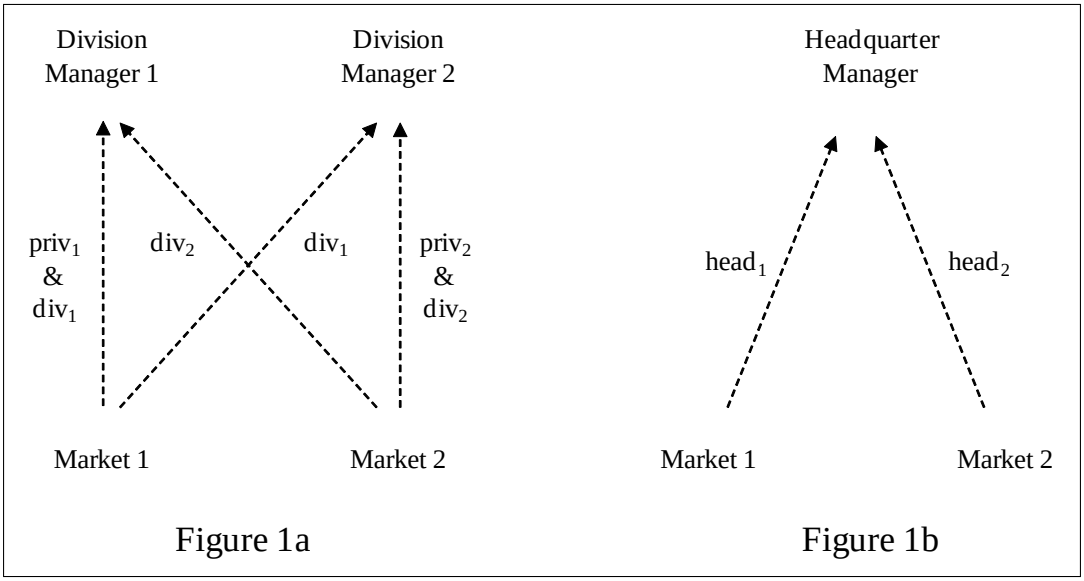
[41] Williamson, Oliver. 1996. *The Mechanisms of Governance.* Oxford University Press. New York, NY.

[42] ———. 2002. "The Theory of the Firm as Governance Structure: From Choice to Contract." *Journal of Economic Perspectives* 16(3): 171-196.

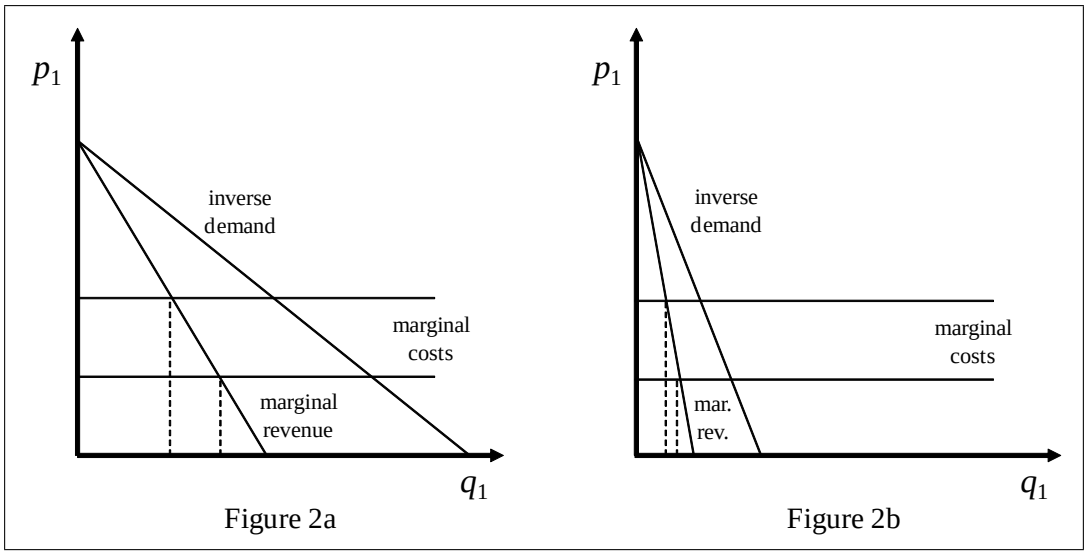
[43] Van Zandt, Timothy. 1999. "Decentralized Information Processing in the Theory of Organizations." In Murat Sertel, ed., Contemporary Economic Issues, 4: Economic Design and Behavior. MacMillan Press, London: 125-160.

