Rational Inattention and Organizational Focus[*]

Wouter Dessein   Andrea Galeotti   Tano Santos

Columbia University   University of Essex   Columbia University

January 8, 2016

Abstract

This paper studies optimal communication flows in organizations. A production process can be coordinated ex-ante, by letting agents stick to a pre-specified plan of action. Alternatively, agents may adapt to task-specific shocks, in which case tasks must be coordinated ex-post, using communication. When attention is scarce, an optimal organization coordinates only a few tasks ex-post. Those tasks are higher performing, more adaptive to the environment, and influential. Hence, scarce attention requires setting priorities, not just local optimization. Our results provide micro-foundations for a central idea in the management literature that firms should focus on a limited set of core competencies.

Organizations exist to coordinate specialized agents, but have a limited communication capacity or 'attention' to do so.[1] Coordination, however, does not necessarily require agents to communicate with each other. According to March and Simon (1958), there are two general ways in which organizations can be coordinated:

> "The *type of coordination* used in the organization is a function of the extent to which the situation is standardized. (...) We may label coordination based on pre-established schedules *coordination by plan*, and coordination that involves transmission of new information *coordination by feedback*. The more stable and predictable the situation, the greater the reliance on coordination by plan (p182)."

Importantly, the type of coordination used is an organizational choice. As pointed out by March and Simon: "(I)t is possible to reduce the volume of communication required from day-to-day by substituting coordination by plan for coordination by feedback." (p.183)

This paper studies the optimal coordination of production when attention is scarce. We posit a simple model of team production in which a number of complementary tasks, such as engineering, purchasing, manufacturing, marketing and selling, must be implemented in a coordinated fashion and where agents privately observe task-specific shocks. While coordination by plan saves on communication costs, it prevents agents from responding to such shocks. Our insight is that scarcity of organizational attention makes it optimal to center all communication around a small number of focal tasks. Agents in charge of those tasks are disproportionally responsive to local information and influence the behavior of others. Agents in charge of non-focal tasks stick closely to pre-established action plans in order to avoid coordination failures. Thus, when attention is scarce, it is optimal to coordinate *ex-ante* one group of tasks (by plan) and coordinate *ex-post* another set of tasks (through communication). In this way, all tasks are well coordinated, but tasks are heterogenous in their adaptiveness and influence. When attention is abundant, this asymmetry in task coordination disappears. It is thus scarce attention that creates organizational focus.

---

[1] As emphasized by Herbert Simon, attention may well be the ultimate scarce resource in the economy. According to Simon "a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources" (Simon 1971, p. 40–41). Similarly, Arrow (1974, p.37) notes that 'the scarcity of information-handling ability is an essential feature for the understanding of both individual and organizational behavior.'

Our results provide micro-foundations for a central idea in the management literature that firms should focus on a limited set of "core competencies" (Prahalad and Hamel 1990) and firms that aim to be "all things to all people," will be "caught in the middle" and fail (Porter 1980;1996). In our model, when attention is scarce, optimal organizations set priorities – they select a number of tasks to focus attention on – even when all tasks are ex ante identical. Organizations which fail to prioritize tasks underperform. Our insights further link bureaucracy and rigid rules to limited organizational attention and the strategic choices of a firm. In the face of conflicting needs for attention, the designer must decide which tasks can be flexible and adaptive and which tasks are coordinated using bureaucratic rules.[2] This creates a stark asymmetry in the performance of various organizational tasks.

We conceptualize the available attention as the time agents spend in meetings to coordinate production. If attention is abundant, all tasks share attention equally. Equal sharing is intuitive, as we posit that there are decreasing marginal returns to attention: the probability with which communication (and coordination) is successful is concave in the attention devoted to a task. Our central result, however, is that if attention is scarce and coordination important, it is optimal to treat tasks asymmetrically and focus all attention on a few tasks. The mechanism underlying the above result is a complementarity between attention and decision-making, which creates a convexity in the value of attention. Agents take initiative by adapting their task to local information. The more attention an agent receives, the more initiative this agent can take as those initiatives are now better coordinated. In turn, the more initiative an agent takes, the more important it is to devote scarce attention to him in order to avoid coordination failures. By the same token, agents who are ignored by others are forced to largely ignore their own private information. Since this makes their behavior predictable, devoting attention to such agents is indeed a waste of time. Because of the above complementarities, agents either communicate intensively about a particular task, or they ignore it. Decreasing marginal returns to attention, however, provide a counterveiling force which dominates when focal tasks receive too much attention. In the latter case, it is optimal to reduce organizational focus by increasing the number of focal tasks.

---

[2]See Powell (2015) for an incentive-based rationale for rigid rules. In his model, the organization designer reduces the adaptiveness of decisions to new information in order to limit influence activities.

We derive a number of comparative static results. The scarcer is attention, the smaller is the number of focal tasks. Improvements in communication technology which relax attention constraints thus result in less focused organizations, with more agents being influential and taking initiative. This is consistent with new trends in organizational design towards more network-like organizations where communication flows are horizontal rather than vertical, and decision-making and influence is broadly shared in the organization.[3] More broadly, such trends can often be usefully interpreted as "a move away from ex ante coordination methods" towards more ex post coordination (i.e. using bilateral communication, rather than bureaucratic rules). Online Appendix H discusses three case studies of organizational change through the lens of our model. The number of focal tasks is also decreasing in task-interdependence and the size of the organization. Larger organizations are thus more bureaucratic and have a more centralized communication network. When tasks are ex-ante symmetric, our theory does not select the tasks that the organization focuses on: there are multiple optimal organizations. When tasks are heterogenous in terms of their importance, however, the optimal organization is unique and the attention devoted to a given task often depends more on the rank-order of a task's importance than its absolute level. A small decrease in task importance may then result in a task becoming non-focal and the agent losing all his influence.

Section 3, finally, describes how our model can be extended to endogenize organizational size, with larger organizations exploiting economies of scope but facing more daunting coordination problems. As we show, organizations in more volatile environments or where adaptation to local shocks is more important are then smaller but less focused. Perhaps counter-intuitively, the impact of communication technology on size is ambiguous.

Literature. Our paper is part of a larger literature on team theory (Marschak and Radner 1972, Radner 1993), which studies games where agents share the same objective, but have asymmetric information. As Garicano and Van Zandt (2013) emphasize, team theory is a natural framework in which to analyze the questions raised in this paper, as these models study interrelated decision-making when information is dispersed and there are limits to

---

[3]See Roberts and Saloner (2013) and references therein.

communication, a canonical organizational problem.[4] Despite the need for coordination in all of those models, as far as we are aware, no paper has identified the corner solutions which are at the center of this paper. Proposition 1 highlights why this is the case: It is when communication flows are endogenously designed to support decision-making, that non-concavities arise in team theory models.

Our paper further belongs to a recent literature in organizational economics, most prominently Dessein and Santos (2006) (DS hereafter), Alonso, Dessein and Matouschek (2008), Rantakari (2008) and Bolton, Brunnermeir and Veldkamp (2013), that emphasizes coordination adaptation trade-offs as a key mechanism determining organizational form.[5] All these papers study how organizations are designed to coordinate production and assume a similar quadratic pay-off structure as our paper. Closest to us is DS, which studies task specialization in organizations, but restricts communication flows to be symmetric; instead the present paper takes task specialization as given and endogenizes communication patterns. Calvo, de Marti and Prat (2014) also endogenizes communication patterns in a framework similar to that of DS. Their focus, however, is on how asymmetries in pay-off externalities result in asymmetric communication flows and differential influence for agents.[6]

Finally, a literature on 'narrow business strategies' and 'vision' (Rotemberg and Saloner 1994, 2000) has argued that the commitment by a principal or leader to select a certain type of projects provides strong incentives for agents to exert effort related to such projects. Similarly, Dewatripont, Jewitt, and Tirole (1999) shows how a lack of focus impairs incentives in a multi-task carreer concerns model. In contrast to the present paper, organizational focus is thus a tool to improve incentives.

1   The model

We posit a team-theoretic model, based on Dessein and Santos (2006) and Alonso et al. (2008), to study decision-making and communication within an organization.

---

[4]See, e.g., Geanakoplos and Milgrom (1991) and Cremer, Garicano and Prat (2007).

[5]See also Van den Steen (2013), which studies the role of an explicitly formulated 'strategy' (a small set of key decisions) in coordinating production.

[6]In contrast to our paper, asymmetric communication patterns do not arise in a symmetric set-up.

Production and Payoffs. Production involves the implementation of n tasks, each performed by one agent $i \in \mathcal{N} = \{1, 2, ..., n\}$. Organizational trade-offs arise because agents need to adapt tasks to privately observed task-specific shocks while maintaining coordination across different tasks. Specifically, agent i observes a piece of information θ_i and must take a primary action, $q_{ii} \in \mathbb{R}$, and a coordinating action, $q_{ij} \in \mathbb{R}$, for each task $j \in \mathcal{N} \setminus \{i\}$. The local information of agent i, θ_i, is a shock with variance σ_θ^2 and mean $\hat{\theta}_i$, and its realization is independent across agents. Given a particular realization $\boldsymbol{\theta} = [\theta_1, \theta_2, ..., \theta_n]$, and a choice of actions, $\mathbf{q} = [q_1, q_2, ..., q_n]$, with $q_i = [q_{i1}, q_{i2}..., q_{ii}, ..., q_{in}]$, the organization's profits are:

$$\pi(\mathbf{q}|\boldsymbol{\theta}) = nP - \sum_{i \in \mathcal{N}} L_i(\mathbf{q}|\boldsymbol{\theta}), \tag{1}$$

where P are the gross profits per task and $L_i(\mathbf{q}|\boldsymbol{\theta})$ are the expected losses related to task i due to maladaptation and miscoordination:

$$L_i(\mathbf{q}|\boldsymbol{\theta}) = \phi_i(q_{ii} - \theta_i)^2 + \beta \sum_{j \in \mathcal{N} \setminus \{i\}} (q_{ji} - q_{ii})^2. \tag{2}$$

The profits of the organization depend, therefore, on (i) how well *adapted agent $i's$ primary action q_{ii}* is to task $i's$ local information θ_i and (ii) how well *coordinated agent $j's$ coordinating action q_{ji}* is to agent $i's$ primary action q_{ii}. The parameter ϕ_i is the importance of adapting task i to its task-specific shock and β is the weight given to miscoordination. Hence β can be interpreted as measuring task-interdependence. We assume that $\phi_i < (n-1)\beta$, for all $i \in \mathcal{N}$, so that coordination losses are non-trivial.

Communication frictions. Our starting point is that organizations design communication flows to coordinate production, but organizational attention is scarce. Formally, we define an attention allocation as $\mathbf{t} = [t_1, t_2, ..., t_n]$, where $t_i \geq 0$ and $\sum_{i \in \mathcal{N}} t_i \leq \tau$, with $\tau < \infty$. We can think of t_i as the "air-time" or "attention" any agent i receives, and τ as the length of time an agent spend in meetings as opposed to production.[7] We assume that with probability $r(t_i)$ agent j learns agent i's primary action,[8] and, with the remaining probability, communication

---

[7]Alternatively, we can interpret t_i as the total number of signals that agents receive about q_{ii}, where each signal correctly reveals q_{ii} with probability $1 - e^{-\lambda}$.

[8]In our framework it is is immaterial whether the probability of successful communication is correlated or

is uninformative, and that $r(t_i)$ follows a poisson process with hazard rate λ:

$$r(t_i) = 1 - e^{-\lambda t_i}. \tag{3}$$

Importantly for what follows, the communication technology (3) exhibits decreasing marginal returns, $r''(t_i) < 0$. Moreover, we have that $\lim_{t_i \to \infty} r(t_i) = 1$ and $\lim_{t_i \to \infty} r'(t_i) = 0$. In what follows, we write the organizational attention constraint $\sum_{i \in \mathcal{N}} t_i \leq \tau$ as

$$\sum_{i \in \mathcal{N}} \ln(1 - r(t_i)) \geq \ln(1 - r(\tau)). \tag{4}$$

Section 1.1 discusses alternative specifications to the above communication technology.

Timing. In stage 0, an organization designer chooses the optimal attention allocation $\mathbf{t}$. In stage 1, agents observe their local information and take their primary action. In stage 2, agents observe each other's primary actions according to the attention network $\mathbf{t}$, and then take their coordinating actions. Following a team-theoretical approach, at stage 0, the designer also specifies the decision rules for both primary and coordinating actions that maximize expected payoffs, and agents follow these rules in stage 1 and 2. Equivalently, we can model the resulting agents' primary and coordinating actions as the equilibrium outcome of a Bayesian Game with common payoffs (see Garicano and Van Zandt 2013).

1.1 Interpretation and assumptions of the model

We comment on two key modelling choices– organizational objectives and communication technology– and the robustness to alternative specifications.

Organizational objectives and trade-offs. Expression (1) captures the notion, going back to at least March and Simon (1958), that it is adaptation to unpredictable contingencies, combined with communication frictions, that create coordination problems in organizations. If the organization had unlimited attention, agents can perfectly adapt to their local information, because, by means of communication, they can *coordinate ex-post*. Com-

---

independent across agents $j \neq i$. The reason is that agent i chooses his primary action without knowledge of whether communication will be informative or not, which simplifies the analysis (see 'Timing' below).

munication frictions, however, make ex-post coordination less effective. Another option is for agents to ignore their private information and always implement their task in the same manner – stick to a pre-agreed course of action. No communication is then needed to achieve coordination. Agents *coordinate ex-ante*, but this comes at the cost of maladaptation.

We have assumed that each agent takes a primary action and a series of complementary actions. In Online Appendix E we study an alternative specification in which each agent chooses a single action, $q_{ii} = q_{ij} = q_i$, in line with Alonso et al. (2008) and Rantakari (2008). We show that qualitatively identical results obtain. Moreover, for our analysis, it does not matter if agent i maximizes, as we have postulated, firm perfomance (common pay-offs) or a mix of firm performance and task performance, as in Alonso et al. (2008) and Rantakari (2008). Finally, the model allows for different task-specific mal-adaptation costs, $\phi_i \neq \phi_j$, but we have kept mis-coordination costs, β, homogeneous across tasks. Online Appendix F extends the model to have different degrees of interdependence $\beta_i \neq \beta_j$.

Communication technology. In our model, there is an aggregate constraint on the information sent by agents. Moreover, we posit that all agents $j \neq i$ devote the same amount of attention to task i, consistent with communication occuring in public meetings. Information flows in organizations are of course more complex. Online Appendix C considers some alternative models of communication such as *bilateral communication,* where agents communicate one-to-one (and, hence, agent 3 may devote more attention to agent 1 than agent 2 does) and *individual attention constraints* where each agent has a fixed capacity τ that can be used for both sending and receiving information (and, hence, the time agents 3 and 4 communicate does not affect the attention constraints of agents 1 and 2). They all result in attention allocations that are equivalent to the ones that are optimal under public communication with an aggregate constraint. Far from being exhaustive, our framework thus captures some modal forms of communication inside organizations.[9] Furthermore, Online Appendix G shows that our findings are robust to an endogenous communication capacity τ. While we posit a specific binary communication technology, Online Appendix D shows that identical results obtain if communication is noisy instead and, following the literature on Rational

---

[9]Interesting avenues for future research include mixed forms of communication, where members attend public meetings but also meet on a bilateral basis to refine messages and verify instructions. We also do no model the trade-off between time devoted to communication and production.

