ONLINE APPENDIX

Managerial Style and Attention

By WOUTER DESSEIN AND TANO SANTOS*

* Dessein: Columbia University, Graduate School of Business (email: wd2179@columbia.edu). Santos: Columbia University, Graduate School of Business (email: js1786@columbia.edu).

ONLINE APPENDIX

A1. Microfoundations of learning technology

A simple microfoundation of (5) can be obtained as follows. Assume the task-specific shock θ_i is the sum of n independently distributed shocks θ_{ik} :

$$\theta_i = \sum_{k=1}^{n} \theta_{ik} \quad \text{with} \quad \theta_{ik} \sim N(0, \sigma_\theta^2/n),$$

Each element θ_{ik} can be interpreted as an "component" of task i to be understood by the manager to have a complete picture of task i . The manager observes two independent signals, s_{ik} and S_{ik} , about each component θ_{ik} and signals are independent across components. Both signals have the same structure: They are either fully informative about θ_{ik} or pure noise. Signal s_{ik} is endogenous in that its precision is a function of the attention t_i that the agent devotes to task i . Specifically the manager learns θ_{ik} with probability $q\,(t_i)$. We assume that learning follows a Poisson process with hazard rate λ:

$$q(t_i) = 1 - e^{-\lambda t_i}.$$

S_{ik} is instead an exogenous signal. Its precision is a function of the manager's expertise T_i, which is exogenous. As in the case of the endogenous signal the manager thus learns θ_{ik} with probability $q\,(T_i)$. Exogenous learning is also assumed to follow a Poisson process with hazard rate λ. The manager thus learns any given component θ_{ik} with probability:[1]

$$q_i \equiv q\,(t_i + T_i) = 1 - e^{-\lambda(t_i + T_i)}. \tag{A1}$$

Notice thus that attention t_i and expertise T_i are substitutes in the learning process.

Denoting $s_i = [s_{i1}, ..., s_{in}]$ and $S_i = [S_{i1}, ..., S_{in}]$, then

$$\widehat{\theta}_i \equiv E(\theta_i | s_i, S_i) = \sum_{k=1}^{n} E(\theta_{ik} | s_{ik}, S_{ik})$$

In the limit as the number of components n goes to infinity, we have that

$$\mathsf{RV}\,(\theta_i) = E\left(\theta_i - \widehat{\theta}_i\right)^2 = (1 - q_i)\sigma_\theta^2, \tag{A2}$$

[1] Notice that the probability that the manager learns θ_{ik} is given by

$$(1 - q\,(T_i))\, q\,(t_i) + (1 - q\,(t_i))\, q\,(T_i) + q\,(T_i)\, q\,(t_i) = q\,(t_i + T_i)\,.$$

as posited in (5). Moreover, the attention constraint (6) can then be rewritten as

$$(1-q_1)(1-q_2) \geq e^{-\lambda(2\tau+T_1+T_2)}.$$

We interpret $1/\lambda$ as reflecting the complexity of the environment. The larger is $1/\lambda$,the more attention and expertise are required to reduce the residual variance $\mathsf{RV}\left(\theta_i\right)$.

A2. Proof of Proposition 2

Without loss of generality, assume that task 1 is affected by the largest perceived shock, that is $\widehat{\theta}_1^2 > \widehat{\theta}_2^2$. Substituting (8) into (7) and manipulating terms, we obtain

$$E\left(\pi|\widehat{\boldsymbol{\theta}}\right) = \sum_{i\in\{1,2\}} \frac{\widehat{\theta}_i^2}{1+\beta(1-p_i))} \tag{A3}$$

Denoting $\widehat{\theta}_2^2 = k\widehat{\theta}_1^2$ with $k<1$, we can rewrite this as

$$E\left(\pi|\widehat{\boldsymbol{\theta}}\right) = \left(\frac{1}{1+\beta e^{-\mu r_1}} + k\frac{1}{1+\beta e^{-\mu(r-r_1)}}\right)\widehat{\theta}_1^2 \tag{A4}$$

where $e^{-\mu r_i} = 1-p_i$ and $r_1+r_2 = r$. Since $k<1$, it is easy to verify that is never optimal to set $r_1 < r/2$. Hence, let $r_1 \in [r/2, r]$. Taking the derivative of (A4) with respect to r_1 we obtain

$$\frac{\partial E\left(\pi|\widehat{\boldsymbol{\theta}}\right)}{\partial r_1} = \beta\mu\left[\frac{e^{-\mu r_1}}{\left(1+\beta e^{-\mu r_1}\right)^2} - k\frac{e^{-\mu(r-r_1)}}{\left(1+\beta e^{-\mu(r-r_1)}\right)^2}\right]\widehat{\theta}_i^2. \tag{A5}$$

Using plain algebra, if $k=1$, it follows that

$$\frac{\partial E\left(\pi|\widehat{\boldsymbol{\theta}}\right)}{\partial r_1} > 0 \Longleftrightarrow 1-\beta^2 e^{-\mu r} < 0 \Longleftrightarrow p < \bar{p}(\beta) \equiv 1-1/\beta^2$$

where, recall, $p = 1-e^{-\mu r}$. Obviously, if $k \leq 1$, then $p < \bar{p}(\beta)$ is a sufficient condition for $\partial E\left(\pi|\widehat{\boldsymbol{\theta}}\right)/r_1 > 0$. Hence if $\widehat{\theta}_1^2 > \widehat{\theta}_2^2$, then $p < \bar{p}(\beta) = 1-1/\beta^2$ implies $(r_1^*, r_2^*) = (r, 0)$ and $(p_1^*, p_2^*) = (p, 0)$. Note further that $\bar{p}(\beta)$ is increasing in β.

A3. Proof of Proposition 4

We first prove Proposition 4, as the proof of Proposition 3 will make use of it.

EXPECTED PROFITS. — Expected Profits conditional on q_1 and q_2 are given by

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

We can make a simple change of variable $\varphi_1 \equiv \widehat{\theta}_1/\sqrt{q_1}$ and $\varphi_2 \equiv \widehat{\theta}_1/\sqrt{q_2}$, so that both φ_1 and φ_2 are normally distributed with variance σ_θ^2. With some abuse of notation let $F(x) \equiv F(x,\sigma_\theta^2)$, then the expected profits can be rewritten as

$$\begin{aligned} \Pi\left(q_1,q_2\right) &= 4\int_0^{+\infty}\left[\int_0^{\sqrt{\frac{q_2}{q_1}}\varphi_2}\frac{q_1\varphi_1^2}{1+\beta}dF(\varphi_1)+\int_{\sqrt{\frac{q_2}{q_1}}\varphi_2}^{+\infty}\frac{q_1\varphi_1^2}{1+(1-p)\beta}dF(\varphi_1)\right]dF(\varphi_2) \\ &+ 4\int_0^{+\infty}\left[\int_0^{\sqrt{\frac{q_1}{q_2}}\varphi_1}\frac{q_2\varphi_2^2}{1+\beta}dF(\varphi_2)+\int_{\sqrt{\frac{q_1}{q_2}}\varphi_1}^{+\infty}\frac{q_2\varphi_2^2}{1+(1-p)\beta}dF(\varphi_2)\right]dF(\varphi_1) \\ &= \frac{4}{1+(1-p)\beta}\int_0^{+\infty}\left[\int_{\sqrt{\frac{q_2}{q_1}}\varphi_k}^{+\infty}q_1\varphi_1^2dF(\varphi_1)+\int_{\sqrt{\frac{q_1}{q_2}}\varphi_k}^{+\infty}q_2\varphi_2^2dF(\varphi_2)\right]dF(\varphi_k) \\ &+ \frac{4}{1+\beta}\int_0^{+\infty}\left[\int_0^{\sqrt{\frac{q_2}{q_1}}\varphi_k}q_1\varphi_1^2dF(\varphi_1)+\int_0^{\sqrt{\frac{q_1}{q_2}}\varphi_k}q_2\varphi_2^2dF(\varphi_2)\right]dF(\varphi_k), \end{aligned}$$

where φ_k is the normally distributed random variable with mean 0 and variance σ_θ^2.

