

## Managerial Style and Attention<sup>†</sup>

By WOUTER DESSEIN AND TANO SANTOS\*

*Is firm behavior mainly driven by its environment or rather by the characteristics of its managers? We develop a cognitive theory of manager fixed effects, where the allocation of managerial attention determines firm behavior. We show that in complex environments, the endogenous allocation of attention exacerbates manager fixed effects. Small differences in managerial expertise then may result in dramatically different firm behavior, as managers devote scarce attention in a way that amplifies initial differences. In contrast, in less complex environments, the endogenous allocation of attention mitigates manager fixed effects. Firm owners prefer “managers with style” only in complex environments. (JEL D21, D23, G34, M10, M31, M54)*

Scholars in management have long emphasized the role of executive leadership on organizational outcomes. According to the “upper echelons theory,” as set forth by Hambrick and Mason (1984), a central requirement for understanding organizational behavior is to identify those factors that direct or orient executive attention. Organizational outcomes, such as strategies and performance, are expected to reflect the values and cognitive biases of top managers in the organization. In this view, a chief executive’s background in operations makes him more inclined to pursue a cost-reduction strategy, whereas a chief executive with a marketing and sales background is more likely to pursue growth strategies. Bounded rationality and biased information processing is seen as playing a central role in this process. According to Finkelstein, Hambrick, and Cannella (2009, 46):

The logic of bounded rationality hinges on the premise that top executives are confronted with far more stimuli—both from inside and outside the organization—that they can fully possibly comprehend, and that those stimuli are often ambiguous, complex, and contradictory.

A substantial body of empirical evidence supports this view.<sup>1</sup> A major concern about studies which attribute observed variation in firm behavior to manager fixed

\*Dessein: Columbia University, Graduate School of Business (email: wd2179@columbia.edu); Santos: Columbia University, Graduate School of Business (email: js1786@columbia.edu). Michael Ostrovsky was coeditor for this article. For helpful comments and discussion, we would like to thank Bob Gibbons, Alexander Guembel, Cecilia Parlato, John Roberts, three anonymous referees, as well as numerous seminar and conference participants. All errors remain our own.

<sup>†</sup>Go to <https://doi.org/10.1257/mic.20190025> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

<sup>1</sup>Barker and Mueller (2002), for example, find that CEO characteristics explain a significant proportion of the variance in R&D spending even when controlling for industry and firm-level attributes. Similarly, following a seminal paper by Bertrand and Schoar (2003), a growing literature in corporate finance has shown that managerial

effects, however, is that managerial knowledge and expertise are largely endogenous. Managers allocate attention in order to learn about strategic choices. Similarly, boards of directors decide whether or not to hire managers with expertise in certain areas. In addition, it is unclear whether biased information processing and other “human” factors are necessarily at the center of the correlation between managerial characteristics and firm behavior, as suggested by the management literature.<sup>2</sup> Unfortunately, the development of theoretical models which study the endogenous expertise of managers, and its correlation with firm behavior, has lagged compared to the growing body of empirical work.

In this paper, we theoretically study to what extent firm behavior is driven by a firm’s environment or by the characteristics of its managers. To do so, we develop a cognitive theory of manager fixed effects, where the allocation of managerial attention determines firm behavior. We show how in complex environments the endogenous allocation of attention exacerbates manager fixed effects. Small differences in managerial expertise may then result in very different firm behavior. Moreover, firm owners (e.g., boards) then optimally hire “managers with style”—that is, managers with specialized expertise in one particular task—even when the board has no preference for a particular strategy. As a result, in complex environments, manager fixed effects are predicted to be pervasive, even when managers are endogenously chosen by unbiased boards out of a large pool of potential candidates. In contrast, in less complex environments, the endogenous allocation of attention mitigates manager fixed effects and boards opt for generalists with a “broad field of vision.” We refer to environment A as being more complex than environment B if more managerial expertise or attention is required to reduce uncertainty in environment A.

In our model, a manager selectively allocates attention in order to learn about two (nonexclusive) strategic choices. Each strategic choice concerns a different task or function. For example, the two tasks may be operations and marketing, and the manager may want to learn about opportunities to reduce unit costs and grow revenues. In order to understand firm behavior, it is then important to understand how the manager allocates her attention. The role of the manager is three-fold. First, she must learn about the nature of two task-specific shocks, which inform the optimal choices pertaining to those tasks. How well she observes a particular shock depends both on her expertise (which may differ across tasks) and how much attention she devotes to each task. In our framework, expertise and attention are substitutes in the learning process, not complements, which allows for a clean interpretation of our results. Second, the manager makes a strategic choice for each task. Finally, she communicates the firm’s strategy (the two strategic choices) to the remainder of the organization, which needs to implement the chosen strategy. We study whether it is optimal for the manager to allocate more attention to certain tasks and how this

---

characteristics are strongly correlated with a variety of corporate policies, such as mergers and acquisitions, debt levels, and growth versus cost-cutting strategies. See Finkelstein, Hambrick, and Cannella (2009) for a comprehensive review.

<sup>2</sup>According to Hambrick (2018, 1782): “The central premise of upper echelons theory is that top executives view their situations—opportunities, threats, alternatives and likelihoods of various outcomes—through their own highly personalized lenses. These individualized construals of strategic situations arise because of executives’ experiences, values, personalities and other human factors. Thus, according to the theory, organizations become reflections of their top executives.”

allocation is related to her expertise and the strategic choices she ends up making. We refer to *managing with style* as biasing managerial attention to one particular task, and mainly implementing strategies related to that same task.

As a benchmark, we consider the case where the manager only maximizes *external alignment*, that is how adaptive her strategic choices are to the task-specific shocks. A generalist manager who has equal expertise about both tasks then divides her attention evenly, whereas a specialist manager, who has more expertise about one task, compensates by devoting more attention to the task she is less knowledgeable about. Intuitively, additional signals about the same shock are partially substitutes, resulting in decreasing marginal returns to devoting attention to the same task. We refer to this as an “unbiased” allocation of attention.

How effective any given strategic choice is, however, also depends on how well it is executed by the organization, referred to as *internal alignment*. The importance of internal and external alignment has long been emphasized in the management literature and has also been very prominent in the recent organizational economics literature (Alonso, Dessein, and Matouschek 2008; Rantakari 2008, 2013; Bolton, Brunnermeier, and Veldkamp 2013; Van den Steen 2017, 2018). Internal alignment of a strategic choice depends on how well this choice is understood by the organization, which must take complementary actions to ensure effective implementation. As we show, when the need for internal alignment is sufficiently important, the manager optimally *communicates only about the task which faces the largest perceived shock*. In our example above, the manager then communicates either about a strategy for cost-reduction *or* about a strategy for revenue growth. By focusing all communication on one strategic choice, this choice can be very responsive to the corresponding task-specific shock (external alignment) without sacrificing internal alignment, as everyone in the organization has a clear understanding of this choice. Internal alignment on the other task, in contrast, is achieved by selecting a strategy that is largely unresponsive to the relevant shock, and, thus, no communication is needed to achieve coordination as the strategic choice is standard.

When internal alignment is important, the organization thus only adapts to one of the two task-specific shocks. The role for the manager, then, is to identify which task should be adaptive, and communicate a strategy to the organization focused around this task. How should she allocate her scarce attention for this purpose?

Since the manager only communicates to the organization about the task affected by what she perceives to be the largest shock, attention devoted to the other task is largely wasted from an *ex post* perspective. From an *ex ante* perspective, however, the manager does not know which shock is largest and should be the focus of the organization. The more expertise she has in one task, however, or the more attention she devotes to a particular task, the more likely it is that she will perceive as largest the shock pertaining to that task. Intuitively, assume that shocks to both marketing and operations are identically distributed but observed with noise. Then a manager, whose information about operations is very noisy, will find it difficult to identify even large shocks to operations. Indeed, her posterior estimate of a shock will typically be close to her prior estimate. In contrast, if her signal about shocks to marketing is precise, she will rarely miss even moderate shocks to marketing: her posterior estimate will typically be close to the realized shock. It follows that

a marketing expert or a manager that devotes more attention to marketing than to operations is more likely to perceive larger opportunities for revenue growth than for cost-minimization.

But a marketing expert, anticipating that she is likely to implement a marketing focused strategy, then realizes that any attention devoted to operations is likely to be a waste of time. Therefore, in complex environments and in contrast to our benchmark, marketing experts optimally devote all their attention to the marketing task, making it even more likely that a marketing-focused strategy is implemented. In contrast to our benchmark, the endogenous allocation of attention thus reinforces the initial “manager fixed effect.” Small differences in the initial expertise of two managers lead to large differences in organizational behavior, as managers specialize their allocation of attention toward areas in which they have already more expertise.

Interestingly, the same logic applies to a generalist manager who has equal expertise in both tasks. The more such a generalist manager devotes attention to, say, marketing, the more likely she is to perceive marketing shocks to be largest and, hence, the more likely she is to implement marketing-focused strategies. This, in turn, makes it optimal for this manager to devote even more attention to marketing, as information learned about operations is likely to be wasted. Because of this complementarity, and in contrast to our benchmark, a generalist manager optimally focuses her scarce attention on one task in complex environments. Thus, even if marketing and operations are equally important to the organization and have, a priori, the same potential for profit improvement, a manager with equal expertise in both marketing and operations should focus all her attention on one task, say marketing. Ex post, such a manager mainly (but not always) selects and communicates about marketing strategies, and she *appears* to be arbitrarily and inefficiently biased toward marketing. Managers may thus have a large impact on firm behavior in a way that cannot be traced back to observable managerial characteristics. Empirical research on the role of managers may therefore underestimate the impact of managers on firm behavior.

While the endogenous allocation of attention may exacerbate manager fixed effects, this is only the case when attention is scarce and the environment is complex. Indeed, when managerial attention is *not* scarce, it is possible for the manager to learn both task-specific shocks reasonably well. There is then little to be gained from focusing all attention on one task, and the manager risks being blindsided if she neglects the task in which she has less expertise. As in our benchmark, the manager thus optimally allocates attention in a way that reduces or eliminates any differences in task knowledge. Similarly, when the environment is *not* complex, the specific realization of the task-specific shocks (and not managerial attention) largely determines which task the manager communicates about to the organization. Because of decreasing marginal returns, a specialist manager then optimally allocates more attention to the task in which she has less expertise, again mitigating manager fixed effects. In the same vein, more capable managers, that is those with more initial expertise in both tasks, are less likely to “manage with style.” Indeed, more capable managers observe the environment with greater precision and it is as if the environment is less complex or uncertain. Finally, we show that “managing with style” is more likely to be optimal when organizational implementation is more important. Intuitively, the more important is coordination and internal alignment, the more the

manager ignores the shock she perceives to be smaller, and the larger are the benefits of managerial focus. In sum, the extent to which firm behavior and strategic choices reflect managerial characteristics depends on the complexity of the environment and the scarcity of attention, the capability of the manager, and the need for coordinated implementation inside the organization.

While most of our paper assumes that the initial expertise of managers is exogenous, we also study what happens when managers are endogenously chosen, for example, by a board of directors. We show that keeping total expertise fixed, specialist managers are preferred over generalists when managing with style is optimal. Our model thus predicts generalist managers to be more prevalent in less complex environments (e.g., slow-moving or mature industries). In contrast, managers with style are more likely to be optimal in more complex environments (e.g., in conditions faced by many start-up firms and technology companies).

*Related Literature.*—Following the Carnegie School (March, Simon, and Guetzkow 1958; Cyert and March 1963), a large management literature has studied limits to human cognition in order to explain organizational behavior. In particular, the Upper Echelons Theory of Hambrick and Mason (1984) focuses on the biased and subjective processing of complex, ambiguous information by managers; these biases are informed by the background and values of managers.<sup>3</sup> Instead, the present paper focuses on how managers optimally allocate scarce attention among alternative sources of information. A point of similarity is that backgrounds may matter. If the manager has some previous expertise in a particular area she is more likely to pursue strategies that emphasize those areas of expertise. Unlike the Upper Echelons Theory though, we show that this is only true if the environment is sufficiently complex. Moreover, in our framework, managers matter even when the manager has no specific background in any particular area.

In the economics literature a number of papers, such as Geanakoplos and Milgrom (1991), do study the optimal allocation of attention in organizations, but almost all are focused on how hierarchies or delegation of decision-making authority can alleviate information-processing constraints or costs.<sup>4</sup> An exception are a series of recent papers that study the optimal allocation of attention in networks, such as Galeotti and Goyal (2010); Akerlof and Holden (2016); Dessein, Galeotti, and Santos (2016); and Herskovic and Ramos (2020). A common finding in those papers is that, even when all agents in the network are *ex ante* identical, attention tends to be concentrated on a select few of them. Another exception is Van den Steen (2018), who studies a strategy formulation game in which a strategist chooses which decision to investigate—among a set of interrelated decisions—and then communicates her preferred action to a group of agents in charge of implementing those decisions. Unlike in our model, there is no trade-off between investigating several decisions versus learning more about one decision, and the strategist is never blindsided—she always communicates about the decision she

<sup>3</sup> See also Ocasio (1997) and the recent survey by Hambrick (2018).