Inattention (Sims 2003), entropy is used to model the cost of more precise communication.

2 Optimal Attention Networks

Without loss of generality, we focus on decision rules that are linear in the agent's information. Consider first stage 2, where agents observe each other's primary actions according to the attention network $\mathbf{t}$, and then take their coordinating actions. With probability $r(t_i)$, communication about task i has been successful and agent j optimally sets $q_{ji} = q_{ii}$; otherwise the optimal coordinating action is given by $q_{ji} = E[q_{ii}]$. In stage 1, agent i's primary action can be written as $q_{ii} = \hat{\theta}_i + \alpha_i(\theta_i - \hat{\theta}_i)$ and, at the optimum, minimizes task i's expected losses $E[L_i(\mathbf{q}|\boldsymbol{\theta})]$. Note that $\alpha_i = Cov(q_{ii}, \theta_i)/\sigma_\theta^2$ and, therefore, is *a measure of agent i's adaptiveness* to his local information. Similarly, a natural *measure of agent i's influence on agent j* is $Cov(q_{ji}, \theta_i)/\sigma_\theta^2 = r(t_i)\alpha_i$. Substituting decision rules into (2), we obtain

$$E[L_i(\mathbf{q}|\boldsymbol{\theta})] = \phi_i(1-\alpha_i)^2\sigma_\theta^2 + \beta(n-1)(1-r(t_i))\alpha_i^2\sigma_\theta^2 \tag{5}$$

2.1 Exogenous Attention Networks

As a benchmark, we derive optimal decision-making under the assumption that communication flows are exogenously given (as in Dessein and Santos, 2006). Minimizing (5) with respect to α_i, we obtain that the optimal degree of adaptiveness equals

$$\alpha_i^*(t_i) = \frac{\phi_i}{\phi_i + \beta(n-1)(1-r(t_i))}. \tag{6}$$

From (6), the adaptiveness and influence of agent i is decreasing in the need for coordination β and the number of tasks n, but is increasing in the importance of task i (i.e. ϕ_i) and the attention devoted to task i. We obtain the following benchmark result:

Proposition 1 *Assume communication flows $\mathbf{t}$ are exogenously given. Agent i's adaptiveness and influence is continuous and decreasing in the cost of mis-coordination, β, and continuous and increasing in the importance of his task ϕ_i. If tasks are ex ante symmetric–*

i.e., $\phi_i = \phi$ *and* $t_i = t$ *for all* $i \in \mathcal{N}$,– *then each agent is equally adaptive and influential*–
i.e., $\alpha_i^* = \alpha_j^*$ *for all* $i, j \in \mathcal{N}$.

Proposition 1 shows that, when communication flows are determined independently from decision-making, the need for coordination by itself does not create any discontinuities or non-concavities in organization design. As we show in the next section, organizations operate very differently when communication flows are designed optimally.

2.2 Endogenous Attention Networks

It is useful to think of a designer choosing both decision rules–i.e., $(\alpha_1, ..., \alpha_n)$– and communication qualities–i.e, $\{r_1, .., r_n\}$, with $r_i = 1 - e^{-\lambda t_i}$– rather than an allocation of attention **t**. The organization design problem is then equivalent to

$$\max_{\alpha_1,..,\alpha_n,r_1,...,r_n} E[\pi(\mathbf{q}|\theta)] = nP - \sum_{i \in \mathcal{N}} \left[\phi_i(1-\alpha_i)^2\sigma_\theta^2 + \beta(n-1)(1-r_i)\alpha_i^2\sigma_\theta^2 \right], \qquad (7)$$

subject to constraint (4). Inspecting (7), it is immediate that agent i's adaptiveness , α_i, and task i's communication quality, r_i, are complementary choices, i.e.,

$$\frac{\partial^2 E[\pi(\mathbf{q}|\theta)]}{\partial r_i \partial \alpha_i} = \beta(n-1)\sigma_\theta^2 > 0. \qquad (8)$$

The more adaptive is an agent's task, the more important it is to communicate effectively regarding this task in order to ensure coordination. Similarly, the better is the communication quality r_i, the higher is the optimal level of adaptiveness α_i. As we illustrate now, this fundamental complementarity between the attention devoted to an agent and his adaptiveness, implies that, when the organizational attention is scarce, attention is concentrated on a few tasks. This, in turn, dictates large ex-post asymmetries across agents in terms of their influence and adaptiveness.

Ex-ante symmetric tasks. Consider first the most striking case where all tasks are ex-ante symmetric– i.e, $\phi_i = \phi$ for all $i \in \mathcal{N}$. From (8), if there were constant marginal returns to attention, that is $r'(t_i) = c$, then whenever task l is more adaptive than task $j \neq l$, profits can always be increased by reallocating attention away from task j towards task l. This, in

9

turn, makes it optimal to further increase the adaptiveness of task l and further reduce the adaptiveness of task j. Similarly, when task j and l are equally adaptive and receive equal attention, profits can always be improved by increasing both α_l and t_l and reducing both α_j and t_j. It follows that whenever $r(\tau) < 1$, task l receives all the attention and task j none.

In our model, however, there are *decreasing marginal returns to attention*, as captured by $r''(t_i) < 0$, and it would require an infinite amount of attention to set $r_i = 1$. Decreasing marginal returns provide a counterveiling force for organizational focus. Indeed, when $t_l > t_j$ and $\alpha_l > \alpha_j$ it may be optimal to re-allocate attention on the margin to task j since $r'(t_j) > r'(t_l)$. As we show in the Appendix, in an *optimal organization*, it must then be that $k \in \{1, ..., n\}$ focal tasks split the organizational attention evenly and obtain communication quality $r(\tau/k) = 1 - e^{-\lambda\tau/k}$, whereas the remaining $n - k$ tasks receive no attention.

Proposition 2 *Assume all tasks are ex ante symmetric– i.e., $\phi_i = \phi$ for all $i \in \mathcal{N}$. Then in an optimal organization, $k^* \in \{1, ..., n\}$ tasks split the organizational attention evenly with $t_i = \tau/k$ and the remaining tasks receive no attention.*

Given Proposition 2, the adaptiveness of focal and non-focal tasks are respectively $\alpha^*(\tau/k)$ and $\alpha^*(0)$, where $\alpha^*(t_i)$ is given by expression (6). Substituting $\alpha^*(\tau/k)$ and $\alpha^*(0)$ in (7) and rearranging terms, expected profits in an organization with k focal tasks can be written as

$$\Pi(k) \equiv nP - k\left[1 - \alpha^*(\tau/k)\right]\phi\sigma_\theta^2 - (n-k)\left[1 - \alpha^*(0)\right]\phi\sigma_\theta^2 \tag{9}$$

and the optimal number of focal tasks, k^*, is given by $k^* = \mathsf{argmax}_{k\in\{1,2,\cdots,n\}}\Pi(k)$. We say that an organization is 'focused' whenever $k^* < n$, and some tasks receive no attention. An organization is 'balanced' when $k^* = n$ and all tasks split attention evenly.

When is a focused organization optimal? From (9), the pay-offs associated with task i are linear in its (optimal) level of adaptiveness $\alpha^*(t_i)$. Hence, a focused organization with $k < n$ focal tasks will be preferred over a balanced organization if and only if,

$$k\left[\alpha^*(\tau/k) - \alpha^*(\tau/n)\right] > (n-k)[\alpha^*(\tau/n) - \alpha^*(0)] \tag{10}$$

Thus, the benefit of moving from a balanced to a focused organization with k focal tasks is

the increase $\alpha^*(\tau/k) - \alpha^*(\tau/n)$ in the adaptiveness of k focal tasks; the cost is the decline $\alpha^*(\tau/n) - \alpha^*(0)$ in the adaptiveness of the $(n-k)$ non-focal tasks. It is easy to verify that the function $\alpha^*(t_i)$ is S-shaped. Because of the complementarity between the attention t_i devoted to a task and the adaptiveness α_i of the same task, $\alpha^*(t_i)$ is convex and increasing in t_i whenever λt_i is small. But because of the decreasing marginal returns to attention, $\alpha^*(t_i)$ becomes concave in t_i whenever λt_i is large. Indeed, we have that

$$\frac{\partial^2 \alpha^*(t_i)}{\partial t_i^2} > 0 \Leftrightarrow \frac{2\beta(n-1)}{\phi + \beta(n-1)(1 - r(t_i))} > -\frac{r''(t_i)}{r'(t_i)^2} \Leftrightarrow (n-1)\beta e^{-\lambda t_i} > \phi \qquad (11)$$

It follows that a focused organization with $k^* < n$ is optimal when effective attention, $\lambda \tau$, is scarce. More generally, fixing $k < n$, a sufficient condition for $k^* \leq k$ is that $\alpha^*(t_i)$ is convex in t_i at $t_i = \tau/k$. We summarize this discussion as follows:

Proposition 3 *Assume all tasks are ex ante symmetric. Fix $k < n$, if $(n-1)\beta e^{-\lambda \tau/k} > \phi$ then the optimal number of focal tasks $k^* \leq k$. It follows that the optimal organization is focused ($k^* < n$) when $\lambda \tau$ is sufficiently small, n is sufficiently large or the ratio β/ϕ is sufficiently large.*

Our assumption that $r(t_i)$ represents a Poisson learning process allows us to derive clean conditions for organizational focus. More generally, from (11), as long as $-r''(t_i)/r'(t_i)^2 < 2$ for t_i small, organizational focus will be optimal whenever attention is scarce and coordination important. This condition on the curvature of $r(t_i)$ captures the notion that decreasing marginal returns to attention must be mild when t_i is small.[10] In the Appendix, we derive the following comparative static results:

Proposition 4 *Assume all tasks are ex ante symmetric. The optimal number of focal tasks k^* is increasing in effective communication capacity $\lambda \tau$ and the importance of adaptation ϕ, and is decreasing in tasks interdependence β, and the size/complexity of the organization n.*

There are some immediate implications from Proposition 4. Over the last decades, there have been enormous technological innovations in communication and coordination technologies (E-mail, wireless communication and computing, intra networks, electronic data

---

[10]For $r(t_i) = 1 - e^{-\lambda t_i}$, we have that $-r''(t_i)/r'(t_i)^2 = 1/e^{-\lambda t_i}$. Note that $-r''(t_i)/r'(t_i)^2$ grows without bound for t_i large, reflecting strongly decreasing marginal returns for t_i large.

interchange (EDI)) which can be interpreted as an exogenous increase in the effective attention capacity $\lambda\tau$. An implication of Proposition 4, therefore, is that such technological improvements result in a shift towards more balanced organizations that pay attention to the task-specific information of a larger number of agents. This is consistent with new trends in organizational design towards more network-like organizations where communication flows are horizontal rather than vertical, and where decision-making and influence is broadly shared in the organization. These novel organizations have been documented in both cases studies and large scale empirical studies (Whittington et al. 1999, Guadalupe, Li and Wulf 2014).[11] While our model abstracts away from important organizational issues (hierarchy, for example, plays no role in our model), those new organizational structures can be usefully interpreted as a shift away from ex ante coordination, through bureaucratic rules and standard operating procedures, towards more ex post coordination, through bilateral communication between agents directly involved in production, and this in response to improvements in communication technology. Online Appendix H discusses three case studies of organizational change through the lens of our model. The first case discusses a novel organizational structure at Proctor and Gamble, introduced in the early 2000's and widely imitated by other global consumer packaged goods companies. The second example studies changes in the apparel industry towards 'lean retailing', which were caused by the interaction of an increased need for adaptation to fashion trends and improvements in IT such as Electronic Data Interchange (EDI). The final example is concerned with an innovation in management called Quality Function Deployment (QFD), which is geared towards solving problems of coordination between different functions, such as marketing and engineering.

Our results further point to a trade-off between organizational *size* and organizational *focus*. Smaller organizations distribute attention more evenly and, hence, the information of more agents is reflected in decision-making. As an organization grows larger, leadership becomes more concentrated as there is more need for coordination. Despite the organization having more members, fewer of them receive attention. An increase in n thus has a similar effect as an increase in β. This is consistent with the experience of many entrepreneurial firms,

---

[11]Guadalupe et al. document how, in recent decades, C-level executive teams in Fortune 500 firms have almost doubled in size, mainly because of the inclusion of more functional managers.

whose culture of joint-decision-making and open lateral communication often disappears as they grow larger and more hierarchical.

Asymmetric tasks. When tasks are ex-ante symmetric, our theory does not select the tasks that the organization focuses on: there are multiple optimal organizations. Our next result considers the case where tasks are heterogenous in terms of their importance. Not only is the optimal organization unique then, the attention devoted to a given task often depends more on the rank-order of a task's importance than its absolute level.

Proposition 5 *Assume that $\phi_i > \phi_{i+1}$ for all $i \in \{1, ..., n-1\}$. In the unique optimal organization, attention is focused on the k^* most important tasks and more important tasks receive more attention–i.e., there exists k^* so that for every $i < j \leq k^*$, $t_i^* > t_j^* > 0$ and $t_i^* = 0$ for every $i > k^*$.*

A corollary of Proposition 5 is that a change in the importance of a task that alters its relative ranking may result in a discrete drop or jump in the attention devoted to that task, and, therefore, to the agent's adaptiveness and influence.[12] In other words, scarce attention requires setting priorities – not just local optimization. In our benchmark with exogenous communication flows (Proposition 1), an agent's adaptiveness and influence was continuous in the importance of its task. In contrast, when communication flows are optimally designed to support decision-making, a small decrease in task importance may result in a task becoming non-focal and the agent losing all his influence. Similarly, an increase in the importance of an already focal task may make it optimal for the organization to focus on less tasks.

3   Concluding remarks

The main contribution of this paper is to shed light on the coordination of production when communication flows are optimally designed. When attention is scarce, organizations optimally set priorities and coordinate only a few tasks ex post, through communication. Those tasks are higher performing, more adaptive to the environment, and influential. The remaining tasks are coordinated ex ante, and the agents in charge of them stick closely to

---

[12]For example, an increase in the importance of task i, with $i > k^*$, will eventually result in a discrete jump from $t_i^* = 0$ to $t_i^* > 0$ and a discrete drop from $t_{k^*}^* > 0$ to $t_{k^*}^* = 0$.

prespecified action plans – standard operating procedures. Beyond this insight, our model can be used to shed light on new trends in organization design, as discussed following Proposition 4 and in Online Appendix H.