Notice that when $q_1 = q_2 = q$ the profit expression $\Pi\left(q_1,q_2\right)$ simplifies to

(A6)

$$\Pi\left(q,q\right) = 2q\int_0^{\infty}\left(\frac{4}{1+\beta}\int_0^{\varphi_k}\varphi_i^2dF(\varphi_i)+\frac{4}{1+\beta(1-p)}\int_{\varphi_k}^{\infty}\varphi_i^2dF(\varphi_i)\right)dF(\varphi_k),$$

where both φ_i and φ_k are normally distributed random variables with mean 0 and variance σ_θ^2. (A6) has the following closed form solution:

$$\Pi\left(q,q\right) = 2q\left(\frac{1}{1+\beta}\frac{\pi-2}{2\pi}+\frac{1}{1+\beta(1-p)}\frac{\pi+2}{2\pi}\right)\sigma_\theta^2,$$

which simplifies in turn to

(A7) $$\Pi\left(q,q\right) = 2qC$$

with

$$C \equiv \left(\tfrac{1}{1+\beta}\right)\left[1+\tfrac{\beta p}{1+\beta(1-p)}\tfrac{\pi+2}{2\pi}\right]\sigma_\theta^2$$

Finally, notice that if $q_1 > q_2$, we can rewrite expected profits as

$$\begin{aligned}\Pi\left(q_1,q_2\right) &= \left(q_1+q_2\right)C \\ &+ \frac{4}{1+(1-p)\beta}\left[\int_{\sqrt{\frac{q_2}{q_1}}\varphi_k}^{\varphi_k} q_1\varphi_1^2 dF(\varphi_1) - \int_{\varphi_k}^{\sqrt{\frac{q_1}{q_2}}\varphi_k} q_2\varphi_2^2 dF(\varphi_2)\right] dF(\varphi_k) \\ &- \frac{4}{1+\beta}\left[\int_{\sqrt{\frac{q_1}{q_2}}\varphi_k}^{\varphi_k} q_1\varphi_1^2 dF(\varphi_1) - \int_{\varphi_k}^{\sqrt{\frac{q_2}{q_1}}\varphi_k} q_2\varphi_2^2 dF(\varphi_2)\right] dF(\varphi_k)\end{aligned}$$

or still

$$\text{(A8)} \qquad \Pi\left(q_1,q_2\right) = \left(q_1+q_2\right)C + D\int_0^\infty \left[B_1\left(q_1,q_2\right) - B_2\left(q_1,q_2\right)\right] dF\left(\varphi_k\right)$$

where

$$D = \frac{4}{1+\beta}\left(\frac{\beta p}{1+\beta\left(1-p\right)}\right)$$

with

$$B_1\left(q_1,q_2\right) = \int_{\sqrt{\frac{q_2}{q_1}}\varphi_k}^{\varphi_k} q_1\varphi_1^2 dF\left(\varphi_1\right) \quad \text{and} \quad B_2\left(q_1,q_2\right) = \int_{\varphi_k}^{\sqrt{\frac{q_1}{q_2}}\varphi_k} q_2\varphi_2^2 dF\left(\varphi_2\right)$$

PROOF OF PROPOSITION 4. — Assume $T_1 > T_2$ and assume that $t_i \in [0,\tau]$ with $t_1 + t_2 = \tau$ with τ small. The proof for $t_i \in [0, 2\tau]$ with $t_1 + t_2 = 2\tau$ is identical, up to a transformation. Expected profits $\Pi\left(q_1,q_2\right)$ conditional on $q_1 = q(T_1+t_1)$ and $q_2 = q(T_2+t_2)$ are given by (A8). Note that

$$\begin{aligned}\frac{\partial B_1\left(q(T_1+\tau),q(T_2)\right)}{\partial\tau} &= \lambda\exp\left(-\lambda\left(T_1+\tau\right)\right)\left[\int_{\sqrt{\frac{q_2}{q_1}}\varphi_k}^{\varphi_k}\varphi_1^2 dF\left(\varphi_1\right) + \frac{1}{2}\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_k\right)\right] \\ \frac{\partial B_2\left(q(T_1+\tau),q(T_2)\right)}{\partial\tau} &= \frac{\lambda}{2}\exp\left(-\lambda\left(T_1+\tau\right)\right)\left(\frac{q_1}{q_2}\right)^{\frac{1}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_1}{q_2}}\varphi_k\right)\end{aligned}$$

and, similarly,

$$\frac{\partial B_1\left(q(T_1),q(T_2+\tau)\right)}{\partial\tau} = -\frac{\lambda}{2}\exp\left(-\lambda\left(\tau_2+\tau\right)\right)\left(\frac{q_2}{q_1}\right)^{\frac{1}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_k\right)$$

$$\frac{\partial B_2\left(q(T_1),q(T_2+\tau)\right)}{\partial\tau} = \lambda\exp\left(-\lambda\left(T_2+\tau\right)\right)\left[\int_{\varphi_k}^{\sqrt{\frac{q_1}{q_2}}\varphi_k}\varphi_2^2 dF\left(\varphi_2\right)-\frac{1}{2}\left(\frac{q_1}{q_2}\right)^{\frac{3}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_1}{q_2}}\varphi_k\right)\right]$$

It follows that

$$\begin{aligned}\frac{\Pi(q(T_1+\tau),q(T_2))}{\partial\tau} &= \lambda\exp\left(-\lambda\left(T_1+\tau\right)\right)C \\ &+ D\int_0^\infty\left[\frac{\partial B_1\left(q(T_1+\tau),q(T_2)\right)}{\partial\tau}-\frac{\partial B_2\left(q(T_1+\tau),q(T_2)\right)}{\partial\tau}\right]dF\left(\varphi_k\right)\end{aligned}$$

$$\begin{aligned}&= \lambda\exp\left(-\lambda\left(T_1+\tau\right)\right)C+D\lambda\exp\left(-\lambda\left(T_1+\tau\right)\right)\\ &\times\int_0^\infty\left(\frac{1}{2}\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_k\right)-\frac{1}{2}\left(\frac{q_1}{q_2}\right)^{\frac{1}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_1}{q_2}}\varphi_k\right)+\int_{\sqrt{\frac{q_2}{q_1}}\varphi_k}^{\varphi_k}\varphi_1^2 dF\left(\varphi_1\right)\right)dF\left(\varphi_k\right)\end{aligned}$$