<sup>4</sup> See Garicano and Prat (2013) and Garicano and Van Zandt (2012) for overviews of the literature.

investigates.<sup>5</sup> In the empirical literature, Bandiera et al. (2017) and Bandiera et al. (2020) employ time use surveys to measure how CEOs allocate their attention. Their main focus, however, is on the time CEOs spend on activities with large private benefit (such as meeting with outsiders) as opposed to activities that mainly improve firm performance (such as meeting with insiders). Finally, our paper contributes to the literature on narrow business strategies and vision (Rotemberg and Saloner 1994, 2000) and organizational focus (Dessein, Galeotti, and Santos 2016) by endogenizing to what extent selected business strategies are contingent on the organizational environment as opposed to managerial characteristics. The result that Boards will pick biased managers even if they themselves are unbiased is also reminiscent of Van den Steen (2005), who shows that firms may optimally hire managers with strong beliefs.<sup>6</sup>

## I. The Model

### A. The Production Process

We posit a team-theoretic model in which there are two tasks  $i \in \{1, 2\}$ , say marketing and operations, one manager, and an organization consisting of a continuum of employees  $j \in [0, 1]$ . Profits of the organization depend on (i) *external alignment*, that is, how well is each task  $i$  adapted to an independently normally distributed task-specific shock  $\theta_i \sim N(0, \sigma_\theta^2)$ , and (ii) *internal alignment*, that is, how well are the two tasks implemented by the organization. Specifically, the manager selects for each task  $i$  a strategic choice  $a_{Mi}$ , which must match as well as possible the task specific shock  $\theta_i$ . In turn, each employee  $j$  must choose complementary actions  $a_{j1}$  and  $a_{j2}$  to match as well as possible the strategic choices selected by the manager. The realized payoffs are given by

$$(1) \quad \pi \equiv \sum_{i \in \{1, 2\}} \left[ h(\theta_i) - (a_{Mi} - \theta_i)^2 - \beta \int_0^1 (a_{Mi} - a_{ji})^2 dj \right],$$

where  $\beta$  captures the relative importance of internal alignment.<sup>7</sup> We refer to  $\mathbf{a}_M = (a_{M1}, a_{M2})$  as the strategic choices or strategies for short and to  $\mathbf{a}_j = (a_{j1}, a_{j2})$  as the implementation of those strategies by agent  $j$ .

<sup>5</sup>Hence, strategic choices are purely a function of managerial characteristics: *ceteris paribus*, the strategist investigates and communicates about the decision about which she expects to receive the most informative signal (that is, the one she has more expertise in). In contrast, our model endogenizes to what extent strategic choices reflect managerial characteristics as opposed to the realization of environmental shocks.

<sup>6</sup>There is of course a vast literature on attention allocation that is delinked from organizational considerations. For instance, a recent literature studies the joint problem of the allocation of attention and portfolio choice (see Van Nieuwerburgh and Veldkamp 2010 and Kacperczyk, Van Nieuwerburgh, and Veldkamp 2016).

<sup>7</sup>The payoff function (1) is similar to the payoff functions considered in a series of organizational economics papers focused on coordination issues in organizations, such as Dessein and Santos (2006); Alonso, Dessein, and Matouschek (2008, 2015); Rantakari (2008, 2013); Bolton, Brunnermeier, and Veldkamp (2013); Calvó-Armengol, de Martí, and Prat (2015); and Dessein, Galeotti, and Santos (2016), among others. All these papers including this one view the trade-off between external and internal alignment (or adaptation and coordination) as the central trade-off in organizations.



Only the manager can learn about  $\theta_1$  and  $\theta_2$ , but she can communicate her strategic choices to the organization. Specifically, we assume that the manager first devotes attention to tasks 1 and 2 in order to learn about  $\theta_1$  and  $\theta_2$ , then the manager chooses her strategy  $\mathbf{a}_M = (a_{M1}, a_{M2})$  and communicates those choices to the organization and, finally, the employees  $j \in [0, 1]$  implement those strategies by choosing complementary actions  $a_{j1}$  and  $a_{j2}$ . Without loss of generality, we will assume that  $h(\theta_i) \equiv \theta_i^2$  so that profits  $\pi$  are normalized to 0 whenever  $\mathbf{a}_M = \mathbf{a}_j = (0, 0)$  for all  $j \in [0, 1]$ .<sup>8</sup> Adapting to  $\theta_i$  can then be interpreted as an opportunity to improve performance in task  $i$ .

We will describe in more detail the learning process and communication technology below. At this point, we want to note that whenever communication is imperfect, there is a trade-off between external and internal alignment. By selecting strategic choices that are responsive to the task-specific shocks  $\theta_i$ , the manager sacrifices some internal alignment, as not all employees may understand her strategy. In contrast, perfect internal alignment can always be achieved by selecting the standard strategic choice  $a_{Mi} = 0$ . As we will show, in the absence of any communication—or when communication fails—employees optimally choose actions that are complementary to the standard strategy  $a_{Mi} = 0$ .

### B. Communication and Implementation inside the Organization

In order to ensure effective implementation, the manager needs to communicate her strategic choices to employees so they can take the appropriate complementary actions. Communication though is imperfect. The more attention devoted to a strategic choice, the more likely an employee understands how to implement it. Formally, we model communication as a Poisson process with a hazard rate  $\mu$  and the stochastic event corresponding to the employee “understanding” a particular strategic choice. Specifically, let  $r_i \geq 0$  be the amount of time devoted to process information related to strategic choice  $i$ , then an employee understands strategic choice  $i$  with probability

$$(2) \quad p_i = 1 - e^{-\mu r_i},$$

which is independent across agents. Given this, employee  $j$ 's choices are given by

$$a_{ji} = \begin{cases} a_{Mi} & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i. \end{cases}$$

Thus, with probability  $p_i$ , worker  $j$  understands  $a_{Mi}$  and sets the complementary action equal to this choice; and with probability  $1 - p_i$ , the worker simply does not, and sets the complementary action equal to the the mean value of the shock  $\theta_i$ , which

<sup>8</sup>Indeed, neither the optimal choices for  $\mathbf{a}_M$  and  $\mathbf{a}_j$ , nor the optimal allocation of attention are affected by the functional form of  $h(\theta_i)$ .

is 0. The manager controls the allocation of attention by workers  $r_i$  subject to an organizational attention constraint

$$(3) \quad r_1 + r_2 \leq r \quad \text{with} \quad r > 0.$$

Alternatively,  $r_i$  can be interpreted as the time the manager devotes to communicate about strategic choice  $a_{Mi}$ . It will be useful to rewrite communication constraint (3) as follows:

$$(4) \quad (1 - p_1)(1 - p_2) \geq 1 - p,$$

where  $p \equiv 1 - e^{-\mu r}$  denotes the probability of understanding  $a_{Mi}$  when an employee's attention is fully dedicated to task  $i$ .

### C. Allocation of Attention and Learning by the Manager

Consider now the learning process of the manager. We assume that the manager observes an *endogenous* signal  $s_i$  about each shock  $\theta_i$  whose informativeness depends on the *managerial attention*  $t_i$  devoted to task  $i$ . In addition, the manager observes an *exogenous* signal  $S_i$  about  $\theta_i$ . How informative  $S_i$  is depends on her *managerial expertise*  $T_i$  in task  $i$ .<sup>9</sup> In the remainder of the paper, we will assume that  $F(\theta_i|S_i, s_i)$ , the manager's posterior about  $\theta_i$  is normally distributed with mean estimate  $\hat{\theta}_i = E[\theta_i|s_i, S_i]$  and variance:

$$(5) \quad RV(\theta_i|t_i, T_i) \equiv \sigma_\theta^2 - q_i \sigma_\theta^2 \quad \text{where} \quad q_i \equiv q(T_i + t_i) = 1 - e^{-\lambda(T_i + t_i)}.$$

Hence, the mean-squared error or residual variance  $RV(\theta_i|t_i, T_i)$  decreases at a logarithmic rate as a function of both attention  $t_i$  and expertise  $T_i$ , where  $t_i$  and  $T_i$  are substitutes in the learning process (see Figure 1). Online Appendix A provides microfoundations for (5). As was the case for organizational attention  $(r_1, r_2)$ , managerial attention is also scarce in that

$$(6) \quad t_1 + t_2 \leq 2\tau.$$

We denote by  $\Upsilon$  the set of feasible allocations of attention:  $\Upsilon = \{(t_1, t_2) : t_1 + t_2 \leq 2\tau\}$ . Our main propositions will deal with the optimal allocation of managerial attention  $(t_1, t_2)$ .

*Discussion of Managerial Learning Technology.*—In (5), the parameter  $\lambda$  characterizes the speed of learning and, hence,  $1/\lambda$  the *complexity* of the environment. A large  $\lambda$ , for instance, is consistent with situations in which ex post uncertainty is small as, say, the market in which the organization operates is mature and strategic choices are well understood. Instead a small  $\lambda$  is associated with environments with

<sup>9</sup>An alternative interpretation of  $T_i$  is as a sufficient statistic for organizational characteristics that facilitate learning in particular dimensions.



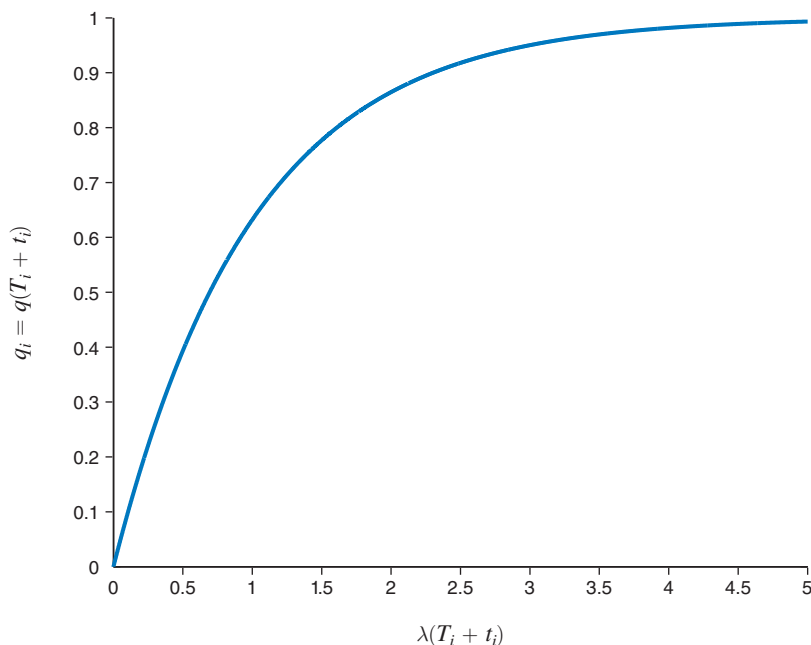


FIGURE 1. MANAGERIAL LEARNING TECHNOLOGY

Note:  $q(T_i + t_i) = 1 - e^{-\lambda(T_i + t_i)}$ , which determines the residual variance  $RV(\theta_i | t_i, T_i) = \sigma_\theta^2 - q(T_i + t_i)\sigma_\theta^2$ .

large ex post uncertainty, perhaps because the organization is operating in a new industry or in a period characterized by a lot of turbulence. We refer to  $\lambda T_i$  as the manager's *effective expertise*, a combination of his managerial expertise  $T_i$  and the complexity of the environment  $1/\lambda$ . What determines the precision of the information on the task specific shock, (5), is the combination of the effective expertise, which is exogenous, and the *effective attention*,  $\lambda t_i$ , which depends on the endogenous allocation of attention. In sum, attention, expertise, and the complexity of the environment combine in  $\lambda(T_i + t_i)$  to determine the precision of the manager's information about  $\theta_i$ .

If  $T_1 > T_2$ , then we speak of a manager who is specialized or is an expert in task 1, whereas we refer to a manager for whom  $T_1 = T_2$  as a generalist. Conceptually, we think of a manager specialized in task  $i$  as having access to more precise information about the shock pertaining to task  $i$  than a generalist or a specialist in task  $j \neq i$ , even before the attention decision is made. However, from the formulation of our learning technology (5), a nonspecialist manager can compensate for her lack in expertise in a specific task by devoting more attention to it. For example, she can consult experts, do extensive research, or simply devote more time to analyze her options in that particular task as she cannot rely on past experience or knowledge. Thus, in our model, expertise and attention are substitutes—a decrease in expertise can be compensated by a corresponding increase in attention.