To conclude, we informally discuss an extension where organizational size is endogenous. In our basic model, the size or "scope" of the organization is fixed. In multi-product firms, however, different tasks may correspond to different types of products or services, and it is natural to think of the number of tasks as being endogenous. In Online Appendix B, we endogenize the number of tasks by introducing economies of scale or scope: certain fixed costs can be shared among tasks (e.g. production facilities or a distribution network), yielding benefits to size. The size of organizations, however, is limited by the need for coordination and limited organizational attention. Two empirically relevant drivers of organizational size are the volatility of the environment and changes in information and communication technologies. Consistent with recent trends in "de-scoping"(Roberts and Saloner 2013), we show that the optimal scope of organizations *decreases* as the environment becomes more volatile and adaptation becomes more important. Intuitively, by reducing the number of tasks that it undertakes, the organization reduces its coordination needs, hence allowing for better adaptation. At the same time, the number of tasks that receive attention increases. Hence, as the environment becomes more volatile, there is a move from large, focussed organizations that maximize scope economies to smaller, but more adaptive and balanced organizations. Improvements in information technology might be conjectured to always increase the size of organizations, as they allow for better coordination. Interestingly, we show that information technology has a decidedly ambiguous impact on firm scope. Intuitively, information technology makes it optimal for organizations to shift towards a strategy that emphasizes adaptation to its environment, but smaller and more balanced firms are better configured to do so. Hence, while for low levels of information technology, large, non-adaptive firms exploiting economies of scale are optimal, we show that for intermediate levels of information technology, smaller, more flexible firms are often preferred.

References

Alonso, Ricardo, Wouter Dessein and Niko Matouschek (2008). "When does Coordination Require Centralization?" American Economic Review, 98 (1), 145-179

Arrow, Kenneth (1974). The Limits of Organization. W. W. Norton & Company, New York.

Bolton, Patrick, Markus Brunnermeier and Laura Veldkamp (2013). "Leadership, Coordination and Corporate Culture." Review of Economic Studies, 80(2), 512-537.

Calvo-Armengol, Antoni, Joan de Marti and Andrea Prat (2015). "Communicaton and Influence." Theoretical Economics 10, 649–690.

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

Dessein, Wouter and Tano Santos (2006). "Adaptive Organizations." Journal of Political Economy, 114 (50), 956-995.

Dewatripont, Mathias, Ian Jewitt, Ian and Jean Tirole (1999). "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies." The Review of Economic Studies, 66 (1), 199-217.

Garicano, Luis and Timothy Van Zandt (2013). "Hierarchies and the Division of Labor." Handbook of Organizational Economics, 604-654, eds. R. Gibbons and J. Roberts. Princeton University Press.

Geanakoplos, John, and Paul Milgrom (1991). "A Theory of Hierarchies Based on Limited Managerial Attention." Journal of the Japanese and International Economies, 5(3), 205–225.

Guadalupe, Maria, Hongyi Li, and Julie Wulf (2014). "Who Lives in the C-Suite? Organizational Structure and the Division of Labor in Top Management." Management Science. 60(4): 824-844

Hellwig, Christian and Laura Veldkamp (2009) "Knowing What Others Know: Coordination Motives in Information Acquisition." Review of Economic Studies, 76(1), 223-251.

March, James G., and Herbert A. Simon (1958). Organizations. Wiley (USA).

Marschak, Jacob and R. Radner (1972). Economic Theory of Teams. Yale University Press.

Prahalad, C.K. and Gary Hamel (1990). "The core competence of the corporation," Harvard Business Review, 68, 79-91.

Porter, Michael (1980). Competitive Strategy: Techniques for Analyzing Industries and Companies. New York: Free Press

Porter, Michael (1996). "What is Strategy?" Harvard Business Review, 74, 61-78.

Powell, Michael (2015). "An Influence-Cost Model of Organizational Practices and Firm Boundaries." Journal of Law, Economics and Organizations, 31, suppl 1, i104-i142.

Radner, Roy (1993). "The organization of decentralized information processing." Econometrica, 61(5), 1109-1146.

Rantakari, Heikki (2008). "Governing Adaptation." Review of Economic Studies, 75, 1257-1285.

Roberts John and Garth Saloner (2013). "Strategy and Organization," Handbook of Organizational Economics, eds. Robert Gibbons and John Roberts, Princeton University Press, 799-849.

Rotemberg, Julio, and Garth Saloner (1994). "Benefits of Narrow Business Strategies," American Economic Review, 84 (5), 1330-1349.

Rotemberg, Julio, and Garth Saloner (2000). "Visionaries, Managers, and Strategic Direction," RAND Journal of Economics,31 (4), 693-716

Simon, Herbert (1971)."Designing Organizations for an Information-Rich World," Computers, Communication, and the Public Interest, ed Martin Greenberger, Baltimore, MD: The Johns Hopkins Press.

Sims, Christopher (2003). "Implications of rational inattention," Journal of Monetary Economics. 50 (3), 665-690.

Van den Steen, Eric (2013). "A Formal Theory of Strategy." Mimeo, Harvard Business School.

Whittington, Richard, Andrew Pettigrew, Simon Peck, Evelyn Fenton, and Martin Conyon (1999). "Change and Complementarities in the New Competitive Landscape: A European Panel Study, 1992-1996." Organization Science, 10(5), 583-600.

Appendix

Proof of Proposition 2. Proposition 2 follows from the fact that if $\mathbf{t}$ is optimal, then $t_i = t_j$ for all i and j such that $t_i > 0$ and $t_j > 0$. We now prove this statement. Using (7) and (6) we write the objective of the organisation as

$$E\left[\pi\left(\mathbf{q}, \mathbf{t}|\boldsymbol{\theta}\right)\right] = nP - n\sigma_\theta^2\phi + \sum_{i=1}^{n}\left[\tfrac{\phi\sigma_\theta^2}{\phi+\beta(n-1)(1-r(t_i))}\right].$$

Suppose, for a contradiction, that $t_i > t_j > 0$. Consider now two alternative organizations. One organization, denoted by $\mathbf{t}'$, is the same as $\mathbf{t}$, but $t_i' = t_i - \epsilon$ and $t_j' = t_j + \epsilon$. The second organization, denoted by $\hat{\mathbf{t}}$, is the same as $\mathbf{t}$, but $\hat{t}_i = t_i + \epsilon$ and $\hat{t}_j = t_j - \epsilon$. These constructions are derived for some small and positive ϵ.

Since $\mathbf{t}$ is optimal, we must have that $E\left[\pi\left(\mathbf{q}, \mathbf{t}|\boldsymbol{\theta}\right)\right] > E\left[\pi\left(\mathbf{q}, \mathbf{t}'|\boldsymbol{\theta}\right)\right]$. This is is equivalent to $\left[e^{-\lambda t_j} - e^{-\lambda(t_i-\epsilon)}\right]\left[\beta^2(n-1)^2 e^{-\lambda(t_i+t_j)} - \phi\right] > 0$, and, since $t_i > t_j$, for small ϵ we have that $e^{-\lambda t_j} - e^{-\lambda(t_i-\epsilon)} > 0$ and therefore optimality of $\mathbf{t}$ requires that $\beta^2(n-1)^2 e^{-\lambda(t_i+t_j)} - \phi > 0$. Similarly, since $\mathbf{t}$ is optimal, we must have that $E\left[\pi\left(\mathbf{q}, \mathbf{t}|\boldsymbol{\theta}\right)\right] > E\left[\pi\left(\mathbf{q}, \hat{\mathbf{t}}|\boldsymbol{\theta}\right)\right]$. This is

equivalent to $-\left[e^{-\lambda(t_j-\epsilon)} - e^{-\lambda t_i}\right]\left[\beta^2(n-1)^2 e^{-\lambda(t_i+t_j)} - \phi\right] > 0$, and, since $t_i > t_j$, we have that $e^{-\lambda(t_j-\epsilon)} - e^{-\lambda t_i} > 0$, and therefore optimality of $\mathbf{t}$ requires that $\beta^2(n-1)^2 e^{-\lambda(t_i+t_j)} - \phi < 0$. We have then reached a contradiction. ∎

Proof of Proposition 3. Inspection of expression (9) implies that

$$k^* = \arg\max_{k\in\{1,...,n\}} \Gamma(k) \equiv k[\alpha(\tau/k) - \alpha(0)],$$

where $\alpha(t)$ is given by expression (6). Simple analysis of the function $\Gamma(k)$ for $k \in [0,\infty]$ reveals that $\Gamma(k)$ is first increasing and then decreasing. Furthermore, the sign of $\Gamma''(k)$ is the same as the sign of $\alpha''(\tau/k)$. By condition (11) we know that $\alpha(\tau/k)$ is concave (resp. convex) at τ/k if $(n-1)\beta e^{-\lambda\tau/k} < \phi$ (resp. $(n-1)\beta e^{-\lambda\tau/k} > \phi$). This implies that there exists a $\bar{k} < \infty$ such that $\Gamma(k)$ is concave for all $k \le \bar{k}$ and convex for all $k > \bar{k}$. Furthermore, since $\Gamma(k)$ is first increasing and then decreasing, it follows that $\bar{k}$ is strictly higher than $k' = \arg\max_{k\in[0,\infty)} \Gamma(k)$. Hence, if we verify that $k \in \{1,...,n\}$ is such that $\alpha(\tau/k)$ is convex at τ/k, then it must be the case that the optimal organization is focused on $k^* \le k$ tasks. ∎

Proof of Proposition 4. We state Lemma A, which is then used to prove Proposition 4. The proof of Lemma A follows by investigating the behaviour of the profit function of an organization with k focal tasks, e.g., the function (9); the proof of Lemma A is relegated in online Appendix A.

Lemma A. There exist $0 < \bar{\beta}(n) < ... < \bar{\beta}(k+1) < \bar{\beta}(k) < ... < \bar{\beta}(2)$ such that the optimal organization has: $k^* = n$ focal tasks if $\beta/\phi < \bar{\beta}(n)$, $k^* \in \{2,...,n-1\}$ focal tasks if $\beta/\phi \in (\bar{\beta}(k^*+1), \bar{\beta}(k^*))$, and $k^* = 1$ if $\beta/\phi > \bar{\beta}(2)$. Furthermore

$$\bar{\beta}(k+1) = \frac{1}{n-1}\left[\frac{e^{\frac{\lambda\tau}{k+1}} + ke^{-\frac{\lambda\tau}{k(k+1)}} - (k+1)}{k + e^{-\frac{\lambda\tau}{k}} - (1+k)e^{-\frac{\lambda\tau}{k(k+1)}}}\right]. \qquad (12)$$

Lemma A implies that the optimal number of focal tasks k^* is decreasing in β and increasing in ϕ. Next, to see that k^* is increasing in $\lambda\tau$, we notice that $\bar{\beta}(k+1)$ is increasing in $\lambda\tau$ because the numerator is increasing in $\lambda\tau$ and the denominator is decreasing in $\lambda\tau$. Finally, that if $k^*(n) < n$ then $k^*(n+1) \le k^*(n)$ follows from the fact that the cut-off $\bar{\beta}(k)$ is decreasing in n. This completes the proof of Proposition 4. ∎

Proof of Proposition 5. We order tasks as follows $\phi_i > \phi_{i+1}$. The problem of the designer specified in (7) subject to the constraint (4) can be rewritten as follow: $\max_{r_1,...,r_n} G(r_1,..,r_n) \equiv \sum_i \phi_i \alpha_i^*(r_i)$, subject to $\sum_i \ln(1-r_i) = \ln(1-r(\tau))$ and $r_i \ge 0$ for all i. Recall also that

$\alpha_i^* = \phi_i/[\phi_i + \beta(n-1)(1-r_i)]$. The Lagrangian is then

$$\mathcal{L} = \sum_i \phi_i \alpha_i(r_i) - \lambda[\ln(1-R) - \sum_i \ln(1-r_i)],$$

and, at the optimum, for each task i so that $r_i > 0$ it must hold

$$\phi_i \frac{\partial \alpha_i(r_i)}{\partial r_i} - \frac{\lambda}{1-r_i} = 0 \text{ iff } \alpha_i^2(r_i)[1-r_i]\beta(n-1) = \lambda.$$

And so, if $r_i > 0$ and $r_j > 0$, then $[1-r_i]\alpha_i^2(r_i) = [1-r_j]\alpha_j^2(r_j)$. We now use this necessary condition to show that at the optimum if $r_i > 0$ then $r_i > r_j$ for all $j > i$. For a contradiction suppose that $r_j \geq r_i > 0$, with $j > i$.

First, consider that $r_j = r_i > 0$. From the necessary condition above we know that $(1-r_i)\alpha_i^2(r_i) = (1-r_j)\alpha_j^2(r_j)$, but this is impossible because $r_i = r_j$ and $\alpha_i(r) > \alpha_j(r)$ for all r. Second, consider that $r_j > r_i$. Optimality implies that $G(r_1,..,r_i,..,r_j,..,r_n) \geq G(r_1,..,r_j,..,r_i,..,r_n)$ if, and only if, $\phi_i \alpha_i(r_i) + \phi_j \alpha_j(r_j) > \phi_i \alpha_i(r_j) + \phi_j \alpha_j(r_i)$, or, equivalently, $\phi_i[\alpha_i(r_j) - \alpha_i(r_j)] < \phi_j[\alpha_j(r_j) - \alpha_j(r_i)]$. Note that $\phi_i[\alpha_i(r_j) - \alpha_i(r_i)] = \alpha_i(r_i)\alpha_i(r_j)\beta(n-1)[r_j - r_i]$ and therefore the above inequality is equivalent to $\alpha_i(r_i)\alpha_i(r_j) < \alpha_j(r_i)\alpha_j(r_j)$, which is a contradiction because $\alpha_i(r) > \alpha_j(r)$ for every r. Hence, in the optimal organization there exists a $k^* \leq n$ so that $r_i > 0$ for all $i \leq k^*$ and $r_i = 0$ otherwise, and that $r_i > r_j$ for every $i < j \leq k^*$. ∎

4 On-line Appendix: Not For Publication

Appendix A contains the proof of Lemma A, which is used to prove Proposition 4 in the paper. Appendix B endogenizes the size of the organization. Appendix C extends the analysis to different models of communication. Appendix D discusses our communication technology and relates it to the literature on Rational Inattention. Appendix E shows that our insights are robust to an alternative model of production that has been widely used in the literature on organizational economics. Appendix F takes the case of tasks that have different coordination costs. Appendix G endogenizes attention capacity. The extensions presented in Appendix E-F-G are developed, for simplicity, for the case of an organization with two agents. Appendix H discusses three examples of organizational change through the lens of our model.