We have that

$$\begin{aligned}\int_0^\infty\frac{1}{2}\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_k\right)dF\left(\varphi_k\right) &= \int_0^\infty\frac{1}{2}\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_k\right)f(\varphi_k)d\varphi_k \\ &= \int_0^\infty\frac{1}{2}\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\varphi_k^3\frac{1}{\sqrt{2\pi}}e^{-\frac{\frac{q_2}{q_1}\varphi_k^2}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{\varphi_k^2}{2}}d\varphi_k \\ &= \left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\frac{1}{\sqrt{2\pi}}\int_0^\infty\frac{1}{2}\varphi_k^3\frac{1}{\sqrt{2\pi}}e^{-\frac{\frac{q_2+q_1}{q_1}\varphi_k^2}{2}}d\varphi_k \\ &= \frac{1}{\sqrt{2\pi}}\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\sqrt{\frac{q_1}{q_2+q_1}}\int_0^\infty\frac{1}{2}\varphi_k^3\frac{1}{\sqrt{\frac{q_1}{q_2+q_1}}\sqrt{2\pi}}e^{-\frac{\varphi_k^2}{\frac{2q_1}{q_2+q_1}}}d\varphi_k \\ &= \frac{1}{\sqrt{2\pi}}\sqrt{\frac{q_1}{q_2+q_1}}\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\int_0^\infty\frac{1}{2}\varphi_k^3 f\left(\varphi_k;0,\sqrt{\frac{q_1}{q_2+q_1}}\right)d\varphi_k,\end{aligned}$$

where

$$f\left(\varphi_k;0,\sqrt{\frac{q_1}{q_2+q_1}}\right)$$

is the normal density function with mean 0 and standard deviation $\sqrt{\frac{q_1}{q_2+q_1}}$. Since

$$\int_0^{\infty} \frac{1}{2} x^3 f\left(x;0,\sigma\right) dx = \frac{\sigma^3}{2}\frac{\sqrt{2}}{\sqrt{\pi}},$$

where $f\left(x;0,\sigma\right)$ is the normal density function when the mean is 0 and the standard deviation is σ, this can be simplified to

$$\begin{aligned}\int_0^{\infty} \frac{1}{2}\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}} \varphi_k^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_k\right) dF\left(\varphi_k\right) &= \frac{1}{\sqrt{2\pi}}\sqrt{\frac{q_1}{q_2+q_1}}\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\sqrt{\frac{q_1}{q_2+q_1}}^{3}\frac{1}{2}\frac{\sqrt{2}}{\sqrt{\pi}}\sigma_\theta^3 \\ &= \frac{1}{2\pi}\left(\frac{q_1}{q_2+q_1}\right)^2\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\sigma_\theta^3\end{aligned}$$

Similarly,

$$\begin{aligned}\int_0^{\infty} \frac{1}{2}\left(\frac{q_1}{q_2}\right)^{\frac{1}{2}} \varphi_k^3 f\left(\sqrt{\frac{q_1}{q_2}}\varphi_k\right) dF\left(\varphi_k\right) &= \frac{1}{2\pi}\left(\frac{q_2}{q_2+q_1}\right)^2\left(\frac{q_1}{q_2}\right)^{\frac{1}{2}}\sigma_\theta^3 \\ &= \frac{1}{2\pi}\left(\frac{q_1}{q_2+q_1}\right)^2\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\sigma_\theta^3\end{aligned}$$

It follows that

$$\frac{\partial\Pi(q(T_1+\tau),q(T_2))}{\partial\tau} = \lambda\exp\left(-\lambda\left(T_1+\tau\right)\right)C + D\lambda\exp\left(-\lambda\left(T_1+\tau\right)\right)\int_0^{\infty}\int_{\sqrt{\frac{q_2}{q_1}}\varphi_k}^{\varphi_k}\varphi_1^2 dF\left(\varphi_1\right)dF\left(\varphi_k\right)$$

Similarly,

$$\begin{aligned}\frac{\partial\Pi(q(T_1),q(T_2+\tau))}{\partial\tau} &= \lambda\exp\left(-\lambda\left(T_2+\tau\right)\right)C + D\lambda\exp\left(-\lambda\left(T_2+\tau\right)\right) \\ &\times \int_0^{\infty}\left[\frac{1}{2}\left(\frac{q_1}{q_2}\right)^{\frac{3}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_1}{q_2}}\varphi_k\right) - \frac{1}{2}\left(\frac{q_2}{q_1}\right)^{\frac{1}{2}}\varphi_k^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_k\right) - \int_{\varphi_k}^{\sqrt{\frac{q_1}{q_2}}\varphi_k}\varphi_2^2 dF\left(\varphi_2\right)\right]dF\left(\varphi_k\right) \\ &= \lambda\exp\left(-\lambda\left(T_2+\tau\right)\right)C - D\lambda\exp\left(-\lambda\left(T_2+\tau\right)\right)\int_0^{\infty}\int_{\varphi_k}^{\sqrt{\frac{q_1}{q_2}}\varphi_k}\varphi_2^2 dF\left(\varphi_2\right)dF\left(\varphi_k\right)\end{aligned}$$

In sum, we have that

$$\left.\frac{\partial\Pi(q(T_1+\tau),q(T_2))}{\partial\tau}\right|_{\tau=0} > \left.\frac{\partial\Pi(q(T_1),q(T_2+\tau))}{\partial\tau}\right|_{\tau=0}$$

$$\Leftrightarrow \quad \lambda \exp\left(-\lambda T_1\right) C + D\lambda \exp\left(-\lambda T_1\right) \left[\int_0^\infty \int_{\varphi_k}^{\sqrt{\frac{q(T_1)}{q(T_2)}}\varphi_k} \varphi_1^2 dF\left(\varphi_1\right) dF\left(\varphi_k\right)\right]$$

$$> \quad \lambda \exp\left(-\lambda T_2\right) C - D\lambda \exp\left(-\lambda T_2\right) \left[\int_0^\infty \int_{\sqrt{\frac{q(T_2)}{q(T_1)}}\varphi_k}^{\varphi_k} \varphi_2^2 dF\left(\varphi_2\right) dF\left(\varphi_k\right)\right]$$

or still

$$\Leftrightarrow \exp\left(-\lambda\left(T_1 - T_2\right)\right)) > \frac{C - D\int_0^\infty \int_{\sqrt{\frac{q_2}{q_1}}\varphi_k}^{\varphi_k} \varphi_2^2 dF\left(\varphi_2\right) dF\left(\varphi_k\right)}{C + D\int_0^\infty \int_{\varphi_k}^{\sqrt{\frac{q_1}{q_2}}\varphi_k} \varphi_1^2 dF\left(\varphi_1\right) dF\left(\varphi_k\right)}$$

Define

$$\rho = \frac{q(T_1)}{q(T_2)}$$

Then

$$\left.\frac{\partial \Pi(q(T_1+\tau), q(T_2))}{\partial \tau}\right|_{\tau=0} > \left.\frac{\partial \Pi(q(T_1), q(T_2+\tau))}{\partial \tau}\right|_{\tau=0}$$

$$\Leftrightarrow \frac{1-q(T_1)}{1-q(T_1)/\rho} > \frac{C - D\int_0^\infty \int_{\varphi_k/\sqrt{\rho}}^{\varphi_k} \varphi_2^2 dF\left(\varphi_2\right) dF\left(\varphi_k\right)}{C + D\int_0^\infty \int_{\varphi_k}^{\sqrt{\rho}\varphi_k} \varphi_1^2 dF\left(\varphi_1\right) dF\left(\varphi_k\right)}$$

or, substituting C and D and simplifying,

$$\text{(A9)} \qquad \Leftrightarrow \frac{1-q(T_1)}{1-q(T_1)/\rho} > \frac{\left(1 + \frac{\pi+2}{2\pi}b\right)\sigma_\theta^2 - 4b\int_0^\infty \int_{\varphi_k/\sqrt{\rho}}^{\varphi_k} \varphi_2^2 dF\left(\varphi_2\right) dF\left(\varphi_k\right)}{\left(1 + \frac{\pi+2}{2\pi}b\right)\sigma_\theta^2 + 4b\int_0^\infty \int_{\varphi_k}^{\sqrt{\rho}\varphi_k} \varphi_1^2 dF\left(\varphi_1\right) dF\left(\varphi_k\right)},$$