Our assumption that expertise and attention are substitutes is “conservative” from a modeling perspective. Indeed, a central result in this paper is that experts often devote more attention to tasks in which they have superior expertise. By ruling out that expertise and attention are complements in the learning technology, we ensure that our main results are not driven by assumptions regarding the learning technology.

#### D. Timing

Figure 2 illustrates the timing of our model, flow of information, and actions taken by both management and employees. The timing in our model is as follows:

- (i) The manager allocates attention  $t_i \in [0, 2\tau]$  to each task  $i = 1, 2$  to learn about  $(\theta_1, \theta_2)$ , with  $t_1 + t_2 \leq 2\tau$ .
- (ii) The manager observes signals  $S_i$  and  $s_i$  about  $\theta_i$  for  $i = 1, 2$  and selects strategic choices  $(a_{M1}, a_{M2})$ . Employees do not observe any signals.<sup>10</sup>
- (iii) The manager chooses the probabilities  $p_1$  and  $p_2$  with which employees will understand her strategic choices  $a_{M1}$  and  $a_{M2}$  with  $(1 - p_1)(1 - p_2) \geq (1 - p)$ .
- (iv) Having learned about management choices, employees select their complementary actions,  $a_{ji}$  for  $j \in [0, 1]$  and  $i = 1, 2$ .

#### E. Benchmark: Attention and External Alignment

As a benchmark, consider the case where the manager only maximizes *external alignment*

$$\sum_{i \in \{1, 2\}} E[h(\theta_i) - (a_{Mi} - \theta_i)^2].$$

One interpretation of this benchmark is that there is no need for implementation ( $\beta = 0$  in the pay-off function (1)). A second interpretation is that while there is a need for implementation, there are no communication frictions and the manager can perfectly communicate her strategic choices to employees. Given signals  $(s_1, S_1, s_2, S_2)$  the manager then optimally sets  $(a_{M1}, a_{M2}) = (\hat{\theta}_1, \hat{\theta}_2)$  and allocates attention in order to minimize

$$\sum_{i \in \{1, 2\}} E[(\hat{\theta}_i - \theta_i)^2 | t_i, T_i] = \sum_{i \in \{1, 2\}} RV(\theta_i | t_i, T_i).$$

<sup>10</sup>In personal communication, Bob Gibbons has expressed to us that a more appropriate definition of strategy in the present framework would be the vector  $(\mathbf{t}, \mathbf{a}_M, \mathbf{p})$ . Strategies then would be described as a combination of attention allocation by managers, managerial choices, and communication protocols inside the organization.

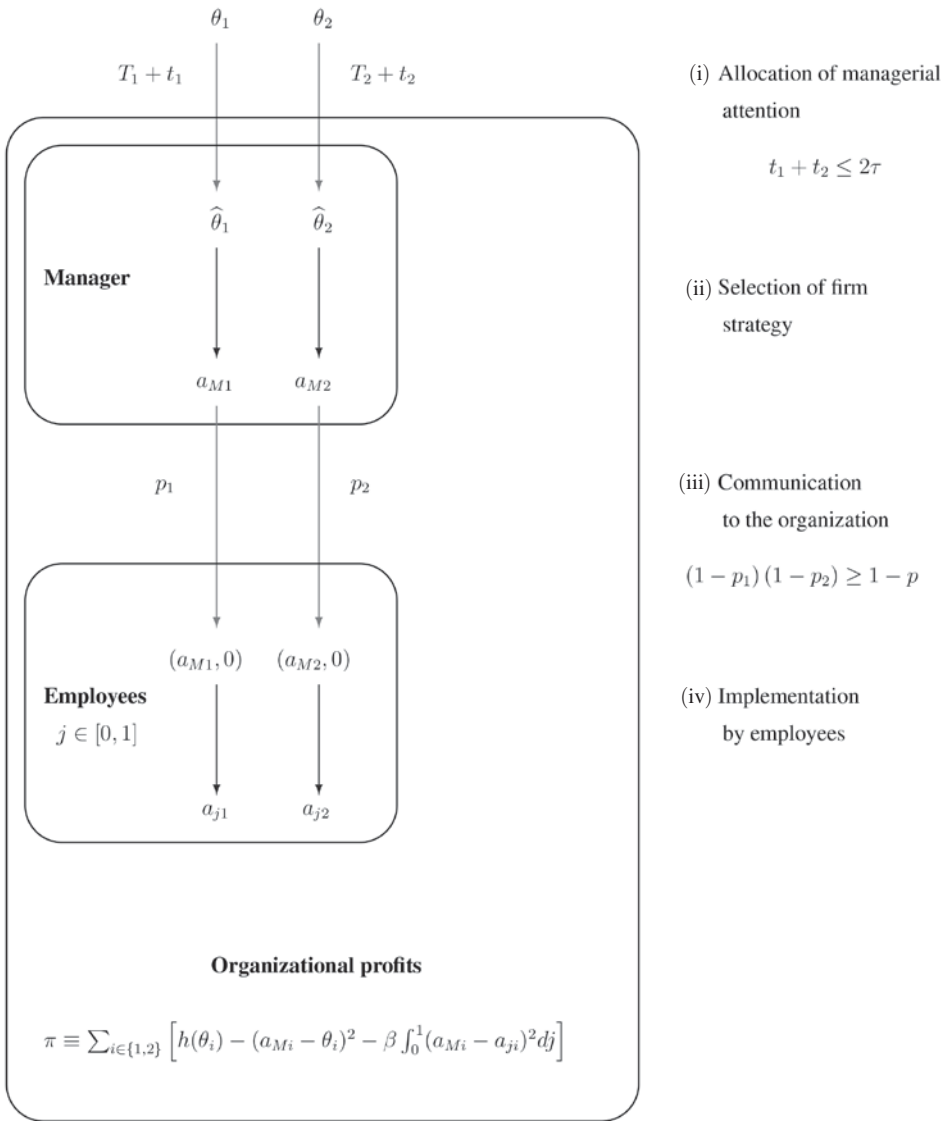


FIGURE 2. A GRAPHICAL REPRESENTATION OF THE MODEL

Given learning technology (5), the marginal returns to devoting attention to task  $i$  are given by

$$\frac{\partial(-RV(\theta_i|t_i, T_i))}{\partial t_i} = \lambda(1 - q_i)\sigma_{\theta}^2,$$

where  $q_i = q(t_i + T_i)$  is increasing in both  $t_i$  and  $T_i$ . It follows that when  $t_1 > t_2$ , the marginal returns to devoting attention to task 2 are higher than to task 1 provided the manager is a generalist ( $T_1 = T_2$ ). Similarly, the marginal returns to devoting attention to any given task  $i$  are decreasing in the manager's expertise  $T_i$  in that task.

Intuitively, our learning technology features decreasing marginal returns to attention and, by assumption, attention and expertise are substitutes.<sup>11</sup> The following result thus obtains.

**PROPOSITION 1:** *The allocation of managerial attention  $\mathbf{t} = (t_1, t_2)$ , which maximizes external alignment ( $\beta = 0$ ) is given by (i)  $t_1 = t_2 = \tau$  if the manager is a generalist ( $T_1 = T_2$ ) and (ii)  $0 \leq t_1 < t_2 \leq 2\tau$  if the manager has more expertise in task 1 ( $T_1 > T_2$ ).*

#### F. A Preview of the Results

We graphically illustrate our main results and contrast them with the benchmark in Section IE. Consider Figure 3 below, panels A and B, which show managerial attention choices when  $p = 0.75$  and  $\beta = 5$ . There we represent managers in the space  $(\lambda(T_1 + t_1), \lambda(T_2 + t_2))$ . The allocation of attention can be visualized as arrows going from a point in the plane  $(\lambda T_1, \lambda T_2)$ , which is the manager's initial expertise, to another point  $(\lambda(T_1 + t_1), \lambda(T_2 + t_2))$ , her final expertise profile once the manager decides on a particular allocation of attention. The complexity of the environment, as summarized by  $\lambda$ , plays a critical role in what follows. The line going through points *A* and *B* in both panels corresponds to a manager with a given profile  $(T_1, T_2)$  under environments of different complexity. Point *A* corresponds to a complex environment (low  $\lambda$ ) and *B* to a less complex environment (high  $\lambda$ ). We refer to *managing with style* as either a situation in which a specialist manager (say with  $T_1 > T_2$ ) devotes all her attention to the task in which she has more expertise, or a situation in which a generalist manager (for whom  $T_1 = T_2$ ) arbitrarily biases her attention to one particular task. Note that in our benchmark above, managing with style is always suboptimal.

Consider first the case of the specialist manager with  $T_1 > T_2$  operating in a complex environment (point *A*) and assume that the amount of endogenous attention to be allocated,  $t_1 + t_2 = 2\tau$ , is not too large. Proposition 4 shows that in this case, the manager optimally “manages with style.” She allocates her attention so as to reinforce her expertise in task 1 and the manager's precision of information in task 1 and 2 is given by  $\lambda(T_1 + 2\tau)$  and  $\lambda T_2$ , respectively. The organization is then disproportionately responsive to shocks regarding task 1. In fact, the manager then often only communicates to the organization about task 1. Similarly, a manager with  $T_2 > T_1$  operating in a complex environment (point *A'*) optimally devotes all her attention to task 2 and mainly communicates to the organization about task 2.

It follows that in complex environments, two identical organizations led by managers with different expertise (*A* versus *A'*) tend to respond very differently in response to identical shocks. In other words, firm behavior is largely driven by initial differences in expertise between managers, which are amplified by the

<sup>11</sup> Obviously, in practice there may exist settings in which there are increasing marginal returns to attention (at least for some parameter ranges) or technological complementarities between attention and expertise. As noted above, our assumptions regarding the learning technology should not be regarded as a positive statement but rather as a modeling device to highlight the organizational trade-offs that lead to “managing with style” even in the presence of technological drivers that push against this possibility.

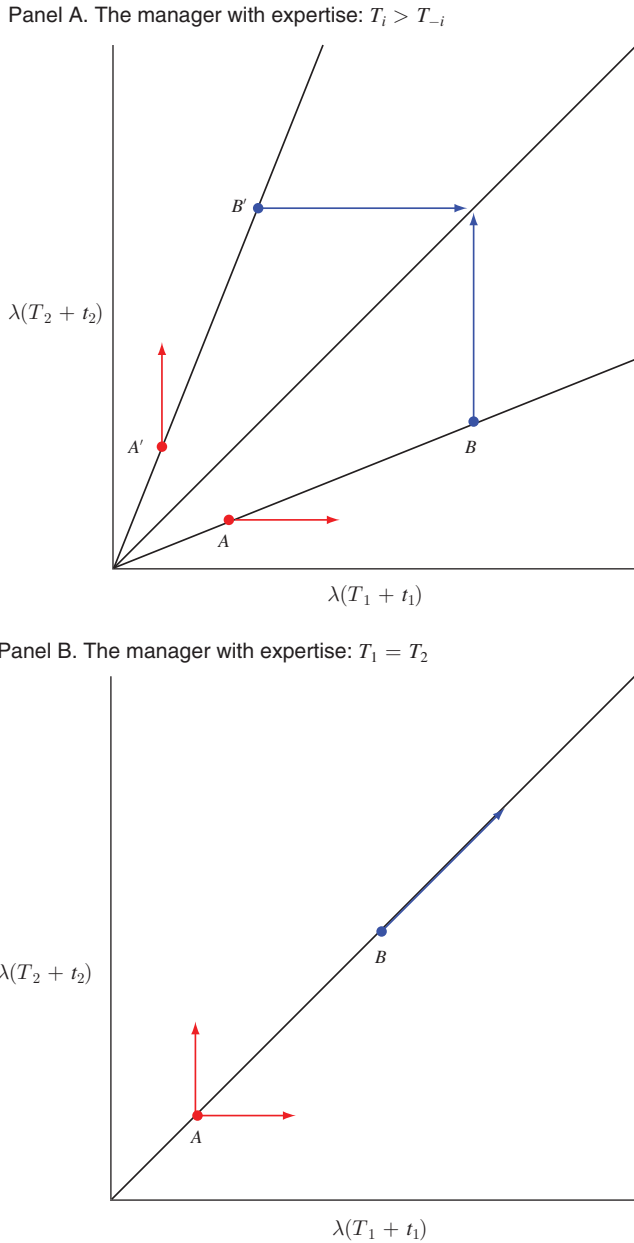


FIGURE 3. A PREVIEW OF THE RESULTS

Notes: Endogenous allocation of attention. The plot shows the endogenous allocation of attention in complex (A and A') and less complex (B and B') environments for the case of the specialist and the generalist manager (panels A and B).

allocation of managerial attention. Note that this stands in sharp contrast with the results obtained in our benchmark.