4.1 Appendix A: Proof of Lemma A.

Lemma A. There exist $0 < \bar{\beta}(n) < ... < \bar{\beta}(k+1) < \bar{\beta}(k) < ... < \bar{\beta}(2)$ such that the optimal organization has: $k^* = n$ focal tasks if $\beta/\phi < \bar{\beta}(n)$, $k^* \in \{2, ..., n-1\}$ focal tasks if $\beta/\phi \in (\bar{\beta}(k^*+1), \bar{\beta}(k^*))$, and $k^* = 1$ if $\beta/\phi > \bar{\beta}(2)$. Furthermore

$$\bar{\beta}(k+1) = \frac{1}{n-1} \left[\frac{e^{\frac{\lambda\tau}{k+1}} + ke^{-\frac{\lambda\tau}{k(k+1)}} - (k+1)}{k + e^{-\frac{\lambda\tau}{k}} - (1+k)e^{-\frac{\lambda\tau}{k(k+1)}}} \right]. \tag{13}$$

Proof of Lemma A. Recall that expected payoffs of an organization with k focal tasks is

$$E(\pi(\mathbf{q}, \mathbf{t}|\boldsymbol{\theta})) = nP - n\sigma_\theta^2\phi + \sigma_\theta^2 \left[\frac{k\phi}{\phi + \beta(n-1)(1-r(\tau/k))} + \frac{(n-k)\phi}{\phi + \beta(n-1)} \right].$$

Denote by $\gamma = \beta/\phi$. Then, we write

$$E(\pi(\mathbf{q}, \mathbf{t}|\boldsymbol{\theta})) = nP - n\sigma_\theta^2\phi + \sigma_\theta^2 \left[\frac{k}{1 + \gamma(n-1)(1-r(\tau/k))} + \frac{(n-k)}{1 + \gamma(n-1)} \right].$$

We obtain that

$$\frac{dE[\pi(\mathbf{q}, \mathbf{t}|\boldsymbol{\theta})]}{dk} = \frac{\lambda\gamma}{(1 + \gamma(n-1))k \left(1 + \gamma(n-1)e^{-\frac{\lambda\tau}{k}}\right)^2} \Phi(k, \gamma, \tau, n),$$

where

$$\Phi(k, \gamma, \tau, n) = k \left[1 - e^{-\frac{\lambda\tau}{k}}\right] \left[1 + \gamma(n-1)e^{-\frac{\lambda\tau}{k}}\right] - \lambda\tau(\gamma(n-1) + 1)e^{-\frac{\lambda\tau}{k}},$$

1

and that
$$\frac{d^2 E\left[\pi\left(\mathbf{q}, \mathbf{t} \mid \boldsymbol{\theta}\right)\right]}{dkdk} = -\frac{\lambda^2 \gamma(n-1)\tau^2 e^{-\frac{\lambda\tau}{k}}}{k^3 \left(1 + \gamma(n-1)e^{-\frac{\lambda\tau}{k}}\right)^3}\left[1 - \gamma(n-1)e^{-\frac{\lambda\tau}{k}}\right].$$

Observation 1. By direct verification, the function $\Phi(k, \gamma, \tau, n)$ is decreasing in γ for all k, τ, n. Note also that the sign of $\frac{dE[\pi(\mathbf{q},\mathbf{t}|\boldsymbol{\theta})]}{dk}$ is the same as the sign of $\Phi(k, \gamma, \tau, n)$.

Denote by $\tilde{\beta}$ the solution to $1 - \tilde{\beta}(n-1)e^{-\frac{\lambda\tau}{n}} = 0$. Also, denote by $\hat{\beta}$ the solution to $1 - \hat{\beta}(n-1)e^{-\lambda\tau} = 0$. Since $1 - \beta(n-1)e^{-\frac{\lambda\tau}{k}}$ is decreasing in β and decreasing in k, the following observation follows:

Observation 2. (2a) $\tilde{\beta} < \hat{\beta}$ for all τ, n; (2b) If $\gamma < \tilde{\beta}$ then $\frac{d^2 E[\pi(\mathbf{q},\mathbf{t}|\boldsymbol{\theta})]}{dkdk} < 0$ for all k; (2c) If $\gamma > \hat{\beta}$ then $\frac{d^2 E[\pi(\mathbf{q},\mathbf{t}|\boldsymbol{\theta})]}{dkdk} > 0$ for all k.

We now show that there exists a $\underline{\beta}(\tau, n) > 0$ such that for all $\gamma < \underline{\beta}(\tau, n)$ the number of focal tasks is $k = n$. Denote by $\underline{\beta}(\tau, n)$ the solution to $\Phi(n, \underline{\beta}(\tau, n), x, n) = 0$. Explicitly,

$$\underline{\beta}(\tau, n) = \frac{n\left(1 - e^{-\frac{\lambda\tau}{n}}\right) - \lambda\tau e^{-\frac{\lambda\tau}{n}}}{\lambda\tau - n\left(1 - e^{-\frac{\lambda\tau}{n}}\right)}\tilde{\beta}.$$

Observation 3. Direct verification implies (3a) $\underline{\beta}(\tau, n) < \tilde{\beta}$ for all τ, n; (3b) $\underline{\beta}(\tau, n)$ is increasing in τ.

Observation 3a together with observation 2b imply that $\frac{dE[\pi(\mathbf{q},\mathbf{t}|\boldsymbol{\theta})]}{dk}$ is declining in k for all $\gamma < \underline{\beta}(\tau, n)$. So, for all $\gamma < \underline{\beta}(\tau, n)$, the lower value of $\frac{dE[\pi(\mathbf{q},\mathbf{t}|\boldsymbol{\theta})]}{dk}$ is obtained when $k = n$, and, at $k = n$ we have

$$\frac{dE\left[\pi\left(\mathbf{q}, \mathbf{t} \mid \boldsymbol{\theta}\right)\right]}{dk}\Big|_{k=n} = \frac{\gamma}{(1 + \gamma(n-1))n\left(1 + \gamma(n-1)e^{-\frac{\lambda\tau}{n}}\right)^2}\Phi(n, \gamma, \tau, n) > 0,$$

because, by observation 1, $\Phi(n, \gamma, \tau, n) > \Phi(n, \underline{\beta}(\tau, n), \tau, n)$, and, by definition, $\Phi(n, \underline{\beta}(\tau, n), \tau, n) = 0$. Hence, for all $\gamma < \underline{\beta}(\tau, n)$ the expected returns of an organisation with k focal tasks are increasing in k, which implies that the optimal organization has $k^* = n$.

We now show that there exists a $\bar{\beta}(\tau, n) > \underline{\beta}(\tau, n)$ such that for all $\gamma > \bar{\beta}(\tau, n)$ in the optimal organization the number of focal tasks is $k^* = 1$. Denote by $\bar{\beta}(\tau, n)$ the solution to $\Phi(1, \bar{\beta}(\tau, n), \tau, n) = 0$. Explicitly

$$\bar{\beta}(\tau, n) = \frac{1 - e^{-\lambda\tau} - \lambda\tau e^{-\lambda\tau}}{\lambda\tau - 1 + e^{-\lambda\tau}}\hat{\beta}.$$

2

Observation 4. Direct verification shows that: 4a. $\tilde{\beta} < \bar{\beta}(\tau,n) < \hat{\beta}$, for all τ and n; 4b. $\bar{\beta}(\tau,n)$ is increasing in τ.

Observation 1 together with $\Phi(1,\bar{\beta}(\tau,n),\tau,n) = 0$ imply that $\Phi(1,\gamma,\tau,n) < 0$ for all $\gamma > \bar{\beta}(\tau,n)$. Similarly, observation 1 together with $\Phi(n,\underline{\beta}(\tau,n),\tau,n) = 0$ and observation 4a, imply that $\Phi(n,\gamma,\tau,n) < 0$ for all $\gamma > \bar{\beta}(\tau,n)$. So, $\frac{dE[\pi(\mathbf{q},\mathbf{t}|\boldsymbol{\theta})]}{dk}$ is negative at $k = 1$ and at $k = n$. Observation 4a and observation 2b implies that $\frac{dE[\pi(\mathbf{q},\mathbf{t}|\boldsymbol{\theta})]}{dk}$ is either first decreasing in k and then increasing in k (when $\gamma \in [\bar{\beta}(\tau,n),\hat{\beta}]$) or it is always increasing in k (when $\gamma > \hat{\beta}]$). Hence, the profits of the organization are decreasing in k for all $\gamma > \bar{\beta}(\tau)$ and therefore the optimal organization has $k^* = 1$.

We now conclude by considering the case where $\gamma \in (\underline{\beta}(\tau,n),\bar{\beta}(\tau,n))$. From the analysis above we infer that the marginal expected profits to k of the organization around $k = 1$ are positive, because $\Phi(1,\gamma,\tau,n) > 0$, and that the marginal expected profits of the organization around $k = n$ are negative, because $\Phi(n,\gamma,\tau,n) < 0$. Furthermore, observation 2b implies that, for all $\gamma \in (\underline{\beta}(\tau,n),\bar{\beta}(\tau,n))$, the marginal expected profits of the organization, $\frac{dE[\pi(\mathbf{q},\mathbf{t}|\boldsymbol{\theta})]}{dk}$, are either always decreasing in k (when $\gamma \in [\underline{\beta}(\tau,n),\tilde{\beta}]$) or they are first decreasing in k and then increasing in k (when $\gamma \in [\tilde{\beta},\bar{\beta}(\tau,n)]$). Hence, there exists a unique $k^* \in [1,n]$ such that $\frac{dE[\pi(\mathbf{q},\mathbf{t}|\boldsymbol{\theta})]}{dk}|_{k=k^*} = 0$; such value of k^* is the solution to $\Phi(k^*,\gamma,x,n) = 0$ and, k^* maximizes the expected profit of the organization. Finally, by applying the implicit function theorem, $dk^*/d\gamma < 0$ if and only if $d\Phi(k^*,\gamma,\tau,n)/dk < 0$. Note that this last inequality holds because the fact that there exists a unique k^* in which $\Phi(k^*,\gamma,\tau,n) = 0$ and the fact that $\Phi(1,\gamma,\tau,n) > 0$ and $\Phi(n,\gamma,\tau,n) < 0$, assure that for all $\gamma \in (\underline{\beta}(\tau,n),\bar{\beta}(\tau,n))$ the function $\Phi(k,\gamma,\tau,n)$ is decreasing around k^*.

We have therefore shown that for every $k \in \{1,...,n-1\}$ there exists a $\beta(k+1) < \beta(k)$ such that: a. if $\gamma = \beta(k+1)$ the optimal organization has $k^* = k+1$; b. if $\gamma \in (\beta(k+1),\beta(k))$ the optimal organization has either $k^* = k$ or $k^* = k+1$, and c. if $\gamma = \beta(k)$ the optimal organization has $k^* = k$.

We now show that for every $k \in \{1,...,n-1\}$ there exists a unique value of $\gamma \in (\beta(k+1),\beta(k))$, say $\bar{\beta}(k)$, such that at $\gamma = \bar{\beta}(k)$ the expected profit of the organization with k focal tasks is the same as the expected profit of the organization with $k+1$ focal tasks. For brevity define $G(x) = e^{-\frac{\lambda\tau}{x}}$ and denote by $\Delta(k,\gamma)$ the difference between the expected profit generated by $k+1$ focal tasks and the expected profit generated by the organization with k focal tasks. We obtain

$$\Delta(k,\gamma) = \sigma_\theta^2 \left[\frac{k+1}{1+\gamma(n-1)G(k+1)} - \frac{k}{1+\gamma(n-1)G(k)} - \frac{1}{1+\gamma(n-1)} \right].$$

Taking the minimum common denominator, we have that $\Delta(k, \gamma) = 0$ if, and only if,

$$(1 + \gamma(n-1))\left[(k+1)(1 + \gamma(n-1)G(k)) - k(1 + \gamma(n-1)G(k+1))\right] -$$
$$- [1 + \gamma(n-1)G(k)][1 + \gamma(n-1)G(k+1)] = 0.$$

This is a quadratic equation in γ and therefore there are only two solutions of γ. Moreover, it is immediate to check that $\gamma = 0$ is one of the solution. Hence, there is only one non-zero solution. Simple algebra shows that the non-zero solution is given by expression 13. This completes the proof of Lemma A. ∎

4.2 Appendix B: Endogenous Organizational Size

We endogenize organizational size n^*. A possible interpretation of our model is that each task corresponds to a different type of product or service that is produced by a multi-product firm. By engaging in multiple tasks, firms can spread out some fixed costs $F > 0$ and realize scope economies (Panzar and Willig, 1981). Doing so, however, increases coordination costs as now more tasks need to be coordinated. We maintain the assumption of our base line model that each agent i has an attention capacity τ to participate in public meetings.[13]

Let $\phi_i = \phi$ for all $i \in \mathcal{N}$ and let $k^*(n)$ is the optimal number of focused tasks given size n (Proposition 2). We assume that pay-offs of an organization of size n are given by

$$\tilde{\Pi}(n) = \Pi(k^*(n)) - F$$

where organizational size is chosen to maximize profits per product-line, i.e.,

$$n^* = \arg\max_n \frac{1}{n} E[\tilde{\Pi}(n)]$$

Our underlying assumption is that firms, whenever profitable, have the option to operate a set of product lines independently as a separate organization. Note that splitting up a single firm, with $n^* = m$ agents participating in one meeting, into two independent firms, each with $n^* = m/2$ agents participating in two different meetings, does not create additional communication capacity. Total agent time spent in meetings remains $m\tau$.

Management scholars have cited many reasons for the rise of new organizational forms, but there are two prominent lines of explanation. The first is the "increased turbulence" that managers face because of rapid technological changes, deregulation, and globalization

---

[13]Further, agent i needs to be present in a public meeting with agent j both to learn about agent $j's$ primary action and for agent j to learn about agent $i's$ primary action.

(Rivkin and Siggelkow, 2005; Roberts and Saloner, 2013). In our model this corresponds to an increase in the volatility of the environment σ_θ^2.

Proposition 6 *Assume $\phi_i = \phi$ for all $i \in \mathcal{N}$. The optimal organization size n^* is decreasing in σ_θ^2. Furthermore, $k^*(n^*)/n^*$ is increasing in σ_θ^2. If $k^*(n^*) < n^*$, the number of focal tasks is increasing in σ_θ^2.*

As σ_θ^2 increases, so do the incentives to adapt, which in turn bring coordination costs. By narrowing firm scope (reducing n^*), and increasing the number of focused tasks, organizations partially reduce these coordination costs, allowing for a better adaptation. Proposition 6 therefore reflects the common idea that smaller organizations are more "nimble" and "flexible." [14] Note while organizational scope n^* depends on the variance σ_θ^2, the decision on how many tasks to focus given n, that is $k^*(n)$, is independent of σ_θ^2 if n is fixed. Intuitively, an increase in the variance does not change the trade-off between adaptation and coordination, but it does affects the benefits of resolving this trade-off (for example by reducing size). Thus, for a given level of adaptiveness to local shocks α_i, an increase in the variance increases both expected adaptation losses (as primary actions are then, on average, further away from the realized shock) and expected coordination losses (as primary actions are then, on average, further away from the uninformed coordinating actions). The optimal level of adaptiveness, however, is not affected and neither is the optimal level of organizational focus. But since coordination and adaptation losses are larger with a larger σ_θ^2, it pays for the organization to invest more in communication technology (as in online Appendix G) or to reduce organizational size and incur higher fixed costs (as in this Appendix). Note, finally, that while $k^*(n)$ does not depend on the variance σ_θ^2 of schocks, $k^*(n)$ is decreasing in the importance of adaptation to those shocks, ϕ.

