Managing with Style? Micro-Evidence on the Allocation of Managerial Attention*

Desmond (Ho-Fu) Lo, Francisco Brahm, Wouter Dessein, Chieko Minami

January 15, 2021

Abstract

How does task expertise affect the allocation of attention? Our theory argues that when attention is scarce, expertise and attention are complements: a manager optimally focuses her attention on tasks in which she has relatively more expertise; she "manages with style." In contrast, when attention is abundant, attention and expertise become substitutes: a manager shifts her attention towards tasks she has less expertise in; she "manages against her style." Using micro-level data on managers from two unrelated companies, and employing various measures of time stress and managerial attention, we find converging and supporting evidence. A manager's attention capacity determines whether she "manages with style," or "against it." While current behavioral approaches view "managing with style" as prevalent and biased, our theory and findings suggest, instead, that it is contingent and optimal.

---

*Lo: Santa Clara University, hlo@scu.edu. Brahm: London Business School, fbrahm@london.edu. Dessein: Columbia Business School, wd2179@columbia.edu. Minami: Kobe University, cminami@kobe-u.ac.jp. We thank Guoli Chen, Donal Crilly, Shi Tang, David Zhu, three anonymous referees at the Strategy Management Society Conference, and seminar participants at Columbia Business School, the London Business School, the University of Maryland, Empirical Management Conference, and the Thought Leadership on the Sales Profession Conference for constructive comments. We are also indebted to company executives who preferred to remain anonymous for helpful discussions and access to data. Chieko Minami acknowledges the financial support from MEXT/the Japan Society for the Promotion of Science (JSPS) Multi-Year Grant # 16H02035. The usual caveat applies.

1 Introduction

Managerial attention in organizations is a scarce resource as humans are bounded in their cognitive capacity (Simon 1947 and 1971; Cyert and March 1963). Upper echelons theory (UET) (Hambrick and Mason 1984; Hambrick 2007) and the attention-based view (ABV) (Ocasio 1997 and 2011; Ocasio and Joseph 2005) have emphasized the role of managerial and organizational attention in shaping an organization's behavior and outcomes. An important proposition of these theories is that (i) given bounded rationality, managers, and the organization in general, have difficulties in processing all the stimuli and information coming from the environment and the organization itself, and (ii) in order to cope with this cognitive deficit, managers, and the organization as a whole, tend to bias their attention towards the set of tasks and strategies in which they possess expertise and knowledge (lodged within managers, the focus of UET, or reflected in a firm's symbols, structures, routines, and roles, the focus of ABV). Consistent with (ii), Bertrand and Schoar (2003) carefully show that executives generate recognizable patterns across several organizational choices which reflect their managerial experience and background; in short, executives "manage with style." Several empirical studies provide convincing evidence of a systematic relationship between expertise-biased attention of managers and firm choices or outcomes (Levy 2005; Cho and Hambrick 2006; Eggers and Kaplan 2009; Kaplan 2008; Nadkarni and Barr 2008; Zhu and Chen 2015; Barker and Mueller 2002; Kaplan, Keblanov, and Sorenson 2012; Kilduff et al. 2000). Broadly supportive empirical reviews of UET can be found in Carpenter et al. (2004), Finkelstein et al. (2009, Ch. 4), and Neely Jr. et al. (2020), and a review on selected papers on ABV in Ocasio (2011).

Nonetheless, the underlying mechanism of this empirical regularity – bounded rationality leading to cognitive limits – requires further scrutiny. Current approaches view "managing with style" as prevalent, based on an assumption of pervasive bounded rationality. However, cognitive capacity – or how much rationality is bounded – may very well vary across situations. Depending on the work environment, family demands, and personal circumstances, managerial cognitive capacity may be abundant or scarce, varying between individuals as well as within individuals over time. It is also possible that, provided with time and resources, managers might regain a more purposeful and rational command of their attention allocation.

Executives and organizations might even attempt to reduce this cognitive limitation, striving to set conditions to fully process most stimuli and information.[1] Thus, some important questions to consider are: Is "managing with style" more prevalent in situations where managers are under considerable time stress? Do managers change their attentional focus when they are less overloaded, perhaps even rationally allocating their attention? Hambrick et al. (2005) and Hambrick (2007) theorize about this, predicting that when "job demands" are low, the expertise bias of attention will become weaker due to enhanced cognitive capacity. However, to the best of our knowledge, no research has tackled this question empirically. For instance, the latest review by Neely Jr. et al. (2020) does not report research in this area.[2]

In this paper we examine this issue by using a formal model and testing its predictions. Building on Dessein and Santos (2020), our model assumes rational decision making and shows that the allocation of managerial attention crucially depends on the degree of attention capacity. In line with previous research, we predict that when attention is scarce, managers focus their attention optimally on tasks in which they have more expertise. However, the model predicts the opposite when attention is abundant: managers optimally go *"against their style" and focus more on the areas on which they have less expertise*. This "flip" in attention allocation when cognitive restrictions are reduced, contradicts the attenuation effect by Hambrick et al. (2005) and Hambrick (2007). Thus, the allocation of attention under abundant cognitive capacity provides a crucial test for unpacking the rational versus behavioral mechanisms underlying our theory and extant approaches, respectively.

We use micro-level data on sales and retail managers in two large, unrelated companies to test this novel theory. Our data contains information on managers' allocation of attention across two tasks, and there is variation in attention capacity faced by managers at work, both between managers and, for one of our companies, within managers across time. As detailed below, our results support the model's predictions: when attention is scarce, "managing with

[1] Indeed, an important focus of ABV is to endogenize organization features or structures so that executives and the organization can manage, reduce, and improve upon its cognitive and attention deficiencies.

[2] Recently, Reina, Peterson, and Zhang (2017) empirically find that family-to-work conflicts reduce the decision-making comprehensiveness of CEOs, presumably by reducing their information processing ability. However, they do not relate the reduced breadth in decision making to prior expertise.

style" prevails; however, when attention is abundant, "managing against one's style" emerges.

Overall, while extant research predicts "managing with style" to be prevalent and attention to be "behaviorally" biased towards expertise, our theory and empirical findings suggest, instead, that the allocation of attention can be contingent and optimal. As such, our theory and findings complement the conventional behavioral approach that emphasizes biased managerial attention, and invite further analysis to unpack the underlying mechanisms driving the relationship between cognition/attention, expertise, and firm behavior.[3]

The setup and intuition of the model is as follows. A manager has to allocate her attention across two tasks with equal importance a priori, but she may have different levels of expertise in each. Expertise and attention combine to yield task-specific information, which may or may not be actionable. Specifically, the manager only adapts to task-specific information which is sufficiently precise and informative. Such adaptation can take the form, for example, of a creative initiative that may fail if not sufficiently thought out. In the absence of precise information, the manager prefers to avoid adaptation and sticks to a standard business strategy, such as maintaining "business as usual." Our main predictions follow from this. When attention is scarce, it is optimal to devote attention to the task in which the manager has relatively more expertise, as attention and expertise are complements. Intuitively, devoting scarce attention to a task in which the manager has limited expertise is likely to be a waste of time. This then yields the classic "managing with style" result, where a manager focuses on the task with which she is more familiar. In contrast, when attention is abundant, the manager optimally devotes attention to both tasks, but more so to the task in which she has relatively less expertise. Intuitively, abundant attention then yields actionable information on both tasks, and is allocated in a way to compensate for the lack of expertise in one of the two tasks. In

[3]Outside of management, experimental psychology studies note that, in a variety of social settings, an individual's attention often focuses on the most salient tasks when resources are limited (Mullainathan and Sharif 2013). Karau and Kelly (1992) show that time pressure leads group members to focus on a restricted range of familiar task features. Shah et al. (2012) also show that resource scarcity in the form of poverty leads to an attentional focus on borrowing to spend on time-pressured items such as food and drinks instead of saving for the long term. However, this literature has not studied what choices subjects make when resources are abundant. This precludes a "microfoundational/experimental" comparison between our theory and bounded rationality approaches.

other words, when attention is abundant, attention and expertise are substitutes and we predict a manager to "manage against her style."[4]

The summary of our empirics is as follows. We use two hand-collected data sources that contain micro-level information on managers' allocation of attention across two equally important tasks in their daily work. Our contexts are appealing since individual managers in the sampled companies possess heterogeneous expertise and face different levels of attention capacity, measured by time stress. Our first data source comes from surveying managers working for a major Japanese retail operator ("the Japanese retailer"). These retail managers have to distribute their attention across the equally important selling and servicing tasks in one of the twelve stores in a Tokyo sales region. Our second data source comes from surveying sales managers at the largest beverage producer in Chile ("the Chilean producer"). These sales managers have to allocate their attention to selling the producer's two major product categories: beer and soft drinks. Each of the two companies has homogenous management practices and compensation structures across managers. This helps to mitigate unobserved heterogeneity in technological, macroeconomic, and cultural factors that might affect managerial attention. With these data sets, our paper is in a unique position to test how varying attention capacity affects the relationship between managerial expertise and attentional focus across tasks.

To measure expertise, for the "Japanese retailer," we use a manager's total work experience on each task; for the "Chilean producer," we use total work experience on each product category. To measure attention, each manager in the two companies reports her working hours and self-rated effort intensity on each of the two tasks (for the Japanese retailer) or product categories (for the Chilean producer). To measure attention capacity, managers reported on a Likert scale their overall time stress at work, stemming from demands from work, personal,

[4]The "slack search" literature studies how excess time or resources drives firm change and innovation (Cyert and March 1963; Levinthal and March 1981; Greve 2015; Agrawal et al. 2018). We differentiate from this literature in two ways. In addition to the rationality assumption, we assume two tasks of equal importance for the firm, both of which need to be executed. Adaptation in our theory is within the confines of the activities that are already performed by the firm. Thus, it is about allocation of "managerial" attention. Instead, the "slack search" literature focuses largely on pursuing new activities and strategies, a more "radical" form of adaptation. It goes beyond management and more towards R&D, experimentation, or entrepreneurship.

and family related issues (Schriber and Gutek 1987; Mittal 1994; Andrew and Smith 1999). To match our theoretical model, we construct ratios of the expertise and the attention measures between the two tasks or products. The main analysis examines how the ratio of expertise, and its interaction with time stress, impacts the ratio of allocated attention.

For the Chilean producer, we also obtained sales data for 32 months that permitted a longitudinal data analysis. In this analysis, we use the ratio of sales volumes across product categories as the proxy of allocation of attention. Attention capacity is measured using seasonality for its two main product categories. Specifically, scarce attention is measured by high-temperature months. In those months, a spike in customer demand greatly impacts the beverages sector on its sales and overall activity, increasing the time stress of sales managers as the salesforce capacity remains unchanged during the period.

Our empirical results support the theoretical predictions of our model. Under high time stress, managers focus their attention on the task (in the Japanese retailer) or product category (in the Chilean producer) in which they have relatively more expertise, whereas under low time stress, managers put their attention toward the task or product in which they have relatively less expertise. When not conditioning for time stress, we find that the relative expertise in task or product by managers, has little correlation with these managers' allocation of attention. The longitudinal data on monthly sales volumes of beer and soft drinks corroborates these survey-based results. Specifically, Chilean managers achieve relatively higher sales volumes in beer than in soft drinks when they have relatively more experience in beer, but only during high seasonal months (and vice versa if they have relatively more experience in soft drinks). These results hold when adding manager fixed effects, and thus controlling for their time-invariant unobservable characteristics. All in all, it is reassuring to observe the convergent evidence from two unrelated companies with very different businesses (retailing vs. distribution), products (general merchandise vs. beverages), and local culture (East Asian vs. Latin American).

Our paper has several managerial implications. First, variations in attention capacity are realistic and common in work settings: it varies within individual managers over time as well as between managers at a given time. Second, our finding that, under low time stress, attention is being diverted away from one's expertise domain has implication for job design, multitasking, hiring, among others. On job design, firms may treat the level of time stress as a decision variable, calibrating a "right" level of stress that balances attention across domains and strategies. Same with multitasking, depending on (endogenous) stress level, multitasking may or may not be desirable. In hiring, relative expertise across tasks and domains might matter as much as absolute levels.

2 Institutional context

This subsection describes the institutional context of the two companies from which we collected data and the key features of their managerial practices. Table 1 provides a summary of the following descriptions.

<insert Table 1 here>

2.1 A large Japanese retailer

The first data source is one of the largest retailers in Japan and the world. It operates a portfolio of retail formats such as general merchandising stores (GMS), shopping malls, and convenience stores. Our sample covers all 12 GMS ("stores") that locate in a Tokyo metropolitan area and report to the same regional headquarters. According to the North American Industry Classification System, establishments in the "general merchandise stores" subsector (NAICS code 452) "are unique in that they have the equipment and staff capable of retailing a large variety of goods from a single location." Target, Wal-Mart, Marks and Spencer, and Tesco are examples of companies that operate similar stores in other parts of the world. Two of our sampled stores are located inside shopping malls while the other ten are standalones. The average floor space of the twelve stores is over 20,000 square meters, each employing about 480 employees and catering to over 11,000 daily shoppers. Annual sales per square footage at the time when we collected data (2017) was US$340, which is slightly higher than that of American retailers (US$325).

Each store operates up to 23 departments. Examples of departments are groceries, womenswear, and kid's apparel. A store is managed by the store manager, with over a dozen department managers formally reported to him. Department managers supervise non-managerial workers in their departments. They are broadly skilled; this is not surprising because Japanese firms are known to be more flexible in task assignments and job rotations than U.S. companies (Aoki 1990). Employees tend to have ample opportunities to develop various skills and expertise through their training and work. For example, a merchandising employee may also work on check-out operations and order replenishment; a grocery staff may also be involved in sales operations, and a sales manager may conduct marketing research in addition to her originally assigned work in sales planning and execution. In this way, Japanese employees display variation in their expertise profiles, being less skewed toward an homogeneous set of skills.

According to executive interviews, the company classifies its ten managerial tasks into two categories. The following five "selling tasks" make up the first category: sales (i.e., sales planning and administration), personal selling (i.e., face-to-face selling), merchandising (i.e., assortment and variety), product management (i.e. freshness and hygiene), and pricing (i.e., discounts and incentives). These are clearly important retailing tasks. The second category is referred to as "servicing tasks:" shop floor (i.e., cleaning, display, and decoration), order placement (i.e., product ordering), checkout operations (i.e., cashiers and receipts), training (e.g., on product knowledge and customer service), and personnel management (e.g., on work shifts and hiring part-timers). These servicing tasks are also vital to meet customer demand. For instance, similar to sales and pricing as part of selling tasks, ordering and shop floor arrangement are time sensitive tasks which keep up product availability and shopping ambience. This entails, among many other things, that managers work to eliminate assortment and variety gaps between an initial headquarters' plan and the store reality. Notably, the store manager rarely gives direct instructions on time spent on a particular task, although the person is in charge of the overall operation at the store. In other words, it is up department managers to decide how to allocate their time across the two equally important kind of tasks in their daily work. Our survey and company supplied personnel data include more than 180 department managers;

however, some of them such as IT or personnel managers indicated zero involvement in either one of the two kinds of tasks. Together with missed entries in our questionnaires and personal information, our effective sample size is reduced to 126.

A manager's daily work routine starts with a morning meeting in which the Store Manager and all managers on duty may review various issues such as sales, profits, expenses, promotional activities, customer complaints, market intelligence, hiring, etc. After the daily morning meeting, managers return to their respective floors (or units) to work on the various tasks defined above. In an afternoon meeting, managers are given further specific instructions on merchandising, sales implementations, and customer care. Finally, a portion of the managers participates in a weekly Sunday evening meeting. Standardized documents and spreadsheets are shared before and during these meetings. The key reference or benchmark document is the monthly sales plan that is pre-agreed at the beginning of the month. Managers may also hold ad hoc, small group meetings during working days. As we observed during our store visits, many managers and employees were busy managing their units or shop floors throughout the day. This suggests that managers often have to be selective in their attention allocation. Finally, the regional headquarters impose a uniform compensation scheme and consistent management practices across the twelve stores. Managers' compensation includes a fixed pay (based on experience, seniority, length of service, and qualification) and three possible discretionary bonuses that may sum up to five months' salary. Performance pay is decided by a panel of senior managers using pre-set criteria handed down from the headquarters. Those criteria are related to factors such as achievement of departmental and/or store-wide sales and gross margin targets, and corporate and store missions (e.g., merchandise development, food waste). Importantly, the structures of ex ante incentives in terms of fixed salary and bonuses are identical across stores and departments (see also Lo et al. 2011).

2.2 A large Chilean beverage producer

The second data source is the largest Chilean producer and distributor of beverages ("the Chilean producer"). The company has four product divisions that control product marketing and production in a decentralized way. These four product categories are beer, soft drinks, spirits, and wine, which have the following volume shares, respectively (sales revenue shares in parentheses): 34% (42%), 60% (42%), 4% (10%), and 2% (8%). The sales force of the company consists of approximately 730 salespeople and 110 sales managers. Most of these employees belong to a centralized sales unit, which provides sales services to the four product divisions (i.e., the centralized sales force sells the entire product portfolio of the company). Other administrative units such as finance, human resources, and logistics are also centralized.