Instead when the environment is less complex (point B in panel A of Figure 3), a manager for whom  $T_1 > T_2$  will instead opt to devote all her attention to task 2, in

which she has less expertise, resulting in a final expertise profile that is almost identical for both tasks. In contrast to the case of a complex environment, two organizations led by managers with different initial expertise now behave very similarly in response to identical shocks. In other words, initial differences in expertise between two managers are much less predictive of organizational behavior.

Proposition 3 shows that, surprisingly, a similar result obtains for the case of the generalist manager ( $T_1 = T_2$ ). In complex environments, a generalist manager optimally specializes in one task by devoting all her attention to one task or the other, which task specifically being a matter of indifference (see panel B of Figure 3). In contrast, in less complex environments, a generalist manager splits her attention equally among both tasks.

## II. Optimal Allocation of Organizational Attention

The benchmark studied in Section IE ignored the organizational implementation of strategic choices. In this section, we exclusively focus on strategy implementation: How are strategic choices optimally communicated or, equivalently, how is organizational attention optimally allocated? And how does this affect optimal strategic choices? We answer those questions taking the allocation of managerial attention and the resulting posterior estimates  $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)$  as given. In Section III we then endogenize the allocation of managerial attention.

Our aim is to characterize strategic choices  $(a_{M1}, a_{M2})$  and organizational attention choices  $(p_1, p_2)$  given posteriors  $\hat{\theta}$ . Simple manipulations of the objective function (1) yield

$$\begin{aligned} E(\pi|\hat{\theta}) &= \sum_{i \in \{1,2\}} \left[ E(\theta_i^2 | \hat{\theta}_i) - (a_{Mi} - \hat{\theta}_i)^2 - E[(\theta_i - \hat{\theta}_i)^2 | \hat{\theta}_i] - \beta(1 - p_i)a_{Mi}^2 \right] \\ &= \sum_{i \in \{1,2\}} \left[ \hat{\theta}_i^2 - (a_{Mi} - \hat{\theta}_i)^2 - \beta(1 - p_i)a_{Mi}^2 \right]. \end{aligned}$$

As is standard in the team theory literature, we limit our attention to equilibria in linear strategies, that is,  $a_{Mi} = \alpha_i \hat{\theta}_i$ . The problem for the manager is then equivalent to choosing  $(p_1, p_2, \alpha_1, \alpha_2)$  in order to maximize expected profits

$$(7) \quad E(\pi|\hat{\theta}) = \sum_{i \in \{1,2\}} \left[ \hat{\theta}_i^2 - (1 - \alpha_i)^2 \hat{\theta}_i^2 - \beta(1 - p_i)\alpha_i^2 \hat{\theta}_i^2 \right] \quad \text{subject to} \quad (4).$$

Direct inspection of (7) shows that  $\alpha_i$  and  $p_i$  are complementary choices. The larger is  $\alpha_i$  and the more responsive is the manager to her posterior about shock  $\theta_i$ , the more organizational attention should be directed to her strategic choice  $i$  in order to ensure internal alignment. Similarly, when  $p_i$  is larger and employees are better at implementing strategic choice  $i$ , then it becomes optimal for the manager to be more responsive to her posterior  $\hat{\theta}_i$ . Formally, taking  $p_1$  and  $p_2$  as given, and maximizing (7) with respect to  $\alpha_1$  and  $\alpha_2$ , yields

$$(8) \quad \alpha_1 = \frac{1}{1 + \beta(1 - p_1)} \quad \text{and} \quad \alpha_2 = \frac{1}{1 + \beta(1 - p_2)}.$$



Hence, the larger is  $p_i$ , the more adaptive is the manager to the posterior  $\hat{\theta}_i$ . Armed with this we can now turn to the optimal choice of  $p_i$ .

**PROPOSITION 2:** *Suppose  $\beta > 1$ . Whenever  $p < \bar{p}(\beta) \equiv 1 - 1/\beta^2$ , the manager directs all organizational attention to one strategic choice:*

$$(p_1^*, p_2^*) = \begin{cases} (p, 0) & \text{if } \hat{\theta}_1^2 > \hat{\theta}_2^2 \\ (0, p) & \text{if } \hat{\theta}_1^2 < \hat{\theta}_2^2. \end{cases}$$

Intuitively, when internal alignment is important ( $\beta$  large), external alignment is very costly unless communication is effective. The manager then optimally communicates intensively about one strategic choice, allowing that strategic choice to be responsive to its task-specific shock without compromising internal alignment (see (8)). Internal alignment on the other strategic choice is then achieved by largely giving up on external alignment regarding that task—in other words, the other task will not be very responsive to the posterior in order to avoid poor internal alignment. Naturally, it is optimal to communicate about the strategic choice that faces the largest shocks, as external alignment is most important for that task.

In contrast, when implementation/internal alignment is not very important ( $\beta$  is small), the manager is optimally responsive to both shocks provided they are sufficiently equal in size. Sacrificing external alignment on one task in order to improve internal alignment is then not worth it. Similarly, if attention is relatively unconstrained ( $p$  is large), then the manager can communicate effectively about both strategic choices, and there is no real trade-off between external and internal alignment. Even when implementation is very important, it is then still optimal to communicate about both tasks.<sup>12</sup>

In what follows, we will assume that there is a tight bound on organizational attention:

$$(A) \quad p \leq \bar{p}(\beta)$$

so that whenever  $\hat{\theta}_i^2 > \hat{\theta}_{-i}^2$ , then the manager only communicates  $a_{Mi}$ , that is,  $p_i = p$  and  $p_{-i} = 0$ , and (8) specializes to

$$(9) \quad a_{Mi} = \frac{\hat{\theta}_i}{1 + \beta(1 - p)} \quad \text{and} \quad a_{M-i} = \frac{\hat{\theta}_{-i}}{1 + \beta}.$$

<sup>12</sup>The result in Proposition 2 is reminiscent of the main result in Dessein, Galeotti, and Santos (2016) (henceforth, DGS). There are three key differences though. First, in DGS, information is dispersed: For each task, there is one manager who observes information pertaining to his and only his task. In contrast, in our model, there is only one manager who decides which tasks to devote (managerial) attention to. Second, in DGS, the allocation of organizational attention *cannot* be made contingent on the realization of task specific shock. In contrast, in our model, the manager directs organizational attention *after* observing the task-specific shocks. Finally, in DGS, managers observe their task-specific shock perfectly. DGS therefore cannot address the issue of managerial style and attention, the subject of this paper.

We maintain (A) throughout the paper in order to simplify the analysis. Note that (A) is a condition that guarantees that the manager only communicates about one task, even when  $|\hat{\theta}_i| = |\hat{\theta}_{-i}|$ . Hence, even when  $p > \bar{p}(\beta)$ , the manager may communicate only about task  $i$  provided that  $|\hat{\theta}_i|$  is sufficiently larger than  $|\hat{\theta}_{-i}|$ .

Before studying the optimal allocation of managerial attention, it is useful to consider the limit case where managerial attention is unconstrained so that  $\hat{\theta}_i = \theta_i$ . Given (A), the manager then directs organizational attention to the largest realized shock. There is no sense, however, in which the organization or the manager are biased towards one particular task. Ex ante, each task is equally likely to be the focus of organizational attention. Furthermore, all organizations faced with the same environment, a given  $\theta_1$  and  $\theta_2$ , will focus attention on the same task. It follows then that scarcity of organizational attention in the absence of scarcity of managerial attention does not result in any systematic bias in organizational strategies.

### III. Optimal Allocation of Managerial Attention

Scarcity of *organizational attention* implies that the manager only communicates about the largest perceived shock and is disproportionately responsive to her posterior about this shock. Anticipating this, how does the manager optimally allocate her scarce *managerial attention*, and how does this affect firm behavior?

#### A. Expected Profits

Before analyzing the optimal allocation of attention we first develop the expected profit function for a given allocation of attention  $\mathbf{t} = (t_1, t_2)$ . For this purpose, we first express expected profits for given posteriors  $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2)$  and communication choices  $(p_1, p_2)$ , and subsequently take expectations over posteriors for an allocation of managerial attention.

Given (8), the expected profit, (7), associated with task  $i$  conditional on  $\hat{\boldsymbol{\theta}}$  and communication choices  $(p_1, p_2)$ , is given by

$$E[\pi_i | \hat{\theta}_i] = \hat{\theta}_i^2 - \left( \frac{\beta(1-p_i)}{1+\beta(1-p_i)} \right)^2 \hat{\theta}_i^2 - \frac{\beta(1-p_i)}{(1+\beta(1-p_i))^2} \hat{\theta}_i^2 = \frac{\hat{\theta}_i^2}{1+\beta(1-p_i)}.$$

Given that  $\hat{\theta}_i^2 > \hat{\theta}_{-i}^2$  and given our assumption on scarce organizational attention, (A), the manager only communicates about task  $i : (p_i, p_{-i}) = (p, 0)$  (see Proposition 2). Hence, the expected profit of the organization conditional on posteriors  $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2)$  equals

$$E(\pi | \hat{\boldsymbol{\theta}}) = E[\pi_i | \hat{\theta}_i] + E[\pi_{-i} | \hat{\theta}_{-i}] = \frac{\hat{\theta}_i^2}{1+\beta(1-p)} + \frac{\hat{\theta}_{-i}^2}{1+\beta}.$$

Note that for large values of  $\beta$  and  $p$ , the expected profitability of task  $i$  conditional on  $\hat{\theta}_i$  is much greater than that of task  $-i$ , even when  $\hat{\theta}_i^2 \sim \hat{\theta}_{-i}^2$ , as the manager only

communicates about task  $i$ . Given attention allocation  $\mathbf{t} = (t_1, t_2)$ , the (unconditional) expected profits are given by

$$(10) \quad \Pi(q_1, q_2) \equiv E[\pi] = \Pr(\hat{\theta}_1^2 \geq \hat{\theta}_2^2) E\left[\frac{\hat{\theta}_1^2}{1 + \beta(1 - p)} + \frac{\hat{\theta}_2^2}{1 + \beta} \mid \hat{\theta}_1^2 \geq \hat{\theta}_2^2\right] \\ + \Pr(\hat{\theta}_1^2 < \hat{\theta}_2^2) E\left[\frac{\hat{\theta}_1^2}{1 + \beta} + \frac{\hat{\theta}_2^2}{1 + \beta(1 - p)} \mid \hat{\theta}_1^2 < \hat{\theta}_2^2\right].$$

From (5), given attention allocation  $\mathbf{t} = (t_1, t_2)$ , we have

$$(11) \quad \hat{\theta}_i \sim \mathcal{N}(0, q_i \sigma_\theta^2),$$

where  $q_i = q(T_i + t_i)$ . Let  $F(x, y)$  denote the normal c.d.f. of random variable  $x$  with mean 0 and variance  $y$ . Then given (11), we can rewrite (10) as<sup>13</sup>

$$\Pi(q_1, q_2) = 4 \int_0^{+\infty} \left[ \int_0^{\hat{\theta}_2} \frac{\hat{\theta}_1^2}{1 + \beta} dF(\hat{\theta}_1, q_1 \sigma_\theta^2) \right. \\ \left. + \int_{\hat{\theta}_2}^{+\infty} \frac{\hat{\theta}_1^2}{1 + \beta(1 - p)} dF(\hat{\theta}_1, q_1 \sigma_\theta^2) \right] dF(\hat{\theta}_2, q_2 \sigma_\theta^2) \\ + 4 \int_0^{+\infty} \left[ \int_0^{\hat{\theta}_1} \frac{\hat{\theta}_2^2}{1 + \beta} dF(\hat{\theta}_2, q_2 \sigma_\theta^2) \right. \\ \left. + \int_{\hat{\theta}_1}^{+\infty} \frac{\hat{\theta}_2^2}{1 + \beta(1 - p)} dF(\hat{\theta}_2, q_2 \sigma_\theta^2) \right] dF(\hat{\theta}_1, q_1 \sigma_\theta^2).$$

We use this expression to investigate the optimal allocation of attention. We proceed by distinguishing between the case of a generalist manager and a specialist manager with superior expertise in one task versus the other.

### B. Allocation of Attention by a Generalist Manager

Consider first a generalist manager with equal expertise about both tasks, that is  $T_1 = T_2 = T$ . Should the generalist manager divide her attention equally among both tasks, or should she focus her attention on one (randomly chosen) task?