The other prominent line of explanation is improvements in information and communication technology. In our model this corresponds to an increase in $\lambda\tau$. One may conjecture that an increase in the effective communication capacity always results in (weakly) larger organizations. The next proposition states that this is not necessarily the case

Proposition 7 *Organizational size n^* may be decreasing in communication capacity $\lambda\tau$ when $\lambda\tau$ is small.*

Intuitively, what restricts organizational size, is the adaptiveness of the organization, not communication capacity. When the communication capacity ($\lambda\tau$) is very limited, organizations often give up on adapting to local shocks, and choose n^* large in order to minimize

---

[14]Rantakari (2013b) obtains a related result in a different setting. He shows how firms operating in more volatile environments decentralize decision-making and reduce task-interdependence, whereas in our model, firms become more balanced and reduce firm scope.

Figure 1: **Endogenous organizational size and focus as a function of τ**

Optimal organizational size, n^* (continuous line), and focal tasks, k^* (dashed line), as a function of τ. In this example the maximum number of tasks is $\bar{n} = 18$, $\sigma_\theta^2 = 1$, $\beta = .25$ and $F = 3$.
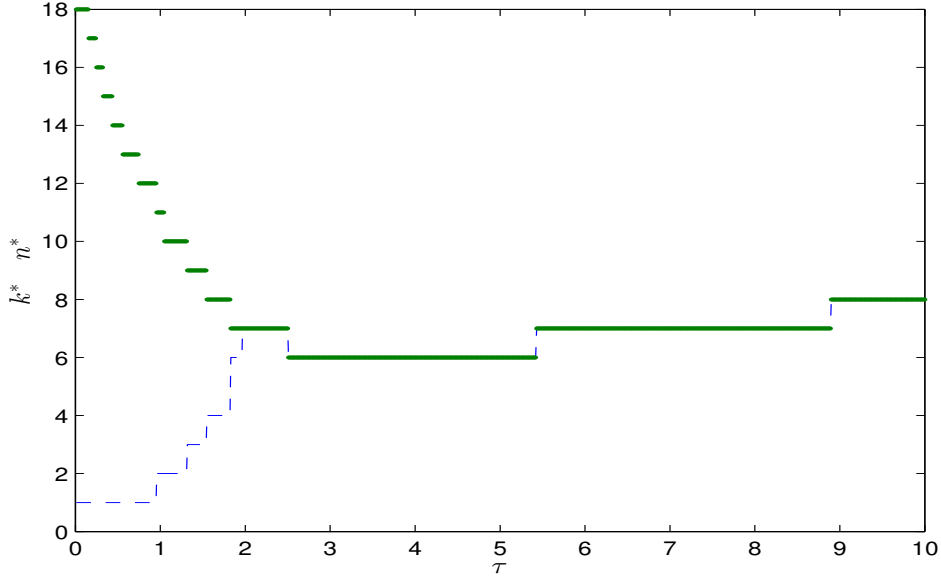
average fixed costs/maximize economies of scope. As $\lambda\tau$ becomes larger, the organization then uses the extra communication capacity to become more adaptive. Doing so without incurring substantial coordination costs, however, requires reducing organizational size n^*, often substantially. As a result of an increase in $\lambda\tau$ the organization then moves from a "large, rigid bureaucracy" into a "nimble, adaptive democracy". For larger values of $\lambda\tau$, organizational size slowly increases again with τ. Figure 1 illustrates changes in organizational size and on the number of focal tasks in response to changes in communication capacity τ. For simplicity, it is assumed that n^* is constrained to $n^* \leq \bar{n} = 18$. For τ very small, the organization size is set at the maximum, $n^* = \bar{n}$. As communication capacity τ increases, the organization is transformed from a "large, rigid bureaucracy" with eighteen tasks but only one focused task into a "nimble, adaptive democracy" with six tasks which share the attention evenly. For larger values of τ, organizational size slowly increases again with τ and attention remains evenly distributed.

Our model thus predicts that improvements in ICT may result in a shift from large inflexible organizations emphasizing economies of scale and scope, towards smaller, more balanced organizations, which are focused on being adaptive to external shocks and emphasize horizontal communication linkages.[15] This is consistent with recent trends in organization design,

---

[15]This prediction stands in contrast with those of obtained in recent team-theory models that model organizations as information-processing (Bolton and Dewatripont 1994) or problem-solving institutions (Gar-

as described by Whittington et al. (1999) and Roberts and Saloner (2013). According to our model, only organizations that are already very adaptive, respond to ICT improvements by increasing organizational scope. Alternatively, observed trends toward de-sizing and de-scoping may have been a response to an increased variability in the environment (Proposition 6), for example because of globalization and increased competition (Rivkin and Siggelkow, 2005; Roberts and Saloner, 2013).

Proof of Proposition 6. We prove that the optimal organization size is decreasing in σ_θ^2. Recall that k_{n+1}^* is the optimal number of focal tasks given $n+1$ tasks and k_n^* is the optimal number of focal tasks given n tasks. Then

$$\frac{E[\tilde{\Pi}(n)]}{n} = P - \sigma_\theta^2 - F/n + \frac{1}{n}\left(k_n^*\frac{1}{1+(n-1)\beta e^{-\lambda\tau/k_n^*}} + (n-k_n^*)\frac{1}{1+(n-1)\beta}\right)\sigma_\theta^2 \quad (14)$$

whereas

$$\frac{E[\tilde{\Pi}(n+1)]}{n+1} = P - \sigma_\theta^2 - F/(n+1) \quad (15)$$
$$+ \frac{1}{n+1}\left[\frac{k_{n+1}^*}{1+(n-1)\hat{\beta}e^{-\lambda\tau/k_{n+1}^*}} + \frac{(n-k_{n+1}^*)}{1+(n-1)\hat{\beta}} + \frac{1}{1+(n-1)\hat{\beta}}\right]\sigma_\theta^2,$$

where $\hat{\beta} = \frac{n}{(n-1)}\beta > \beta$.

Suppose first that $k_{n+1}^* \leq n$. Then, Proposition 3 implies that $k_n^* \geq k_{n+1}^*$. To prove the proposition is then sufficient to show that

$$\Delta \equiv \frac{E[\tilde{\Pi}(n+1)]}{n+1} - \frac{E[\tilde{\Pi}(n)]}{n}$$

is decreasing in σ_θ^2. Since $\hat{\beta} > \hat{\beta}e^{-\lambda\tau/k^*}$, a sufficient condition for Δ to be decreasing in σ_θ^2 is that

$$\frac{k_n}{1+(n-1)\beta e^{-\lambda\tau/k_n^*}} + \frac{n-k_n^*}{1+(n-1)\beta} > \frac{k_{n+1}^*}{1+(n-1)\hat{\beta}e^{-\lambda\tau/k_{n+1}^*}} + \frac{n-k_{n+1}^*}{1+(n-1)\hat{\beta}}$$

Since $k_n^* \geq k^*$ and $\hat{\beta} > \beta$, this is indeed satisfied.

icano, 2000; Garicano and Rossi-Hansberg, 2006). While these papers also characterize optimal information flows in organizations, improvements in communication technology unambiguously result in larger and more centralized organizations.

Next, assume that $k_{n+1}^* = n + 1$; We then have hat $k_n^* = n$. Hence

$$\Delta = \left[\frac{1}{1 + (n-1)\hat{\beta}e^{-\lambda\tau/(n+1)}} - \frac{1}{1 + (n-1)\beta e^{-\lambda\tau/n}} \right] \sigma_\theta^2 + F/n - F/(n+1).$$

Since $\hat{\beta} > \beta$, it follows that Δ is decreasing in σ_θ^2. The second part of the proposition follows from this result and proposition 3. ∎

4.3 Appendix C: Alternative communication models

This Appendix extends the result of Proposition 2 and Proposition 3 to alternative models of communication. Without loss of generality we set, hereafter, $\phi = 1$.

B.1. Bilateral communication with aggregate organizational constraints.

We now consider that communication is bilateral and that the constraint is at the organizational level. Formally, the allocation of attention is $\mathbf{t} = \{t_{ji}\}_{ji \in \mathcal{N}}$, where t_{ji} denotes the amount of communication between agent i and agent j about local information θ_i. Let τ be the total communication capacity of the organization. Then, we require that $\mathbf{t}$ satisfies

$$\sum_i \sum_j t_{ij} \leq \tau.$$

We maintain the assumption that $r(t_{ij}) = 1 - e^{-\lambda t_{ij}}$. The following equivalent result obtains:

Result 1. In an optimal organization under bilateral communication and constraint τ, the allocation of attention $\mathbf{t} = \{t_{ji}\}$ satisfies

$$t_{ji} = t_i^P \text{ for all } i, j \in \mathcal{N},$$

where $\mathbf{t}^P = \{t_1^P, ..., t_n^P\}$ is the allocation of attention in an optimal organization under public communication and constraint $\tau^P = \tau/(n-1)$.

Proof of Result 1. The key step for this equivalence result is the proof of the following Lemma

Lemma 8 *Consider bilateral communication and constraint τ. In an optimal organization all agents devote the same attention to a particular agent, that is, for all $i \in \mathcal{N}$, $t_{ji} = t_{ki}$ for all $j, k \in \mathcal{N} \setminus \{i\}$.*

Proof of Lemma 8. Suppose that $\mathbf{t}$ is optimal and, for a contradiction, assume that there exists some agent i such that $t_{ji} > t_{ki} \geq 0$. Define a new organization $\mathbf{t}'$, which is the

same as $\mathbf{t}$ with the exception that $t'_{ji} = t_{ji} - \epsilon$ and $t'_{ki} = t_{ki} + \epsilon$, for some small and positive ϵ. Using the expression for expected payoffs, it is easy to verify that

$$E\left[\pi\left(\mathbf{q}, \mathbf{t}|\boldsymbol{\theta}\right)\right] - E\left[\pi\left(\mathbf{q}, \mathbf{t}'|\boldsymbol{\theta}\right)\right] \geq 0,$$

if, and only if,

$$e^{-\lambda t'_{ji}} + e^{-\lambda t'_{ki}} \geq e^{-\lambda t_{ji}} + e^{-\lambda t_{ki}}. \tag{16}$$

Since $t'_{ji} = t_{ji} - \epsilon$ and $t'_{ki} = t_{ki} + \epsilon$, after some algebra we obtain that condition 16 is equivalent to

$$e^{-\lambda t_{ki}} \leq e^{-\lambda(t_{ji} - \epsilon)} \iff t_{ki} \geq t_{ji} - \epsilon,$$

which, for ϵ sufficiently small, contradicts our initial hypothesis that $t_{ji} > t_{ki}$. This completes the proof of Lemma 8. ∎

Note that under bilateral communication and arbitrary capacity τ, Lemma 8 implies that the optimal allocation of attention $\mathbf{t}$ satisfies $t_{ji} = t_{li}$ for all $j, l \neq i$. Hence, in the optimal organization every agent $j \neq i$ devotes the same attention to agent i, that is the restriction imposed by public communication. It is immediate to see the relation between τ and τ^P. ∎

B.2. Individual Communication Constraints

So far we have assumed that the communication constraint is determined at the organizational level. Alternatively, each agent may have a limited communication capacity τ^I. Formally, let each agent have access to an individual communication channel, whose finite capacity τ^I can be used to broadcast information to all other agents and/or to process information broadcasted by others. Each agent i then optimally decides on a vector $t_i = [t_{i1}, t_{i2}, ..., t_{ii}, ..., t_{in}]$, where

$$\sum_{j \in N} t_{ij} \leq \tau^I \qquad \forall i \in \mathcal{N}, \tag{17}$$

and where t_{ii} is the capacity devoted to broadcast information about θ_i, and t_{ij} is the capacity devoted to listen to the information broadcasted by agent $j \neq i$. We maintain the assumption that $1 - r(t_{ij}, t_{jj}) = e^{-\lambda \max\{t_{ij}, t_{ii}\}}$ We now proof the following equivalence result, which again implies that the optimal organization is focused on k^* tasks with $k \in \{1, 2, \cdots, n\}$ and that the same comparative statics hold as in Proposition 3.

Result 2. Under individual communication and individual capacity constraint τ^I, in the

optimal organization the allocation of attention $\mathbf{t} = \{t_{ij}\}_{i,j}$ satisfies

$$t_{jj} = t_{ij} = t_{ij}^b \ \forall i, j \in \mathcal{N}$$

where $\mathbf{t}^b = \{t_{ij}^b\}_{i \neq j}$ is the allocation of attention in the optimal organization under bilateral communication and capacity constraint $\tau = (n-1)\tau^I$.

Proof of Result 2. Consider the case of individual communication with individual capacity constraint τ^I. Suppose that $\mathbf{t}$ is an optimal organization. It is immediate to see that $\mathbf{t}$ satisfies: a. $t_{ji} \leq t_{ii}$ for all $i, j \in \mathcal{N}$ and b. $\sum_j t_{ji} = \tau^I$ for all $ij \in \mathcal{N}$. Now note that if $\mathbf{t}^b$ is an optimal organization under bilateral communication and constraint $\tau = (n-1)\tau^I$, then organization $\mathbf{t}^*$ with $t_{ji}^* = t_{ii}^* = t_{ji}^b$ is a feasible organization under individual communication and satisfies property a. and b. above. We now claim that $\mathbf{t}^*$ is optimal under individual communication and individual capacity constraint τ^I. Suppose there is another organization $\mathbf{t}$ that does strictly better than $\mathbf{t}^*$. First, note that the expected profit of an organization, for a given $\mathbf{t}$, can be written in terms of residual variances as follows

$$E[\pi(\mathbf{t}|\theta)] = -n\sigma_\theta^2 + \sum_{i=1}^n \frac{\sigma_\theta^2}{\sigma_\theta^2 + \beta(n-1)\sum_{j=1}^n \mathsf{RV}(t_{ji}, t_{ii})},$$

where $\mathsf{RV}(t_{ji}, t_{ii}) = \sigma_\theta^2(1 - r(t_{ji}, t_{ii}))$. Second, $\mathbf{t}$ must satisfy property a and property b and therefore $\min\{t_{ji}, t_{ii}\} = t_{ji}$, and so the residual variance that agent j has about task i is $\mathsf{RV}(t_{ji})$. Since $\mathbf{t}$ is strictly better than $\mathbf{t}^*$ is follows that the profile of residual variances $\{\mathsf{RV}(t_{ji})\}_{ji}$ is better than $\{\mathsf{RV}(t_{ji}^*)\}_{ji}$. But then, construct $\hat{\mathbf{t}}^b$ as follows: $\hat{t}_{ji}^b = t_{ji}$. Note that $\hat{\mathbf{t}}^b$ is feasible under bilateral communication and capacity τ. Furthermore since the profile of residual variances $\{\mathsf{RV}(t_{ji})\}_{ji}$ is better than $\{\mathsf{RV}(t_{ji}^*)\}_{ji}$, it must also be true that profile of residual variances $\{\mathsf{RV}(\hat{t}_{ji}^b)\}_{ji}$ is better than $\{\mathsf{RV}(t_{ji}^b)\}_{ji}$, and so $\hat{\mathbf{t}}^b$ must be strictly better than $\mathbf{t}^b$, which contradicts our initial hypothesis that $\mathbf{t}^b$ is optimal. $\blacksquare$

4.4 Appendix D: Information Theory.

While we posit a specific binary communication technology, this Appendix shows that identical results obtain if communication is noisy instead and, following the literature of Rational Inattention, entropy is used to model the cost of more precise information.