In the large metropolitan areas of the country, around 100 salespersons and 15 sales managers work in two independent sales units that are lodged in the spirit and wine divisions.[5] Given that in our analysis, we analyze the allocation of attention across product categories, we do not consider these two, small single-category sales units. This reduces the number of eligible sales managers to 95. Finally, given that the number of clients that have permits to sell alcoholic beverages is significantly smaller than those that have no such permits, there are 29 sales managers that supervise salespersons that only sell soft drinks. Therefore, our effective sample is reduced to 66 sales managers who would face the trade off of their attention across the two product lines. Notably, these managers have the flexibility to allocate their time on product lines during a typical work day. According to company executives and sales data, beer and soft drinks are the two product categories that have the same seasonality and are of the equal importance in terms of sales volumes and revenues. Selling either product category involves similar steps and processes, but the difference lies in knowing the product portfolio of each category and the idiosyncratic demand of specific clients in each one of the categories.

The source of expertise variation across beer and soft drinks comes from the history of the centralized sales service unit. This unit was created in 2005 to cater to the large and high-density regions of Chile by merging the product-specific and independent sales forces that were operating in these regions before 2005. After July 2016, the centralized unit was expanded to the whole country by merging the product specific sales forces that were still operating in the large regions until that date (with the exception, as indicated above, of wine and spirits). As of 2017, roughly half of the sales force in the centralized unit comes from the "2005 wave" and the other half from the "2016 wave." Thus, heterogeneous expertise

[5] In the remaining areas of the country, the centralized sales services unit carries the spirits and wine products. But as mentioned earlier, their sales are minimal.

between supervisors is generated by the different historical background as part of the product-specific sales forces. Further, the merger nature of the waves provides exogenous variation for expertise across the two categories.

The sales services unit is organized as a five-layer hierarchy. Below the Chief Sales Officer, there are four major sales areas (e.g., north or south of Chile), each with its Area Sales Manager. Then come the 18 regions, each with an appointed Regional Sales Manager. The next layer consists of the 30 sales districts, each one with its own office. These sales offices lodge the final two layers: sales managers and salespeople. On average, there are roughly three sales managers per office and 6.5 salespeople per sales manager.

Each sales office has a marketing manager who provides commercial guidelines and serves as a liaison with the marketing areas under the jurisdiction of the four product divisions. In addition, a sales intelligence analyst at each office feeds information to the sales force and takes care of administrative work at the location. The sales manager holds daily sales meetings with her salespeople. This hour-long meeting occurs first thing in the morning and allows the manager to instruct, monitor, coordinate, and communicate with her salespeople. Of the remaining workday, about 45% is devoted to planning-for and coaching-to her salespeople; 35% is devoted to visiting and developing customer accounts; and the balance is split between data analysis, planning, and administrative work such as credit evaluation and equipment maintenance.

At the start of each month, sales managers receive a volume sales goal for each category for that month. This goal is defined at the central office of the sales unit and is a combination of two subgoals: (i) a weighted average of sales in the previous three months (approximately a 50% weight) and the same month last year (the remaining 50%), to clients that are under the managers' jurisdiction (i.e., the clients whom its salespeople serve), (ii) a goal that is obtained from disaggregating across managers the sales' target at the office level and that is defined by the Area Managers based on yearly budgets and other high-level strategic considerations. Managers plan the activities of salespeople based on these monthly volume goals and adjust selling activities accordingly as the month progresses.

The structure of compensation for sales managers is uniform across sales districts and product lines, with approximately 60% of their pay contingent on performance. As indicated above, sales managers have a monthly volume goal for each product category. The ratios of realized sales to sales goal for each category are then added to an overall quota achievement (in percentage terms) using weights that capture the categories' share of volume in the past year. This overall goal achievement is then translated into a sale commission using a sigmoidal function which starts paying at 75% of quota achievement and is capped at 125% of quota achievement. Approximately 70% of the sales force is unionized.

3 A cognitive theory of manager fixed effects

To generate testable hypotheses to our context, we adapt the model of Dessein and Santos (2020).[6] Consider a manager who must divide her attention between two tasks $i \in \{1, 2\}$. In our empirical analysis, a task corresponds to a category of managerial tasks at the Japanese retailer and a product category at the Chilean producer. Profits depend on how well each task i is adapted to a random task-specific shock θ_i with mean 0 and variance 1. In particular, for each task, the manager can either choose:

(i) An adaptive strategy $a_i \in R$ which yields a pay-off

$$K - (1 + \phi)(a_i - \theta_i)^2 \tag{1}$$

where $\phi > 0$.

(ii) A standard strategy $\bar{a}_i$ which yields a pay-off $K - \theta_i^2$.

[6] Our model simplifies the set-up in Dessein and Santos (DS), while maintaining the crucial difference between adaptive and standard strategies. Importantly, our model focuses on comparative statics with respect to attention capacity, whereas DS studies the role of task complexity. Concretely, DS considers a set-up where a manager imperfectly observes shocks θ_1 and θ_2 and communicates this information to a team of agents, who must carry out tasks 1 and 2 in a coordinated way. In the presence of communication and coordination frictions, the manager optimally communicates only about the largest perceived shock to her team. The agents, in turn, choose an adaptive strategy only for the task on which there is communication. As in our model, the manager must decide how to allocate her attention to learn about the shocks θ_1 and θ_2, anticipating that she will only communicate about one shock. Focusing attention on the task in which the manager is an expert is optimal if and only if tasks are sufficiently complex.

Note that the pay-offs of an adaptive strategy depend on how well the manager can observe the shock θ_i. Indeed, if the manager could perfectly observe θ_i, she would always choose an adaptive strategy $a_i = \theta_i$ and obtain a pay-off $K > K - \theta_i^2$. As we show below, the manager will choose an adaptive strategy if and only if her information about θ_i is sufficiently precise and the expected shock is sufficiently large.

In our empirical context of the Japanese retailer and the Chilean producer, a standard strategy can be interpreted as the pre-agreed activities based on the monthly sales plan, whereas an adaptive strategy corresponds to what the manager revises and executes as her selling conditions evolve.

3.1 Information precision, task expertise and managerial attention.

A manager's information precision about θ_i depends on her task expertise and the attention she devotes to task task i. Formally, Let $F(.)$ be the distribution of θ_i and let s_i be the signal the manager receives. We assume that with probability p_i the manager observes the shock, that is $s_i = \theta_i$, and with probability $1 - p_i$ her signal s_i is pure noise with the same distribution $F(.)$. Since $E(\theta_i) = 0$, it follows that

$$E(\theta_i | s_i) = p_i s_i.$$

Following Dessein and Santos (2020), the manager's precision p_i is increasing in the *managerial attention* t_i devoted to task i and her *managerial expertise* T_i in task i. We assume that the manager's knowledge about task i is an additive function of her expertise in task i and the attention devoted to task i

$$\omega_i = T_i + t_i$$

where p_i is an increasing function of $\omega_i : p_i \equiv p(\omega_i)$.[7] Conceptually, we think of a manager specialized in task i as having access to more precise information about the shock pertaining to task i. However, a non-specialist manager can compensate for her lack of expertise in a specific task by devoting more attention to it. For example, she can consult experts, do

[7] In the Appendix, we specify a specific functional form for $p(\omega_i)$ such that $p'(\omega_i) > 0$ is a continuous function, $p''(\omega_i) \leq 0$ (decreasing marginal returns) and $\lim_{\omega_i \to \infty} p(\omega_i) = 1$.

extensive research, or simply devote more time to analyze her options in that particular task as she cannot rely on experience or knowledge.

Rather than imposing a cost for attention, we assume that managerial attention is scarce in that

$$t_1 + t_2 \leq \bar{t}. \tag{2}$$

Our main propositions will deal with the optimal allocation of managerial attention (t_1, t_2) given some total cognitive or attention capacity $\bar{t}$.

3.2 Main Results

The timing in our model is as follows. (1) The manager allocates attention $t_i \in [0, \bar{t}]$ to each task $i = 1, 2$ to learn about (θ_1, θ_2), with $t_1 + t_2 \leq \bar{t}$. (2) Having observed the corresponding signals s_1 and s_2, the manager either chooses a standard strategy $\bar{a}_i$ yielding a pay-off of $K - \theta_i^2$, or an adaptive strategy $a_i \in R$ whose pay-off depends on how well a_i matches θ_i.

In Appendix, we show that the manager will choose an adaptive strategy if and only if $|E(\theta_i|s_i)|$ is above some threshold, where this threshold is increasing in the noisiness of her information. In other words, a key feature of our model is that a manager with noisy information is less likely to chose an adaptive strategy than a manager with more precise information. In turn, the precision of a manager's information depends on how much expertise T_i she has in task i, which is assumed to be exogenous, and how much attention t_i she devotes to task i, which is endogenous.

Our main results, summarized in Proposition 1 below, pertain to the optimal allocation of attention when the manager has an cognitive capacity - or attention - constraint: $t_1 + t_2 \leq \bar{t}$.

First, *when attention is scarce*, the manager devotes all her attention to the task in which she has more expertise. Intuitively, this is the task for which she is more likely to choose an adaptive strategy and, hence, for which more precise information is more valuable. Indeed, if she is not an expert in a particular task, she is unlikely to choose an adaptive strategy, and devoting scarce attention to this task is largely a waste of time. In this way, attention and expertise act as complements.

In contrast, *when attention is abundant*, it will be possible for the manager to learn both shocks θ_1 and θ_2 with great precision. In equilibrium, the manager is then very adaptive to both tasks. While attention is then valuable for both tasks, the manager optimally devotes more attention to the task in which she has less expertise. Intuitively, the returns to attention are then larger on the task in which the manager has less expertise, as attention and expertise are substitutes. In Appendix, we show the following result:

Proposition 1 *Assume $T_2 < T_1 < \tau$. (1) When attention is scarce ($\bar{t}$ is sufficiently small), the manager devotes all her attention to the task in which she has more effective expertise:*

$$(t_1^*, t_2^*) = (\bar{t}, 0).$$

(2) In contrast, when attention is not scarce,($\bar{t}$ is sufficiently large), the manager devotes more attention to the task in which she has less effective expertise:

$$(t_1^*, t_2^*) \textit{ is such that } t_2^* > t_1^*.$$

3.3 Task complexity

To conclude, we investigate the impact of task complexity on managerial attention. To do so, we add an ingredient to our model. The probability p_i with which the manager understands the local shock θ_i not only depends on her expertise T_i and attention t_i, but also on the complexity λ_i of task i.[8] Concretely, we posit that $p_i = p(\omega_i)$ where

$$\omega_i = T_i - \lambda_i + t_i$$

with λ_i is the complexity of task i or, similarly, the uncertainty surrounding task i. We are interested how task complexity λ_i affects the impact of task expertise T_i on the allocation of attention t_i. In Appendix, we show the following result:

Proposition 2 *Consider the impact of λ_i (the complexity of task i) on attention t_i :*

- *If attention is scarce, that is $\bar{t}$ is small, then an increase in λ_i*
  (i) (weakly) reduces t_i if T_i is small.
  (ii) (weakly) increases t_i if T_i is sufficiently large.

[8] A similar assumption is made in Dessein and Santos (2020).

- *If attention is not scarce, $\bar{t}$ is large, then an increase in λ_i always increases t_i.*

In sum, whether or not an increase in the complexity of a task increases the returns to attention depends on how knowledgeable the manager is about that task. If she used to be very knowledgeable at the optimum, then an increase in complexity increases the returns to attention on that task. If the manager used to be not very knowledgeable about a task, then an increase in complexity reduces the returns to attention on that task. The intuition is similar to that for Proposition 1 (Our Appendix provides more discussion).

4 Empirical model and results

4.1 Data and measures

4.1.1 Selection of survey participants and data collection procedure

To investigate the effect of attention capacity and relative expertise on managerial attention, secondary data are unlikely to come by. Instead, we chose to use survey instruments to collect primary data from the Japanese retailer and the Chilean producer. To design our questionnaire at the Japanese retailer, we conducted two rounds of meetings with company executives and managers. Initial meetings involved executives of the strategic planning function of the company's President Office. These meetings and follow-up email exchanges provided information on company organization, managerial practices, and store operations. The company permitted our survey to be conducted at all the twelve stores belonging to a regional sales district in Tokyo. We designed a list of pilot questions and interviewed about twenty managers at two stores, including store managers and department managers. These onsite pilot interviews provided valuable information on managerial tasks and local markets, which in turn helped us to design our questionnaire. The finalized questionnaire copies were distributed to all department managers in the twelve stores (excluding store managers).[9] Managers at each store returned their completed questionnaires in a sealed envelope (printed with one of our universities' name

[9]Since there are only twelve store managers, we excluded their participation in the survey to avoid being identified and hence to maintain the integrity of our data.

and logo) and then put this envelope into a box designated to be used for our survey. In the process, we ensured that the content of each questionnaire remained confidential to company executives who would only receive a store-level overview. All department managers filled out the questionnaire. We removed the managers who indicated zero involvement in either selling or servicing tasks, since the trade-off between tasks in our theory is not applicable to them. Further, due to missing entries, the final sample size becomes 126. To supplement our survey data, the retailer also provided information on managers' education and gender.

We followed a similar process for the Chilean producer. First, we interviewed company executives and four former sales managers of the centralized sales services unit to get a detailed picture of the role of sales managers in this company (e.g., tasks, functions, performance metrics, etc.). The company also shared confidential documents such as guidelines and practices. Then, we designed and pilot-tested the questionnaire. The survey was announced to all the 95 eligible sales managers of the company by their direct superior – the regional sales managers – in the first sales meeting of April 2017. The managers were told that this was part of a research collaboration between the company and one of our coauthors' affiliation at the time. The questionnaire was conducted via email using the software Qualtrics and stayed accessible for responses for three weeks' time. Confidentiality was assured by indicating, verbally by the company manager and on the front page of the online survey, that the individual responses would be handled only by the researchers and that the company would receive only a report with aggregate statistics. To further motivate truthful responses, the sales managers were promised to receive a personalized and private report that provided a comparison to their colleagues and international benchmarks. We obtained responses from 92 sales managers (out of 95). Of the 66 sales managers who sell both the beer and soft drink product categories, three sales managers did not reply the survey and two had missing information in their responses. Therefore, we use 61 sales managers when analyzing the survey data. The company further provided personnel information on the managers such as education, residence, and marital status, as well as transaction data for 32 months on all sales managers, from January 2015 to August 2017. The longitudinal data on sales enable us to corroborate the results based on the survey using within manager variance and seasonality as a time stress shock. Beyond its role of tackling endogeneity concerns, this longitudinal data analysis also studies whether the self-reported behavior on effort translates into objective performance outcomes.

We list our empirical measures and their descriptions in Table 2. While measures such as expertise and working hours are cardinal, other measures such as effort intensity and time stress are reported as scales. Summary statistics of our variables are shown in Table 3. In the following subsection, we begin by describing our measures on the Japanese retailer and then on the Chilean producer.

<insert Tables 2 and 3 here>

4.1.2 Measures: Japanese retailer

We summarize our measures below:

Working hours. This is a self-reported measure of the number of working hours a retail manager spends on her department's (i) selling tasks: personal selling, sales, pricing, product management, and merchandising; and (ii) servicing tasks: human resources, training, floor arrangement, ordering, and check-out operations. To measure relative attention, we use the ratio between the hours spent on selling and servicing tasks. Notice that the ratio equals 1 when the manager evenly allocates her attention across the two kind of tasks. However, the value of the ratio is between (0, 1) when she spends more time in selling than servicing tasks but between (1, ∞) otherwise. In other words, the ratio is *not* symmetric around 1 at which both tasks receive the same attention. To make the value of the ratio symmetric around an equal attention to both tasks, we use the logarithm value of the original ratio in our regression analysis. The log value of the ratio, log(*Working hours ratio*), generates the range of values symmetrically between ($-\infty, \infty$) in which the mid-value of 0 denotes an even allocation of attention.[10]

Effort intensity. This scale measures the intensity of the time and effort that a retail manager devotes to selling and servicing tasks. Similarly to working hours ratio, we use the

[10] We checked whether the regression results would change if we: (i) use the original values of the ratios rather than their log values, and (ii) flip the numerator and denominators in those ratios. The results are robust, both in the Japanese and Chilean contexts. These estimates are available from the authors upon request.

log value of the ratio between the two task categories as a measure of attention allocation in our analysis.