From (11), when the manager allocates more attention to task 1, then she is more likely to communicate to the organization about task 1, as she is more likely to perceive the shock affecting task 1 to be the largest. Indeed, from (11), even though  $\Pr(\theta_1^2 > \theta_2^2) = \Pr(\theta_2^2 > \theta_1^2)$ , we have that

$$(12) \quad \Pr(\hat{\theta}_1^2 > \hat{\theta}_2^2) > \Pr(\hat{\theta}_2^2 > \hat{\theta}_1^2) \Leftrightarrow T_1 + t_1 > T_2 + t_2.$$

<sup>13</sup>The integral further uses the fact that the probability density of  $|\hat{\theta}_i|$  is twice the probability density of  $\hat{\theta}_i = |\hat{\theta}_i|$ , as reflected in the number 4 in front of the expression.

To see this more intuitively, consider the extreme case where  $T_1 = T_2 = 0$  and  $t_1 > t_2 = 0$ . The posterior estimate of  $\theta_2$  then always equals the prior estimate, that is  $\hat{\theta}_2 = 0$ , so that we always have that  $\hat{\theta}_1^2 > \hat{\theta}_2^2$ .

The allocation of attention thus influences how adaptive the organization is to an external shock through two channels: how well the manager observes the relevant external shock and how likely the manager is to communicate about the strategic choice. It follows that the more attention a manager devotes to task 1, the more (on average) the organization is adaptive to  $\hat{\theta}_1$  and, hence, the more valuable it is to improve the precision of  $\hat{\theta}_1$  by devoting even more attention to task 1. Formally, if  $t_1 > t_2$ , then  $\Pr(\hat{\theta}_1^2 > \hat{\theta}_2^2) > \Pr(\hat{\theta}_2^2 > \hat{\theta}_1^2)$ , and from expression (10), profits are more sensitive to  $E(\hat{\theta}_1^2)$  than to  $E(\hat{\theta}_2^2)$ . By devoting all attention to one task, the manager is then very good at responding to the shock affecting this task, and she is likely to communicate about this task.

If there were to be constant returns to attention, for example, if  $q_i = \delta t_i$ , then, using (10), it is easy to show that the manager optimally focuses all attention on one task up to the point where the shock is perfectly observed, that is  $q_i = 1$ . Indeed, because of the complementarity between (i) how much attention a task receives and (ii) how likely a task is to be communicated to the organization, the profit function (10) is then convex in  $t_i$ .

If instead, as in this paper, there are decreasing marginal returns to attention, then whether or not focused attention is optimal depends on the scarcity of managerial attention  $\tau$  and the manager's effective expertise,  $\lambda T$ . Proposition 3 characterizes the optimal allocation of attention, but restricts the analysis to the case where managerial attention is either abundant or scarce.

**PROPOSITION 3 (Generalist Manager):** *Consider a generalist manager ( $T_1 = T_2 = T$ ) and assume (A) and  $t_1 + t_2 \leq 2\tau$ .*

- (i) *If managerial attention is abundant ( $\tau$  large), then balanced attention is optimal, that is,  $(t_1^*, t_2^*) = (\tau, \tau)$ .*
- (ii) *If managerial attention is scarce ( $\tau$  small), then focused attention (managing with style) is optimal whenever the effective expertise,  $\lambda T$ , is low enough. Thus, there exists a  $\Lambda > 0$ , such that*

$$(t_1^*, t_2^*) = \begin{cases} (2\tau, 0) \text{ or } (0, 2\tau) & \text{if } \lambda T < \Lambda \\ (\tau, \tau) & \text{if } \lambda T > \Lambda. \end{cases}$$

- (iii) *An increase in the importance of internal alignment ( $\beta$ ) may result in a shift from balanced managerial attention to focused managerial attention (managing with style), but never the other way around.*

Proposition 3 shows that when attention is scarce ( $\tau$  small) and the environment is sufficiently complex ( $\lambda T < \Lambda$ ), a generalist manager optimally specializes in one task by devoting all her attention to one task (which task specifically being

a matter of indifference). Thus, even if two tasks (say marketing and operations) are equally important to the organization and have, a priori, the same potential for profit improvement, a manager with equal expertise in both marketing and operations should focus all her attention on one task, say marketing. Ex post, such a manager mainly (but not always) selects and communicates about marketing strategies, and she *appears* to be arbitrarily and inefficiently biased toward marketing. Note that while the manager optimally focuses attention on one task, ex post she may be forced to communicate and implement a strategy about which she has “poor visibility.” The organization, thus, may be forced ex post to do things based on very imprecise information, resulting in lower profitability and the appearance of a manager focused on the wrong things.

As discussed above, the benefits of managerial focus stem from the complementarity between (i) how much attention task  $i$  receives and (ii) how likely the shock affecting task  $i$  is perceived to be larger than the shock affecting task  $-i$ . When attention is not scarce ( $\tau$  large) or when the environment is not complex ( $\lambda$  large), however, the above complementarity is overwhelmed by decreasing marginal returns to attention. As Proposition 3 shows, balanced attention is optimal in those instances.

Intuitively, when  $\tau$  is large, then by dividing her attention evenly, the manager learns both shocks with great precision. In contrast, by focusing all attention on task 1, the manager runs a high risk of being blindsided. Because of decreasing marginal returns to attention, the additional knowledge gained about task 1 is then much less than the information lost on task 2. Similarly, when the environment is not very complex, so that  $\lambda T > \Lambda$ , the effective managerial expertise  $\lambda T$  goes a long way in improving knowledge about the realized shocks  $\theta_1$  and  $\theta_2$ . The manager then typically identifies the largest shock correctly, and which task the manager communicates about to the organization is largely driven by the realization of  $\theta_1$  and  $\theta_2$  rather than by the allocation of managerial attention. The above complementarity between (i) how much attention a task receives and (ii) which task is communicated to the organization, is then very minor. Because of decreasing marginal returns to attention, it is then optimal to allocate an equal amount of attention to both tasks as in our benchmark.

Beyond the complexity of the environment and scarcity of attention, Proposition 3 yields two other comparative static results. First, ceteris paribus, more competent managers (managers with a higher  $T$ ) are less likely to focus attention on one task. This suggests a career path where managers, when they are younger, manage with style and focus attention on one area, whereas they act more as generalists when they are older and have accumulated more experience. Second, from Proposition 3(iii), when the importance of internal alignment  $\beta$  is large, focused attention is more likely to be optimal. Indeed, when  $\beta$  is large, the manager is largely unresponsive to the shock affecting the task on which no communication occurs. Devoting attention to both tasks and learning both shocks is then mainly valuable to learn which shock is largest but effectively half of the information the manager collects is “wasted.” In contrast, when internal alignment is not very important ( $\beta$  small), the manager wants to be responsive to both shocks, even when she only communicates about one task. Learning both shocks is then much more valuable.

C. Allocation of Attention by a Specialist Manager

Consider next a specialist manager, one that has more expertise about, say, task 1 than task 2, that is,  $T_1 > T_2$ . Formally, define

$$\rho(T_1, T_2) \equiv \frac{q(T_1)}{q(T_2)} = \frac{1 - e^{-\lambda T_1}}{1 - e^{-\lambda T_2}},$$

which we refer to as the *relative specialization ratio*. Thus,  $\rho$  captures how much better the manager is at observing shock  $\theta_1$  than observing shock  $\theta_2$ . A ratio of  $\rho = 1$  characterizes a generalist manager and a ratio  $\rho = +\infty$  or  $\rho = 0$  a fully specialized manager in task 1 and 2, respectively.

*Marginal Allocations of Attention.*—Consider first the case where attention is scarce, that is  $\tau$  is small and  $t_i \in [0, 2\tau]$  with  $t_1 + t_2 \leq 2\tau$ . Should the specialist manager reinforce her expertise by devoting her scarce attention to the task she is already an expert on (“managing with style”) or should she instead compensate for her lack of expertise in task 2?

The logic here is similar to the case of a generalist manager. When the manager has more expertise in task 1, she is more likely to communicate about task 1 to the organization as she is more likely to perceive the shock affecting task 1 to be largest. Improving the precision of  $\hat{\theta}_1$  is then more valuable than improving the precision of  $\hat{\theta}_2$ . Similarly, the higher is  $T_1$ , the less likely it is that the manager will communicate about task 2 and, hence, the more likely it is that devoting scarce attention to task 2 is a waste of time. Formally, if  $T_1 > T_2$ , then from (11),  $\Pr(\hat{\theta}_1^2 > \hat{\theta}_2^2) > \Pr(\hat{\theta}_2^2 > \hat{\theta}_1^2)$ . From expression (10), profits are then more sensitive to  $\hat{\theta}_1^2$  than to  $\hat{\theta}_2^2$ . It is then optimal to set  $(t_1, t_2) = (2\tau, 0)$  to further improve the precision of  $\hat{\theta}_1$  provided  $\Pr(\hat{\theta}_1^2 > \hat{\theta}_2^2) - \Pr(\hat{\theta}_2^2 > \hat{\theta}_1^2)$  is sufficiently large and  $\tau$  is small.

In environments that are less complex ( $\lambda$  large), however, the realization of both shocks will be learned with great precision. In this case, the manager communicates with almost equal probability about both tasks even when  $T_1$  is larger than  $T_2$ . Indeed, from (11), when  $\lambda$  is large,  $\Pr(\hat{\theta}_1^2 > \hat{\theta}_2^2) \approx \Pr(\hat{\theta}_2^2 > \hat{\theta}_1^2)$ . Given that there are decreasing marginal returns to attention, the expected profit function (10) is then maximized by devoting attention to the shock in which the manager has less expertise.

Formally, fixing the relative specialization ratio  $\rho > 1$  and defining

$$(13) \quad b \equiv \frac{\beta p}{1 + \beta(1 - p)} \in \mathbb{R}_+,$$

we show in the online Appendix that there exists a unique cut-off  $\Lambda(\rho, b) > 0$  such that

$$\frac{\partial}{\partial t} \Pi(q(T_1 + t), q(T_2 - t)) \Big|_{t=0} > 0 \Leftrightarrow \lambda T_1 < \Lambda(\rho, b),$$



where  $\Lambda(\rho, b)$  is a continuous function of  $\rho$  and  $b$ , implicitly defined in the online Appendix. Hence, keeping the relative specialization ratio  $\rho$  constant, a manager strictly prefers to “manage with style” and devote her scarce attention to task 1 if and only if  $\lambda T_1$  is below  $\Lambda(\rho, b)$ .

The cutoff  $\Lambda(\rho, b)$  is uniquely determined by two parameters: the relative specialization ratio,  $\rho$ , and the parameter  $b$ , which reflects the magnitude of the adaptation-coordination trade-off. As we show in the online Appendix, the cutoff  $\Lambda(\rho, b)$  is strictly increasing in  $b$ . We can rewrite  $b$  as  $(1 + \beta)|\alpha_1^* - \alpha_2^*|$ , where  $\alpha_i^*$  is the adaptiveness of the manager’s action to the local shock  $\theta_i$ , as given by (9). Hence, an increase in the importance of internal alignment  $\beta$  or an increase in the ability of the manager to communicate his strategy  $p$  (which increases  $|\alpha_1^* - \alpha_2^*|$ ) increases  $\Lambda(\rho, b)$  and makes managing with style more likely to be optimal.