In particular, we now consider that messages m_{ji} and local information θ_i are normally distributed and the attention constraint $\sum_i t_i \leq \tau$ is modelled as a constraint on the total reduction in entropy, as in Information Theory (Cover and Thomas 1991) and the literature on Rational Inattention (Sims 2003). This specification leads to the same residual variance

that is obtained in the binary communication technology adopted in the paper. Since, in an equilibrium with linear strategies, the expected profits for a given attention allocation $\mathbf{t}$ can be written as a function of the residual variance, identical results obtain with this alternative communication technology. We now develop these arguments formally.

For simplicity, we focus on the two-task case. Let m_i be a message about θ_i and let $m = (m_1, m_2)$. The mutual information between m and θ, denoted by $I(\theta; m)$, equals the average amount by which the observation of m reduces uncertainty about the state θ, where the ex ante uncertainty is measured by the (differential) *entropy* of θ,

$$H(\theta) = - \int f(\theta) \log f(\theta) d\theta,$$

and the uncertainty after observing m is measured by the corresponding entropy

$$H(\theta|m) = - \int f(\theta|m) \log f(\theta|m) d\theta.$$

Denoting by τ the (Shannon) capacity of the communication channel, the constraint on information conveyed by m about θ is given by [16]

$$I(\theta; m) = H(\theta) - H(\theta|m) \leq \tau. \tag{18}$$

Following Sims (2003) and the subsequent literature on rational inattention, we assume that θ_1 and θ_2 are (independently) normally distributed, and communicated through a Gaussian communication channel which contaminates its inputs with independent normally distributed noise, e.g., $m_i = \theta_i + \epsilon_i$, where ϵ_i is normally distributed. As a result, also m_1 and m_2 and the conditional distributions $F(\theta_1|m_1)$ and $F(\theta_2|m_2)$ are independently normally distributed. As noted by Sims, Gaussian communication channels minimize the variance of $F(\theta_i|m_i)$ given the constraint (18) on the mutual information between θ_i and m_i. Hence, they maximize the correlation between m_i with θ_i.[17] Given that θ_1 and θ_2 are independently distributed, we have

$$I(\theta; m) = I(\theta_1; m_1) + I(\theta_2; m_2), \tag{19}$$

where $I(\theta_i; m_i) = H(\theta_i) - H(\theta_i|m_i)$. Moreover, since the entropy of a normal variable with

---

[16] The capacity of a channel is a measure of the maximum data rate that can be reliably transmitted over the channel. Shannon capacity has proven to be an appropriate concept for studying information flows in a variety of disciplines: probability theory, communication theory, computer science, mathematics, statistics, as well as in both portfolio theory and macroeconomics.

[17] This follows from a well known result in information theory that among all distributions with the same level of entropy, the normal distribution minimizes the variance.

variance σ^2 is given by $\frac{1}{2} \ln(2\pi e \sigma^2)$, we obtain

$$I(\theta_i, m_i) = \frac{1}{2} \left(\ln \sigma_\theta^2 - \ln \text{Var}(\theta_i | m_i) \right). \tag{20}$$

It follows that the constraint (18) on the mutual information between θ and m can be rewritten as

$$\ln \sigma_\theta^2 - \ln \text{Var}(\theta_1 | m_1) + \ln \sigma_\theta^2 - \ln \text{Var}(\theta_2 | m_2) \leq 2\tau. \tag{21}$$

We can now re-interpret the *mutual information between m_i and θ_i* as the *attention devoted by the organization to task i*. Denoting $t_1 \equiv I(\theta_1, m_1)$ and $t_2 \equiv I(\theta_2, m_2)$, the constraint on mutual information (18) imposed by the Shannon capacity becomes equivalent to our attention constraint $t_1 + t_2 \leq \tau$.

Using the above formalization, we obtain a tractable expression for $\text{RV}(t_i) \equiv Var(\theta_i | m_i)$. Indeed, from (20) and $t_i \equiv I(\theta_i, m_i)$, we have

$$\ln \text{RV}(t_i) = \ln \sigma_\theta^2 - 2t_i, \quad i = 1, 2. \tag{22}$$

or still

$$\text{RV}(t_i) = \sigma_\theta^2 e^{-2t_i}, \quad i = 1, 2, \tag{23}$$

where $t_1 + t_2 \leq \tau$. To conclude, it is easy to show that, for a given $\mathbf{t}$, the expected profits in an equilibrium with linear strategies can be written as:

$$E[\pi(\mathbf{t}|\boldsymbol{\theta})] - n\sigma_\theta^2 + \sum_{i=1}^{n} \frac{\phi \sigma_\theta^2}{\phi \sigma_\theta^2 + \beta(n-1)\text{RV}(t_i)}.$$

4.5 Appendix E: Technological trade-offs between adaptation and coordination.

In this Appendix we show that our insights hold in a model of coordination a la Alonso, Dessein, Matouschek (2008), Rantakari (2008) and Calvo et al. (2011). We consider the case for two agents, but everything can be generalized to n agents. In these class of models, instead of having the distinction between primary action and complementary action, each agent chooses one single action. We posit that agent i chooses q_i. Given a particular realization of the string of local information, $\boldsymbol{\theta} = [\theta_1, \theta_2]$, and a choice of actions, $\mathbf{q} = [q_1, q_2]$, the realized profit of the organization is:

$$\pi(\mathbf{q}|\boldsymbol{\theta}) = K - \phi(q_1 - \theta_1)^2 - \phi(q_2 - \theta_2)^2 - \beta(q_1 - q_2)^2, \tag{24}$$

12

where β is some positive constant. Without loss of generality we normalize $\phi = 1$. The communication technology follows the description in our basic model.

Standard computation allows us to derive agents' best replies, for a given network $\mathbf{t} = (t, \tau - t)$. We obtain:

$$q_1 = \frac{1+\beta}{1+2\beta+\beta^2 e^{-\lambda t_1}}\theta_1 + \frac{\beta}{1+2\beta+\beta^2 e^{-\lambda t_2}}E[\theta_2|\mathcal{I}_1] \qquad (25)$$

$$q_2 = \frac{1+\beta}{1+2\beta+\beta^2 e^{-\lambda t_2}}\theta_2 + \frac{\beta}{1+2\beta+\beta^2 e^{-\lambda t_1}}E[\theta_1|\mathcal{I}_2] \qquad (26)$$

where $E[\theta_2|\mathcal{I}_1]$ is θ_2 if communication is successful, otherwise it equals $\hat{\theta}_2$.

Substituting (25) and (26) into (24) and taking unconditional expectations we find that the problem

$$\max_{\mathbf{t}} E\pi(\mathbf{q}|\boldsymbol{\theta}) \ s.t. t_1 + t_2 = \tau$$

is equivalent to

$$\max_{t\in[0,\tau]} \frac{1}{1+2\beta+\beta^2 e^{-\lambda t}} + \frac{1}{1+2\beta+\beta^2 e^{-\lambda(\tau-t)}}$$

where $t = t_1$ and $t_2 = \tau - t$.

It is easy to verify that

$$\frac{\partial E\pi(\mathbf{q}|\boldsymbol{\theta})}{\partial t} > 0 \iff (1+2\beta)^2 - \beta^4 e^{-\lambda\tau} > 0.$$

We then obtain a result that is qualitatively the same as the one stated in Proposition 2 and Proposition 3. For every τ there exists a $\beta(\tau) > 0$, so that for all $\beta < \beta(\tau)$ the optimal organization has $t = \tau/2$, whereas for every $\beta > \beta(\tau)$ the optimal organization has $t = \{0, \tau\}$. Furthermore, $\beta(\tau)$ is increasing in τ.

4.6   Appendix F: Asymmetric Coordination Costs.

In this appendix we consider tasks that are asymmetric in terms of their potential coordination costs. That is, some tasks impose larger coordination costs (delays, low product quality) should other tasks not take the appropriate coordinating actions. For example, in designing a car, important changes made to how the engine works, may have important consequences for the remainder of the design. Should attention be focused on those highly interdependent tasks? We show that this is not necessarily the case. For conciseness of the argument, we consider the two-task case and set $\phi = 1$.

Let the coordination parameters be β_1 and β_2 for task 1 and 2, respectively. We assume that coordination problems are not trivial, i.e., $\beta_1 > \beta_2 \geq 1$. Define $\beta = \sqrt{\beta_1\beta_2}$, the geometric mean of β_1 and β_2 and consider situations where

$$\beta_1 = \beta(1+\epsilon) \qquad \text{and} \qquad \beta_2 = \beta(1+\epsilon)^{-1}.$$

The parameter ϵ thus determines the "spread" between the coordination costs across tasks: An increase in $\epsilon > 0$ increases the coordination costs associated with task 1 and decreases that of task 2, leaving the geometric average, a sufficient statistic for how costly lack of coordination is to the organization, unchanged. When $\epsilon = 0$ the case collapses to the one of ex-ante symmetric tasks.

Proposition 9 *There exists $\top(\beta) > 0$:*

1. *If $\lambda\tau < \top(\beta)$, the optimal organization is focused on task 2, i.e., $(t_1^*, t_2^*) = (0, \tau)$.*

   If $\lambda\tau \geq \top(\beta)$, let $\widehat{\epsilon}$ be the solution to $(1+\widehat{\epsilon})^2 e^{-2\tau} = 1$:

   (a) *If $\epsilon < \widehat{\epsilon}$ then $\tau > t_1^* > t_2^* > 0$.*
   (b) *If $\epsilon \geq \widehat{\epsilon}$, then $(t_1^*, t_2^*) = (\tau, 0)$.*

If attention is limited, $\lambda\tau < \top(\beta)$, then all attention is focused on the task which is *least* interdependent: Task 2. The reason is that allocating limited attention to task 1 is essentially not worth it as it would translate into limited adaptation given the large coordination costs the organization would bear. Instead, it is better to provide all attention to task 2 and let task 2 be adaptive. Task 1 is then coordinated by restricting its adaptiveness.

Instead when the attention capacity is larger and the asymmetry ϵ is not too large, both tasks receive attention but task 1 receives more than task 2. Intuitively, if both tasks are allowed to be adaptive, more attention needs to be devoted to that task that is more interdependent. If asymmetries between both tasks are sufficiently large, task 2 may even receive no attention for $\lambda\tau > \top(\beta)$. At the threshold $\lambda\hat{\tau} = \top(\beta)$, the organization then switches from being fully focussed on task 2 to being fully focussed on task 1.

Proof of Proposition 9.

We can express expected profit for a given $\mathbf{t}$ as

$$E[\pi(\mathbf{q}|\theta] = -(1-\alpha_{11})^2\sigma_\theta^2 - (1-\alpha_{22})^2\sigma_\theta^2 - \beta_1(1-r_1)\alpha_{11}^2\sigma_\theta^2 - \beta_2(1-r_2)\alpha_{22}^2\sigma_\theta^2, \qquad (27)$$

where $\alpha_{ii} = 1/(1+\beta_i(1-r(t_i)))$. Hence, the organizational problem is to choose $t_1 = t \in [0, \tau]$ to maximize expression (27). We obtain that the profits of the organization are decreasing in t, if, and only if,

$$-[1 - \beta_1\beta_2 e^{-\lambda\tau}][\beta_1 e^{-\lambda t} - \beta_2 e^{-\lambda(\tau-t)}] > 0. \tag{28}$$

It is convenient to divide the analysis in two cases. Recall that we are assuming that $\beta > 1+\epsilon$ (which is equivalent of assuming $\beta_2 > \hat{\beta} = 1$).

Case 1. Assume that $\beta_1 e^{-\lambda\tau} - \beta_2 > 0$, or $\epsilon > \hat{\epsilon}$. This assumption implies that $\beta_1 e^{-\lambda t} - \beta_2 e^{-\lambda(\tau-t)} > 0$ for all $t \in [0, \tau]$. This in turn implies that the objective function is decreasing in t if, and only if,

$$1 - \beta_1\beta_2 e^{-\lambda\tau} < 0 \iff \lambda\tau < \ln\beta$$

which is always satisfied because $\beta > 1 + \epsilon$. So, if $\lambda\tau < \ln\beta$ and $\epsilon > \hat{\epsilon}$, it is optimal to set $t = 0$ and there is focus on task 2.

Case 2. Assume now that $\beta_1 e^{-\lambda\tau} - \beta_2 < 0$, or $\epsilon < \hat{\epsilon}$. Since $\beta_1 - \beta_2 e^{-\lambda\tau} > 0$ and since $\beta_1 e^{-\lambda t} - \beta_2 e^{-\lambda(\tau-t)}$ declines in t, it follows that there exists a t^* so that $\beta_1 e^{-\lambda t^*} - \beta_2 e^{-\lambda(\tau-t^*)} = 0$. Indeed, such t^* solves $\beta_1/\beta_2 = e^{-\lambda(\tau-t^*)}/e^{-\lambda t^*}$ and since $\beta_1 > \beta_2$ and $e^{-\lambda t}$ is decreasing in t, it follows that $t^* > \lambda\tau/2$. The next two observations complete the proof:

First, if $1 - \beta_1\beta_2 e^{-\lambda\tau} > 0$, then the objective function is increasing in t for $t \leq t^*$ and it is decreasing in t for all $t > t^*$. Hence, in the optimal organization $t = t^*$. Second, if $1 - \beta_1\beta_2 e^{-\lambda\tau} < 0$, then the objective function is decreasing in t for all $t \leq t^*$ and increasing in t for all $t \geq t^*$. Hence, there are two candidates for the minimum: either $t = 0$ or $t = \tau$. Comparing the two organizations it reveals that since $1 - \beta_1\beta_2 e^{-\lambda\tau} < 0$ the optimal organization has $t = 0$, and so there is focus on task 2. Note also that $1 - \beta_1\beta_2 e^{-\lambda\tau} > 0$ and $\beta_1 e^{-\lambda\tau} - \beta_2 < 0$, are mutually compatible, if and only if, $\beta > 1 + \epsilon$, which holds by assumption. This concludes the proof of Proposition 9. ∎

4.7 Appendix G: Endogenous Attention Capacity

So far we have taken τ to be a hard constraint in the amount of time agents can devote to communication with each other. In practice this is another margin that organizations can use to improve performance, by, for example, allowing more time for meetings and communication between teams. Equivalently, the organization can increase the effective communication capacity τ, by cross-training and rotating employees, by hiring employee with higher cognitive abilities, or by investing in communication technology. Assume thus that an organization can acquire a capacity τ at a cost $C(\tau)$. $C(\tau)$ represents for example

15

the costs of having team members engaged in communications activities rather than in production. We assume that this cost has the following properties:

$$C\left(0\right) = C'\left(0\right) = 0 \qquad C'\left(\tau\right) > 0 \qquad C''\left(\tau\right) \geq 0 \quad \text{and} \quad C'''\left(\tau\right) \geq 0.$$

The problem of organizational design is now

$$\max_{\tau, \mathbf{t}} E\pi\left(\mathbf{q}|\boldsymbol{\theta}\right) - C\left(\tau\right) \qquad \text{subject to} \quad \sum_i t_i \leq \tau. \tag{29}$$

The following proposition characterizes the optimal organization in the case of two ex-ante identical agents. Without loss of generality we set $\phi = 1$.