Time stress. Adapting Mittal (1994), Good et al. (1996), and Andrews and Smith (1999), our eight-item scale measures the degree to which respondents are under time pressure at work by combining existing ones such that it is inclusive of work-, personal-, and family-related items. Our scale has a Cronbach's alpha of 0.838 which is in line with the previous studies.[11] We use the mean value of the eight items in our analysis. Notice that we treat time stress as a moderator (and hence a determinant) of managerial attention. Its exogeneity to the latter are justified by two reasons. First, the previous studies treat their original scales of time stress also exogenously as a determinant of work tasks (Andrews and Smith 1999), job attitudes (Good, et al. 1996), or shopping behavior (Mittal 1994). For instance, Andrews and Smith (1999)'s original scale is to examine if respondents are under time stress to complete their marketing programs. Second, our scale is a global assessment of a person's time stress and does not direct toward a domain-specific assessment in terms of attention allocation at work, tasks/products, or expertise. We view the items in this scale as indicators of the respondents's perceived overall time stress due to the multifaceted nature in one's work-personal-family life (Saris and Gallhofer 2014, pp.23-26). In Section 4.2.2 on empirical results below, we discuss - and discard - potential concerns about reverse causality. In particular, one might wonder whether time stress is a consequence rather than a determinant of time allocation.

Expertise is measured by the number of years of work experience and training the retailer manager has in selling and servicing tasks. We also use the log value of the ratio between the two tasks as relative expertise in our analysis. The correlations coefficients of time stress with expertise are all very small and not statistically significant at the 10%-level: selling tasks (0.054) and no-selling tasks (0.032) in our Japanese retailer, and, beer (0.172) and soft drinks (-0.080). These low correlations are indicative that a high level of expertise in a task (product) does not necessarily translate into a high level of overall time stress, and vice versa.

Demand unpredictability. This scale measures the effect of the unpredictability of cus-

[11] Mittal (1994) uses a three-item measure on time stress and reports a construct reliability of 0.76. Andrews and Smith (1999) use a six-item scale. Their scale has a reliability of 0.81. Good et al. (1996) use a 13-item scale to measure work-family tension, with a reliability of 0.85-0.86.

tomer demand on departmental sales and profits.

Competition. Managers rate the intensity of competition in the area where their store is located.

Education and *Gender* are the two pieces of demographic information provided by the company. We include them as control variables.

Summary statistics in Table 3 show that managers working for the Japanese retailer on average put in 4.13 and 4.08 hours of selling and servicing tasks respectively, whereas for *a given manager*, selling tasks on average occupy 26.4% more time than nonselling tasks.[12] On the other hand, the means of effort intensity across the two kinds of tasks engaged by all managers show that selling tasks are slightly less engaged than servicing tasks across managers. For a given manager, however, the effort intensity ratio is 0.98, a number very close to 1. Sampled managers are on average slightly more experienced in selling tasks than servicing tasks; but for a given manager, she is 27.5% more experienced in the former. However, the median values of the three variables all equal to one. This implies that Japanese managers are pretty balanced in their expertise and attention loads when outliers are not considered, which is consistent with the broad skills and work for a typical Japanese manager (Aoki 1990). Regarding time stress, Japanese managers reported a wide range of stress levels, with the mean at about the mid-point (i.e., 4) of the 7-point scale. The average demand unpredictability is below the mid-point of the scale while the intensity of competition is perceived to be medium (mean=3.96).

4.1.3 Measures: Chilean producer

Survey Data The survey measures for our Chilean producer are:

Working hours are the self-reported daily time spent by the sales manager on the beer or soft drink product categories. To measure relative attention, we first take the ratio between the hours spent on beer and soft drinks and then use its log value in our regressions.

---

[12]Notice that the mean of the ratio between the variables $x_1$and x_2 is $\frac{\sum_{i=1}^{n}\frac{x_{1i}}{x_{2i}}}{n}$. In general, this is not equal to the ratio of the means of the two variables, $\frac{\left(\sum_{i=1}^{n}x_{1i}\right)/n}{\left(\sum_{i=1}^{n}x_{2i}\right)/n}$.

Effort allocation measures the percentage of effort expended by the sales manager on selling beer or soft drinks. Like working hours, we use the ratio between the two to measure relative attention. Again, we use its log value in our analysis. [13]

Time stress. The eight-item scale used in the Chilean questionnaire is identical to the one used in the Japanese context. This scale in the Chilean data set has a Cronbach's alpha, or scale reliability, of 0.823.

Expertise measures the number of years of work experience a manager has in selling beer or soft drinks. As a proxy for relative expertise, we use the log value of the ratio between the two products in our empirical analysis. Although the raw ratio of this expertise ratio is somewhat skewed (skewness = 3.21) as shown in Table 3, its log value is not. The log value of the raw ratio has a range of -3.689 to 3.296, with its median and mean being 0 and -0.045 respectively. The very low skewness, -0.096, and the test of normality joint based on skewness (asymmetry) and kurtosis (long tails) (p-value=0.884) confirms the almost symmetric and short-tail nature of the log transformed expertise ratio.

Product-demand unpredictability. This scale measures the perceived unpredictability of consumer demand in the overall market of beer or soft drinks. Demand uncertainty is also view as the complexity of managerial tasks (Campbell 1988). We use the log value of the ratio between the two product categories in our analysis.

Competition is the number of direct, competing beverage producers of beer or soft drink products. We use the log value of the ratio between the two values as relative competitive pressure in our regressions.

Assortment breadth. This scale measures the breadth of SKUs carried by the manager's direct customer accounts. If these customer accounts carry more SKUs, they tend to be more complex and larger retail formats.

We also include *Education* and *Marital status* as control variables in our analysis.

Summary statistics in Table 3 show that sales managers at the Chilean producer spend

[13]Notice that since the managers also spend time on other products (wine and spirits), so the sum of a manager's effort allocation between the two key product lines does not equal 100 percentage points. However, these two products lines capture only 9% of effort, which is consistent with their small share in terms of sales volume (6%) and sales (18%).

on average 5.28 and 4.47 hours, respectively, on beer and soft drinks. For a given manager, the average ratio of working hours of beer to soft drinks is 1.27. Beer also has a higher effort intensity than soft drinks, both in terms of the average intensity of effort across all managers and in terms of the average ratio of beer to soft drinks for a given sales manager. The average experience in selling beer and soft drinks is fairly similar across managers; but for a given manager, the mean of the ratio of experience in beer to experience in soft drinks is 3.21, although its median is 1. Sales managers perceive only a fair level of product-demand unpredictability and competition, possibly due to the producer's leading market position in the country.

Transaction (Longitudinal) Data. In addition to the above survey data, we also obtained *monthly sales volume* data on beer and soft drinks for sales managers between January 2015 and August 2017. These transaction data allow for longitudinal analyses, where we use the (log value of the) ratio between the sales volumes of beer and soft drinks as a (time-varying) measure of the relative managerial attention, and high seasonality months as a (time-varying) measure of time stress.

High season. Beverage sales are highly dependent on weather, especially temperature. In central Chile, where most sales occur, temperatures vary from an average high/low of 30/13 C$^\circ$ in summer to 14/3 C$^\circ$ in winter. This generates strong seasonality in beverage sales. Company executives disclosed that they have an informal but well established notion of "high season" and "low season." The high season typically starts in September when a national holiday generates a big spike in sales in the first half of the month in the early Spring. The high season ends in March when the summer ends. December is the busiest month since it is the start of the summer and covers Christmas and New Year celebrations as well. We also checked for the robustness of our high seasonality window by using the top four months in terms of sales (December, January, February, and September) or the top three months (December, January, and September). Importantly, seasonalities of beer and soft drinks are both driven by the same weather pattern (i.e., temperature and holidays) and hence are identical.

Compared to the average monthly volume sold across the year, the high season has a 14% higher volume and the low season has a 19% lower volume (in our data time frame).

High seasonality months create significant time stress for sales managers because of spiked market demand and hectic operational and marketing activities with customer accounts. At the same time, there is a very limited supply of experienced sales people who are willing to work temporarily in the high season, and thus the salesforce is largely fixed. All these make seasonality a good measure for time stress and attention capacity.

4.2 Empirical results

4.2.1 Econometric Specifications

We use OLS in our analysis on the survey data collected from the Japanese and Chilean companies. Our main regression is specified as:

$$A_i = \alpha_1 E_i + \alpha_2 S_i + \alpha_3 E_i \cdot S_i + \alpha_4 X_i + \epsilon_i, \tag{3}$$

where A_i is the log value of the ratio between the attention spent on the two different tasks or product lines by manager i (as measured by working hours and effort intensity), E_i is the log value of her expertise ratio, S_i is her time stress, and X_i is the vector of control variables including the constant. Recall that our key theoretical predictions are (i) the positive interaction effect between E_i and S_i in (3), or $\alpha_3 > 0,$ and (ii) the negative effect of E_i when S_i is low, or $\alpha_1 < 0.$ In other words, high time stress amplifies the positive impact of relative expertise in tasks (products) on the relative attention to those tasks (products). In contrast, and importantly, low time stress *flips* the impact of relative expertise on the relative attention. For comparison purposes, we omit the interaction term in (3) in some of our regressions.[14]

Our panel data of monthly sales volumes, generated by sales managers at the Chilean producer, also allow for a difference-in-difference approach ("DID"). Based on our theory, we hypothesize a different impact of experience on sales volumes in high seasonality months, when time stress is elevated and thus attention is scarce, when compared to low seasonality months. Denoting $H = \{t | t \text{ is during high season}\}$, we posit that

$$Y_{it} = \hat{a} + \delta_H E_i + I_i + \lambda_H + M_t + \epsilon_{it}, \text{ if } t \in H$$

[14]Hambrick et al. (2005) hypothesize low job demands attentuates - but does not filp - the expertise bias of attention. This implies they predict α_1 and α_3 are both greater than 0.

and

$$Y_{it} = \hat{a} + \delta_L E_i + I_i + \lambda_L + M_t + \epsilon_{it}, \text{ if } t \notin H$$

where Y_{it} is the log value of the ratio of sales volumes of beer to soft drinks, generated by manager i and in month t, E_i is log value of manager i's expertise ratio of beer to soft drinks, a is the constant, and λ_k, I_i, and M_t are season, sales manager, and month fixed effects. This then yields the following continuous difference-in-difference ("DID") model:

$$Y_{it} = a + \delta_L E_i + \lambda h_t + (\delta_H - \delta_L) E_i \cdot h_t + I_i + M_t + \epsilon_{it}, \tag{4}$$

where $a \equiv \hat{a} + \lambda_L$, $\lambda \equiv \lambda_H - \lambda_L$ and h_t equals 1 when month t is in the high season and 0 otherwise,

The "treatment" in regression (4) is “high season” when time stress is elevated for sales managers. The difference between (i) the treatment of high seasonality months and the non-treatment of low seasonality months and between (ii) the managerial expertise ratio, form the basis of the DID estimator $\delta \equiv \delta_H - \delta_L$ (Cameron and Trivedi 2005). Our theoretical model predicts that, during high (low, respectively) seasonality months, a higher expertise ratio of beer to soft drinks correlates with relative more attention to beer (soft drinks) and thus a higher (lower) sales volume ratio of beer to soft drinks. This implies that the interaction term in regression (4) is positive, i.e., $\delta \equiv \delta_H - \delta_L > 0$. We exclude the interaction term $E_i \cdot h_t$ and the two fixed effects in some of our regressions for comparison purposes. Note that in our regressions, the manager fixed effect I_i absorbs the direct effect of expertise $\delta_L E_i$. Similarly, the month fixed effect M_t absorbs the direct effect of high seasonality λh_t.

In what follows, we first discuss our empirical results that use the survey data obtained from the Japanese retailer. Then, we show our results that use the survey and the longitudinal sales data obtained from the Chilean producer.

4.2.2 Results: Japanese retailer

Table 4 shows our results on the Japanese retailer, with log(*Working hours ratio*) in columns 1-3 and log(*Effort intensity ratio*) in columns 4-6 as dependent variables. In the Japanese data, we always put selling to servicing tasks in our ratios. Among the regressors, the interaction term between log(*Expertise ratio*) and *Time stress* is omitted in Columns 1 and 4 for comparison purposes. Columns 3 and 6, our most complete model, include store fixed effects to control for unobserved store-specific factors such as location and management style of the store manager. Using alternative measures for managerial attention as the dependent variable, the two sets of results (columns 1-3 and 4-6) are qualitatively similar.

Without the interaction term between log(expertise ratio) and time stress, more expertise in a task is associated with more attention to that task: $\alpha_1 =$ 0.182 and 0.091 in columns 1 and 4 respectively. This implies that a manager spends more attention on the task in which she has more expertise. More time stress correlates with more attention to servicing tasks, but only the coefficients in column 3 ($\alpha_2 = -0.055$) is significant.

<insert Table 4 here>

When we include the interaction term between log(expertise ratio) and time stress (columns 2-3 and 5-6), the primary effect of the log(expertise ratio) becomes negative. Whereas that of time stress remains negative, its interaction effect with time stress is positive and statistically significant. Notably, the result of $\alpha_1 < 0$ is novel and consistent with what our model predicts. These results imply that when time stress is low, more managerial experience in a task is associated with the manager devoting *less* attention to that task. For example, under low time stress, a manager who is very experienced in selling tasks (i.e., log(*Expertise ratio)* $>>$0) *flips* her attention such that she spends more time on servicing tasks than on selling tasks. When time stress is high, however, more managerial experience in a task is associated with more attention on that task: the positive interaction effect eventually overcomes the negative primary effect and renders the marginal effect of the expertise ratio positive. In other words, when times stress is elevated, a manager who is very experienced in selling tasks focuses more on selling tasks than on servicing tasks. All in all, these empirical results are consistent with our main theoretical predictions.

One might argue that our results may be subject to reverse causality: the overall time stress a manager experiences may be a consequence of the attention allocated to a task rather than a determinant. But if this were true, then a manager who was an expert in selling and devoted most of her attention to selling tasks would experience higher time stress. However, this would be strange, because people tend to be more stressed when they devote effort to a task with which they are *not* familiar. In other words, if there was a reverse causality from attention allocation to time stress, we would expect to see a negative interaction effect between time stress and task expertise (the opposite of what we find).

We use the results in the most complete regression model in column 3 to illustrate the interaction effect in Figure 1. For ease of graphical interpretation, we denote 5th- and 95th-percentile values as low and high levels of the raw expertise ratio (instead of its log value as used in the regression) and time stress in the figure. We assume the mean values of the control variables in our calculations. Under high time stress, the working hours ratio of selling to servicing tasks increases from 0.27 when the manager is an expert in servicing tasks to 1.70 when the manager is an expert in selling tasks. In other words, the manager "manages with style." In contrast, when times stress is low, the working hours ratio of selling to servicing decreases from 2.47 when the manager is an expert in servicing to 0.76 when the manager is an expert in selling. In other words, the manager "manages against style."

$<$insert Figure 1 here$>$

Regarding the control variables in Table 4, the negative coefficients on gender in columns 1-3 are statistically significant. This implies that female managers employed by the Japanese retailer work relatively more on servicing tasks than male managers. Columns 4-5 shows that a more competitive environment displays a statistically significant correlation with more efforts spent on selling rather than servicing tasks. Finally, only four store fixed effects are statistically significant at the 10% level in column 6.

4.2.3 Results: Chilean producer

Survey data In the same format, Table 5 shows the results obtained on the Chilean producer, with log(*Working hours ratio*) and log(*Effort allocation ratio*) as dependent variables in columns 1-3 and 4-6 respectively. For the Chilean producer, we always put beer to soft drinks in our ratio expressions. Among the regressors, with separate product-line information on beer and soft drinks, we are able to construct ratios using product-line specific measures of demand unpredictability and competition.

First, in columns 1 and 4, the estimates of both log(expertise ratio) and time stress are not statistical significant and hence, on average, they do not affect attention allocation. When we include their interaction term in the rest of the columns, time stress remains having little effect on the relative attention spent on beer versus soft drinks. However, like what we see in the previous table, the primary effect of log(expertise ratio) turns to be negative and the interaction effect with time stress is positive, with both effects statistically significant. Again, the novel result of $\alpha_1 < 0$ is as what our theory predicts. In other words, when time stress is low, more managerial experience in a product category is associated with a manager devoting *less* attention that product line. Hence, if a manager is very experienced in selling beer (i.e., log(*Expertise ratio*)$>>$0), then under low time stress, she *flips* her attention to spend more time on selling soft drinks. On the other hand, when time stress is high, she focuses her attention primarily on beer: the positive interaction effect eventually overcomes the negative direct effect and renders the marginal effect of log(expertise ratio) positive. As such, the data on the Chilean producer not only supports our main predictions, but are also qualitatively similar to what we have found in the analysis on our Japanese retailer.

Figure 2 graphically shows the interaction effect obtained from column 3. Again, for ease of graphical interpretation, we denote 5th- and 95th-percentile as low and high levels of time stress and we use the raw expertise ratio (instead of its log value). We assume the mean values of the control variables in our calculations. The visual contrast between high versus low time stress is as pronounced as what we see with the Japanese retailer in Figure 1. Under high time stress, the working hours ratio of beer to soft drinks increases from 0.84 when the manager is a relative expert in soft drinks to 1.85 when she is a relative expert in beer. Thus, the manager "manages with style" under high time stress. In contrast, under low time stress, the working hours ratio of beer to soft drinks decreases from 1.44 when the manager is a relative expert in soft drinks to 0.91 when the she is a relative expert in beer. In other words, under low time stress, the manager *flips* her focal attention to soft drinks when she is very experienced in beer (and vice versa).