where

$$b \equiv \frac{\beta p}{1+\beta(1-p)} \in (0, \infty),$$

Fix $\rho > 1$, then on the one hand, the RHS is strictly smaller than 1 and independent of $q(T_1)$. On the other hand, the LHS is strictly decreasing in $q(T_1)$, and equals 1 as $q(T_1)$ goes to 0 and goes to 0 as $q(T_1)$ goes to 1. Hence, keeping ρ fixed, if $q(T_1)$ is sufficiently small, then managing with style $(t_1^*, t_2^* = (\tau, 0))$ is always optimal. Similarly, fixing ρ as $q(T_1)$ goes to 1, then for $q(T_1)$ sufficiently large, rebalancing attention $((t_1^*, t_2^*) = (0, \tau)$ is optimal. It follows that there exists a unique cut-off q_1^* given by

$$\text{(A10)} \qquad \frac{1-q_1^*}{1-q_1^*/\rho} = \frac{\left(1 + \frac{\pi+2}{2\pi}b\right)\sigma_\theta^2 - 4b\int_0^\infty \left(\int_{\theta_j/\sqrt{\rho}}^{\theta_j} \theta_i^2 dF(\theta_i)\right) dF(\theta_j)}{\left(1 + \frac{\pi+2}{2\pi}b\right)\sigma_\theta^2 + 4b\int_0^\infty \left(\int_{\theta_j}^{\sqrt{\rho}\theta_j} \theta_i^2 dF(\theta_i)\right) dF(\theta_j)},$$

so that if $q(T_1) < q_1^*$, we have $(t_1^*, t_2^*) = (\tau, 0)$, and for $q(T_1) > q_1^*$ we have $(t_1^*, t_2^*) = (0, \tau)$. Note further that q_1^* is continuous in ρ and continuous and increasing in b. Defining $\Lambda(\rho, b)$ as

$$q_1^* = 1 - e^{-\Lambda} \tag{A11}$$

It follows that $(t_1^*, t_2^*) = (\tau, 0)$ if $\lambda T_1 < \Lambda(\rho, b)$ whereas $(t_1^*, t_2^*) = (0, \tau)$ if $\lambda T_1 > \Lambda(\rho, b)$. Moreover, $\Lambda(\rho, b)$ is continuous in ρ and continuous and increasing in b. □

A4. Proof of Proposition 3

Assume now that $T_2 = T_1 = T$. We first prove Proposition 3 for the case where $t_i \in \{0, \tau, 2\tau\}$ with $t_1 + t_2 = 2\tau$ (Discrete Attention). We subsequently generalize the result for any $t_i \in [0, 2\tau]$ with $t_1 + t_2 = 2\tau$ (Continuous Attention).

PROPOSITION 3: DISCRETE ATTENTION. — We first establish some preliminary results. First, if the generalist manager opts to balance attention evenly among tasks:

$$q_1 = q_2 \equiv q = 1 - \exp\left(-\lambda\left(T + \tau\right)\right) \tag{A12}$$

then, given (A8), the profits under balanced attention equal

$$\Pi\left(q, q\right) = 2qC. \tag{A13}$$

Second, if the manager focuses all her attention on one task, say, task 1,

$$q_1 = q\left(2\tau + T\right) = 1 - \exp\left(-\lambda\left(T + 2\tau\right)\right) \tag{A14}$$

$$> q_2 = q(T) = 1 - \exp\left(-\lambda T\right) \tag{A15}$$

and, given (A8), expected profits equal

$$\Pi\left(q_1, q_2\right) = \left(q_1 + q_2\right) C + D \int_0^{\infty} \left[B_1\left(q_1, q_2\right) - B_2\left(q_1, q_2\right)\right] dF\left(\varphi_k\right) \tag{A16}$$

Finally, we compute the expressions for

$$\frac{\partial^2 B_2\left(q_1, q_2\right)}{\partial t^2} \quad \text{and} \quad \frac{\partial^2 B_1\left(q_1, q_2\right)}{\partial t^2},$$

which are useful in what follows. First notice that

$$\frac{\partial}{\partial \tau}\sqrt{\frac{q_1}{q_2}} = \frac{\lambda}{\left(q_1 q_2\right)^{\frac{1}{2}}} \exp\left(-\lambda\left(T + 2\tau\right)\right).$$

Hence

$$\frac{\partial B_2(q_1,q_2)}{\partial \tau} = \lambda \left(\frac{q_1}{q_2}\right)^{\frac{1}{2}} \varphi_2^3 f\left(\sqrt{\frac{q_1}{q_2}}\varphi_2\right) \exp\left(-\lambda\left(T+2\tau\right)\right),$$

which yields

$$\begin{aligned}\frac{\partial^2 B_2}{\partial \tau^2} &= \frac{\lambda^2}{(q_1 q_2)^{\frac{1}{2}}}\varphi_2^3 f\left(\sqrt{\frac{q_1}{q_2}}\varphi_2\right)\exp\left[-2\lambda\left(T+2\tau\right)\right] \\ &+ \frac{\lambda^2}{q_2}\varphi_2^4 f'\left(\sqrt{\frac{q_1}{q_2}}\varphi_2\right)\exp\left[-2\lambda\left(T+2\tau\right)\right] \\ &- 2\lambda^2\left(\frac{q_1}{q_2}\right)^{\frac{1}{2}}\varphi_2^3 f\left(\sqrt{\frac{q_1}{q_2}}\varphi_2\right)\exp\left(-\lambda\left(T+2\tau\right)\right).\end{aligned}$$

Notice in particular

$$\begin{aligned}\left.\frac{\partial^2 B_2}{\partial \tau^2}\right|_{\tau=0} &= \frac{\lambda^2}{q_2}\varphi_2^3 f\left(\varphi_2\right)\exp\left(-2\lambda T\right) \\ &+ \frac{\lambda^2}{q_2}\varphi_2^4 f'\left(\varphi_2\right)\exp\left(-2\lambda T\right) - 2\lambda^2\varphi_2^3 f\left(\varphi_2\right)\exp\left(-\lambda T\right).\end{aligned}$$

Next, notice that

$$\frac{\partial}{\partial \tau}\sqrt{\frac{q_2}{q_1}} = -\lambda\left(\frac{q_2^{\frac{1}{2}}}{q_1^{\frac{3}{2}}}\right)\exp\left(-\lambda\left(T+2\tau\right)\right)$$

and thus

$$\begin{aligned}\frac{\partial B_1(q_1,q_2)}{\partial \tau} &= \int_{\sqrt{\frac{q_2}{q_1}}\varphi_2}^{\varphi_2} 2\lambda\exp\left(-\lambda\left(T+2\tau\right)\right)\varphi_1^2 dF\left(\varphi_1\right) \\ &+ \lambda\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\varphi_2^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_2\right)\exp\left(-\lambda\left(T+2\tau\right)\right).\end{aligned}$$

Define

$$P(q_1,q_2) = \lambda\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\varphi_2^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_2\right)\exp\left(-\lambda\left(T+2\tau\right)\right).$$

Then

$$\begin{aligned}\frac{\partial^2 B_1}{\partial \tau^2} &= -4\lambda \exp\left(-\lambda\left(T+2\tau\right)\right)\int_{\sqrt{\frac{q_2}{q_1}}\varphi_2}^{\varphi_2}\varphi_1^2 dF\left(\varphi_1\right) \\ &\quad +2\lambda^2\left(\frac{q_2^{\frac{3}{2}}}{q_1^{\frac{5}{2}}}\right)\exp\left(-2\lambda\left(T+2\tau\right)\right)\varphi_2^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_2\right)+\frac{\partial P}{\partial \tau}.\end{aligned}$$