Proposition 10 *The optimal communication capacity τ^* is increasing in σ_θ^2. Furthermore, there exists $\bar{\sigma}_\theta^2 > \underline{\sigma}_\theta^2 > 0$ such that $t_1^* \in \{0, \tau^*\}$ if $\sigma_\theta^2 \leq \underline{\sigma}_\theta^2$ and $t_1^* = \frac{\tau^*}{2}$ if $\sigma_\theta^2 > \bar{\sigma}_\theta^2$.*

Proof of Proposition 10. We first show that the optimal capacity τ^* is increasing in σ_θ^2 in the focused organization and in the balanced organization. We consider the focused organization first. Recall that the expected profits in the focused organization are

$$E\left[\pi^c\left(\mathbf{q}|\boldsymbol{\theta}\right)\right] = -\beta\sigma_\theta^2 \left[\frac{1}{1+\beta} + \frac{e^{-\lambda\tau}}{1+\beta e^{-\lambda\tau}}\right] - C(\tau).$$

Taking the derivative with respect to τ we have

$$\frac{\partial E\left[\pi^c\left(\mathbf{q}|\boldsymbol{\theta}\right)\right]}{\partial \tau} = \frac{\lambda\beta\sigma_\theta^2 e^{-\lambda\tau}}{[1+\beta e^{-\lambda\tau}]^2} - C'(\tau).$$

We now observe that, since $C'(0) = 0$, it follows that $\frac{\partial E[\pi^c(\mathbf{q}|\boldsymbol{\theta})]}{\partial \tau}|_{\tau=0} > 0$, and that, since $C'(\cdot) > 0$, it follows that $\frac{\partial E[\pi^c(\mathbf{q}|\boldsymbol{\theta})]}{\partial \tau}|_{\tau=\infty} < 0$. Moreover

$$\frac{\partial^2 E\left[\pi^c\left(\mathbf{q}|\boldsymbol{\theta}\right)\right]}{\partial \tau \partial \tau} = -\left[\frac{\lambda^2 \beta\sigma_\theta^2 e^{-\lambda\tau}}{[1+\beta e^{-\lambda\tau}]^3}\left[1 - \beta e^{-\lambda\tau}\right] + C''(\tau)\right].$$

Since $C''(\cdot) \geq 0$, $C'''(\cdot) \geq 0$ and $1 - \beta e^{-\lambda\tau}$ is negative for small value of τ (recall that $\beta > \hat{\beta} = 1$) and, as τ increases, $1 - \beta e^{-\lambda\tau}$ becomes eventually positive, it follows that $\frac{\partial^2 E[\pi^c(\mathbf{q}|\boldsymbol{\theta})]}{\partial \tau \partial \tau}$ is either negative for all $\tau > 0$, or it is positive for small value of τ and negative otherwise. Summarizing, we have shown that the function $\frac{\partial E[\pi^c(\mathbf{q}|\boldsymbol{\theta})]}{\partial \tau}$ is (i) positive at $\tau = 0$, (ii) negative at $\tau = \infty$ and (iii) it is either decreasing in τ or it is first increasing and then decreasing in τ. As a consequence of (i)-(iii) we obtain that the optimal capacity τ^c uniquely

16

solves

$$\frac{\partial E\left[\pi^c\left(\mathbf{q}|\boldsymbol{\theta}\right)\right]}{\partial \tau} = \frac{\lambda\beta\sigma_\theta^2 e^{-\lambda\tau^c}}{[1+\beta e^{-\lambda\tau^c}]^2} - C'(\tau^c) = 0.$$

Since $\frac{\partial E[\pi^c(\mathbf{q}|\boldsymbol{\theta})]}{\partial\tau}$ is increasing in σ_θ^2 and since, from above, $\frac{\partial^2 E[\pi^c(\mathbf{q}|\boldsymbol{\theta})]}{\partial\tau\partial\tau}|_{\tau=\tau^c} < 0$, an application of the implicit function theorem implies that τ^c is an increasing function of σ_θ^2. From investigation of the optimality condition of τ^c and the assumptions that $C'(0) = 0$, it follows that $\tau^c \to 0$ as $\sigma_\theta^2 \to 0$ and that $\tau^c \to \infty$ as $\sigma_\theta^2 \to \infty$.

We now consider the case in which the organization is balanced. The expected profits in the balanced organization are

$$E\left[\pi^d\left(\mathbf{q}|\boldsymbol{\theta}\right)\right] = -\frac{2\beta\sigma_\theta^2 e^{-\frac{\lambda\tau}{2}}}{1+\beta e^{-\frac{\lambda\tau}{2}}} - C(\tau).$$

Taking the derivative with respect to τ we obtain

$$\frac{\partial E\left[\pi^d\left(\mathbf{q}|\boldsymbol{\theta}\right)\right]}{\partial \tau} = \frac{\lambda\beta\sigma_\theta^2 e^{-\frac{\lambda\tau}{2}}}{\left[1+\beta e^{-\frac{\lambda\tau}{2}}\right]^2} - C'(\tau).$$

We can now proceed in the same fashion as in the case for the balanced organization to conclude that the optimal capacity τ^d uniquely solves

$$\frac{\partial E\left[\pi^d\left(\mathbf{q}|\boldsymbol{\theta}\right)\right]}{\partial \tau} = \frac{\lambda\beta\sigma_\theta^2 e^{-\frac{\lambda\tau^d}{2}}}{\left[1+\beta e^{-\frac{\lambda\tau^d}{2}}\right]^2} - C'(\tau^d) = 0,$$

and that τ^d is an increasing function of σ_θ^2, $\tau^d \to 0$ as $\sigma_\theta^2 \to 0$ and $\tau^d \to \infty$ as $\sigma_\theta^2 \to \infty$.

Since the optimal capacity in the focused and balanced organization are both increasing in σ_θ^2 and since the optimal organization is either focused or balanced, it follows that the optimal capacity of the optimal organization is increasing in σ_θ^2.

We now prove the second part of the proposition. First note that for a given common τ

$$\frac{\partial E\left[\pi^c\left(\mathbf{q},\tau|\boldsymbol{\theta}\right)\right]}{\partial \tau} - \frac{\partial E\left[\pi^d\left(\mathbf{q},\tau|\boldsymbol{\theta}\right)\right]}{\partial \tau} > 0,$$

if, and only if,

$$\frac{e^{-\lambda\tau}}{[1+\beta e^{-\lambda\tau}]^2} - \frac{e^{-\frac{\lambda\tau}{2}}}{[1+\beta e^{-\frac{\lambda\tau}{2}}]^2} > 0,$$

and, after plain algebra, this condition is equivalent to

$$-\left[e^{-\frac{\lambda\tau}{2}} - e^{-\lambda\tau}\right]\left[1 - \beta^2 e^{-\frac{3\lambda\tau}{2}}\right] > 0 \qquad \Longleftrightarrow \qquad 1 - \beta^2 e^{-\frac{3\lambda\tau}{2}} < 0.$$

Since $\tau^c(\sigma_\theta^2)$ is increasing in σ_θ^2 ranging from 0 to ∞, there exists a unique $\hat{\sigma}_\theta^2$ that solves $1 - \beta^2 e^{-\frac{3\lambda\tau^c(\hat{\sigma}_\theta^2)}{2}} = 0$. By construction, if $\sigma_\theta^2 = \hat{\sigma}_\theta^2$, then $\tau^c(\hat{\sigma}_\theta^2) = \tau^d(\hat{\sigma}_\theta^2)$. The next observation is used in the rest of the proof.

Observation 1. $\tau^d(\sigma_\theta^2) < \tau^c(\hat{\sigma}_\theta^2)$ if, and only if, $\sigma_\theta^2 < \hat{\sigma}_\theta^2$.

To see this note that since τ^c is increasing in σ_θ^2, it follows that $1 - \beta^2 e^{-\frac{3\lambda\tau^c(\hat{\sigma}_\theta^2)}{2}} < 0$ for all $\sigma_\theta^2 < \hat{\sigma}_\theta^2$. Hence, $\frac{\partial E\left[\pi^d(\mathbf{q}|\boldsymbol{\theta})\right]}{\partial\tau}|_{\tau^c(\sigma_\theta^2)} < 0$, which implies that $\tau^d(\sigma_\theta^2) < \tau^c(\sigma_\theta^2)$. Analogously, since τ is increasing in σ_θ^2, it follows that $1 - \beta^2 e^{-3\tau^c(\sigma_\theta^2)} > 0$ for all $\sigma_\theta^2 > \hat{\sigma}_\theta^2$. Hence, $\frac{\partial E\left[\pi^d(\mathbf{q}|\boldsymbol{\theta})\right]}{\partial\tau}|_{\tau^c(\sigma_\theta^2)} > 0$, which implies that $\tau^d(\sigma_\theta^2) > \tau^c(\sigma_\theta^2)$.

Define now $\underline{\sigma}_\theta^2$ as the solution to $1 - \beta^2 e^{-\lambda\tau^d(\underline{\sigma}_\theta^2)} = 0$ and define $\bar{\sigma}_\theta^2$ be such that $1 - \beta^2 e^{-\lambda\tau^c(\bar{\sigma}_\theta^2)} = 0$. We now show that $\underline{\sigma}_\theta^2 > \hat{\sigma}_\theta^2$. By definition of $\hat{\sigma}_\theta^2$ and $\underline{\sigma}_\theta^2$, we have that

$$1 - \beta^2 e^{-\frac{3\lambda\tau^d(\hat{\sigma}_\theta^2)}{2}} = 0 = 1 - \beta^2 e^{-\lambda\tau^d(\underline{\sigma}_\theta^2)},$$

which implies that $\tau^d(\underline{\sigma}_\theta^2) > \tau^d(\hat{\sigma}_\theta^2)$, and since τ^d is increasing in σ_θ^2 it follows that $\underline{\sigma}_\theta^2 > \hat{\sigma}_\theta^2$.

We next show that $\bar{\sigma}_\theta^2 > \underline{\sigma}_\theta^2$. By definition of $\bar{\sigma}_\theta^2$ and $\underline{\sigma}_\theta^2$ we have that

$$1 - \beta^2 e^{-\lambda\tau^d(\underline{\sigma}_\theta^2)} = 0 = 1 - \beta^2 e^{-\lambda\tau^c(\bar{\sigma}_\theta^2)},$$

which implies that $\tau^d(\underline{\sigma}_\theta^2) = \tau^c(\bar{\sigma}_\theta^2)$. Since $\underline{\sigma}_\theta^2 > \hat{\sigma}_\theta^2$ and since $\tau^d(\sigma_\theta^2) > \tau^c(\sigma_\theta^2)$ for all $\sigma_\theta^2 > \hat{\sigma}_\theta^2$, we have that $\tau^d(\underline{\sigma}_\theta^2) > \tau^c(\underline{\sigma}_\theta^2)$. Hence, in order for $\tau^d(\underline{\sigma}_\theta^2) = \tau^c(\bar{\sigma}_\theta^2)$ to hold we must have that $\bar{\sigma}_\theta^2 > \underline{\sigma}_\theta^2$.

We now complete the proof of the second part of Proposition 10. If $\sigma_\theta^2 \leq \underline{\sigma}_\theta^2$, then $1 - \beta^2 e^{-\lambda\tau^d(\sigma_\theta^2)} \leq 0$ and $1 - \beta^2 e^{-\lambda\tau^c(\sigma_\theta^2)} < 0$. We know that for all τ such that $1 - \beta^2 e^{-\lambda\tau} \leq 0$ the optimal organization is focused. Hence, if $\sigma_\theta^2 \leq \underline{\sigma}_\theta^2$ the optimal organization is focused. Finally, if $\sigma_\theta^2 \geq \bar{\sigma}_\theta^2$, then $1 - \beta^2 e^{-\lambda\tau^c(\sigma_\theta^2)} \geq 0$ and $1 - \beta^2 e^{-\lambda\tau^d(\sigma_\theta^2)} > 0$ and therefore it follows that the balanced organization is optimal. ∎

From the first part of the Proposition, it pays to invest more in communication capacity when the environment becomes more volatile. Intuitively, the cost of not being adapted is then larger and a better communication capacity allows for better adaptation. From Part 2, a focused organization is optimal in environments for which adaptation is not very important. Intuitively, a focused organizations is optimal when the communication capacity is limited,

and the organization does not invest much in communication capacity when adaptation is not very important. Similarly, balanced organizations are optimal when adaptation to the environment is very important, and the organization invests heavily in communication capacity.

4.8 Appendix H: Examples of organizational Change

Management scholars argue that improvements in communication technology, the increased importance of adaptation to consumer needs as well as our better understanding of the principles of management have led to a profound change in the organization of production.[18] There is indeed clear and mounting evidence of organizational change[19] but the causes behind it remain murkier.[20] In our framework several sources of exogenous variation can result in different organizational forms, such as an increase in the importance of adaptation, as measured by ϕ_i, or improvements in communication technology as proxied by either an increase in λ or investments in τ, perhaps driven by a drop in the costs of IT (see Appendix F in the new version of the paper.)

We next describe, briefly, three examples of organizational change and argue that our model helps illuminate the drivers of these changes. The first case discusses organizational changes at Procter & Gamble and, more broadly, global consumer packaged goods firms. The second example studies changes in the apparel-retail industries, which were caused by the interaction of an increased need for adaptation to fashion trends and improvements in IT such as Electronic Data Interchange (EDI). The final example is concerned with a particularly successful innovation in management called Quality Function Deployment (QFD), which is geared towards solving problems of coordination between different functions, such as marketing and engineering, in, for instance, product development and design.

4.8.1 The organization of global consumer packaged good companies.

As our first example, we discuss changes in the organization of global consumer packed good (CPG) companies in the last few decades.[21] To make the link with our model concrete, one can think of global CPG companies, such as Nestle, P&G or Unilever, as having two

---

[18]Consultants have also embraced the mantra of organizational change encouraging, for instance, the adoption of flatter organizational forms as well as the blurring of traditional hierarchical relations. See, for example, Boston Consulting Group (2006).

[19]In the literature in economics two classic references have become Rajan and Wulf (2010) and Caroli and van Reenen (2011).

[20]There are though some notable exceptions. For instance, Guadalupe and Wulf (2010) show the causal effect of competition on delayering and broader task allocation.