$<$insert Table 5 and Figure 2 here$>$

We find *Assortment breadth* to be negatively related to attention to beer. This is understandable. The presence and importance of beer in a client's retail store –and other alcoholic beverages– is negatively correlated with more complex and larger retail formats which carry a wider assortment of products and brands in the soft drink category. There is a relatively larger and more complex presence of soft drinks in these stores and, thus, more relative managerial attention devoted to it (the converse holds for liquor stores). We also notice that log(*Competition ratio)* shows a positive and statistically significant coefficient throughout: it is intuitive that attention increases in a product line's market competition.

Transaction (longitudinal) data Our survey data on the Japanese retailer and the Chilean producer allow for an analysis of managerial attention across managers who differ among each other in terms of expertise (in tasks or product categories) as well as levels of self-reported time stress. Unfortunately, the survey data do not allow us to observe how a given manager shifts her attention over time in response to changes in time stress. To remedy this, the next analysis takes advantage of longitudinal data at the Chilean producer of more than two years worth of monthly sales volumes (in beer and soft drinks) for each of our sales managers.

As explained in subsection 4.1.3, we use high seasonality months and the sales volume ratio of beer to soft drinks as *time varying measures* of, respectively, time stress and managerial attention. We keep years of experience in each product category as the measure of expertise. Our definition of the duration of high seasonality months ranges from long to brief: 7, 4, or 3 months. The company conventionally views the 7-month duration as the high season, but, as a robustness check, we also operationalize it by the 3 or 4 months when the highest monthly company sales occur. The spike in customer demand during high seasonality months is due to climate and holiday patterns, but the salesforce capacity is kept largely constant. This creates a period of stressful working conditions for existing sales managers at the company. We use log values of ratios in this panel data analysis. Our key findings are qualitatively similar among the three alternative definitions of the high versus low seasonality months. Using the longer and shorter definitions of high seasonality, Table 6 summarizes those results respectively.

Column 1 of Table 6 uses the model of "pooled cross sections over time" (Wooldridge 2010, p.146) in which we omit manager fixed effect and the interaction term of log(expertise ratio) and high season. Columns 2 to 5 are our DID analysis in which we introduce the interaction term. In the last three columns (4-6), we sequentially introduce monthly fixed effects, manager fixed effects, and both types of fixed effects. All models use robust standard errors which are further clustered at the manager level. Notice that coefficients of high season and those of non-time-varying variables such as log(expertise ratio) are absent when, respectively, monthly and manager fixed effects are included. All columns in the table show a consistent result on our key prediction: each regression model generates a significant, positive coefficient on the interaction term. The strongest statistical result on the interaction term is obtained in columns 4 and 5 where manager fixed effects are included.

$<$insert Table 6 here$>$

Since both the "pooled" model in column 1 and the DID models in columns 2 and 3 do not include manager fixed effects, they only take advantage of between-subject variations in the analysis. While these columns do not exploit the within-subject variations of the transaction data, by using alternative, transaction-based measures for time stress and managerial attention, they provide a robustness check for the survey-based measures in our analysis above. The panel-data results corroborate the survey results shown above in Table 5.

Columns 1-3 show a tiny (positive, non-significant) correlation between log(expertise ratio) and log(sales volume ratio). However, the former result is somewhat different from that in Table 5. In combination with a positive interaction term, this implies that as in the survey results, in high seasonality months (when time stress is high), a higher expertise ratio of beer to soft drinks is associated with achieving a higher sales volume of beer to soft drinks; however, in low seasonality months (when time stress is low), a high expertise ratio of beer to soft drinks has almost no correlation with the relative sales volume of the two product lines. The disappearance of the negative, individual effect of log(expertise ratio) is consistent with our theory: sales is generated both *by attention and experience*. While on the one hand more experience in beer should result in less attention to beer when time stress is low (and thus lower beer sales), on the other hand more experience in beer generates, by itself, more sales in beer. Therefore, the combination of both effects should not result in a relative lower sales volume in beer.

The DID specification discussed in columns 4 and 5 exploits within-subject variations of our panel data by including manager fixed effects. Consistent with our theory and previous results, the interaction effect between log(*Expertise ratio)* and *High Season* is again positive and highly significant. Since manager fixed effects are included, these results show that managers effectively shift their attention during the course of the year as a function of their relative expertise in beer and soft drinks. In particular, a given manager tends to shift more of her attention to beer during high seasonality months – that is, she achieves a higher beer to soft drinks volume ratio – if she has a higher experience ratio of beer to soft drinks. Conversely, during low seasonality months, a manager who is more experienced in beer will shift more of her attention to soft drinks, and a manager who is more experienced in soft drinks will shift more of her attention to beer. Column 4 omits the month fixed effects, and uncovers a positive and significant primary effect of the variable *High season*, similar to columns 2 and 3.

Table A1 of our online Appendix replicates the exercise but operationalizes high seasonality as 4- and 3- month durations respectively in columns 1-5 and 6-10. The results of pooled cross section time varying model in columns 1-3 and the DID models (omitting manager fixed effects) in 6-8 are qualitatively similar to those in the previous table. In our DID analysis with manager fixed effects in columns 4-5 and 9-10, results are similar as well. The slighter smaller coefficients in these columns as compared to columns 4 and 5 of Table 6 are expected: by excluding two or three months of high demand (perceived as such by the salesforce), it becomes harder to find an effect.

All in all, our data sets from the Japanese retailer and Chilean producer, with both survey and transaction (longitudinal) data for the latter company, provide converging evidence for our key hypothesis about the interaction between expertise and time stress on the allocation of managerial attention.

Impact of (product line) complexity on attention Having studied the impact of time stress on attentional focus, we now turn our analysis to a related but different aspect of our theory: the effect of task complexity (the parameter λ_i in our model). While it is infeasible to measure task-specific complexity in the Japanese context, we use product-demand unpredictability in the beer and soft drink categories to gauge task complexity at the Chilean producer.

Unpredictability of demand is an important dimension of task complexity (Campbell 1988), particularly in a sales environment (Lo et al. 2011).

According to our theory, whether or not an increase in task complexity increases the returns to attention depends on how knowledgeable the manager is about a task. If she has substantial task knowledge, then an increase in task complexity increases her returns to attention on that task. In contrast, if she is not very knowledgeable, an increase in task complexity decreases her returns to attention. In the context of our Chilean producer, highly unpredictable customer demand will make the managerial tasks of sales planning and management more complex to handle. Our theory predicts that a more unpredictable demand for beer will induce a sales manager who is a beer expert to devote more attention to beer (in order to better resolve this task complexity). In contrast, if the sales manager is not very knowledgeable in beer, then an increase in product-line unpredictability which makes demand information about beer harder to capture and interpret, may instead cause her to shift her attention towards soft drinks (which are easier for her to understand). Intuitively, an increase in the product-demand unpredictability –that is, task complexity– of beer makes it less likely that attention to beer will generate actionable information for a manager who has a poor understanding of the beer category to start with. In sum, as shown in Proposition 2, our model predicts the same interaction effect between expertise and product-demand unpredictability on attentional focus as between expertise and time stress.

We show our results in Table 7. The table uses the same format as in Tables 4 and 5 but treats *Time stress* as a control variable in the last columns of 3 and 6. First, excluding the interaction term between expertise and product-demand unpredictability ratios (columns 1 and 4), the average effects of both the expertise ratio and product-demand unpredictability are not statistically significant. When the interaction term is included in columns 2 and 5, the primary effects of both log(*Expertise ratio*) remains not significant. Nonetheless, the interaction effects in both models are positive as predicted by our model. Including *Time stress* and sales-region fixed effects in the most complete models in columns 3 and 6 show similar results. Although these results do not show a flip in attention, product-demand unpredictability do attenuates the manager's expertise in her attention allocation. Intuitively, both time stress and task complexity increase the returns at the margin to attention for experienced managers, but decrease the returns to attention for inexperienced managers.

<insert Table 7 here>

4.2.4 Additional Analysis

The main hypothesis of our theory focuses on the positive interaction effect between relative expertise and time stress. For a given ratio of expertise, nonetheless, the absolute value of expertise can be very different. For example, a manager with ten-year experience in each task has the same expertise ratio as another manager with only one year experience in each of the two task. One may wonder if our main results continue to hold if the *absolute value* of expertise - instead of its ratio - is used to interact with time stress. Tables A2, A3, and A4 of our online Appendix show the results of this additional analysis.

The first two rows in Table A2 show that, under low time stress, managers at the Japanese retailer who have more expertise in selling, or less expertise in servicing tasks, correlate with more working hours in and effort allocated to selling tasks. Per the positive and negative signs of its interaction terms with selling and servicing expertise in rows 4 and 5 respectively, managers put forth more time and effort in selling tasks but less so in servicing tasks as time stress increases. Table A3 show the survey results for the sales managers working for our Chilean producer. Although the coefficients of beer expertise and its interaction with time stress (rows 1 and 4) are not statistically significant, those of soft drink expertise (rows 2 and 5) are consistent with our previous results: sales managers who are experienced in soft drinks diversify their working hours and allocate effort to beer under low time stress, but shift to beer under high time stress. In columns 1, 3, and 4 in Table A4, our Chilean longitudinal data show that sales volume shifts from beer to soft drinks only when sales managers are more experienced in soft drinks in high seasonality months (i.e., high time stress). Column 2 shows that managers with more experience in beer also generates relatively more sales volume in beer, but their sales volume also shift toward more soft drinks during high seasons (i.e., high times stress). In sum, raw values in expertise generate similar patterns in both the Japanese and Chilean companies with the ones using expertise ratios when they are interacting with time stress.

5 Conclusion

Scholars in management and economics have long recognized the notion of "managing with style" in which manager fixed effects such as expertise and experience determine managerial attention and strategic decisions (Bertrand and Schoar 2003; Finkelstein et al. 2009). Our theory predicts that "managing with style," or not, depends on how scarce managerial attention is. Micro-evidence on sales and retail managers working for two unrelated, large companies shows converging support for our theory. On the one hand, when time stress is high, managers do manage with style by focusing their attention on the tasks or product line in which they have more expertise. On the other hand, we obtain a novel finding under attention abundance: managers move their attention toward the tasks or product line in which they have less expertise – put simply, they "manage against style." We obtain these results both in a cross section of managers facing different attention scarcity as well as, for one of our companies, in longitudinal data where managers experience time-varying attention capacity and shift their attention across tasks in response to those constraints.

Our findings have important managerial implications for organizational attention when managers engage in multiple tasks. While it is common for companies to direct managerial attention to a particular task by exerting a sense of urgency or time pressure, our analysis shows that this is effective only when their employees have relatively more expertise in the intended, focal task. If managers are relatively more experienced in a non-focal task, imposing time pressure may backfire since scarcity in cognitive resources will shift their attention away from the intended task. To direct attention to the task in which employees have relatively less expertise, companies should instead provide a working environment that is less time pressured. Our analysis provides strong evidence that less time stress helps diversify managerial attention to the relatively inexperienced task.[15] As such, one could examine the effect of attention

[15]In our model, the manager and the firm share the same objective. As such, the allocation of managerial attention is always optimal. It would be easy, however, to introduce an incentive conflict between the manager and the firm. For example, managers may be reluctant to adapt to a shock unless they have very precise information.

capacity when it is used as a job-design variable. For instance, companies such as Google builds time buffers for their workforce (Schrage 2013). Providing time abundance may also be helpful to explore what is outside one's existing expertise and hence facilitate, for instance, product innovation. Last but not the least, while the common practice of company hiring is to focus on the absolute level of expertise in pre-defined task areas, relative expertise across areas may have unintended consequences on managerial attention and hence should be factored in. We believe such potential extensions to other areas of organization and managerial studies are a fruitful avenue for future research.

References

[1] Agrawal, Ajay, Christian Catalini, Avi Goldfarb, and Hong Luo. 2018. Slack time and innovation. *Organization Science* 29 (6): 1056-1073

[2] Andrews, Jonlee, and Daniel C. Smith. 1999. In Search of the Marketing Imagination: Factors Affecting the Creativity of Marketing Programs for Mature Products. *Journal of Marketing Research* 33(May): 174-87.

[3] Aoki, Masahiko.1990. *Information, Incentives, and Bargaining in the Japanese Economy*. Cambridge University Press.

[4] Barker, Vincent, and George Mueller. 2002. CEO Characteristics and Firm R&D Spending. *Management Science* 48(6): 782-801.

[5] Bertrand, Marianne, and Antoinette Schoar. 2003. Managing with Style: The Effect of Managers on Firm Policies. *Quarterly Journal of Economics* 118(4): 1169-1208.

[6] Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics*. New York, NY: Cambridge University Press

This would result in managers being *too* focused on tasks in which they have more expertise whenever attention is scarce. In our model, this would correspond to $\phi_m > \phi$ where ϕ_m is the adaptation cost internalized by the manager.

[7] Campbell, Donald J. 1988. Task complexity: A review and analysis. *Academy of Management Review* 13(1) (1988): 40-52.

[8] Carpenter, Mason A., Marta A. Geletkanycz, and W. Gerard Sanders. 2004. Upper Echelons Research Revisited: Antecedents, Elements, and Consequences of Top Management Team Composition. *Journal of Management*, 30(6): 749-778.

[9] Cho, T.S., and Hambrick, D.C. 2006. Attention as the Mediator Between Top Management Team Characteristics and Strategic Change: The Case of Airline Deregulation. *Organization Science* 17(4): 453–469.

[10] Cyert, Richard M., and James G. March. 1963. *A Behavioral Theory of the Firm.* Englewood, Cliffs, NJ: Prentice Hall.

[11] Dessein, Wouter, Andrea Galeotti, and Tano Santos. 2016. Rational Inattention and Organizational Focus. *American Economic Review* 106 (6): 1522-1536.

[12] Dessein, Wouter, and Tano Santos. 2020. Managerial Style and Attention. *American Economic Journal: Micro-economics* Forthcoming.

[13] Eggers, J.P., and Sarah Kaplan. 2009. Cognition and Renewal: Comparing CEO and Organizational Effects on Incumbent Adaptation to Technical Change. *Organization Science* 20(2): 461–477.

[14] Finkelstein, Sydney, Donald C. Hambrick, and Albert A. Cannella. 2009. *Strategic leadership: Theory and Research on Executives, Top Management Teams, and Boards*. New York, NY: Oxford University Press.

[15] Hambrick, Donald. 2007. Upper Echelons Theory: An update. *Academy of Management Review* 32(2): 334-343

[16] 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.

[17] Hambrick, Donald C., Sydney Finkelstein, and Ann C. Mooney. 2005. Executive Job Demands: New Insights for Explaining Strategic Decisions and Leader Behaviors. *Academy of Management Review* 30(3): 472-491.

[18] Good, Linda K., Thomas J. Page Jr., and Clifford E. Young. 1996. Assessing Hierarchical Differences in Job-Related Attitudes and Turnover Among Retail Managers. *Journal of the Academy of Marketing Science* 24(2): 148-56.

[19] Greve, Henrich. 2015. Investment and the Behavioral Theory of the Firm: Evidence from Shipbuilding. *Industrial and Corporate Change* 12 (5): 1051-1076.

[20] Karau, Steven J., and Janice R. Kelly. 1992. The Effects of Time Scarcity and Time Abundance on Group Performance Quality and Interaction Process. *Journal of Experimental Social Psychology* 28: 542-571.

[21] Kaplan, S. 2008. Cognition, Capabilities and Incentives: Assessing Firm Response to the Fiber-Optic Revolution. *Academy of Management Journal* 51(4): 672–694.

[22] Kaplan, Steven N., Mark M. Klebanov, and Morten Sorensen. 2012. Which CEO characteristics and abilities matter? *The Journal of Finance* 67(3): 973-1007.

[23] Kilduff, Martin, Reinhard Angelmar, and Ajay Mehra. 2000. Top Management-Team Diversity and Firm Performance: Examining the Role of Cognitions. *Organization Science* 11(1): 21-34.

[24] Levinthal, Daniel, and James March JG. 1981. A model of adaptive organizational search. *Journal of Economic Behaviour and Organization* 2(4):307–333.

[25] Levy, O. 2005. The Influence of Top Management Team Attention Patterns on Global Strategic Posture of Firms. *Journal of Organizational Behavior* 26(7): 797–819.

[26] Lo, Desmond, Mrinal Ghosh, and Francine Lafontaine. 2011. The Incentive and Selection Roles of Sales Force Compensation Contracts. *Journal of Marketing Research* 48(August): 781–98.

[27] Mittal, Banwari. 1994. An Integrated Framework for Relating: Diverse Consumer Characteristics to Supermarket Couple Redemption. *Journal of Marketing Research* 31(4): 533-44.