**PROPOSITION 4 (Specialist Manager):** *Assume (A), and consider a specialist manager with  $T_1 > T_2$ . There exists a cutoff  $\Lambda(\rho, b)$  such that if  $\tau$  is sufficiently small, then*

(i) *If  $\lambda T_1 < \Lambda(\rho, b)$ , managing with style is optimal:  $(t_1^*, t_2^*) = (2\tau, 0)$ .*

(ii) *If  $\lambda T_1 > \Lambda(\rho, b)$ , managing with style is suboptimal:  $(t_1^*, t_2^*) = (0, 2\tau)$ .*

(iii) *The cutoff  $\Lambda(\rho, b)$  is increasing in the need for internal alignment,  $\beta$ , and the manager’s ability to communicate strategic choices,  $p$ .*

Figure 4 illustrates Proposition 4.<sup>14</sup> The upward-sloping (blue) lines indicate effective expertise combinations  $(\lambda T_1, \lambda T_2)$  for which the relative specialization ratio  $\rho$  is constant (the plot is obviously symmetric around the 45 degree line). For each value of  $\rho$  there is a unique threshold  $\Lambda(\rho, b)$ , such that if  $\lambda T_1 < \Lambda(\rho, b)$ , then on the margin, the manager prefers to “manage with style” and devote attention to task 1 rather than task 2. The downward-sloping (black) line links all those points to plot the function  $\Lambda(\rho, b)$  in the  $(\lambda T_1, \lambda T_2)$  space. Hence, for all points  $(\lambda T_1, \lambda T_2)$  below the downward-sloping (black) line, the optimal allocation of the marginal unit of attention is  $(t_1^*, t_2^*) = (2\tau, 0)$ , whereas it is  $(t_1^*, t_2^*) = (0, 2\tau)$  above it. Note that when managing with style is suboptimal and, thus,  $(t_1^*, t_2^*) = (0, 2\tau)$ , an upper-bound on  $2\tau$  is that  $2\tau \leq T_1 - T_2$ . Indeed, once  $2\tau > T_1 - T_2$ , it is optimal to allocate attention such that  $T_1 + t_1 = T_2 + t_2$ .<sup>15</sup>

Proposition 4 emphasizes the interaction between managerial expertise  $(T_1, T_2)$ , the complexity of the environment  $1/\lambda$  in which the manager operates, and the optimal allocation of scarce attention  $(t_1^*, t_2^*)$ . From Figure 4, for any expertise combination

<sup>14</sup>The numerical values used to generate this plot are the same as the ones used in the example introduced in Section IF,  $\beta = 5$  and  $p = 0.75$  such that  $b = 5/3$ . We have plotted the function  $\Lambda(\rho, b)$  for  $b = \beta p / (1 + \beta(1 - p))$  ranging from  $b \approx 0$  to  $b$  large ( $b = 100$ ). Notice that  $b$  is only a function of the primitives of the model. As discussed in online Appendix A.A5, for all those parameter values, we obtain similar shapes for  $\Lambda(\rho, b)$  as for  $b = 5/3$ . Thus, the black curve representing  $\Lambda(\rho, b)$  in the  $(\lambda T_1, \lambda T_2)$  space is decreasing, and any linear function  $\lambda T_2 = Z - \lambda T_1$ , with  $Z \in \mathbb{R}^+$ , crosses  $\Lambda(\rho, b)$  at most once, and always from above.

<sup>15</sup>See the next subsection on “intermediate allocations of attention.”

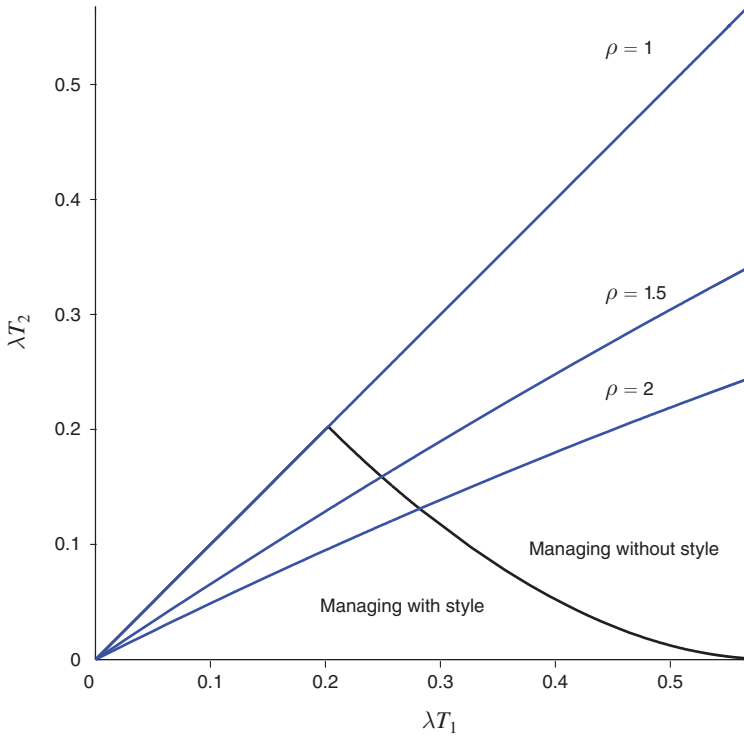


FIGURE 4. OPTIMAL ALLOCATION OF THE MARGINAL UNIT OF ATTENTION GIVEN EFFECTIVE EXPERTISE  $(\lambda T_1, \lambda T_2)$  WITH  $T_1 > T_2$

Notes: The downward-sloping line is the function  $\Lambda(\rho, b)$  plotted in the space  $(\lambda T_1, \lambda T_2)$  (see Proposition 4). Below the downward-sloping curve the manager allocates the marginal unit of attention to task 1, whereas above she allocates the marginal unit of attention to task 2. The upward-sloping (blue) lines show iso- $\rho$ , where  $\rho$  is the relative specialization ratio for  $\rho \in \{1, 1.5, 2\}$ .

$(T_1, T_2)$  with  $T_1 > T_2$ , we have that  $(t_1^*, t_2^*) = (2\tau, 0)$  if  $\lambda$  is sufficiently small (complex environments), whereas  $(t_1^*, t_2^*) = (0, 2\tau)$  if  $\lambda$  is large (less complex environments). Note that this is also the result shown in Figure 3, panel A, which provided a preview of our results.

*Intermediate Allocations of Attention.*—Proposition 4 also has implications for when  $2\tau$  is “intermediate” or “large.” Consider the set  $S$  of effective expertise configurations where the manager has more expertise in task 1 than task 2:<sup>16</sup>

$$S = \{(\lambda T_1, \lambda T_2) : T_1 > T_2 \geq 0\}.$$

Let  $S^-$  be the subset where “*managing with style*” is optimal, that is

$$S^- = \{(\lambda T_1, \lambda T_2) \in S : \lambda T_1 < \Lambda(\rho, b)\}$$

<sup>16</sup>If the manager has more expertise in task 2, we can simply relabel the tasks.

and  $S^+$  the subset of  $S$  where managing with style is suboptimal, that is

$$S^+ = \{(\lambda T_1, \lambda T_2) \in S : \lambda T_1 > \Lambda(\rho, b)\}.$$

In Figure 4 the subset  $S^-$  is the area below the black curve and  $S^+$  is the area above the black curve. It can be shown that both  $S^-$  and  $S^+$  are connected sets (see online Appendix A5, Remark 1).

Let us further denote by

$$\lambda\omega_i = \lambda(T_i + t_i)$$

the “final” effective expertise of the manager in task  $i$ , that is her effective expertise after the allocation of attention. We denote  $(\lambda\omega_1, \lambda\omega_2) \succ (\lambda\omega'_1, \lambda\omega'_2)$  whenever payoffs are higher given final expertise  $(\lambda\omega_1, \lambda\omega_2)$  than final expertise  $(\lambda\omega'_1, \lambda\omega'_2)$ .

Proposition 4 has the following corollary, formally proven in the online Appendix.

**COROLLARY 5:** Denote  $\omega_i = T_i + t_i$ . For  $\varepsilon$  sufficiently small:

(i) If  $(\lambda\omega_1, \lambda\omega_2) \in S^-$ , then  $(\lambda(\omega_1 + \varepsilon), \lambda(\omega_2 - \varepsilon)) \succ (\lambda\omega_1, \lambda\omega_2)$ .

(ii) If  $(\lambda\omega_1, \lambda\omega_2) \in S^+$ , then  $(\lambda(\omega_1 - \varepsilon), \lambda(\omega_2 + \varepsilon)) \succ (\lambda\omega_1, \lambda\omega_2)$ .

In other words, if the allocation of attention  $(t_1, t_2)$  is such that the manager’s final effective expertise  $(\lambda\omega_1, \lambda\omega_2)$  is below the black curve in Figure 4 (area  $S^-$ ), then on the margin, profits can be improved by shifting attention away from task 2 toward task 1. Intuitively, this is the area where managing with style is optimal. Similarly, if the allocation of attention were to be such that the final effective expertise is above the black curve in Figure 4 (area  $S^+$ ), then on the margin, profits can be improved by shifting attention away from task 1 towards task 2, where the manager is less of an expert. Intuitively, this is the area where managing with style is suboptimal. A direct implication of the Corollary 5 is that any attention allocation which result in a final effective expertise  $(\lambda(T_1 + t_1^*), \lambda(T_2 + t_2^*))$  belonging to  $S^-$  must have  $t_2^* = 0$ . Similarly, any optimal allocation of allocation which result in a final effective expertise belonging to  $S^+$  must have  $t_1^* = 0$ .

**COROLLARY 6:** Consider an optimal allocation of attention  $(t_1^*, t_2^*)$ , and denote  $\omega_i^* = T_i + t_i^*$ :

(i) If  $(\lambda\omega_1^*, \lambda\omega_2^*) \in S^-$ , then  $(t_1^*, t_2^*) = (2\tau, 0)$ .

(ii) If  $(\lambda\omega_1^*, \lambda\omega_2^*) \in S^+$ , then  $(t_1^*, t_2^*) = (0, 2\tau)$ .

To understand the implications of the corollaries, consider Figure 5, panels A–C, which plot, again, the function  $\Lambda(\rho, b)$ .<sup>17</sup> The area below  $\Lambda(\rho, b)$  is the set  $S^-$ . The

<sup>17</sup> Figure 5, panels A–C, plots  $\Lambda(\rho, b)$  for  $b = 5/3$ . Online Appendix A.A5 provides a more formal analysis and discussion and shows that our insights carry through for other values of  $b$  as well.

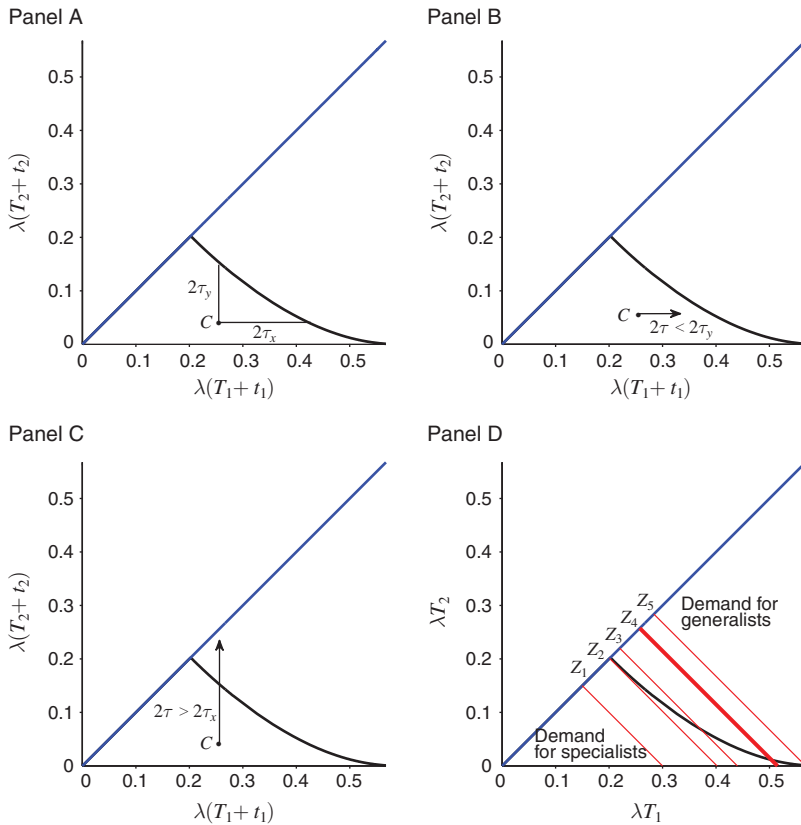


FIGURE 5

Notes: Panel A: Attention cutoff levels,  $2\tau_x$  and  $2\tau_y$ , given effective expertise  $C$ . Panels B–C: Endogenous allocation of attention  $2\tau$ . Panel D: Optimal choice of managerial expertise  $(T_1, T_2)$  given  $T_1 + T_2 \leq Z_i$ . Specialists optimal for  $Z_i \leq Z_4$  and generalists optimal for  $Z_i > Z_4$ .

area above  $\Lambda(\rho, b)$ , but below the 45 degree line, represents  $S^+$ . Consider now any expertise configuration  $(\lambda T_1, \lambda T_2) \in S^-$ , such as the one represented by point  $C = (\lambda T_1^C, \lambda T_2^C)$ .

As long as  $2\tau$  is not too large, a manager then wants to devote all her attention to task 1, on which she is already an expert. Indeed, assume  $2\tau < 2\tau_y$ , where

$$\tau_y \equiv \sup \{ \tau \in \mathbb{R}^+ : (\lambda T_1, \lambda(T_2 + 2\tau)) \in S^- \}.$$

From Figure 5, panel B, if  $2\tau < 2\tau_y$ , then regardless of the choice of  $(t_1, t_2)$ , we have that  $(\lambda\omega_1, \lambda\omega_2) \in S^-$ . But from Corollary 5, any attention allocation  $(t_1, t_2)$  with  $t_2 > 0$  can then be strictly improved upon by shifting attention from task 2 to task 1. It follows that for  $2\tau < 2\tau_y : (t_1^*, t_2^*) = (2\tau, 0)$ , as in panel B.