[21]In this section we draw heavily on HBS case 9-707-519 "Proctor & Gamble Organization 2005"

primary and equally important functions: sales and marketing/product development. The sales organization develops the firm's short-term sales strategy and is responsible for adapting the firm's product portfolio to trends in regional markets. The sales team relies on close contacts with the distribution channel for decision-relevant information. The marketing and product development team, in contrast, is responsible for the firm's long-term marketing strategy, and relies on focus groups and market tests to develop and launch new products. In order to be effective, each functional team must be responsive to its task-specific local information and undertake steps to coordinate short-term sales and long-term marketing strategies. For the latter purpose, the head of the sales organization and the head of the marketing organization hold regular conference calls, exchange emails, in addition to face-to-face meetings. Thus, as in our model, both functions require the organization to be responsive to (different) task-specific shocks and both tasks are highly interdependent.

Our model predicts that if organizational attention is scarce, it is optimal to prioritize one of these two functions, even when both are equally important for competitive succes. Thus, global consumer good companies should prioritize either global marketing and product developement – and dedicate most of the inter-task communication to discuss and coordinate new initiatives in product development or, alternatively, prioritize the regional sales organizations, and spend most of the meetings and communication on how to customize products to local tastes. Trying to excell at both functions, local customization and new product development, is bound to produce an organization which is good at neither. Improvements in communication technology, however, may change this and allow for dual objectives.

Consistent with this prediction, global CPG companies have in the past alternated between architectures that are organized along regional lines, and prioritize local customization, and structures which are organized along product lines, and favor global product development. For example, until recently, P&G used to be organized along regional lines, with global marketing managers having limited power, and each region having its own marketing function directly reporting to regional management. By the late 1990's, however, P&G was lagging behind some of its competitors in product development and new product introductions. In response to this, P&G launched a new organizational architecture, dubbed "Organization 2005." In the new organizational chart, Market Development Organizations (MDO's) responsible for sales and tailoring global strategies to local markets, and Product Divisions, responsible for global marketing initiatives and new product development, would "sit next to each other" in the organizational chart, with no hierarchical reporting relationship between them. According to P&G's legendary CEO, A.G. Lafley, the MDO's were responsible for "the first moment of truth, where the customer sees the product in the store." The product divisions were responsible for "the second moment of truth, where the customer

uses the product at home." Rather than having one function reporting to the other, as in the past, coordination in the new organization purely relied on horizontal communication. In fact, "mutual interdependence" became the new moto at P&G, and employees were given intensive training in interpersonal skills and building social networks. After some initial adjustment, the structure met with substantial success, and several of its competitors, such as Unilever, put similar structures into place. Beyond falling behind on competitors, what prompted organizational change (and its widespread imitation) is unclear. A combination of improvements in communication technology and an increase in the importance of adaptation to consumer needs (arguably because of increased global competition) seem to be the most likely drivers.

4.8.2   The apparel industry and lean retailing.

The drivers of organizational change are more transparent in our second example: the apparel industry. Apparel is perhaps the quintessential example of a fashion good and apparel retailers compete furiously to match current trends. Forecasting fashion trends though is notoriously hard. As a result, many apparel retailers have recently adopted lean management methods[22] that allows them to avoid the "curse of forecasting" and instead adapt to current trends through the rapid replenishment of inventories. There was a time though when fashion, at least in some segments, was not as volatile and the need to adapt to fashion trends less acute. Men's fashion is a case in point.

The introduction of the sewing machine, the standardized body-size measurement system and the need to produce military uniforms for the Civil War led to a remarkable revolution in the production of men's clothes. Whereas in 1880 less than half of the men's suits were ready made, by 1920 that had become the norm.[23] Standardization of men's clothes extended to many other pieces of garments such as shirts, which, to use the Model T aphorism, men could buy in any color as long as it was white. Indeed men's white shirts accounted for about 72% of the market in 1962. But the social changes of the postwar period led to a new taste for fashion also amongst men. By 1972, a decade later, white shirts accounted for only 19% of the sales and "fancy shirts," anything that was not white, and sport shirts came to dominate the market.[24] The need to adapt to men's new fashion consciousness put

---

[22]According to Abernathy et al. (1995) "[T]he term lean retailing ... refers to a cluster of inter-related practices undertaken by retail channels to achieve the objective of matching consumer demand and retail supply.

[23]See Abernathy et al. (1995, p. 180).

[24]See Pashigian (1988) for a wonderful study of these changes in men's fashion. He attributes these changes to "the dramatic transition toward more casual clothing where there is greater opportunity for individual expression and creativity through product selections."

considerable pressure on traditional retailers, which in the case of maladaptation were forced to offer considerable markdowns with the consequent loss in revenue.

Simultaneously, there were considerable advances during this period in communication technology. Two were the innovations that greatly increased the ability to communicate in the apparel business. First the introduction of the Uniform Product Code (UPC; the barcode) in the mid 1970s and its widespread adoption[25] in the 1980s allowed retailers to keep track of the enormous growth of different products (or SKUs, Stock Keeping Units).[26] Second the introduction of the Electronic Data Interchange (EDI) made possible for apparel manufacturers to receive information directly from the point of sales, which transmit information about what is selling or not. Whereas the adoption of UPC is an industrywide phenomenon, EDI requires specific investments by firms to connect directly points of sales to apparel manufacturers.[27]

Our model speaks to these issue as follows. Consider Figure 2. There we show the case of the organization of apparel production, which is comprised of three agents: Headquarters (HQ), where managerial and other decisions are taken, the Shop, in contact with customers, and the Supplier in charge of producing the garments. Traditional retailers are structured as in (a). In this case both the Shop and Supplier direct their attention to HQ, in charge of establishing product design and quality standards to suppliers and supplying the shop directly. Instead the lean retailing model is as in (b). There the company invests in EDI (and the adoption of the UPC) and now HQ and the Shop both communicate with each other and the supplier directs its attention to both.[28]

There is considerable anecdotal evidence that this is what happened with some of the apparel retailers in the early 1970s. Many of the large department stores failed to meet the increased need for adaptation (an increase in ϕ) which opened the door to new, more specialized, retailers with new lean management techniques and thus more adaptive (such as Esprit, founded in 1968 or The Gap, founded in 1969).

There is also more systematic evidence. Hwang and Weil (1998) look at a sample of 103 apparel business units between 1988 and 1992. Together they comprise 20% of all apparel shipments in the US. They find that the business units that adopted the lean retailing

---

[25]Barcodes were extended to the products sold by the mass retailers such as Walmart between 1983 and 1987. See Abernathy et al. (2000).

[26]For instance an average food store has gone from offering about 6000 SKUs to customers in the 1960s, to about 25,000 in the early 1990s, to almost 40,000 a decade later; see Abernathy et al. (2000).

[27]For a survey of the adoption of lean practices in the apparel business see Aberthany et al. (1995).

[28]There are also changes in the organization of production in the supplier as documented in the literature on lean retailing. Suppliers supplying lean retailers abandon traditional production methods, the Progressive Bundle System, in favor of methods of production that emphasize team work and job rotation; for these implications the model of Dessein and Santos (2006) is more appropriate.

Figure 2: EDI and retailing

(a) Traditional retailer (b) Lean retailer
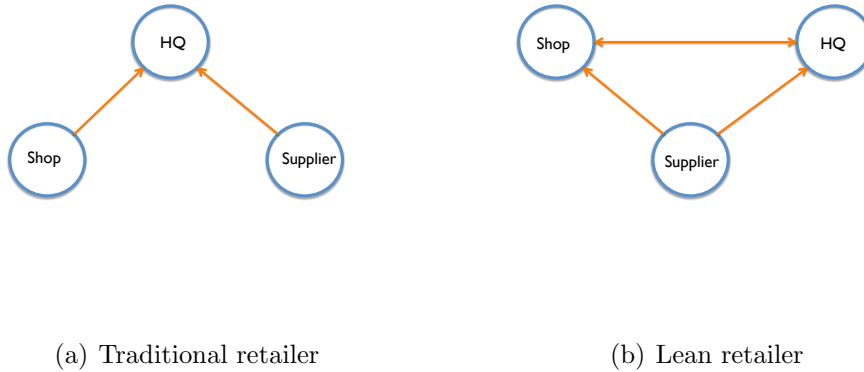
manufacturing practices and transitioned from Figure 2 (a) to (b) invested heavily in EDI. The drop in information costs were of course industry wide. They show that what explains the adoption of these practices was precisely the need for quick replenishment of inventories, which they take as a proxy for the increased need to adapt to customer tastes.[29] In sum, the adoption of these more horizontal communication networks allows retailers to adjust the supply of products offered to consumers to match actual levels of demand for different products: "By using daily demand information arising from point-of-sale data collected at the store-level to govern supply, modern retailers change the flow of information and goods with apparel suppliers." (see Hwang and Weil, 1998, p. 7).

4.8.3 An innovation in management: Quality Function Deployment

The adoption of lean manufacturing in the apparel business has the striking characteristic that the short production cycle of garment allows for the direct connection between the customer and the manufacturer. In other sectors such a direct connection between the customer and the manufacturer is simply not feasible given the length of the production cycle. Here different solutions have to be found to the problem of adaptation without direct customer contact. One such famous solution is the Quality Function Deployment (QFD) framework, which tries to integrate customer needs at all stages of the design and production processes.[30] As explained by Hauser and Clausing (1988) "[a] set of planning and communication rou-

---

[29]The measure is constructed by the percentage of sales provided by apparel business units to the retailers on a daily and weekly basis.

[30]The literature on QFD is simply staggering and the number of cases studied as well. Here we discuss QFD in the context of design or product developments but there are indeed many other applications. For an overwhelming survey of the literature see Chan and Wu (2002).

tines, quality function deployment focuses and coordinates skills within an organization, first to design, then to manufacture and market goods that customers want to purchase and will continue to purchase. The foundation of the house of quality[31] is the belief that products should be designed to reflect customers' desires and tastes - so marketing people, design engineers, and manufacturing staff must work closely together from the time a product is first conceived." This technique, pioneered in the early 70s in Japan,[32] stands in contrast to the traditional phase review process, where each stage in the design process is reviewed by management before it proceeds to the next one. Instead under the QFD framework, marketing, engineering and R & D are supposed to collaborate and communicate actively to integrate customer needs from the start. In terms of our model, QFD can be seen as an innovation in management which makes ex post coordination and communication more effective. As such, it can be interpreted as an increase in λ.

Most relevant with respect to our model, QFD results in clear communication patterns inside the organization that differ from the communications patterns of other organizational arrangements. Griffin and Hauser (1992) compare communication patterns in two new-product teams working on parallel component projects in the automobile industry, one working under QFD and the other subject to the phase review process described above. They find that "QFD enhances communication levels within the core team (marketing, engineering, manufacturing). QFD changes communication patterns from "up-over-down" flows through management to more horizontal routes where core team members communicate directly with one another." Interestingly the QFD team communicates *less* with members which are external to the core team (whereas in the phase review process everyone communicates through management in each of the stages, including those parties outside the team). Thus, consistent with the model, the adoption of QFD led to a starch dichotomy between two types of tasks: Internal ones engaged in active communication and external ones with limited input into the core activities.

Finally, a hallmark of the house of quality, the main tool to implement QFD, is the identification and prioritization of engineering and production targets in order to adapt to particular customer demands. Indeed, the house of quality features a double entry chart in

---

[31]The house of quality is a particular technique for the implementation of QFD. See Hauser and Clausing (1988) for an example of a house of quality applied to the design of car doors.

[32]According to Chan and Wu (2002), the first application of QFD techniques was in the Kobe Dockyard of Mitsubishi Heavy Industries in 1971, followed shortly afterwards by its adoption by Toyota, first Toyota's Hino Motor in 1975, then in Toyota Autobody in 1977 with impressive results, and finally into the whole Toyota group. In the US, and always according to Chan and Wu (2002), the first recorded case study in QFD was probably in 1986 when Kelsey Hayes used QFD to develop a coolant sensor, "which fulfilled critical customer needs such as "easy-to-add coolant," "easy-to-identify unit," and "provide cap removal instructions," " Other early adopters included 3M, AT&T, Ford, ...

which customer attributes are matched to engineering targets and are marked to reflect their relative importance, which seems to correspond well with the version of the model where adaptation to particular tasks differ in their importance.[33]

References

Aberthany, F. H. J. T. Dunlop, J. H. Hammond and D. Weil (1995) "The information-integrated channel: A study of the US apparel industry in transition," Brookings Papers: Microeconomics, 175-246.

Aberthany, F. H., J. T. Dunlop, J. H. Hammond, and D. Weil (2000) "Retailing and supply chains in the information age," Technology in Society, 22 , 5-31.

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

Boston Consulting Group (2006) "Global Delayering for Competitive Advantage," BCG Opportunities for Action, April.

Caroli, E. and J. van Reenen (2011) " Skill-biased organizational change? Evidence from a panel of British and French establishments," Quarterly Journal of Economics 116, 1449-1492.

Chan, L-K., and M-L. Wu (2002) "Quality Function Deployment: A Literature Review," European Journal of Operational Research 143, 463-497.

Cover, Thomas M., and Joy A. Thomas (1991). *Elements of information theory.* John Wiley & Sons, New York.

Garicano, Luis (2000). "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economy,* 108, 874-904.

Garicano, Luis, and Esteban Rossi-Hansberg (2006). "Organization and Inequality in a Knowledge Economy." *Quarterly Journal of Economics* 121(4): 1383-1435.

Griffin, A and J. R. Hauser (1992) "Patterns of Communication Among Marketing, Engineering and Manufacturing - A Comparison Between Two New-Product Teams," Management Science, vol. 38 no. 3, March.

Guadalupe, M. and J. Wulf (2010) "The Flattening firm and prodct market competition," American economic Journal: Applied Economics, 2, 105-127.

---

[33]The entries in the house of quality are, for instance, marked with red to denote the critical aspect of meeting targets for a particular feature of the product development or the engineering target. In the example provided by Hauser and Clausing (1988) of the use of the house quality for car door design the importance of meeting particular targets are marked strong and medium positive and strong and medium negative.

Hauser, J. R. and D. Clausing (1988) "The House of Quality," Harvard Business Review, May - June, 63-73.

Hwang, M. and D. Weil (1998) "The Diffusion of Modern Manufacturing Practices: Evidence from the Apparel-Retail Sectors," working paper, available in SSRN.

Panzar, John C., and Robert D. Willig (1981). "Economies of scope," *American Economic Review*, 268-272.

Pashigian, B. P. (1988) "Demand Uncertainty and Sales: A Study of Fashion and Markdown Pricing," American Economic Review, December.

Rajan, R. and J. Wulf (2010) "The flattening firm: evidence from panel data on the changing nature of corporate hierarchies," Review of Economics and Statistics 88, 759-773.

Rantakari, Heikki (2013b), "Organizational Design and Environmental Volatility." *Journal of Law, Economics, and Organization,* 29(3), 569-607.

Siggelkow, Nicolaj, and Jan W. Rivkin (2005). "Speed and Search: Designing Organizations for Turbulence and Complexity." *Organization Science*, 16: 101-122.