[28] Mullainathan, Sendil, and Eldar Sharif. 2013. *Scarcity: Why Having Too Little Means So Much*. Allen Lane Publishers.

[29] Nadkarni, S., and Barr, P.S. 2008. Environmental Context, Managerial Cognition, and Strategic Action: An Integrated View. *Strategic Management Journal* 29(13): 1395–1427.

[30] Neely, Brett H. Jr., Jeffrey B. Lovelace, Amanda P. Cowen, and Nathan J. Hiller. 2020. Metacritiques of Upper Echelons Theory: Verdicts and Recommendations for Future Research. *Journal of Management* 46(6): 1029-1062.

[31] Ocasio, William. 1997 Towards an attention-based view of the firm. *Strategic Management Journal* 18(S1): 187-206.

[32] Ocasio, William. 2011. Attention to Attention. *Organization Science* 22(5):1286-1296

[33] Ocasio, William, and John Joseph. 2005. An Attention-Based Theory of Strategy Formulation: Linking Micro and Macroperspectives in Strategy Processes. *Advances in Strategic Management* 22(18): 39-61

[34] Reina, Christopher S., Suzanne J. Peterson, Zhen Zhang. 2017. Adverse Effects of CEO Family-to-Work Conflict on Firm Performance. *Organization Science* 28(2):228-243.

[35] Saris, Willem E., and Irmtraud N. Gallhofer. 2014. *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. 2nd Edition. Hoboken, NJ: John Wiley & Sons.

[36] Schriber, Jacquelyn, B. and Barbara A. Gutek. 1987. Some Time Dimensions of Work: Measurement of an Underlying Aspect of Organization Culture. *Journal of Applied Psychology* 72(4): 642-650.

[37] Shah, Anuj K., Sendhil Mullainathan, and Eldar Shafir. 2012. Some Consequences of Having Too Little. *Science* 338: 682-685.

[38] Schrage, M. 2013. Just How Valuable is Google's "20% Time"? *Harvard Business Review*.

[39] Simon, Herbert A. 1947. *Administrative Behavior*. New York, NY: Free Press.

[40] Simon, Herbert A. 1971. Designing Organizations for an Information-Rich World. M. Greenberger, ed. *Communications, Computers, and the Public Interest*. Baltimore, MD: Johns Hopkins Press.

[41] Wooldridge, Jeffrey. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

[42] Zhu, David H., and Guoli Chen. 2015. CEO Narcissism and the Impact of Prior Board Experience on Corporate Strategy. *Administrative Science Quarterly* 60(1): 31-65.

6 Appendix: Theoretical Analysis

In this Appendix, we derive our main theoretical propositions, Propositions 1 and 2. For completeness, we first briefly restate the main assumptions of our model.

6.1 Model

A manager must divide her attention between two tasks $i \in \{1,2\}$. For each task i, the manager can either choose:
(i) An adaptive strategy $a_i \in R$ which yields a pay-off $K - (1+\phi)(a_i - \theta_i)^2$, where $\phi > 0$ and θ_i is a random task-specific shock with mean 0 and variance 1.
(ii) A standard strategy $\bar{a}_i$ which yields a pay-off $K - \theta_i^2$.

Let $F(.)$ be the distribution of θ_i and let s_i be the signal the manager receives. A manager's information precision about θ_i will depend on the *managerial attention* t_i devoted to task i, which is endogenous, and her *managerial expertise* T_i in task i, which is exogenous. Rather than imposing a cost for attention, we assume that managerial attention is scarce in that $t_1 + t_2 \leq \bar{t}$. Our main propositions will deal with the optimal allocation of managerial attention (t_1, t_2) given some attention capacity $\bar{t}$ and expertise T_1 and T_2.

Denote the manager's knowledge by $\omega_i \equiv T_i + t_i$, then with probability $p_i \equiv p(\omega_i)$ the manager observes the shock, that is $s_i = \theta_i$, and with probability $1 - p_i$ her signal s_i is pure noise with the same distribution $F(.)$. We assume that $p'(\omega_i) > 0$ and continuous, $p''(\omega_i) \leq 0$ (weakly decreasing marginal returns) and $\lim_{\omega_i \to \infty} p(\omega_i) = 1$.

To make the analysis tractable, we specify the following functional form which satisfies these properties: there exists a $\tau > 0$ such that

$$p(\omega_i) = \omega_i \quad \text{if } \omega_i \leq \tau$$

and

$$p(\omega_i) = 1 - (1-\tau) * e^{-\dfrac{\omega_i - \tau}{1-\tau}} \quad \text{if } \omega_i \geq \tau.$$

Note that there are constant returns to attention for $\omega_i < \tau$, but decreasing marginal returns to attention for $\omega_i > \tau$. Intuitively, beyond some point, it becomes increasingly difficult to further increase the precision of one's signal. Figure 3 illustrates $p(\omega_i)$ when $\tau = 0.5$.

<insert Figure 3 here>

6.2 Optimal action choice

We start our analysis by characterizing decision-making for each task $i \in \{1, 2\}$, given a signal s_i and a precision of information p_i. In the next section, we endogenize this information precision p_i by studying the optimal allocation of scarce attention. Indeed, as discussed above, the precision p_i of the manager's information is a function of the attention t_i devoted to task i and his expertise T_i in task i : $p_i = p(\omega_i)$ with $\omega_i = T_i + t_i$.

The manager must choose whether to choose a standard strategy $\bar{a}_i$ which ignores her local information or an adaptive strategy a_i which she optimally sets equal to the expected state of nature, $a_i = E(\theta_i|s_i)$.

Given a signal s_i and given the pay-offs of an adaptive and a standardized strategy, the manager prefers to ignore her signal and implement a standard strategy $\bar{a}_i$ if and only if

$$E(\theta_i^2|s_i) < (1 + \phi) * E[(\theta_i - E(\theta_i|s_i))^2|s_i].$$

As $E(\theta_i|s_i) = p_i s_i$, this is equivalent to

$$p_i s_i^2 + (1 - p_i)\sigma_\theta^2 < (1 + \phi)\left((1 - p_i)(\sigma_\theta^2 + (p_i s_i)^2) + p_i((1 - p_i)s_i)^2\right).$$

It follows that the manager will choose a standard strategy whenever

$$[1 - (1 + \phi)(1 - p_i)]\, p_i s_i^2 < \phi(1 - p_i)\sigma_\theta^2.$$

Hence, if $(1 + \phi)(1 - p_i) > 1$, the manager always chooses a standard strategy. In contrast, if $(1 + \phi)(1 - p_i) < 1$, the manager will choose a standard strategy whenever

$$|s_i| < \bar{s}_i \equiv \bar{s}(p_i) = \sqrt{\frac{\phi}{1 - (1 + \phi)(1 - p_i)} \frac{(1 - p_i)}{p_i} \sigma_\theta^2} \qquad (5)$$

and an adaptive strategy $a_i = p_i s_i$ whenever $s_i^2 > \bar{s}_i^2$.

Recall, finally, that the (unconditional) distribution of s_i is the same as the distribution of θ_i, and this regardless of p_i. Since one can verify that $\bar{s}(p_i)$ is decreasing in p_i, it follows that the probability that the manager chooses an adaptive strategy is increasing in the precision p_i of her signal. We summarize as follows:

Lemma 1 (Action choices) Given a signal s_i with precision p_i about local shock θ_i :

- If $p_i < \phi/(1+\phi)$, the manager always chooses a standard strategy $a_i = \bar{a}_i$.
- If $p_i > \phi/(1+\phi)$, the manager chooses an adaptive strategy $a_i = E(\theta_i) = p_i s_i$ whenever she perceives the shock to be large, that is $|s_i| > \bar{s}_i$, and a standard strategy otherwise.
- $\bar{s}_i$ is decreasing in p_i : the probability that the manager chooses an adaptive strategy for task i is increasing in the precision of her information p_i about θ_i.

6.3 Optimal allocation of attention

Having characterized decision-making for each task $i \in \{1, 2\}$, given a signal s_i and precision of information p_i, we now study the optimal allocation of scarce *managerial attention.* Thus, given a constraint $t_1 + t_2 \leq \bar{t}$ and given initial expertise T_1 and T_2 about each task $i \in \{1, 2\}$, how does the manager optimally allocate her attention t_1 and t_2?

6.3.1 Expected profits

We first derive expected profits for a given precision (p_1, p_2) of the manager's information on task 1 and 2. If $p_i < \phi/(1+\phi)$, the manager always chooses a standard strategy and expected pay-offs for task i equal $K - \sigma_\theta^2$. Assume therefore that $p_i > \phi/(1+\phi)$. From Proposition 6.2, a standard strategy $\bar{a}_i$ is implemented whenever $s_i \in [-\bar{s}_i, \bar{s}_i]$. For a given s_i and p_i, this yields an expected task payoff of

$$\pi^S(p_i, s_i) = K - E(\theta_i^2 | s_i).$$

In contrast, an adaptive strategy $a_i = p_i s_i$ is implemented whenever $|s_i| \geq \bar{s}_i$. For a given s_i and p_i, this yields an expected task payoff of

$$\pi^T(p_i, s_i) = K - (1+\phi) * E\left[(\theta_i - p_i s_i)^2 | s_i\right].$$

Note that regardless of p_i, the distribution of s_i is given by $F(.)$. Hence, for a given information precision p_i, expected task payoffs equal

$$\Pi(p_i) = K - \sigma_\theta^2 + 2\int_{\bar{s}_i}^{+\infty} \left(\pi^T(p_i, s_i) - \pi^S(p_i, s_i)\right) dF(s_i).$$

Lemma 2 (Expected Profits) Given information precision p_1 and p_2, expected pay-offs are given by $\Pi(p_1) + \Pi(p_2)$ where

$$\Pi(p_i) = K - \sigma_\theta^2 + 2\int_{\bar{s}(p_i)}^{+\infty} \left[(1-(1+\phi)(1-p_i))\, p_i s_i^2 - \phi(1-p_i)\sigma_\theta^2\right] dF(s_i)$$

if $p_i > \phi/(1+\phi)$ and

$$\Pi(p_i) = K - \sigma_\theta^2 \qquad \text{if } p_i < \phi/(1+\phi).$$

Proof See Subsection 6.3.4.

The firm allocates attention t_1 and t_2 in order to maximize the above payoffs subject to the constraint $t_1 + t_2 \leq \bar{t}$.

6.3.2 Optimal allocation of attention

Above we have characterized task payoffs $\Pi(p_i)$, as a function of the precision of the manager's information about θ_i. This information precision, p_i, in turn, depends on the manager's (exogenous) expertise T_i, the complexity/uncertainty λ_i surrounding task i and, crucially, her (endogenous) attention t_i to task i.

We now characterize the optimal allocation of managerial attention t_i, given the constraint $t_1 + t_2 \leq \bar{t}$. Consider first the marginal returns to attention on each task.

Obviously, if $p_i < \phi/(1+\phi)$ then $d\Pi(p_i)/dt_i = 0$ as the manager always chooses a non-adaptive action. Assume therefore that $p_i > \phi/(1+\phi)$ so that the manager sometimes chooses an adaptive action.

Lemma 3 (marginal returns to attention) Assume $p_i > \phi/(1+\phi)$. Then the marginal return to attention, t_i, on task i is given by

$$\frac{\partial \Pi(p_i)}{\partial t_i} = \frac{\partial p_i}{\partial t_i} \cdot \frac{\partial \Pi(p_i)}{\partial p_i}$$

where

$$\frac{\partial \Pi(p_i)}{\partial p_i} = 2 \int_{\bar{s}(p_i)}^{+\infty} \phi \left[\left(2 \frac{1+\phi}{\phi} p_i - 1 \right) s^2 + \sigma_\theta^2 \right] dF(s). \tag{6}$$

Proof See Subsection 6.3.4.

The central question in our model is whether the marginal return to attention on task i is increasing or decreasing in the manager's knowledge p_i on task i. Thus, we want to know the sign of

$$\frac{\partial}{\partial p_i} \left(\frac{\partial \Pi(p_i)}{\partial t_i} \right) = \frac{\partial}{\partial p_i} \left(\frac{\partial p_i}{\partial t_i} \cdot \frac{\partial \Pi(p_i)}{\partial p_i} \right).$$

If the above expression is *positive*, then the returns to attention are higher on the task in which the manager has more expertise T_i or to which the manager already devotes more attention. Indeed, a manager's knowledge p_i depends on her expertise T_i in task i and her attention t_i. If the above expression is *negative*, then it is optimal, on the margin, to devote more attention to a task in which the manager is less knowledgeable (p_i is lower). The following trade-off arises:

On the one hand, the more knowledgeable a manager is about task i, the more valuable it is to further increase her knowledge:

$$\frac{\partial}{\partial p_i} \left(\frac{\partial \Pi(p_i)}{\partial p_i} \right) > 0.$$

It follows that all else equal, it is more valuable to increase knowledge on task 1 when $p_1 > p_2$. There are two reasons for this. First, on the extensive margin, the manager is more likely to adapt to her information when she is an expert: $\bar{s}(p_i)$ is decreasing in p_i and the manager adapts if and only if $s_i > \bar{s}(p_i)$. In turn, this makes more precise information about θ_i more valuable when p_i is high. Second, conditional on $s_i > \bar{s}_i$ and choosing an adaptive strategy, a manager is more sensitive to her information s_i when she has more precise information (p_i is higher). Indeed, we have that $a_i = p_i s_i$. As a result, the marginal returns to improving the quality of that information (increasing p_i) are again larger for task 1 than task 2 when $p_1 > p_2$.

On the other hand, for $T_i + t_i > \tau$, the more knowledgeable the manager is about task i, the more attention is requires to further improve her knowledge:

$$\frac{\partial}{\partial p_i}\left(\frac{\partial p_i}{\partial t_i}\right) < 0.$$

In other words, there are decreasing marginal returns to attention as far as knowledge production itself is concerned.

Taken together, we obtain that

$$\frac{\partial}{\partial p_i}\left(\frac{\partial \Pi(p_i)}{\partial t_i}\right) = \underbrace{\frac{\partial p_i}{\partial t_i}\cdot\frac{\partial}{\partial p_i}\left(\frac{\partial \Pi(p_i)}{\partial p_i}\right)}_{>0} + \underbrace{\frac{\partial}{\partial p_i}\left(\frac{\partial p_i}{\partial t_i}\right)\cdot\frac{\partial \Pi(p_i)}{\partial p_i}}_{\leq 0}$$

where the first term is positive and the second term is negative. When attention is sufficiently scarce, the first term dominates: on the margin, the manager wants to devote attention to the task in which she is already more knowledgeable (task 1). In contrast, when attention is not scarce at all, that is $\bar{t}$ is sufficiently large, the second term dominates. On the margin, it is then more valuable to devote attention to the task in which she is less knowledgeable (task 2).

Proposition 1 *Assume $T_2 < T_1 < \tau$. (1) When attention is scarce ($\bar{t}$ is sufficiently small), the manager devotes all her attention to the task in which she has more effective expertise:*

$$(t_1^*, t_2^*) = (\bar{t}, 0).$$

(2) In contrast, when attention is not scarce,($\bar{t}$ is sufficiently large), the manager devotes more attention to the task in which she has less effective expertise:

$$(t_1^*, t_2^*) \text{ is such that } t_2^* > t_1^*.$$

Proof See Subsection 6.3.4

We summarize the intuition for Proposition 1 as follows. When attention is scarce, the manager is more likely to choose an adaptive strategy for the task in which she is an expert. As a result, the marginal returns to improving the quality of her information on her task of expertise are higher as well. In contrast, when attention is abundant, the manager has the capacity to devote a lot of attention to both tasks. At the optimum, she then chooses to become knowledgeable in both and, hence, is likely to choose an adaptive strategy for both. Since attention and expertise are substitutes, the marginal returns to attention are then higher on the task in which she has less expertise.

6.3.3 Task complexity

To conclude, we investigate the impact of task complexity on managerial attention. To do so, we add an ingredient to our model. The probability p_i with which the manager understands the local shock θ_i not only depends on her expertise T_i and attention t_i, but also on the complexity λ_i of task i.[16] Concretely, we posit that $p_i = p(\omega_i)$ where

$$\omega_i = T_i - \lambda_i + t_i$$

with λ_i is the complexity of task i or, similarly, the uncertainty surrounding task i. We are interested how task complexity λ_i affects the impact of task expertise T_i on the allocation of attention t_i.

Proposition 2 *Consider the impact of* λ_i *(the complexity of task* i*) on attention* t_i *:*

- *If attention is scarce, that is* $\bar{t}$ *is small, then an increase in* λ_i
  (i) (weakly) reduces t_i *if* T_i *is small* $(T_i < \tau)$.
  (ii) (weakly) increases t_i *if* T_i *is sufficiently large.*

- *If attention is not scarce,* $\bar{t}$ *is large, then an increase in* λ_i *always increases* t_i.

Proof See Subsection 6.3.4.