When  $2\tau$  is sufficiently large, however, it will eventually become optimal to devote all or most attention to task 2 in which the manager is not an expert. Indeed, let  $2\tau > 2\tau_x$ , with

$$\tau_x \equiv \sup\{\tau \in \mathbb{R}^+ : (\lambda(T_1 + 2\tau), \lambda T_2) \in S^-\}.$$

From Figure 5, panel A, if  $2\tau > 2\tau_x$  and  $T_1 + t_1 > T_2 + t_2$ , then  $(\lambda\omega_1, \lambda\omega_2) \in S^+$ . But by Corollary 5, profits can then be increased by shifting some attention from task 1 to task 2. It follows that for  $2\tau > 2\tau_x$ , either  $(t_1^*, t_2^*) = (0, 2\tau)$  as in panel C or  $(t_1^*, t_2^*)$  is such that  $T_1^C + t_1^* = T_2^C + t_2^*$ .<sup>18</sup>

Finally, in less complex environments where  $(\lambda T_1, \lambda T_2) \in S^+$ , we always have that  $(t_1^*, t_2^*) = (0, 2\tau)$  or  $T_1^C + t_1^* = T_2^C + t_2^*$ . This corresponds, for example, to the effective expertise configuration *B* in Figure 3 (preview of results). The logic is again that on the margin, profits can then always be improved by shifting attention away from task 1 to task 2.

*D. Endogenous Managerial Expertise: Specialists versus Generalists*

So far, we have taken managerial expertise as given and endogenized the allocation of managerial attention. Managers though are typically appointed by boards (or, interchangeably in this paper, firm owners) who select them depending on their expertise. We are interested in the board’s decision to appoint a generalist or a specialist manager, even when firm owners are indifferent about the organization’s strategic direction.

Assume therefore that a board can choose any manager whose expertise  $(T_1, T_2)$  belongs to some “opportunity set”  $\Gamma$ . To fix ideas we will assume that

$$\Gamma = \Gamma(Z) \equiv \{(T_1, T_2) : T_1 + T_2 \leq Z \text{ and } T_1 \geq T_L, T_2 \geq T_L\},$$

where  $Z$  is the total “expertise budget” and  $T_L \geq 0$  is the minimum expertise of any manager.

Notice that since there are no agency problems between firm owners and the manager, they agree as to the optimal allocation of managerial attention  $t_1$  and  $t_2$ , with  $t_1 + t_2 \leq 2\tau$ . We can, thus, think of firm owners choosing  $\tilde{T}_1 = T_1 + t_1$  and  $\tilde{T}_2 = T_2 + t_2$  with

$$(\tilde{T}_1, \tilde{T}_2) \in \Gamma(Z + 2\tau),$$

where  $\tilde{T}_i$  is the final expertise in task  $i$  after the optimal allocation of managerial attention  $t_i$ : An attention budget  $2\tau > 0$  paired with an expertise budget  $Z$  is formally equivalent to an attention budget  $2\tau = 0$  paired with an expertise budget  $Z + 2\tau$ . Without any loss of generality, we therefore simplify the problem by setting  $\tau = 0$  so that  $T_i = \tilde{T}_i$ . Moreover, we further posit, again, without any loss

<sup>18</sup>It is also easy to see that we cannot have  $\omega_2^C > \omega_1^C$ , as then it would be optimal to shift attention from task 2 to task 1.

of generality, that  $\lambda = 1$ . An increase in  $\lambda$  is formally equivalent with a proportional increase in the expertise budget  $Z$ . Comparative statics with respect to  $\lambda$  are therefore equivalent to comparative statics with respect to  $Z$ . We will thus associate a more complex environment with a smaller expertise budget for the board.

Figure 5, panel D, is useful in illustrating board choices regarding a specialist or generalist manager.<sup>19</sup> There, in addition to the iso- $\rho$  lines, we have drawn five straight (red) iso- $Z$  lines, where the total expertise budget  $Z = T_1 + T_2$  is kept constant. Recall that the downward-sloping curve (in black) represents, again, all expertise combinations  $(T_1, T_2)$  for which

$$\left. \frac{d\Pi(q(T_1 + t), q(T_2 - t))}{dt} \right|_{t=0} = 0,$$

with this derivative being positive for values of  $(T_1, T_2)$  below this line and negative above it.

It follows from Proposition 4 that keeping the total expertise budget  $Z = T_1 + T_2$  fixed, profits are decreasing in  $T_1$  above the downward-sloping curve but increasing in  $T_1$  below the curve. Moreover, since the curve in Figure 5 has a derivative that is larger than  $-1$ , it follows that moderate specialists for which  $0 < T_2 < T_1$  are always dominated by either extreme specialists (for which  $T_2 = 0$ ) or by complete generalists (for which  $T_1 = T_2$ ). Obviously, extreme specialists are preferred for the budgets  $Z_1$  and  $Z_2$  corresponding to the first two straight (red) lines, whereas complete generalists are preferred for the budget  $Z_5$  corresponding to the last budget line. More generally, one can show that extreme specialists are preferred over complete generalists if and only if  $Z \leq Z_4 \approx 0.515$ , where  $Z_4$  is the budget corresponding to the thick red line.

Sometimes, however, extreme specialists are not available to the board, as any manager has some minimal expertise  $T_L$  on either task. For example,  $T_L$  can be considered as basic knowledge any manager has or, equivalently, the minimum attention a manager must devote to either task. For expertise budgets  $Z < Z_2 \approx 0.402$  or budgets  $Z > Z_4$ , the minimal task knowledge  $T_L$  does not affect the choice between generalist and specialists. But for  $Z \in (Z_2, \bar{Z})$ , the board may prefer a complete generalist  $(T_1, T_2) = (Z/2, Z/2)$  over a moderate specialist  $(T_1, T_2) = (Z - T_L, T_L)$  when  $T_L$  is sufficiently large. Moreover, the larger the basic knowledge  $T_L$  any manager has about both tasks, the more likely a generalist manager is optimal.<sup>20</sup>

In sum, our analysis shows that generalist managers are preferred in less complex environments—that is in environments with larger expertise budgets, where even a generalist will have a precise estimate of both task-specific shocks. In contrast, when

<sup>19</sup> Figure 5, panel D, assumes  $b = 5/3$ . Online Appendix A.A5 provides a more formal discussion and analysis and shows that our insights carry through for other values of  $b$  as well.

<sup>20</sup> Indeed, consider the red budget line, which represents an expertise budget  $Z_3 = 0.44$  and crosses the (black) downward-sloping curve at  $T_2 = 0.065$ . Since  $Z_3 < Z_4$ , the board strictly prefers an extreme specialist  $(Z_3, 0)$  over a complete generalist  $(Z_3/2, Z_3/2)$ . But whenever  $T_L \geq 0.065$ , profits are even minimized by choosing a manager with expertise  $(Z_2 - T_L, T_L)$ .



expertise is in short supply, that is for smaller expertise budgets or more complex environments, boards prefer hiring specialist managers.

An interesting extension of our model would be to endogenize the expertise budgets of firms (and, potentially, managerial wages). If, as seems intuitive, larger firms have larger expertise budgets in equilibrium, a prediction of our model would be that, all else equal, generalist managers are more likely to be hired by larger, more established firms, whereas specialist managers are more likely to be hired by smaller firms.

Another question is when and how a board would decide on replacing a manager in a dynamic version of our model. As long as  $\theta_1$  and  $\theta_2$  are i.i.d., the board has no incentive to do so. A shift in the distribution of shocks, however, may make it optimal to replace a manager. Consider an extension where one shock, say  $\theta_1$  has a larger variance. A board then strictly prefers a specialist in task 1 rather than a specialist in task 2. But this also implies that when the variance of task 2 shocks becomes larger, the board may optimally replace the manager. Of course, this begs the question as to whether boards have the required expertise to identify shifts in the distribution of shocks and whether there are any board fixed effects in identifying such shifts. We leave this question open for future research.

#### IV. Discussion and Conclusions

The goal of this paper was to shed light on when we should expect firm behavior and strategic choices to reflect managerial characteristics as opposed to the realization of environmental shocks. In our model, the allocation of managerial attention partially determines firm behavior, but it can either exacerbate manager fixed effects or mitigate them. We now discuss our findings and their implications in more detail.

##### A. Magnitude of Manager Fixed Effects

*Performance Differences between Seemingly Similar Managers.*—A first implication of our model is that *small* initial differences in managerial expertise may result in dramatically different firm behavior, as managers devote their scarce attention in a way that amplifies initial differences in expertise. It follows that assessing the impact of managerial expertise on firm strategies and outcomes requires precise measures of that expertise; otherwise *small* errors in measurement can produce the standard attenuation bias and, thus, lead the researcher to underestimate the effect of expertise on firm strategies.

*Environmental Complexity and Manager Fixed Effect.*—Ours is a theory of managerial style where differences in firm behavior are driven by differences in information processing (or cognition) between managers. An implication from cognition-based theories of managerial style is that the magnitude of manager fixed effects should depend on environmental complexity. As noted above, in more complex settings, the endogenous allocation of managerial attention exacerbates initial differences in task expertise. But in environments with less complexity (corresponding to high values of  $T$  or  $\lambda$  in our model), our model predicts that

managers with different expertise will make similar or even identical choices as they allocate attention in a manner that reduces or eliminates differences in task expertise. The predictions of our cognition-based model can therefore be empirically distinguished from alternative theories where manager fixed effects reflect differences in the capability of managers to execute certain strategies or differences in managerial preferences. Indeed, under these alternative theories, we still expect to see strong manager fixed effects in less complex environments (e.g., settings with limited uncertainty), whereas this is not the case in our framework.

In the same vein, our results indicate that studies that focus on tenure and educational background (a potential proxy for  $T$ ) to understand managerial strategies need to interact those managerial characteristics with industry measures that capture the complexity of the environment in which the manager operates to explain the cross section of managerial strategies.

### *B. Managers and Performance Differences between Firms*

*Persistent Performance Differences between Seemingly Similar Firms.*—While a large literature has established the impact of managers on firm behavior, another literature has been interested in persistent performance differences between seemingly similar firms (see, for example, Gibbons and Henderson, 2013, and Li, Matouschek, and Powell 2017). In many settings, differences in behavior and performance between seemingly similar firms are arguably related to differences in managerial style. But this creates another, related, question: Why do seemingly similar firms hire managers with different managerial characteristics to begin with? Our model implies that boards—when complexity is large—optimally hire managers with specialized expertise, but do not necessarily care about the particular area or function the manager has expertise in (say marketing versus operations). In other words, while the area of expertise is an important predictor of firm behavior and (ex post) firm performance, it is not necessarily an important criterion of choice for a board of directors. Instead, boards are likely to choose managers based on leadership ability, general cognitive ability, availability, and other factors. Different boards are therefore likely to hire very different managers (in terms of functional expertise), even when faced with an identical economic environment and even when boards are themselves very similar. Given that the average tenure of a CEO of a S&P 500 firm was 9.7 years in 2013, such “random” choices may have long-lasting effects on firm behavior and performance (see also Dessein and Prat 2019, for a related argument).

*Environmental Complexity and Performance Differences between Firms.*—To the extent that differences in how managers allocate attention and process information are at the source of manager fixed effects, as posited by the present paper, we expect to see a conformity of firm strategies and firm behavior in less complex environments (e.g., those characterized by low uncertainty) but a large dispersion of firm behavior (correlated with managerial backgrounds) in highly complex environments. In contrast, if differences in managerial capabilities are at the source of differences in firm behavior, one should expect to see a similar dispersion in firm behavior and firm strategies in highly and moderately complex environments.

*Managerial Style versus Organizational Culture.*—In this paper, we have assumed that the manager can direct the attention of workers toward a particular task. Indeed, one of the important roles of a CEO is to communicate a vision/strategy for the firm. There may also be environments, however, where worker attention is, to some extent, outside of the manager's control. Equilibria may then exist where workers coordinate all their attention on, say, task 1 and, in return, the manager is mainly responsive to her posterior  $\hat{\theta}_1$  (but largely ignores  $\hat{\theta}_2$ ). In this case, managerial style is not driven by managerial characteristics but, instead, by an (inflexible) organizational culture. Performance differences between firms may thus also reflect different equilibria (or cultures) rather than manager fixed effects.