Finally

$$\begin{aligned}\frac{\partial P}{\partial \tau} &= -3\lambda^2\left(\frac{q_2^{\frac{3}{2}}}{q_1^{\frac{5}{2}}}\right)\varphi_2^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_2\right)\exp\left[-2\lambda\left(T+2\tau\right)\right] \\ &\quad - \lambda^2\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\left(\frac{q_2^{\frac{1}{2}}}{q_1^{\frac{3}{2}}}\right)\varphi_2^4 f'\left(\sqrt{\frac{q_2}{q_1}}\varphi_2\right)\exp\left[-2\lambda\left(T+2\tau\right)\right] \\ &\quad - 2\lambda^2\left(\frac{q_2}{q_1}\right)^{\frac{3}{2}}\varphi_2^3 f\left(\sqrt{\frac{q_2}{q_1}}\varphi_2\right)\exp\left[-\lambda\left(T+2\tau\right)\right]\end{aligned}$$

It follows that

$$\begin{aligned}\left.\frac{\partial^2 B_1}{\partial \tau^2}\right|_{\tau=0} &= 2\frac{\lambda^2}{q_2}\varphi_2^3 f\left(\varphi_2\right)\exp\left(-2\lambda T\right)-3\frac{\lambda^2}{q_2}\varphi_2^3 f\left(\varphi_2\right)\exp\left(-2\lambda T\right) \\ &\quad -\frac{\lambda^2}{q_2}\varphi_2^4 f'\left(\varphi_2\right)\exp\left(-2\lambda T\right)-2\lambda^2\varphi_2^3 f\left(\varphi_2\right)\exp\left(-\lambda T\right) \\ &= -\left[\frac{\lambda^2}{q_2}\varphi_2^3 f\left(\varphi_2\right)+\frac{\lambda^2}{q_2}\varphi_2^4 f'\left(\varphi_2\right)\right]\exp\left(-2\lambda T\right)-2\lambda^2\varphi_2^3 f\left(\varphi_2\right)\exp\left(-\lambda T\right).\end{aligned}$$

Given this we are ready to prove the following Lemma.

LEMMA 7:

The profit function is such that

$$\left.\frac{\partial^2 \Pi\left(q,q\right)}{\partial \tau^2}\right|_{\tau=0} = -2\lambda^2\exp\left(-\lambda T\right)C \tag{A17}$$

and

$$\left.\frac{\partial^2 \Pi\left(q_1,q_2\right)}{\partial \tau^2}\right|_{\tau=0} = -4\lambda^2\exp\left(-\lambda T\right)C+2\left(\frac{\lambda^2\exp(-2\lambda T)}{1-\exp(-\lambda T)}\right)\left(\frac{1}{4\pi}\right)D \tag{A18}$$

PROOF: Expression (A17) follows directly from the observation that $\Pi(q,q) = 2qC$.
From (A8), we have that

$$\left.\frac{\partial^2 \Pi(q_1,q_2)}{\partial \tau^2}\right|_{\tau=0} = -4\lambda^2 \exp(-\lambda T) C + D \int_0^\infty \left[\left.\frac{\partial^2 B_1}{\partial t^2}\right|_{\tau=0} - \left.\frac{\partial^2 B_2}{\partial t^2}\right|_{\tau=0} \right] dF(\varphi_2)$$

where

$$\left.\frac{\partial^2 B_1}{\partial \tau^2}\right|_{\tau=0} - \left.\frac{\partial^2 B_2}{\partial \tau^2}\right|_{\tau=0} = -2\left(\frac{\lambda^2 \exp(-2\lambda T)}{1-\exp(-\lambda T)}\right) \varphi_2^3 \left[f(\varphi_2) + \varphi_2 f'(\varphi_2)\right]$$

and hence

$$\begin{aligned}\left.\frac{\partial^2 \Pi(q_1,q_2)}{\partial \tau^2}\right|_{\tau=0} = \quad & -4\lambda^2 \exp(-\lambda T) C \\ & -2\left(\frac{\lambda^2 \exp(-2\lambda T)}{1-\exp(-\lambda T)}\right) D \int_0^\infty \varphi_2^3 \left[f(\varphi_2) + \varphi_2 f'(\varphi_2)\right] dF(\varphi_2)\end{aligned}$$

Since

$$\int_0^\infty \varphi_2^3 f(\varphi_2)\, dF(\varphi_2) = \frac{1}{4\pi}\sigma_\theta^2$$

and

$$\int_0^\infty \varphi_2^4 f'(\varphi_2)\, dF(\varphi_2) = \int_0^\infty x^4 \left(-\frac{1}{2}\frac{\sqrt{2}}{\sqrt{\pi}} x e^{-\frac{1}{2}x^2}\right) f(\varphi_2)\, d\varphi_2 = -\frac{1}{2\pi}\sigma_\theta^2$$

Hence

$$\left.\frac{\partial^2 \Pi(q_1,q_2)}{\partial \tau^2}\right|_{\tau=0} = -4\lambda^2 \exp(-\lambda T) C + 2\left(\frac{\lambda^2 \exp(-2\lambda T)}{1-\exp(-\lambda T)}\right)\left(\frac{1}{4\pi}\right)\sigma_\theta^2 D$$

which concludes the proof of Lemma 7. □

We are now ready to prove Proposition 3 for the case of discrete attention.
Proof of Proposition 3(a). In the limit as τ goes to infinity, the manager observes both θ_1 and θ_2 perfectly under balanced attention ($q_1 = q_2 = q = 1$) whereas she observes shock θ_2 imperfectly under focused attention ($q_2 < q_1 = 1$). It follows that for τ sufficiently large, balanced attention is strictly preferred over focussed attention.
Proof of Proposition 3(b). We need to show that there exists a $\bar{T}$ such that for τ sufficiently small, if $T < \bar{T}$, then

$$\Pi(q_1,q_2) = \Pi(q(T,2\tau), q(T,0)) > \Pi(q,q) = \Pi(q(T,\tau), q(T,\tau))$$

and if $T > \bar{T}$, then $\Pi(q_1,q_2) < \Pi(q,q)$.

First notice that

$$\Pi(q_1,q_2)|_{\tau=0} = \Pi(q,q)|_{\tau=0} \quad \text{and} \quad \left.\frac{\partial\Pi(q_1,q_2)}{\partial\tau}\right|_{\tau=0} = \left.\frac{\partial\Pi(q,q)}{\partial\tau}\right|_{\tau=0} = 2\lambda\exp(-\lambda T)\,C$$

From Lemma 7,

$$\begin{aligned}\left.\frac{\partial^2\Pi(q_1,q_2)}{\partial\tau}\right|_{\tau=0} &= -4\lambda^2\exp(-\lambda T)\,C + \frac{\lambda^2}{2\pi}\left(\frac{\exp(-2\lambda T)}{1-\exp(-\lambda T)}\right)\sigma_\theta^2 D\\ \left.\frac{\partial^2\Pi(q,q)}{\partial\tau^2}\right|_{\tau=0} &= -2\lambda^2\exp(-\lambda T)\,C\end{aligned}$$

Define $\bar{T}$ as the (unique) solution of

$$\left.\frac{\partial^2\Pi(q_1,q_2)}{\partial\tau^2}\right|_{\tau=0} = \left.\frac{\partial^2\Pi(q,q)}{\partial\tau^2}\right|_{\tau=0}$$

which after some trivial manipulations boils down to the solution to

$$\left(\frac{1+\frac{\beta p}{1+\beta(1-p)}\left(\frac{\pi+2}{2\pi}\right)}{\frac{\beta p}{1+\beta(1-p)}}\right)\pi = \frac{\exp(-\lambda T)}{1-\exp(-\lambda T)}.$$

Then clearly for $T<\bar{T}$

$$\left.\frac{\partial^2\Pi(q_1,q_2)}{\partial\tau^2}\right|_{\tau=0} < \left.\frac{\partial\Pi(q,q)}{\partial\tau}\right|_{\tau=0}$$

and for $T>\bar{T}$

$$\left.\frac{\partial^2\Pi(q_1,q_2)}{\partial\tau^2}\right|_{\tau=0} > \left.\frac{\partial\Pi(q,q)}{\partial\tau}\right|_{\tau=0}$$

which concludes the proof.