The intuition is similar to that for Proposition 1. When the manager's expertise in, say task 1, is limited, then an increase in task complexity may result in a shift of scarce attention from task 1 to task 2. Intuitively, there are then increasing returns to knowledge and the manager focuses all her attention on the task which she understands best. While she may

[16] A similar assumption is made in Dessein and Santos (2020).

be more of an expert in task 1, the complexity of task 1 may induce her to focus on task 2, which is less complex. Formally, whenever $T_1 - \lambda_1 < T_2 - \lambda_2$, the manager will focus her attention on task 2, even though $T_1 > T_2$.

In contrast, when the manager's expertise on task 1 is very extensive, the returns to focus even more attention on task 1 are low. She may, therefore, want to focus her scarce attention on task 2, in which she is less knowledgeable. An increase in the complexity of task 1, however, may change this calculation and induce her to refocus on task 1, Intuitively, such added complexity increases the returns of attention on task 1. Similarly, if attention is not scarce at all, the manager focuses extensively on both tasks. An increase in complexity of a particular task, say task 1, then always increases the marginal returns of devoting attention to this task.

6.3.4 Proofs lemma and propositions

Proof of Lemma 2 (Expected Profits):

We derive expected profits for a given precision (p_1, p_2) of the manager's information on task 1 and 2. We have that

$$\Pi(p_i) = K - \sigma_\theta^2 + 2\int_{\bar{s}(p_i)}^{+\infty} \left[\pi^T(p_i, s_i) - \pi^S(p_i, s_i)\right] dF(s_i)$$

where

$$\begin{aligned} \pi^S(p_i, s_i) &= K - E(\theta_i^2 | s_i) \\ &= K - p_i s_i^2 - (1 - p_i)\sigma_\theta^2 \end{aligned}$$

and

$$\begin{aligned} \pi^T(p_i, s_i) &= K - (1+\phi) * E\left[(\theta_i - E(\theta_i|s_i))^2 | s_i\right] \\ &= K - (1+\phi)\left[(1-p_i)(\sigma_\theta^2 + p_i^2 s_i^2) + p_i(1-p_i)^2 s_i^2\right] \\ &= K - (1+\phi)(1-p_i)\left(\sigma_\theta^2 + p_i s_i^2\right) \end{aligned}$$

It follows that

$$\Pi(p_i) \equiv K - \sigma_\theta^2 + 2\int_{\bar{s}(p_i)}^{+\infty} \left[p_i s_i^2 + (1-p_i)\sigma_\theta^2 - (1+\phi)(1-p_i)\left(\sigma_\theta^2 + p_i s_i^2\right)\right] dF(s_i)$$

which can be simplified as

$$\Pi(p_i) = K - \sigma_\theta^2 + 2\int_{\bar{s}(p_i)}^{+\infty} \left[(1-(1+\phi)(1-p_i))\, p_i s_i^2 - \phi(1-p_i)\sigma_\theta^2\right] dF\left(s_i\right)$$

QED.

Proof of Lemma 3 (Marginal Returns to Attention): Following Liebniz rule, we have that

$$\begin{aligned} \frac{\partial \Pi(p_i)}{\partial p_i} &= -2\frac{\partial \bar{s}(p_i)}{\partial p_i}\left[(1-(1+\phi)(1-p_i))\, p_1 \bar{s}_i^2 - \phi(1-p_i)\sigma_\theta^2\right] f(\bar{s}_i) \\ &\quad +2\int_{\bar{s}(p_i)}^{+\infty} \frac{\partial}{\partial p_i}\left[(1-(1+\phi)(1-p_i))\, p_i s^2 - \phi(1-p_i)\sigma_\theta^2\right] dF\left(s\right) \end{aligned}$$

Substituting $\bar{s}_i$, the first term disappears. Hence, we have that

$$\begin{aligned} \frac{\partial \Pi(p_i)}{\partial p_i} &= 2\int_{\bar{s}(p_i)}^{+\infty} \left[(1-(1+\phi)(1-2p_i))\, s^2 + \phi\sigma_\theta^2\right] dF\left(s\right) \\ &= 2\int_{\bar{s}(p_i)}^{+\infty} \phi\left[\left(2\tfrac{1+\phi}{\phi}p_i - 1\right) s^2 + \sigma_\theta^2\right] dF\left(s\right) \end{aligned}$$

QED.

Proof of Proposition 1. We first proof Part (1). Assume attention is scarce, that is $T_1 + \bar{t} < \tau$. Then $T_i + t_i < \tau$ for $i = 1, 2$ and

$$\frac{d\Pi(p_i)}{dt_i} = \frac{\partial p_i}{\partial t_i}\frac{\partial \Pi(p_i)}{\partial p_i} = \frac{\partial \Pi(p_i)}{\partial p_i}$$

Recall that $T_1 > T_2$ (the manager is an expert in task 1). Hence, whenever $t_1 \geq t_2$, then also $p(\omega_1) > p_2(\omega_2)$ and from (5), $\bar{s}(p_1) < \bar{s}(p_2)$. It follows from (5) that, whenever $t_1 \geq t_2 \geq 0$,

$$\frac{d\Pi(p(\omega_1))}{dt_1} > \frac{d\Pi(p_2(\omega_2))}{dt_2}$$

In turn, this implies that at the optimum $(t_1, t_2) = (\bar{t}, 0)$.

Part (2) of Proposition (1) follows directly from Lemma A1 and A2, proven below. Assume that $T_1 > T_2$ and $\bar{t}$ is sufficiently large. Lemma A1 shows that there then always exists an $x > 0$ such that $(t_1, t_2) = (\frac{\bar{t}}{2} - x, \frac{\bar{t}}{2} + x)$ is strictly preferred over $(\frac{\bar{t}}{2}, \frac{\bar{t}}{2})$. Thus, when $\bar{t}$ is sufficiently large, the manager prefers to devote more attention to task 2, in which she has less expertise, rather than splitting attention evenly. We subsequently show, in Lemma A2, that for $\bar{t}$ large, $(\frac{\bar{t}}{2}, \frac{\bar{t}}{2})$ is preferred over $(\frac{\bar{t}}{2}+x, \frac{\bar{t}}{2}-x)$ for any $x \in (0, \bar{t}/2]$: when attention is not scarce, the manager never wants to bias her attention towards the task she is an expert in. It follows that at the optimum, (t_1, t_2) much be such that $t_2 > t_1$.

Lemma A1 *Assume $\bar{t}$ is sufficiently large and $T_1 > T_2$. Then there always exists an $x > 0$ such that $(t_1, t_2) = (\frac{\bar{t}}{2}-x, \frac{\bar{t}}{2}+x)$ is strictly preferred over $(\frac{\bar{t}}{2}, \frac{\bar{t}}{2})$.*

Proof: We have that

$$\frac{\partial \Pi(p_i)}{\partial t_i} = \frac{\partial p_i}{\partial t_i}\frac{\partial \Pi(p_i)}{\partial p_i} = e^{-\dfrac{\omega_i - \tau}{1-\tau}} 2\int_{\bar{s}(p_i)}^{+\infty} \left[(1-(1+\phi)(1-2p_i))\, s^2 + \phi\sigma_\theta^2\right] dF(s)$$

And thus

$$\begin{aligned}\frac{\partial^2 \Pi(p_i)}{\partial T_i \partial t_i} \quad &= \quad -\frac{1}{1-\tau} e^{-\dfrac{\omega_i - \tau}{1-\tau}} 2\int_{\bar{s}(p_i)}^{+\infty} \left[(1+(1+\phi)(2p_i-1))\, s^2 + \phi\sigma_\theta^2\right] dF(s) \\ &\quad + e^{-\dfrac{\omega_i - \tau}{1-\tau}} \left(-\frac{\partial \bar{s}(p_i)}{\partial T_i}\right) 2\left[(1-(1+\phi)(1-2p_i))\, \bar{s}_i^2 + \phi\sigma_\theta^2\right] f(\bar{s}_i) \\ &\quad + e^{-\dfrac{\omega_i - \tau}{1-\tau}} \left(\frac{\partial p_i}{\partial T_i}\right) 2\int_{\bar{s}(p_i)}^{+\infty} (1+\phi) 2s^2 dF(s)\end{aligned}$$

Multiplying both sides by

$$\frac{1}{1-p_i} = \frac{1}{1-\tau}\left(e^{-\dfrac{\omega_i - \tau}{1-\tau}}\right)^{-1} > 0,$$

we obtain that

$$\begin{aligned}\frac{1-\tau}{1-p_i}\frac{\partial^2 \Pi(p_i)}{\partial T_i \partial t_i} \quad &= \quad -\frac{1}{(1-\tau)^2} 2\int_{\bar{s}(p_i)}^{+\infty} \left[(1+(1+\phi)(2p_i-1))\, s^2 + \phi\sigma_\theta^2\right] dF(s) \\ &\quad + \frac{1}{1-\tau}\left(-\frac{\partial \bar{s}(p_i)}{\partial T_i}\right) 2\left[(1-(1+\phi)(1-2p_i))\, \bar{s}_i^2 + \phi\sigma_\theta^2\right] f(\bar{s}_i) \\ &\quad + \frac{1}{1-\tau}\left(\frac{\partial p_i}{\partial T_i}\right) 2\int_{\bar{s}(p_i)}^{+\infty} (1+\phi) 2s^2 dF(s)\end{aligned}$$

For ω_i sufficiently large, the first term on the RHS converges to

$$-\frac{1}{(1-\tau)^2}\int_0^{+\infty} \left[(1+(1+\phi)(2p_i-1))\, s^2 + \phi\sigma_\theta^2\right] dF(s)$$

At the same time, for ω_i sufficiently large, we have that

$$\frac{\partial \bar{s}(p_i)}{\partial T_i} \approx 0 \text{ and } \frac{\partial p_i}{\partial T_i} \approx 0$$

so that the the second and third term on the RHS become arbitrarily small. Hence, for ω_i sufficiently large

$$\frac{\partial^2 \Pi(p_i)}{\partial T_i \partial t_i} \approx -\frac{1-p_i}{(1-\tau)^2} \int_{\bar{s}(p_i)}^{+\infty} \left[(1+(1+\phi)(2p_i-1))\, s^2 + \phi \sigma_\theta^2\right] dF(s) < 0$$

It follows that for $\omega_i = T_i + t_i$ sufficiently large, we have that

$$\frac{\partial^2 \Pi(p(\omega_i))}{\partial T_i \partial t_i} < 0$$

Since $T_1 > T_2$, it follows that for $\bar{t}$ sufficiently large

$$\frac{\partial \Pi(p(\omega_1))}{\partial t_1} |t_1 = \bar{t}/2 < \frac{\partial \Pi(p(\omega_2))}{\partial t} |t_2 = \bar{t}/2 \tag{7}$$

Hence, there exists an $x > 0$ such that $(t_1, t_2) = (\frac{\bar{t}}{2} - x, \frac{\bar{t}}{2} + x)$ is strictly preferred over $(\frac{\bar{t}}{2}, \frac{\bar{t}}{2})$. QED

Lemma A2 *Assume $\bar{t}$ is sufficiently large and $T_1 > T_2$. Then for any $x \in (0, \bar{t}/2]$, we have that $(t_1, t_2) = (\bar{t}/2, \bar{t}/2)$ is preferred over $(t_1, t_2) = (\bar{t}/2 + x, \bar{t}/2 - x)$.*

Proof: To show the above result, it will be sufficient to show that, for $\bar{t}$ sufficiently large

$$\frac{\partial \Pi(t_2, T_2)}{\partial t_2} |t_2 = t_2' > \frac{\partial \Pi(t_2, T_2)}{\partial t_2} |t_2 = \bar{t}/2 \text{ for all } t_2' < \bar{t}/2. \tag{8}$$

and

$$\frac{\partial \Pi(t_1, T_1)}{\partial t_1} |t_1 = t_1' < \frac{\partial \Pi(t_1, T_1)}{\partial t_1} |t_1 = \bar{t}/2 \text{ for all } t_1' > \bar{t}/2. \tag{9}$$

Together with (7), inequalities (7) and (7) imply that for any $x \in (0, \bar{t}/2]$,

$$\frac{\partial \Pi(t_2, T_2)}{\partial t_2} |t_2 = \bar{t}/2 - x > \frac{\partial \Pi(t_1, T_1)}{\partial t_1} |t_1 = \bar{t}/2 + x$$

A direct implication is that attention allocation $(t_1, t_2) = (\bar{t}/2, \bar{t}/2)$ is preferred over any attention allocation $(t_1, t_2) = (\bar{t}/2 + x, \bar{t}/2 - x)$. QED

Proof of Proposition 2

Consider first the case where attention is not scarce, then from the proof of Lemma A2, we know that for ω_i sufficiently large, we have that

$$\frac{\partial^2 \Pi(p(\omega_i))}{\partial T_i \partial t_i} < 0$$

Since $p_i = p(\omega_i)$ with $\omega_i = T_i - \lambda_i + t_i$, it follows that also

$$\frac{\partial^2 \Pi(p(\omega_i))}{\partial \lambda_i \partial t_i} > 0$$

At the optimum, if $\bar{t}$ is large, then also ω_1 and ω_2 large, with

$$\frac{\partial^2 \Pi(p(\omega_1))}{\partial t_1} = \frac{\partial^2 \Pi(p(\omega_2))}{\partial t_2}$$

and for both t_1 and t_2

$$\frac{\partial^2 \Pi(p(\omega_i))}{\partial \lambda_i \partial t_i} > 0$$

It follows that an increase in λ_i will result in an increase in t_i and a decrease in t_{-i}.

Consider next the case where attention is scarce ($\bar{t}$ is small). It T_i is sufficiently large, then we have again that

$$\frac{\partial^2 \Pi(p(\omega_i))}{\partial T_i \partial t_i} < 0 \text{ and } \frac{\partial^2 \Pi(p(\omega_i))}{\partial \lambda_i \partial t_i} > 0$$

so that an increase in λ_i increases the marginal returns to increase t_i. Hence, t_i is weakly increasing in T_i. For example, for T_1 large, an increase in λ_1 may result in a shift from $(t_1, t_2) = (0, \bar{t})$ to $(t_1, t_2) = (\bar{t}, 0)$ as an increase in complexity of task 1 increases the returns to devote attention to task 1. On the other hand, if $T_i + \bar{t} < \tau$, then

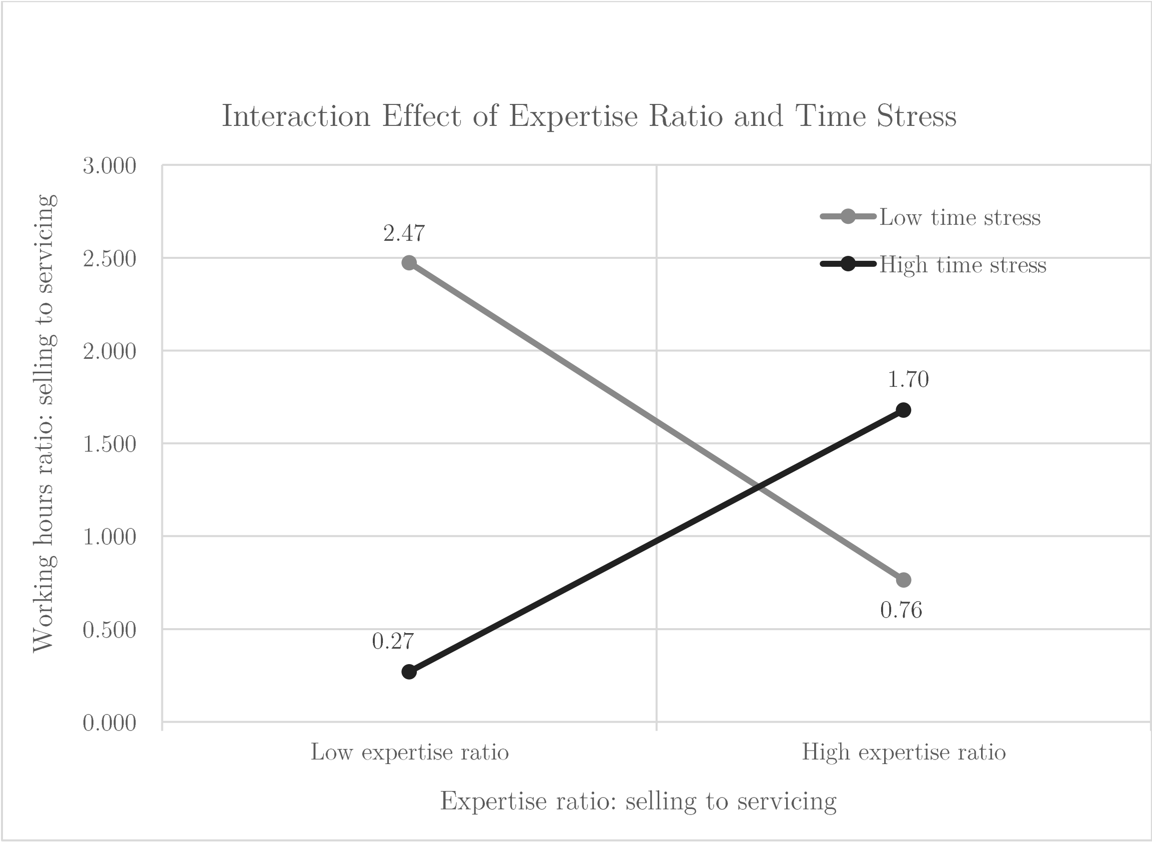
$$\frac{\partial^2 \Pi(p(\omega_i))}{\partial T_i \partial t_i} > 0$$