### C. Why Do Managers Matter? Practices versus Strategic Choices

We view our paper as shedding light on the channels through which managers matter for firm performance. Following a seminal paper by Bloom and Van Reenen (2007), much of the economics' literature on this topic has been focused on "management practices." The premise of much of this literature is that many companies are not run efficiently, and better management can improve operational effectiveness (see, for example, Syverson 2004; Foster, Haltiwanger, and Syverson 2008; Bloom and Van Reenen 2010; Bloom et al. 2013). Instead, much of the management literature on why managers matter has focused on the impact of managers on strategic choices as opposed to operational efficiency.<sup>21</sup> As argued by Finkelstein, Hambrick, and Cannella (2009, 5):

But where does the company's strategy come from? (...) To be sure, strategic actions are sometimes due to imitation, inertia, and careful, objective decision making. But a wealth of research and everyday observation indicates that strategy and other major organizational choices are made by humans who act on the basis of idiosyncratic experiences, motives and dispositions. If we want to understand strategy, we must understand strategists.

Our view is that to assess the impact of managers on firms, empirical work on managers should move beyond management practices and focus on specific strategic choices in particular industries. For example, Kaplan, Murray, and Henderson (2003) analyze the responses of 15 large, incumbent pharmaceutical firms responses to the emergence of biotechnology. Cho and Hambrick (2006) study strategic responses to airline deregulation. Finally, Kaplan (2008) studies how CEOs of 71 communications firms responded to the fiber-optic revolution in the communications technology industry. Another promising approach is to measure managerial attention to functional areas directly by employing time use surveys, as in Bandiera et al. (2017) and Bandiera et al. (2020), and correlate this with specific strategic choices.

While better management practices almost always improve performance, the difference between what is optimal *ex ante* and *ex post* is key when analyzing strategic choices made by managers. For example, if a firm is late to the fiber-optic

<sup>21</sup> See Roberts and Saloner (2013) for a discussion of how the management literature has defined "business strategy" and "strategic choices."

revolution in telecommunications, then this is not necessarily evidence of bad management. Indeed, as our paper shows, it is optimal for managers from an *ex ante* perspective to focus all attention on one area. If unlucky, the manager may then be blindsided and *ex post* be forced to compete along dimensions to which she optimally did not devote much attention. To put it differently, management evaluation should be done along two dimensions: strategic choices and practices. Both obviously affect performance. For strategic choices the difference between *ex ante* and *ex post* is very relevant, while this is less so for the case of management practices.<sup>22</sup> In addition, cross-sectional dispersion in strategic choices within an industry should not be taken to be a sign of inefficiency, whereas diversity in practices is more likely to indicate that some firms could improve performance through the adoption of better practices. Similarly management may stick to adopted strategies particularly in the presence of temporary spikes in uncertainty (because of industry wide regulatory overhauls, for example) and volatility (because of innovations waves for instance), even at the risk of being blindsided and being forced to act on imprecise information. The empirical challenge is to distinguish between strategies and practices when assessing management.

#### *D. Cognitive Constraints as a Source of Manager Fixed Effects*

In our model, managerial fixed effects arise because of constraints on managerial cognition (and the need for organizational alignment). Managers who face the same economic environment and the same facts may come to different strategic choices as they devote their scarce attention to different sources of information. Cognitive limitations are also seen as the primary source of manager fixed effects in the management literature,<sup>23</sup> but, in contrast to our model, emphasis is put on behavioral biases in decision-making as a central argument as to why managers matter. In particular, the management literature follows the logic of the Carnegie School (March, Simon, and Guetzkow 1958; Cyert and March 1963) according to which complex choices are largely determined by behavioral factors, rather than by calculations of optimal actions. According to the dominant stream in this literature:

[I]n arriving at their own rendition of a strategic situation, or “construed reality” (Sutton 1987), executives distill and interpret the stimuli that surround them. This occurs through a three stage filtering process. Specifically, executive orientations affect their field of vision (the directions in which they look and listen), selective perception (what they actually see and hear), and interpretation (how they attach meaning to what they see and hear).<sup>24</sup>

A contribution of our paper has been to show how manager fixed effects may arise even when managers optimally (and rationally) allocate attention and process

<sup>22</sup>There are of course practices, say, sound accounting, that are strategy neutral whereas other practices are naturally strategy dependent.

<sup>23</sup>See Finkelstein, Hambrick, and Cannella (2009, chapter 2) for an overview.

<sup>24</sup>Finkelstein, Hambrick, and Cannella (2009, 46).

information. From a normative point of view, our results thus show that managerial biases in information processing are not necessarily pathological. Indeed, a key insight of our model is that, given the presence of cognitive limits to attend to all possible information, boards or firm owners often prefer managers whose field of vision is narrow.

From a conceptual point of view, we show that managers may matter even when they are rational optimizing creatures. Rather than exogenously posit that the “field of vision” of a manager is determined by her expertise and past experiences, managers in our model optimally choose their “field of vision,” but behave largely as predicted (and, indeed, observed) by the management literature. We further link a manager’s “field of vision” to organizational factors, such as the need for organizational alignment around simple strategies and environmental factors, such as the degree of complexity and the scarcity of attention.

## REFERENCES

- Akerlof, Robert, and Richard Holden. 2016. “Movers and Shakers.” *Quarterly Journal of Economics* 131 (4): 1849–74.
- Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. “When Does Coordination Require Centralization?” *American Economic Review* 98 (1): 145–79.
- Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2015. “Organizing to Adapt and Compete.” *American Economic Journal: Microeconomics* 7 (2): 158–87.
- Bandiera, Oriana, Renata Lemos, Andrea Prat, and Raffaella Sadun. 2017. “Managing the Family Firm: Evidence from CEOs at Work.” *Review of Financial Studies* 31 (5): 1605–53.
- Bandiera, Oriana, Andrea Prat, Stephen Hansen, and Raffaella Sadun. 2020. “CEO Behavior and Firm Performance.” *Journal of Political Economy* 128 (4): 1325–69.
- Barker, Vincent L., III, and George C. Mueller. 2002. “CEO Characteristics and Firm R&D Spending.” *Management Science* 48 (6): 782–801.
- Bertrand, Marianne, and Antoinette Schoar. 2003. “Managing with Style: The Effect of Managers on Firm Policies.” *Quarterly Journal of Economics* 118 (4): 1169–208.
- Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. “Does Management Matter? Evidence from India.” *Quarterly Journal of Economics* 128 (1): 1–51.
- Bloom, Nicholas, and John Van Reenen. 2007. “Measuring and Explaining Management Practices across Firms and Countries.” *Quarterly Journal of Economics* 122 (4): 1351–408.
- Bloom, Nicholas, and John Van Reenen. 2010. “Why Do Management Practices Differ across Firms and Countries?” *Journal of Economic Perspectives* 24 (1): 203–24.
- Bolton, Patrick, Markus K. Brunnermeier, and Laura Veldkamp. 2013. “Leadership, Coordination, and Corporate Culture.” *Review of Economic Studies* 80 (2): 512–37.
- Bolton, Patrick, and Mathias Dewatripont. 1994. “The Firm as a Communication Network.” *Quarterly Journal of Economics* 109 (4): 809–39.
- Calvo-Armengol, Antonio, Joan de Martí, and Andrea Prat. 2015. “Communication and Influence.” *Theoretical Economics* 10 (2): 649–90.
- Cho, Theresa S., and Donald C. Hambrick. 2006. “Attention as the Mediator between Top Management Team Characteristics and Strategic Change: The Case of Airline Deregulation.” *Organization Science* 17 (4): 453–69.
- Cyert, Richard M., and James G. March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.
- Dessein, Wouter, Andrea Galeotti, and Tano Santos. 2016. “Rational Inattention and Organizational Focus.” *American Economic Review* 106 (6): 1522–36.
- Dessein, Wouter, and Andrea Prat. 2019. “Organizational Capital, Corporate Leadership, and Firm Dynamics.” Columbia Business School Research Paper 17-89.
- Dessein, Wouter, and Tano Santos. 2006. “Adaptive Organizations.” *Journal of Political Economy* 114 (50): 956–95.
- Ferreira, Daniel, and Marcelo Rezende. 2007. “Corporate Strategy and Information Disclosure.” *RAND Journal of Economics* 38 (1): 164–84.



- Finkelstein, Sydney, Donald C. Hambrick, and Albert A. Cannella.** 2009. *Strategic Leadership: Theory and Research on Executives, Top Management Teams, and Boards*. New York: Oxford University Press.
- Foster, Lucia, John Haltiwanger, and Chad Syverson.** 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability." *American Economic Review* 98 (1): 394–425.
- Galeotti, Andrea, and Sanjeev Goyal.** 2010. "The Law of the Few." *American Economic Review* 100 (4): 1468–92.
- Garicano, Luis, and Andrea Prat.** 2013. "Chapter 11—Organizational Economics with Cognitive Costs." In *Advances in Economics and Econometrics: Tenth World Congress*, Vol. 1, edited by Daron Acemoglu, Manuel Arellano, and Eddie Dekel, 342–88. Cambridge, UK: Cambridge University Press.
- Garicano, Luis, and Timothy Van Zandt.** 2012. "Chapter 15—Hierarchies and the Division of Labor." In *Handbook of Organizational Economics*, edited by Robert Gibbons and John Roberts, 604–54. Princeton, NJ: Princeton University Press.
- Geanakoplos, John, and Paul Milgrom.** 1991. "A Theory of Hierarchies Based on Limited Managerial Attention." *Journal of the Japanese and International Economics* 5 (3): 205–25.
- Gibbons, Robert, and Rebecca Henderson.** 2013. "Chapter 17—What Do Managers Do? Exploring Persistent Performance Differences among Seemingly Similar Enterprises." In *Handbook of Organizational Economics*, edited by Robert Gibbons and John Roberts, 680–731. Princeton, NJ: Princeton University Press.
- Hambrick, Donald C.** 2018. "Upper Echelons Theory." In *The Palgrave Encyclopedia of Strategic Management*, edited by Augier M., Teece D.J. London: Palgrave Macmillan. [https://doi.org/10.1057/978-1-137-00772-8\\_785](https://doi.org/10.1057/978-1-137-00772-8_785).
- Hambrick, Donald C., and Phyllis A. Mason.** 1984. "Upper Echelons: The Organization as a Reflection of Its Top Managers." *Academy of Management Review* 9 (2): 193–206.
- Herskovic, Bernard, and João Ramos.** 2020. "Acquiring Information through Peers." *American Economic Review* 110 (7): 2128–52.
- Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp.** 2016. "A Rational Theory of Mutual Funds' Attention Allocation." *Econometrica* 84 (2): 571–626.
- Kaplan, Sarah.** 2008. "Cognition, Capabilities, and Incentives: Assessing Firm Response to the Fiber-Optic Revolution." *Academy of Management Journal* 51 (4): 672–95.
- Kaplan, Sarah, Fiona Murray, and Rebecca Henderson.** 2003. "Discontinuities and Senior Management: Assessing the Role of Recognition in Pharmaceutical Firm Response to Biotechnology." *Industrial and Corporate Change* 12 (2): 203–33.
- Li, Jin, Niko Matouschek, and Michael Powell.** 2017. "Power Dynamics in Organizations." *American Economic Journal: Microeconomics* 9 (1): 217–41.
- March, James G., Herbert A. Simon, and Harold Guetzkow.** 1958. *Organizations*. New York: John Wiley and Sons.
- Ocasio, William.** 1997. "Towards an Attention-Based View of the Firm." *Strategic Management Journal* 18: 187–206.
- Rantakari, Heikki.** 2008. "Governing Adaptation." *Review of Economic Studies* 75 (4): 1257–85.
- Rantakari, Heikki.** 2013. "Organizational Design and Environmental Volatility." *Journal of Law, Economics, and Organization* 29 (3): 569–607.
- Roberts, John, and Garth Saloner.** 2013. "Chapter 20—Strategy and Organization." In *Handbook of Organizational Economics*, edited by Robert Gibbons and John Roberts, 799–850. Princeton, NJ: Princeton University Press.
- Rotemberg, Julio J., and Garth Saloner.** 1994. "Benefits of Narrow Business Strategies." *American Economic Review* 84 (5): 1330–49.
- Rotemberg, Julio J., and Garth Saloner.** 2000. "Visionaries, Managers, and Strategic Direction." *RAND Journal of Economics* 31 (4): 693–716.
- Syverson, Chad.** 2004. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics* 86 (2): 534–50.
- Van den Steen, Eric.** 2005. "Organizational Beliefs and Managerial Vision." *Journal of Law, Economics, and Organization* 21 (1): 256–83.
- Van den Steen, Eric.** 2017. "A Formal Theory of Strategy." *Management Science* 63 (8): 2616–36.
- Van den Steen, Eric J.** 2018. "Strategy and the Strategist: How It Matters Who Develops the Strategy." *Management Science* 64 (10): 4533–51.
- Van Nieuwerburgh, Stijn, and Laura Veldkamp.** 2010. "Information Acquisition and Under-Diversification." *Review of Economic Studies* 77 (2): 779–805.