Proof of Proposition 3(c). From (A7) and (A8), focused attention is preferred over balanced attention if and only if

$$\text{(A19)}\quad \Pi(q,q) < \Pi(q_1,q_2)$$

$$\iff 2q-q_1-q_2 \le \left(\frac{\beta p}{\left[1+\beta(1-p)+\beta p\frac{\pi+2}{2\pi}\right]\sigma_\theta^2}\right)$$

$$\text{(A20)}\quad \times 4\int_0^\infty \left[B_1(q_1,q_2)-B_2(q_1,q_2)\right]dF(\varphi_k)$$

where q, q_1 and q_2 are given by (5). $3(c)$ follows from the observation that the RHS of

(A20) is strictly increasing in β. □

PROPOSITION 3: CONTINUOUS ATTENTION. — In order to prove the continuous case, we first introduce some additional notation. Denote by

$$\omega_i \equiv T_i + t_i$$

the agent's final expertise about task i (after attention allocation) . Note that the organization's pay-offs are completely determined by the final effective expertise vector $(\lambda\omega_1, \lambda\omega_2)$. We denote

$$(\lambda\omega_1, \lambda\omega_2) \succ (\lambda\omega_1', \lambda\omega_2')$$

whenever pay-offs are higher given $(\lambda\omega_1, \lambda\omega_2)$ than $(\lambda\omega_1', \lambda\omega_2')$. To simplify notation, but wlog, we provide the proof for $\lambda = 1$. The generalization to any $\lambda > 0$ is direct.

1) Assume $t_i \in [0, 2\tau]$ with $t_1 + t_2 = 2\tau$. We first show that if $T > \Lambda$, then for τ small, $(t_1^*, t_2^*) = (\tau, \tau)$. From Proposition 3 (Discrete Attention), there exists a $\bar{\tau} > 0$, such that for $\tau < \bar{\tau}$ and $T > \Lambda$,

$$(T + \tau, T + \tau) \succ (T + 2\tau, T)$$

We now show that for any $\varepsilon \in (0, \tau)$, we also have that

$$(T + \tau, T + \tau) \succ (T + \tau + \varepsilon, T + \tau - \varepsilon)$$

The proof goes by contradiction. Assume that

$$(T + \tau, T + \tau) \prec (T + \tau + \varepsilon, T + \tau - \varepsilon)$$

Denoting $\tilde{T} = T + \tau - \varepsilon$, then this is equivalent to

$$(\tilde{T} + \varepsilon, \tilde{T} + \varepsilon) \prec (T + 2\varepsilon, \tilde{T})$$

But since $\tilde{T} > T > \Lambda$ and $\varepsilon < \tau < \bar{\tau}$, this is a contradiction of the original Proposition 3. Given the symmetry of our setting, we can show in the same manner that for any $\varepsilon \in (0, \tau)$

$$(T + \tau, T + \tau) \succ (T + \tau - \varepsilon, T + \tau + \varepsilon)$$

It follows that $(t_1^*, t_2^*) = (\tau, \tau)$.

2) Assume next that $T < \Lambda$. From Proposition 3 (Discrete Attention), for τ sufficiently small,

$$(T + \tau, T + \tau) \prec (T + 2\tau, T) \quad \text{(A21)}$$

But from Proposition 4, we then must have that

$$T + \tau < \Lambda(\rho', \beta)$$

with

$$\rho' \equiv \rho(T + \tau, T)$$

Indeed, relabel $T_1 = T + \tau$ and $T_2 = T$, and let τ be the attention that must be optimally allocated. From Propositon 4, it is optimal to allocate τ to task 1 if $T_1 < \Lambda(\rho', \beta)$ and to task 2 if $T_1 > \Lambda(\rho', \beta)$. Since $(A21)$ is equivalent to,

$$(T_1, T_2 + \tau) \prec (T_1 + \tau, T_2),$$

it is optimal to focus attention on task 1. It follows that we must have that $T_1 = T + \tau < \Lambda(\rho', \beta)$.
From Proposition 4, since $T + \tau < \Lambda(\rho', \beta)$ and relabeling $T_1 = T + \tau$ and $T_2 = T$, we have that for any $\varepsilon \in (0, \tau)$

$$(T_1 + \tau, T_2) \succ (T_1 + \tau - \varepsilon, T_2 + \varepsilon)$$

or, equivalently, for any $\varepsilon \in (0, \tau)$

$$(T + 2\tau, T) \succ (T + 2\tau - \varepsilon, T + \varepsilon)$$

Given the symmetry of our setting (task 1 and 2 are interchangeable), we also have that for any $\varepsilon \in (0, \tau)$

$$(T, T + 2\tau) \succ (T + \varepsilon, T + 2\tau - \varepsilon)$$

Since $(t_1, t_2) = (2\tau, 0)$ and $(t_1, t_2) = (0, 2\tau)$ yield the same pay-off, we obtain that given $T < \Lambda$,

$$(t_1^*, t_2^*) \in \{(2\tau, 0), (0, 2\tau)\} \tag{A22}$$

even when $t_i^* \in [0, 2\tau]$.

3) Finally, we show that for τ sufficiently large, $(t_1^*, t_2^*) = (\tau, \tau)$, even when $t_i \in [0, 2\tau]$. We prove by contradiction. Consider an attention allocation (t_1, t_2) where $t_1 > \tau > t_2$. Let us further denote

$$\bar{\Lambda} \equiv \max_{\rho} \Lambda(\rho, b),$$

where $\Lambda(\rho, b)$ is defined in Proposition 4.[2] For τ sufficiently large, we have that $T+\tau > \bar{\Lambda}$. Hence, since $t_1 > \tau$, also $T + t_1 - \varepsilon > \bar{\Lambda}$ for ε sufficiently small. But from Proposition 4, it then follows that for ε sufficiently small,

$$(T + t_1 - \varepsilon, T + t_2 + \varepsilon) \succ (T + t_1, T + t_2) \tag{A23}$$

Indeed, just relabel $T_1 = T + t_1 - \varepsilon$ and $T_2 = T + t_2$, and let ε the attention that must be optimally allocated. From Propositon 4, it is then better to focus attention ε on task 2 rather than on task 1. But from (A23), (t_1, t_2) with $t_1 > \tau > t_2$ then cannot be an optimal allocation of attention.
Consider now an attention allocation (t_1, t_2) where $t_2 > \tau > t_1$. Given the symmetry of the set-up (the two tasks are identical ex ante), we then have also that

$$(T + t_1 + \varepsilon, T + t_2 - \varepsilon) \succ (T + t_1, T + t_2) \tag{A24}$$

for ε sufficiently small. It follows that we must have $t_1 = t_2 = \tau$.