and thus also

$$\frac{\partial^2 \Pi(p(\omega_i))}{\partial \lambda_i \partial t_i} > 0$$

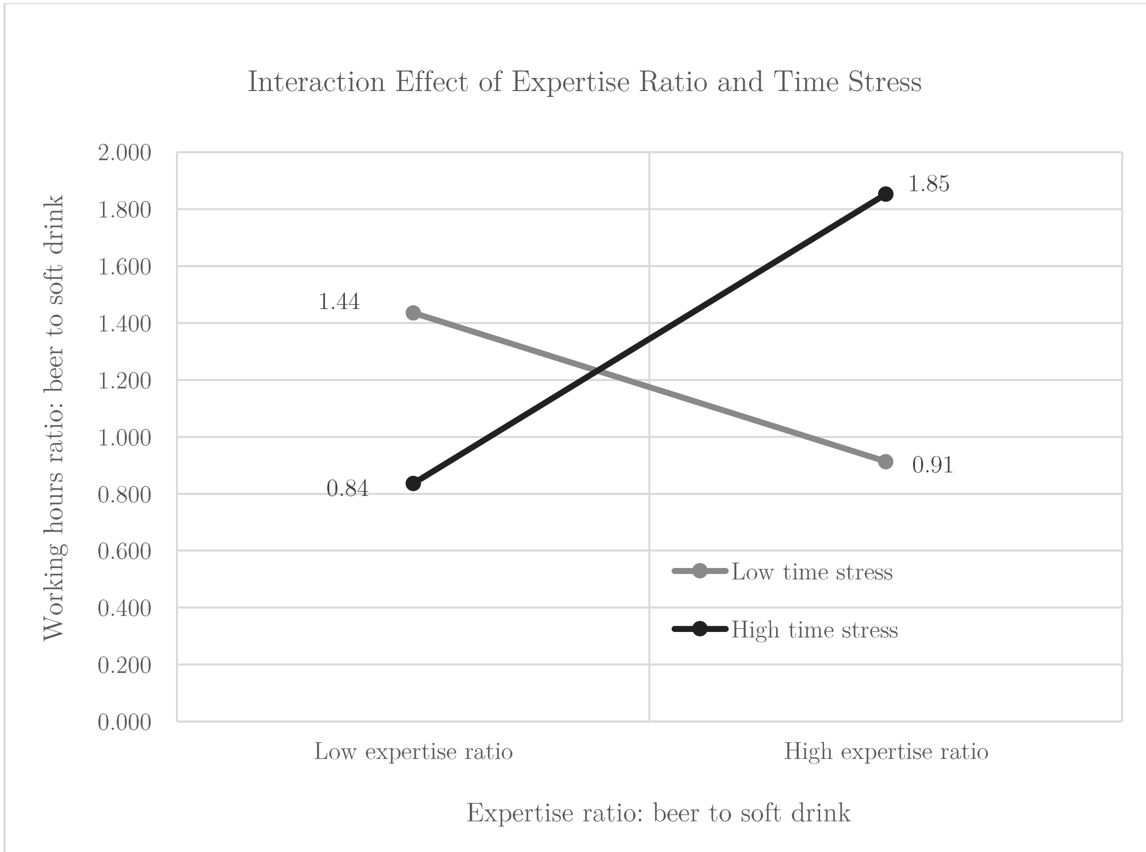
In equilibrium, then $(t_1, t_2) = (\bar{t}, 0)$ whenever $T_1 - \lambda_1 > T_2 - \lambda_2$ and $(t_1, t_2) = (0, \bar{t})$ if $T_1 - \lambda_1 < T_2 - \lambda_2$. It follows that an increase λ_i (weakly) decreases t_i. QED

Figure 1. Japanese Retailer: Interaction Effect of Expertise Ratio and Time Stress



Note: The figure uses estimates obtained from Model 3 in Table 3. For ease of interpretation, we use raw ratios in the y-axis. Low and High values of *Time Stress* are at its 5^{th}-percentile (=2.00) and 95^{th}-percentile (=5.63) respectively. Low and High values of *Expertise Ratio* are at its 5^{th}-percentile (=0.33) and 95^{th}-percentile (=1.89) respectively. Notice that we use the log value of expertise ratio in the regression. Control variables are set at their mean values.

Figure 2. Chilean Producer: Interaction Effect of Expertise Ratio and Time Stress



Note: The figure uses estimates obtained from Model 3 in Table 4. For ease of interpretation, we use raw ratios in the y-axis. Low and High values of *Time Stress* are taken at its 5$^{\text{th}}$-percentile (=1.13) and 95$^{\text{th}}$-percentile (=4.75) respectively. Low and High values of *Expertise Rati*o are taken at its 5th-percentile (=0.05) and 95th-percentile (=22) respectively. Notice that we use the log value of expertise ratio in the regression. Control variables are set at their mean values.

Figure 3. Manager's Precision as a Function of ω when $\tau = 0.5$.

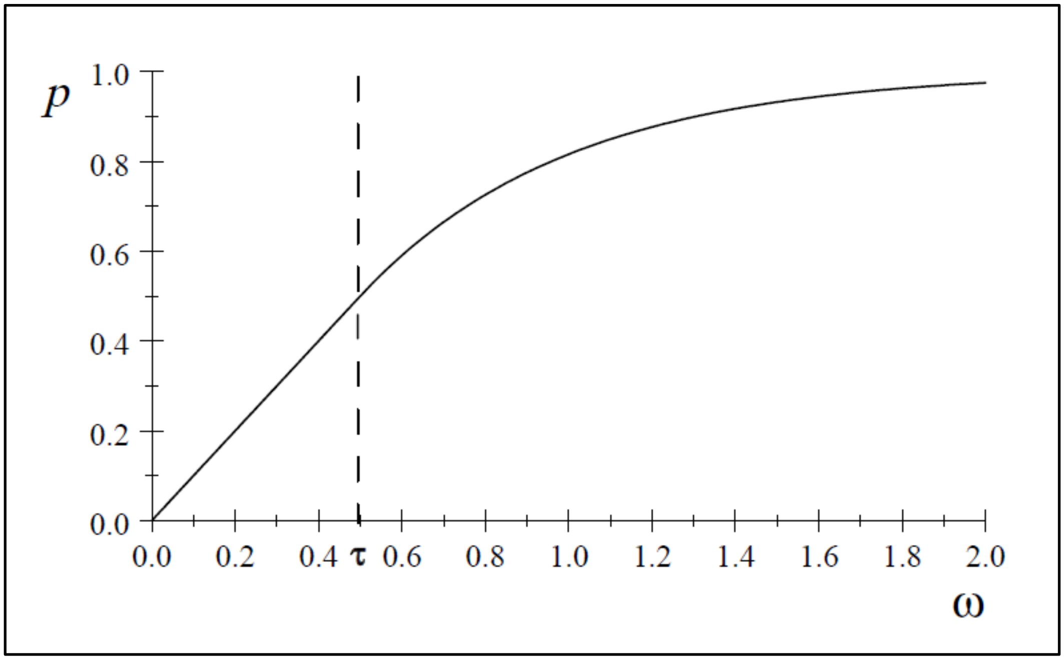


Table 1 Institutional Contexts

	Japanese Retailer	Chilean Producer
Main business	Retailing, in Japan and other Asian countries	Beverage manufacturing and distribution, in Chile and other South American countries
Founding of corporation	Founded over 250 years ago; modern firm established in 1920's	Founded 170 years ago; modern firm established in 1900's
Total number of employees (year 2017)	Over 500,000	Over 8200
Sampled subsidiary	General merchandise stores, belonging to a sales region in Tokyo, Japan	Centralized sales services units in Chile
Managerial staff in our samples	The department manager, who manages non-managerial staff in a department and reports to one of the store managers	The sales manager, who supervises sales people and reports to one of the regional sales managers
Tasks / Products needed managerial attention	2 department tasks: Selling and servicing tasks, both are essential in daily operations	2 product lines: Beer and soft drinks, the two most important product categories
Discretion on allocation of attention	The department manager has full discretions to allocate time and effort across tasks in daily operations	The sales manager has full flexibility to allocate time and effort across the product categories in daily operations
Compensation structure	Same structure across all sampled stores: fixed pay, and up to 3 discretionary bonus (up to five months fixed pay). One bonus relates to monthly revenue and gross margin targets but only for a department that generate revenue	Same structure across all sampled sales regions and product lines: fixed pay (~40% of total pay), and performance pay (~60% of total pay) based on monthly sales volume targets

Table 2. Descriptions of Variables and Measures

Variable	Measure
Japan Retailer	
Working hours in selling tasks	Number of hours in a typical day you spend on the following tasks: personal selling, sales, pricing, product management, and merchandising.
Working hours in servicing tasks	Number of hours in a typical day you spend on the following tasks: human resources, training, floor arrangement, order operation, and checkout operation.
Effort intensity in selling tasks (1-7 scale)	Intensity of effort and time you have to work with your superior(s), subordinates, and other departments in the following tasks: personal selling, sales, pricing, product management, and merchandising.
Effort intensity in servicing tasks (1-7 scale)	Intensity of work and effort you have to work with your superior(s), subordinates, and other departments in the following tasks: human resources, training, floor arrangement, order operation, and checkout operation.
Time stress (8 items; 1-7 scale) [We use the mean value] Cronbach's alpha = 0.838	The degree to which you believe yourself to be under pressure to complete your responsibilities at the store. (1 Completely Disagree – Completely Agree 7) 1. I often have to find extra work time to get my tasks done. 2. I feel like I am always "fighting fire." 3. Personal concerns often reduce my productivity at work. 4. Family issues and matters often cause loss of time for me. 5. I often feel nervous, tense, or frustrated when I get home from work. 6. I am too busy to relax. 7. I am often juggling my time between too many things. 8. I never have time to think and plan ahead.
Expertise in selling tasks [We divide the total number of years by five]	Total number of years your experience and training in the following tasks: personal selling, sales, pricing, product management, and merchandising.
Expertise in servicing tasks [We divide the total number of years by five]	Total number of years your experience and training in the following tasks: human resources, training, floor arrangement, order operation, and checkout operation.
Demand unpredictability (1-7 scale)	The effect of the unpredictability of local, proximate customers' sales and profit to your unit (1 Very Stable – Very Volatile 7).
Competition	Many of our competitors in the area offer similar products and services as our store. (1 Completely Disagree – Completely Agree 7)
Education	1 middle school; 2 high school; 3 technical school; 4 university; 5. master's degree
Gender	0 male; 1 female

Chilean Producer	
Working hours in selling beer	Average number of working hours per day in the last 6 months you spent on all activities to sell the beer product category.
Working hours in selling soft drinks	Average number of working hours per day in the last 6 months you spent on all activities to sell the soft drink product category.
Effort intensity in selling beer (percentage points)	Percentage of time and effort in the last 6 months you spent on all activities to sell the beer product category.
Effort intensity in selling soft drinks (percentage points)	Percentage of time and effort in the last 6 months you spent on all activities to sell the soft drink product category.
Time stress (8 items; 1-7 scale) [We use the mean value] Cronbach's alpha = 0.823	The degree to which you believe yourself to be under pressure to complete your work tasks. (1 Completely Disagree – Completely Agree 7) 1. I often have to find extra work time to get my tasks done. 2. I feel like I am always "fighting fire." 3. Personal concerns often reduce my productivity at work. 4. Family issues and matters often cause loss of time for me. 5. I often feel nervous, tense, or frustrated when I get home from work. 6. I am too busy to relax. 7. I am often juggling my time between too many things. 8. I never have time to think and plan ahead.
Expertise in selling beer	Total number of years in selling or market the beer product category.
Expertise in selling soft drinks	Total number of years in selling or market the soft drink product category.
Product-demand unpredictability – beer (1-7 scale)	Consumer demand of the beer product category in the overall market is very unpredictable. (1 Completely Disagree – Completely Agree 7)
Product-line uncertainty – soft drinks (1-7 scale)	Consumer demand of the soft drink product category in the overall market is very unpredictable. (1 Completely Disagree – Completely Agree 7)
Competition - beer	Number of direct competing manufacturers in beer.
Competition - soft drinks	Number of direct competing manufacturers in soft drinks.
Customer assortment	Your direct customers typically carry a large product assortment from various brands and manufacturers and resell to consumers. (1 Completely Disagree – Completely Agree 7)
Education (1-14)	1: elementary school; 14: completely university. Other numbers indicate intermediate levels of formal education.
Marital status	0 single; 1 married
Sales volume - beer	Total volume of sales in terms of hectoliter for the month in the beer category.
Sales volume - soft drinks	Total volume of sales in terms of hectoliter for the month in the soft drink category.

Table 3. Summary Statistics: Retype: sample size changed

Japan Retailer (N=126)					Chilean Producer (N=61)				
	Task	**Mean; Median**	**Std. Dev.**	**Range**		**Product**	**Mean; Median**	**Std. Dev.**	**Range**
Working hours	**Selling tasks**	4.134; 4	1.708	(0.5,9)	**Working hours**	**Beer**	5.279; 5	1.349	(2,8)
	Servicing tasks	4.084; 4	1.496	(1, 8.33)		**Soft drinks**	4.467; 4	1.399	(2,8)
	Ratio	1.264; 1	0.981	(0.09,7)		**Ratio**	1.268; 1.200	0.504	(0.56,4)
Effort intensity	**Selling tasks (1-7)**	3.781; 4	1.003	(1,5.8)	**Effort allocation (percentage points)**	**Beer**	50.623; 50	9.256	(30,70)
	Servicing tasks (1-7)	3.932; 4	0.993	(1,6)		**Soft drinks**	41.066; 40	8.996	(30,60)
	Ratio	0.977; 1	0.315	(0.25,3)		**Ratio**	1.305; 1.316	0.414	(0.6,2.3)
Time stress (8 items; 1-7)		4.011 4.125	1.026	(1,6.125)	**Time stress (8 items; 1-7)**		2.609; 2.375	1.052	(1,5.625)
Experience (years)	**Selling task**	11.357; 10.567	7.444	(0.367,39)	**Experience (years)**	**Beer**	14.323; 12	10.472	(0.75,35)
	Servicing tasks	10.943; 10.000	6..863	(0.35,30)		**Soft drinks**	13.270; 10	9.788	(1,40)
	Ratio	1.275; 1	1.813	(1,20.57)		**Ratio**[#]	3.206; 1	6.159	(0.025,27)
Demand unpredictability (1-7)					**Product-demand unpredictability (1-7)**	**Beer**	3.016; 2	1.821	(1,7)
		3.056; 3	1.248	(1,6)		**Soft drinks**	3.525; 3	1.689	(1,7)
						Ratio	0.930; 1	0.417	(0.14,2)
Competition		3.960; 4	1.317	(1,7)	**Competition**	**Beer**	3.541; 3	2.270	(1,10))

(1-7)						**Soft drinks**	2.492; 2	1.748	(1,8)
						Ratio	1.661; 1.333	0.977	(0.5,5)
					Client product assortment		6.098; 6	1.028	(3,7)
Education		3.373; 4	0.994	(1,5)	**Education**		8.197; 10	5.498	(1,14)
Gender (0=male)		0.238; 0	0.428	(0,1)	**Marital status**		0.509; 1	0.504	(0,1)
					Sales volume* **(hectoliter)**	**Beer**	4,131; 3,540	2,498	(496, 16,456)
						Soft drinks	4,221; 3,636	2,301	(1,198, 18,044)
						Ratio	1.070; 1.016	0.562	(0.082, 3.468)

[#]We use log value of this expertise ratio in our regressions. The raw ratio of this expertise ratio is somewhat skewed. Its skewness is 3.206 and the joint test of skewness (asymmetry) and kurtosis (long tails) (p-value=0.000) rejects the null hypothesis of normality. However, its log value is not. The log value of the raw ratio has a range of -3.689 to 3.296, with its median and mean being 0 and -0.045 respectively. The very low skewness, -0.096, and the test of normality jointly based on skewness and kurtosis (p-value=0.884) confirm the almost symmetric and short-tail nature of the log transformed ratio. Therefore, an effect of unbalanced data points or outliers would not be a concern in our regression analysis below. * Sales volume has an N=66.