3 Figures

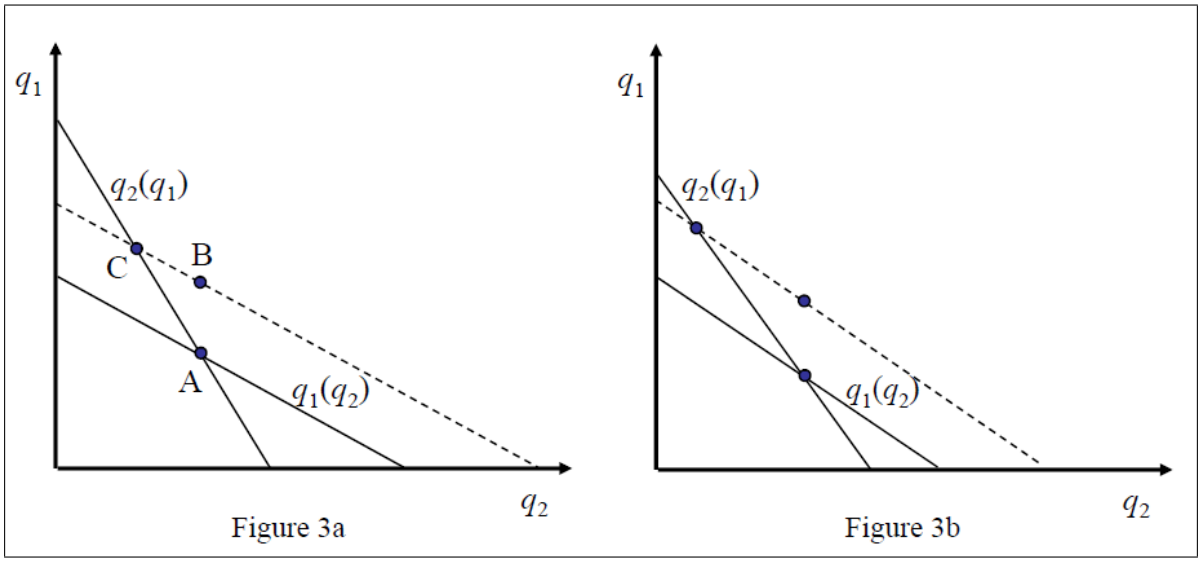
3.1 Figure 1



3.2 Figure 2



3.3 Figure 3



3.4 Figure 4

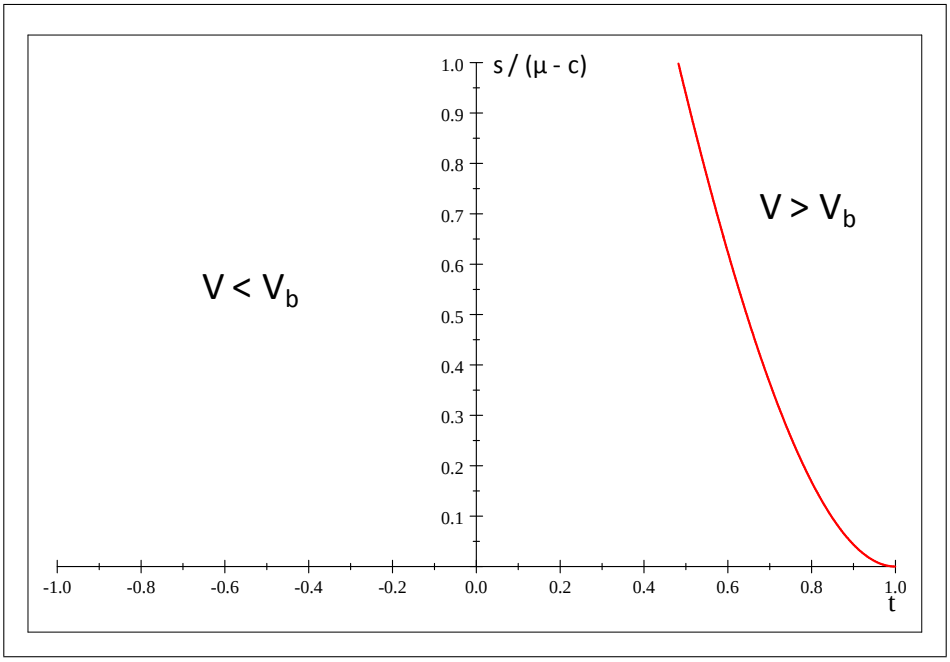


Figure 4: The Informativeness of Communication for Small Own-Division Bias

3.5 Figure 5

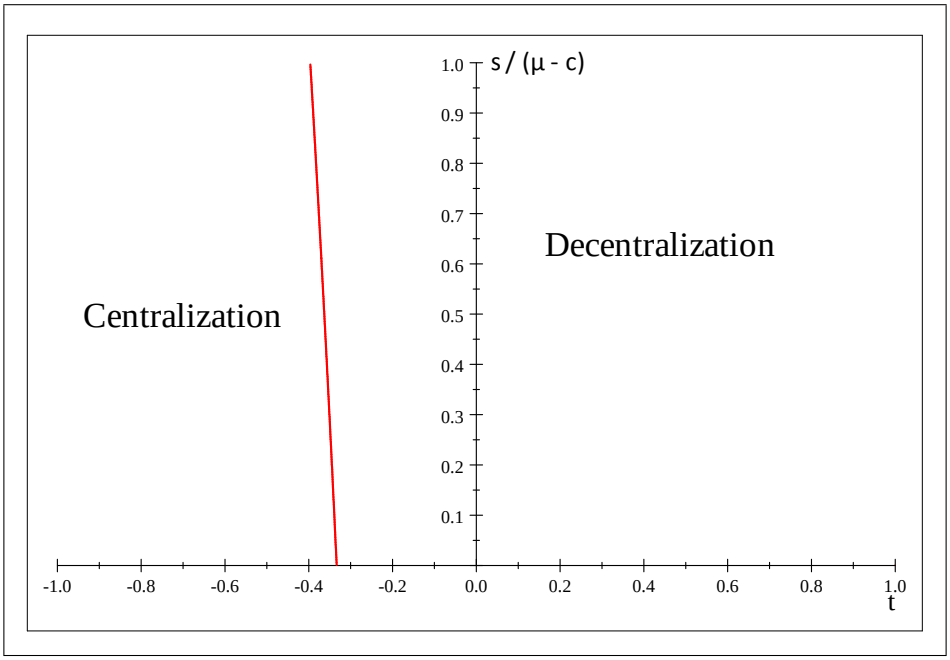


Figure 5: Optimal Organizational Structure for Small Own-Division Bias