A5. Intermediate Allocations of Attention

We now discuss more formally the results regarding the intermediate allocation of attention in Section 4.3. For this purpose, it will be useful to introduce some additional notation. Consider the curve $\xi(., b) : \rho \in (1, +\infty) \to (\lambda T_1, \lambda T_2) \in \mathbb{R}^+ \times \mathbb{R}^+$ where $\xi(\rho, b) = (\xi_1(\rho, b), \xi_2(\rho, b))$ is given by

$$\begin{aligned} \xi_1(\rho, b) &\equiv \Lambda(\rho, b) \\ \hat{q}(\xi_2(\rho, b)) &\equiv \hat{q}\left(\Lambda(\rho, b)\right)/\rho \end{aligned}$$

with $\hat{q}(x) = 1 - e^{-x}$. Since $\Lambda(\rho, b)$ is a continuous function of ρ, both $\xi_1(\rho, b)$ and $\xi_2(\rho, b)$ are continuous functions of ρ and, hence, $\xi(., b)$ is a continuous mapping from $\rho \in (1, +\infty)$ to $(\lambda T_1, \lambda T_2) \in \mathbb{R}^+ \times \mathbb{R}^+$. Abusing notation slightly, we will also use $\xi(., b)$ to refer to the image of the curve, that is, the set

$$\xi(., b) = \left\{(\lambda T_1, \lambda T_1) : (\lambda T_1, \lambda T_1) = \xi(\rho, b) \text{ for } \rho > 1\right\}.$$

In Figure 4 and 5, Panels A-C, we plot $\xi(., b)$ for $b = 5/3$, as represented by the black downward-sloping curve.

REMARK 1: *For any $b > 0$, the curve $\xi(., b)$ divides the set*

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

in two connected subsets S^- and S^+, where $\{S^+, S^-, \xi(., b)\}$ is a partition of S, and

[2] From the proof of Proposition 4, $\Lambda(\rho, b)$ is finite.

where S^+ has no points in common with the closure of S^- (and vice versa). We denote by S^- the subset of S who is to the 'left' of ξ (that is, whose set closure contains $(0,0)$).

PROOF: Note that $\xi(.,b)$ is a 'simple' curve which does not cross itself. Indeed, $\rho = (1-e^{-\xi_1})/(1-e^{-\xi_2})$, so that $\rho' \neq \rho$ implies that $\xi(\rho,b) \neq \xi(\rho',b)$. Moreover, since $\rho = (1-e^{-\xi_1})/(1-e^{-\xi_2})$, we have that

$$\xi^- = \left(\xi_1^-, \xi_2^-\right) \equiv \lim_{\rho\to 1} \xi(\rho, b) \tag{A25}$$

is on the 45 degree line (where $\xi_1 = \xi_2$) and

$$\xi^+ = \left(\xi_1^+, \xi_2^+\right) \equiv \lim_{\rho\to +\infty} \xi(\rho, b) \tag{A26}$$

is on the $x-$axis (where $\xi_2^+ = 0$). Both ξ_1^- and ξ_1^+ can be shown to be finite. □

We are now ready to prove Corollaries 5 and 6 in Section 4.3.

Proof of Corollary 5 (1): Assume that $(\lambda\omega_1, \lambda\omega_2) \in S^-$. From Proposition 4, for $\varepsilon > 0$ sufficiently small, we then have that

$$(\lambda(\omega_1+\varepsilon), \lambda(\omega_2-\varepsilon)) \succ (\omega_1, \omega_2)$$

Indeed, relabel $T_1' = \omega_1 - \varepsilon$ and $T_2' = \omega_2 - \varepsilon$, and let 2ε be the attention to be allocated. Given this relabeling, for ε sufficiently small, also $(\lambda T_1', \lambda T_2') \in S^-$. Hence, from Proposition 4, for ε sufficiently small, it is then preferred to allocate attention 2ε to task 1, resulting in $(\omega_1', \omega_2') = (\omega_1+\varepsilon, \omega_2-\varepsilon)$ rather than splitting attention 2ε evenly, resulting in $(\omega_1', \omega_2') = (\omega_1, \omega_2)$.

Proof of Corollary 5 (2): Assume now that $(\lambda\omega_1, \lambda\omega_2) \in S^+$. From Proposition 4, for $\varepsilon > 0$ sufficiently small, we then have that

$$(\lambda(\omega_1-\varepsilon), \lambda(\omega_2+\varepsilon)) \succ (\omega_1, \omega_2)$$

Indeed, relabel $T_1' = \omega_1 - \varepsilon$ and $T_2' = \omega_2 - \varepsilon$, and let 2ε be the attention to be allocated. Given this relabeling, for ε sufficiently small, also $(\lambda T_1', \lambda T_2') \in S^+$. Hence, from Proposition 4, for 2ε sufficiently small, it is then preferred to allocate all attention to task 2, resulting in $(\omega_1', \omega_2') = (\omega_1-\varepsilon, \omega_2+\varepsilon)$ rather than splitting attention 2ε evenly, resulting in $(\omega_1', \omega_2') = (\omega_1, \omega_2)$.

Proof of Corollary 6: Consider now an optimal attention allocation (t_1^*, t_2^*) given T_1 and T_2. If

$$(\lambda(T_1+t_1^*), \lambda(T_2+t_2^*)) = (\lambda\omega_1, \lambda\omega_2) \in S^-$$

and $t_2^* > 0$, then corollary 1 implies that (t_1^*, t_2^*) is an inferior attention allocation relative to $(t_1^*+\varepsilon, t_2^*-\varepsilon)$. Hence, we cannot have $t_2^* > 0$. If, on the other hand,

$(\lambda\omega_1, \lambda\omega_2) \in S^+$ and $t_1^* > 0$, then corollary 1 implies that (t_1^*, t_2^*) is an inferior attention allocation relative to $(t_1^*-\varepsilon, t_2^*+\varepsilon)$. Hence, we cannot have $t_1^* > 0$.

A6. Endogenous choice of managerial expertise

To conclude, we establish the results informally discussed in Section 4.4, on endogenous managerial expertise. First, we define the function $g(Z)$ for $Z > 0$ as

$$g(Z) \equiv \{(\lambda T_1, \lambda T_2) \in S : \lambda T_2 + \lambda T_1 = Z\}.$$

Assumption A2. (i) $\xi(., b)$ is downwards sloping: $\xi_2(., b)$ is decreasing in ρ and $\xi_1(., b)$ is increasing in ρ;
(ii) the function $g(Z)$ cuts the $\xi(., b)$ at most once and always from above.[3]

We have been unable to prove that Assumption A2 is met for all $b > 0$, but also unable to generate any example where the assumption isnot met. Figure A1 at the end of this appendix plots the curve $\xi(., b)$ for a set of values of b, ranging from $b = .01$ to $b = 100$. All the curves in the figure (and every other we have tried) satisfy Assumption A2.

We are now ready to provide a more rigorous treatment of the results discussed in Section III.C regarding the intermediate allocation of attention. For a given expertise configuration $(\lambda T_1, \lambda T_2) \in S^-$, define

$$\begin{aligned} \tau_y &\equiv \sup\left\{\tau \in \mathbb{R}^+ : (\lambda T_1, \lambda(T_2 + 2\tau)) \in S^-\right\} \\ \tau_x &\equiv \sup\left\{\tau \in \mathbb{R}^+ : (\lambda(T_1 + 2\tau), \lambda T_2)) \in S^-\right\} \end{aligned}$$

Note that if $\xi(., b)$ satisfies Assumption A2, then $0 < \tau_y < \tau_x$.