Table 4

Japanese Retailer: Effect of Time Stress on Allocation of Attention: Selling to Servicing Tasks

	log(Working Hours Ratio)			log(Effort Allocation Ratio)		
	1	2	3	4	5	6
log(Expertise ratio)	0.182†	-1.267†	-1.631*	0.091*	-0.313	-0.355
	(0.131)	(0.836)	(0.915)	(0.051)	(0.278)	(0.278)
Time stress	-0.049	-0.090†	-0.086	-0.055*	-0.067**	-0.056*
	(0.070)	(0.066)	(0.069)	(0.033)	(0.033)	(0.029)
log(Expertise ratio) × Time stress		0.393*	0.477**		0.110†	0.129*
		(0.221)	(0.238)		(0.078)	(0.078)
Demand unpredictability	0.009	0.020	0.009	0.013	0.016	0.016
	(0.050)	(0.050)	(0.053)	(0.019)	(0.019)	(0.023)
Competition	0.052	0.037	0.027	0.059***	0.055***	0.059***
	(0.049)	(0.050)	(0.058)	(0.020)	(0.020)	(0.022)
Education	0.021	0.019	-0.015	0.016	0.015	0.027
	(0.066)	(0.060)	(0.062)	(0.027)	(0.028)	(0.029)
Gender	-0.416**	-0.343**	-0.339**	-0.023	-0.003	0.048
	(0.162)	(0.157)	(0.170)	(0.069)	(0.062)	(0.079)
Constant	-0.030	0.159	0.361	-0.153	-0.100	-0.049
	(0.458)	(0.406)	(0.457)	(0.140)	(0.150)	(0.184)
Store fixed effects#	No	No	Yes (0)	No	No	Yes (4)
R^2	0.112	0.164	0.275	0.109	0.130	0.267
F statistic	1.772	2.169	2.685	1.679	1.636	1.439
N	126	126	126	126	126	126

*p < 0.10; **p < 0.05; ***p < 0.01, for a two-tail test. Robust standard errors in parentheses. †p<0.10 for a one-tail test. # Store fixed effects that are statistically significant at p<0.10 or smaller are in parentheses.

Table 5

Chilean Producer: Effect of Time Stress on Attention Allocation: Beer to Soft Drinks

	log(Working Hours Ratio)			log(Effort Allocation Ratio)		
	1	2	3	4	5	6
log(Expertise ratio)	0.007	-0.106†	-0.137*	-0.025	-0.153*	-0.183**
	(0.037)	(0.080)	(0.079)	(0.035)	(0.085)	(0.089)
Time stress	0.007	0.008	0.022	0.022	0.023	0.036
	(0.036)	(0.036)	(0.034)	(0.031)	(0.031)	(0.029)
log(Expertise ratio) × Time stress		0.041*	0.056**		0.046*	0.062**
		(0.024)	(0.025)		(0.024)	(0.028)
log(Product-demand unpredictability ratio)	0.017	0.000	0.026	-0.081	-0.100	-0.077
	(0.083)	(0.085)	(0.087)	(0.081)	(0.083)	(0.081)
log(Competition ratio)	0.179*	0.183*	0.180*	0.191**	0.196**	0.189**
	(0.105)	(0.104)	(0.100)	(0.094)	(0.093)	(0.088)
Assortment breath	-0.106*	-0.116*	-0.122**	-0.081	-0.092*	-0.097**
	(0.062)	(0.062)	(0.060)	(0.050)	(0.049)	(0.048)
Education	0.007	0.006	0.002	0.005	0.004	-0.001
	(0.007)	(0.007)	(0.009)	(0.007)	(0.007)	(0.009)
Marital status	-0.030	-0.040	-0.068	0.007	-0.003	-0.037
	(0.112)	(0.110)	(0.105)	(0.094)	(0.090)	(0.082)
Constant	0.705	0.763	0.986**	0.519	0.584*	0.818**
	(0.470)	(0.473)	(0.487)	(0.329)	(0.332)	(0.358)
Sales area fixed effects#	No	No	Yes (2)	No	No	Yes (1)
R^2	0.117	0.145	0.220	0.143	0.179	0.258
F statistic	1.049	1.604	2.053	1.571	1.936	2.624
N	61	61	61	61	61	61

*p < 0.10; **p < 0.05; ***p < 0.01, for a two-tail test. Robust standard errors in parentheses. †p<0.10 for a one-tail test. # Sales area fixed effects that are statistically significant at p<0.10 or smaller are in parentheses.

Table 6

Chilean Producer: Effect of Expertise and High Season on Sales Achievement: Beer to Soft Drinks

High Season: September to March

	log(Sales Volume Ratio)				
	1	2	3	4	5
	Pooled	DID	DID	DID	DID
log(Expertise ratio)	0.059	0.039	0.041	(omitted)	(omitted)
	(0.056)	(0.058)	(0.060)		
High season	0.043*	0.047**	(omitted)	0.040**	(omitted)
	(0.024)	(0.023)		(0.020)	
log(Expertise ratio) × High season		0.035*	0.035*	0.024***	0.023***
		(0.018)	(0.019)	(0.008)	(0.008)
log(Product-demand unpredictability ratio)	-0.156*	-0.156*	-0.159*	(omitted)	(omitted)
	(0.087)	(0.087)	(0.088)		
log(Competition ratio)	0.299***	0.298***	0.300***	(omitted)	(omitted)
	(0.099)	(0.099)	(0.102)		
Assortment breath	-0.116**	-0.116**	-0.119**	(omitted)	(omitted)
	(0.055)	(0.055)	(0.056)		
Education	-0.011	-0.011	-0.013	(omitted)	(omitted)
	(0.018)	(0.018)	(0.018)		
Marital status	0.054	0.055	0.046	(omitted)	(omitted)
	(0.127)	(0.127)	(0.130)		
Constant	0.463	0.463	0.400	-0.098***	-0.201**
	(0.342)	(0.341)	(0.348)	(0.011)	(0.093)
Manager fixed effect	No	No	No	Yes	Yes
Month fixed effect	No	No	Yes	No	Yes
R^2	0.117	0.119	0.149	0.791	0.816
N	1,174	1,174	1,174	1,206	1,206
No. of sales managers	63	63	63	66	66
Mean (standard deviation)	-0.074 (0.533)	-0.074 (0.533)	-0.074 (0.533)	-0.077 (0.565)	-0.077 (0.565)

*p < 0.10; **p < 0.05; ***p < 0.01. Standard errors in parentheses are robust and clustered at the manager level. Mean and standard deviation in the last row refer to those of the dependent variable.

Table 7

Chilean Producer: Effect of Task Complexity on Attention Allocation: Beer to Soft Drinks

	log(Working Hours Ratio)			log(Effort Allocation Ratio)		
	1	**2**	**3**	**4**	**5**	**6**
log(Expertise ratio)	**0.008**	**0.009**	**0.021**	**-0.024**	**-0.022**	**-0.009**
	(0.036)	**(0.035)**	**(0.036)**	**(0.035)**	**(0.033)**	**(0.033)**
log(Product-demand unpredictability ratio)	**0.019**	**-0.002**	**0.018**	**-0.076**	**-0.102**	**-0.087**
	(0.081)	**(0.082)**	**(0.087)**	**(0.081)**	**(0.082)**	**(0.082)**
log(Expertise ratio) × log(Product-demand unpredictability ratio)		**0.116***	**0.129***		**0.143****	**0.153****
		(0.062)	**(0.072)**		**(0.072)**	**(0.074)**
Time stress			0.005			0.017
			(0.037)			(0.032)
log(Competition ratio)	0.183*	0.181*	0.194*	0.203**	0.200**	0.205**
	(0.097)	(0.098)	(0.105)	(0.089)	(0.085)	(0.089)
Assortment breath	-0.107*	-0.119*	-0.130**	-0.082*	-0.097**	-0.107**
	(0.061)	(0.062)	(0.062)	(0.050)	(0.049)	(0.048)
Education	0.007	0.004	0.000	0.005	0.002	-0.003
	(0.007)	(0.007)	(0.008)	(0.007)	(0.006)	(0.008)
Marital status	-0.029	-0.009	-0.030	0.009	0.035	0.007
	(0.110)	(0.110)	(0.107)	(0.091)	(0.088)	(0.086)
Constant	0.725*	0.803*	1.069**	0.580*	0.677**	0.922***
	(0.427)	(0.429)	(0.493)	(0.318)	(0.308)	(0.331)
Regional fixed effects#	No	No	Yes (2)	No	No	Yes (2)
R^2	0.117	0.149	0.208	0.139	0.188	0.250
F statistic	1.194	1.611	1.814	1.634	2.844	3.245
N	61	61	61	61	61	61

*p < 0.10; **p < 0.05; ***p < 0.01, for a two-tail test. Robust standard errors in parentheses. †p<0.10 for a one-tail test. # Regional fixed effects that are statistically significant at p<0.10 or smaller.

Online Appendix Table A1

Chilean Producer: Effect of Expertise and High Season on Sales Achievement: Beer to Soft Drinks

(Alternative Definitions of High Seasonality Months)

	log(Sales Volume Ratio)					**log(Sales Volume Ratio)**				
	High season using the top 4 months in company sales (December, January, February, and September)					High season using the top 3 months in company sales (December, January, and September)				
	1	2	3	4	5	6	7	8	9	10
	Pooled	DID	DID	DID	DID	Pooled	DID	DID	DID	DID
log(Expertise ratio)	0.059	0.049	0.050	(omitted)	(omitted)	0.059	0.053	0.055	(omitted)	(omitted)
	(0.056)	(0.056)	(0.058)			(0.056)	(0.056)	(0.058)		
High season	0.043	0.030	(omitted)	0.026	(omitted)	0.041*	0.042**	(omitted)	0.044**	(omitted)
	(0.027)	(0.020)		(0.018)		(0.021)	(0.021)		(0.019)	
log(Expertise ratio) × High season		0.033***	0.031***	0.018***	0.017**		0.024**	0.023**	0.014**	0.013*
		(0.011)	(0.011)	(0.007)	(0.007)		(0.011)	(0.012)	(0.007)	(0.008)
log(Product-demand unpredictability ratio)	-0.156*	-0.156*	-0.159*	(omitted)	(omitted)	-0.156*	-0.156*	-0.159*	(omitted)	(omitted)
	(0.087)	(0.087)	(0.088)			(0.087)	(0.087)	(0.088)		
log(Competition ratio)	0.299***	0.298***	0.300***	(omitted)	(omitted)	0.299***	0.299***	0.300***	(omitted)	(omitted)
	(0.099)	(0.099)	(0.102)			(0.099)	(0.099)	(0.102)		
Assortment breath	-0.116**	-0.115**	-0.119**	(omitted)	(omitted)	-0.116**	-0.116**	-0.119**	(omitted)	(omitted)
	(0.054)	(0.054)	(0.056)			(0.054)	(0.055)	(0.056)		
Education	-0.011	-0.011	-0.013	(omitted)	(omitted)	-0.011	-0.011	-0.013	(omitted)	(omitted)
	(0.018)	(0.018)	(0.018)			(0.018)	(0.018)	(0.018)		
Marital status	0.054	0.053	0.047	(omitted)	(omitted)	0.054	0.054	0.047	(omitted)	(omitted)
	(0.127)	(0.127)	(0.130)			(0.127)	(0.127)	(0.130)		
Constant	0.477	0.475	0.399	-0.085***	-0.202**	0.476	0.476	0.399	-0.087***	-0.202**
	(0.342)	(0.342)	(0.348)	(0.006)	(0.093)	(0.342)	(0.342)	(0.348)	(0.004)	(0.093)
Manager fixed effect	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Month fixed effect	No	No	Yes	No	Yes	No	No	Yes	No	Yes
R^2	0.116	0.117	0.149	0.790	0.815	0.116	0.117	0.148	0.791	0.815
N	1,174	1,174	1,174	1,206	1,206	1,174	1,174	1,174	1,206	1,206
No. of sales managers	63	63	63	66	66	63	63	63	66	66

*p < 0.10; **p < 0.05; ***p < 0.01. Standard errors in parentheses are robust and clustered at the manager level. Mean and standard deviation in the last row refer to those of the dependent variable.

Online Appendix Table: A2

Japanese Retailer: Effect of Time Stress on Allocation of Attention: Original Expertise Value on Ratio of Selling-to-Servicing tasks

	log(Working Hours Ratio)	**log(Effort Allocation Ratio)**
	1	2
log(Selling expertise)	-1.500†	-0.633**
	(0.978)	(0.314)
log(Servicing expertise)	1.623*	0.401†
	(0.910)	(0.276)
Time stress	-0.015	-0.173*
	(0.258)	(0.103)
log(Selling expertise) × Time stress	0.444*	0.195**
	(0.251)	(0.086)
log(Servicing Expertise) × Time stress	-0.476**	-0.143*
	(0.237)	(0.078)
Demand unpredictability	0.008	0.018
	(0.053)	(0.023)
Competition	0.029	0.056***
	(0.059)	(0.021)
Education	-0.015	0.025
	(0.061)	(0.030)
Gender	-0.353**	0.053
	(0.169)	(0.085)
Constant	0.080	0.476
	(1.051)	(0.333)
Store fixed effects#	Yes (0)	Yes (4)
R^2	0.276	0.284
F statistic	2.521	1.437
N	126	126

*p < 0.10; **p < 0.05; ***p < 0.01, for a two-tail test. Robust standard errors in parentheses. †p<0.10 for a one-tail test. # Store fixed effects that are statistically significant at p<0.10 or smaller are in parentheses.

Online Appendix Table: A3

Chilean Producer: Effect of Time Stress on Attention Allocation: Original Expertise Value on Ratio of Beer-to-Soft-Drinks

	log(Working Hours Ratio)	**log(Effort Allocation Ratio)**
	1	2
log(Beer expertise)	0.011	-0.063
	(0.115)	(0.116)
log(Soft-drink expertise)	0.359***	0.357**
	(0.130)	(0.140)
Time stress	0.301**	0.274*
	(0.153)	(0.148)
log(Beer expertise) × Time stress	0.004	0.016
	(0.039)	(0.038)
log(Soft drink expertise) × Time stress	-0.117***	-0.113***
	(0.038)	(0.039)
log(Product-demand unpredictability ratio)	0.029	-0.068
	(0.077)	(0.078)
log(Competition ratio)	0.196*	0.201**
	(0.101)	(0.093)
Assortment breath	-0.111**	-0.087*
	(0.056)	(0.046)
Education	0.007	0.003
	(0.009)	(0.010)
Marital status	-0.016	0.007
	(0.106)	(0.088)
Constant	-0.076	-0.032
	(0.625)	(0.635)
Sales area fixed effects#	Yes (1)	Yes (1)
R^2	0.281	0.297
F statistic	1.674	2.356
N	61	61

*p < 0.10; **p < 0.05; ***p < 0.01, for a two-tail test. Robust standard errors in parentheses. †p<0.10 for a one-tail test. # Sales area fixed effects that are statistically significant at p<0.10 or smaller are in parentheses.

Online Appendix: Table A4

Chilean Producer: Effect of Expertise and High Season on Sales Achievement: Original Expertise Value on Ratio of Beer-to-Soft-Drinks

High Season: September to March

	log(Sales Volume Ratio)			
	1	2	3	4
	DID	DID	DID	DID
log(Beer expertise)	0.098	0.119*	(omitted)	(omitted)
	(0.061)	(0.062)		
log(Soft-drink expertise)	0.055	0.080	(omitted)	(omitted)
	(0.075)	(0.078)		
High season	0.121	(omitted)	0.065	(omitted)
	(0.116)		(0.093)	
log(Beer expertise) ×	0.025	0.022	0.019	0.014
High season	(0.026)	(0.027)	(0.019)	(0.020)
log(Soft-drink expertise) ×	-0.055*	-0.060*	-0.030	-0.036*
High season	(0.032)	(0.033)	(0.020)	(0.021)
log(Product-demand	-0.165*	-0.173*	(omitted)	(omitted)
unpredictability ratio)	(0.084)	(0.082)		
log(Competition ratio)	0.319***	0.330***	(omitted)	(omitted)
	(0.099)	(0.101)		
Assortment breath	-0.123**	-0.133***	(omitted)	(omitted)
	(0.050)	(0.050)		
Education	-0.001	-0.001	(omitted)	(omitted)
	(0.020)	(0.020)		
Marital status	0.055	0.041	(omitted)	(omitted)
	(0.129)	(0.129)		
Constant	0.089	-0.010	-0.098***	-0.147
	(0.414)	(0.429)	(0.011)	(0.125)
Manager fixed effect	No	No	Yes	Yes
Month fixed effect	No	Yes	No	Yes
R^2	0.151	0.197	0.791	0.816
N	1,174	1,174	1,206	1,206
No. of sales managers	63	63	66	66
Mean (standard deviation)	-0.074	-0.074	-0.077	-0.077
	(0.533)	(0.533)	(0.565)	(0.565)

*p < 0.10; **p < 0.05; ***p < 0.01. Standard errors in parentheses are robust and clustered at the manager level. Mean and standard deviation in the last row refer to those of the dependent variable.