PROPOSITION 8:

Assume that $\xi(., b)$ satisfies Assumption A2, then

- *If $(\lambda T_1, \lambda T_2) \in S^-$ and $2\tau < 2\tau_y$, then $(t_1^*, t_2^*) = (2\tau, 0)$*
- *If $(\lambda T_1, \lambda T_2) \in S^+$ or $2\tau > 2\tau_x$, then $(t_1^*, t_2^*) = (0, 2\tau)$ if $2\tau < T_1 - T_2$, and $T_1 + t_1^* = T_2 = t_2^*$ otherwise*

PROOF: (i) First assume that $(\lambda T_1, \lambda T_2) \in S^-$ and $2\tau < 2\tau_y$. Then given Assumption A2, $(\lambda(T_1 + t_1), \lambda(T_2 + t_2)) \in S^-$ whenever $t_1 + t_1 \leq 2\tau$. Given Corollary 5, then $(t_1^*, t_2^*) = (2\tau, 0)$.

[3] Formally, for any $Z > 0$, the line $g(Z)$ intersects $\xi(., b)$ at most once, that is, the set $\xi(., b) \cap g(Z)$ is empty or a singleton. Let $\left(\lambda\tilde{T}_1, \lambda\tilde{T}_2\right) = \xi(., b) \cap g(Z)$ then for all $(\lambda T_1, \lambda T_2) \in g(Z)$, (i) if $\lambda T_1 < \lambda\tilde{T}_1$, then $(\lambda T_1, \lambda T_2) \in S^+$ and (ii) if $\lambda T_1 > \lambda\tilde{T}_1$, then $(\lambda T_1, \lambda T_2) \in S^-$.

(ii) Next assume that $(\lambda T_1, \lambda T_2) \in S^-$ but $2\tau > 2\tau_x$. Given Assumption A2, if $2\tau < T_1 - T_2$, then $(\lambda(T_1 + t_1), \lambda(T_2 + t_2)) \in S^+$ whenever $t_1 + t_1 = 2\tau$. Given Corollary 5, it follows that $(t_1^*, t_2^*) = (0, 2\tau)$. If $(\lambda T_1, \lambda T_2) \in S^-$ but $2\tau > T_1 - T_2$, then it is possible that $(t_1, t_2) = (\hat{t}_1, \hat{t}_2)$ is such that $T_2 + \hat{t}_2 > T_1 + \hat{t}_1$. But expected profits in the latter case are identical as when $(t_1, t_2) = (\tilde{t}_1, \tilde{t}_2)$ with $(T_1 + \tilde{t}_1, T_2 + \tilde{t}_2) = (T_2 + \hat{t}_2, T_1 + \hat{t}_1)$. It is therefore without loss of generality to only consider allocations of attention (t_1, t_2) for which $T_2 + t_2 \leq T_1 + t_1$. Given $2\tau > 2\tau_x$ and given $2\tau > T_1 - T_2$, then whenever $t_1 + t_1 = 2\tau$, either $T_1 + t_1 = T_2 + t_2$ or $(\lambda(T_1 + t_1), \lambda(T_2 + t_2)) \in S^+$. But whenever $(\lambda(T_1 + t_1), \lambda(T_2 + t_2)) \in S^+$ profits can be improved by shifting attention from task 1 to task 2. It follows that (t_1^*, t_2^*) is such that $T_1 + t_1^* = T_2 + t_2^*$.
(iii) Finally, if $(\lambda T_1, \lambda T_2) \in S^+$, the same arguments apply as in (ii). □

To conclude, assume as in section III.D, that $2\tau = 0$, but the expertise of managers (T_1, T_2) is endogenously chosen under the constraint $\lambda(T_1 + T_2) \leq Z$. The following proposition generalizes the insights of section III.D.

PROPOSITION 9:

If managerial expertise (T_1, T_2) is optimally chosen under the constraint $\lambda(T_1 + T_2) \leq Z$ and $\xi(., b)$ satisfies Assumption A2, then

- *If $Z < 2\xi_1^-$, with ξ_1^- defined in (A25), then the optimal manager is a specialist: $(\lambda T_1^*, \lambda T_2^*) = (Z, 0)$ or $(\lambda T_1^*, \lambda T_2^*) = (0, Z)$.*
- *If $Z > \xi_1^+$, with ξ_1^+ defined in (A26), then the optimal manager is a generalist: $(\lambda T_1^*, \lambda T_2^*) = (Z/2, Z/2)$.*

PROOF: If $Z < 2\xi^-$, then Assumption A2 guarantees that $g(Z) \subset S^-$. Given $T_1 \geq T_2$, Corollary 5 then implies that profits can always be improved by shifting expertise from task 2 to task 1. Similarly, if $Z > \xi^+$, then under Assumption A2, $g(Z) \subset S^+$. Given $T_1 \geq T_2$, Corollary 5 then implies that profits can always be improved by shifting expertise from task 1 to task 2 up to the point where $T_1 = T_2$. □

For completeness, we end this section by providing expressions for ξ^- and ξ^+ and show that we always have that $2\xi^- < \xi^+$:
(i) From (A10) and (A11), we have that ξ_1^+ is implicitly given by

$$\text{(A27)} \qquad e^{-\xi_1^+} = \frac{\left(1 + \frac{\pi+2}{2\pi} b\right) \sigma_\theta^2 - 4b \int_0^\infty \left(\int_0^{\theta_j} \theta_i^2 dF(\theta_i) \right) dF(\theta_j)}{\left(1 + \frac{\pi+2}{2\pi} b\right) \sigma_\theta^2 + 4b \int_0^\infty \left(\int_{\theta_j}^\infty \theta_i^2 dF(\theta_i) \right) dF(\theta_j)},$$

$$\text{(A28)} \qquad = \frac{\left(1 + \frac{\pi+2}{2\pi} b\right) - \frac{\pi+2}{2\pi} b}{\left(1 + \frac{\pi+2}{2\pi} b\right) + \frac{\pi-2}{2\pi} b} = \frac{1}{1+b}$$

from which

$$\xi_1^+ = -\ln\left(\frac{1}{1+b}\right)$$

(ii) From Proposition 4, ξ_1^- is finite for ρ close to 1 and is continuous in ρ. Hence, by continuity, also $\xi_1^- = \lim_{\rho\to 1}\xi_1(\rho, b)$ is finite. Moreover, ξ_1^- is implicitly given by

$$\frac{e^{-\xi_1^-}}{1-e^{-\xi_1^-}} = \left(\frac{1+b\left(\frac{\pi+2}{2\pi}\right)}{b}\right)\pi$$

(iii) Finally, note that ξ_1^- is smaller than $\xi_1^+/2$. Indeed, we have that

$$\frac{e^{-\xi_1^+/2}}{1-e^{-\xi_1^+/2}} = \frac{\sqrt{e^{-\xi_1^+}}}{1-\sqrt{e^{-\xi_1^+}}} = \frac{\sqrt{\frac{1}{1+b}}}{1-\sqrt{\frac{1}{1+b}}} < \left(\frac{1+b\left(\frac{\pi+2}{2\pi}\right)}{b}\right)\pi = \frac{e^{-\xi_1^-}}{1-e^{-\xi_1^-}}$$

Since $\frac{e^{-x}}{1-e^{-x}}$ is decreasing in x, it follows that $\xi_1^+/2 > \xi_1^-$.

FIGURE A1. *The downward sloping lines are the function $\Lambda(\rho, b)$ plotted in the space $(\lambda T_1, \lambda T_2)$ for different values of b (see expression (15)), ranging from 10^{-2} (the first downward sloping curve from the right) to 10^{2}. Below the downward sloping curves the manager allocates the marginal unit of attention to task 1, whereas above she allocates the marginal unit of attention to task 2.*