Supplementary Appendices

(to be posted online only)

Appendix S1: Sequential Decision Making

Decision Making: In the last stage of the game Manager 2 chooses d_2 to maximize his expected utility $\mathrm{E}\left[(1-\lambda)\pi_1 + \lambda\pi_2 \middle|\, \theta_2,\, d_1\right]$. The optimal decision that solves this problem is given by

$$d_2^S \equiv \frac{\lambda}{\lambda+\delta}\theta_2 + \frac{\delta}{\lambda+\delta}d_1^S. \tag{S1.1}$$

At the previous stage Manager 1 chooses d_1 to maximize his expected profits $\mathrm{E}\left[\lambda\pi_1 + (1-\lambda)\pi_2 \middle|\, \theta_1,\, m\right]$. The optimal decision is given by

$$d_1^S \equiv \frac{\lambda\left(\lambda+\delta\right)^2}{\lambda^3+3\lambda^2\delta+\delta^2}\theta_1 + \delta\frac{\lambda^2+\delta\left(1-\lambda\right)}{\lambda^3+3\lambda^2\delta+\delta^2}\mathrm{E}\left[\theta_2 \mid \theta_1,\, m\right]. \tag{S1.2}$$

Communication: Let $\mu_2\left(m_2 \mid \theta_2\right)$ be the probability with which Manager 2 sends message m_2, let $d_1(m_2)$ and $d_2(m_2)$ be the decision rules that map messages into decisions and let $g_1\left(\theta_2 \mid m_2\right)$ be the belief function which gives the probability of θ_2 conditional on observing m_2. We can now state the following proposition which characterizes the finite communication equilibria when $\delta > 0$.

PROPOSITION S1.1 (Communication Equilibria). *If $\delta \in (0,\infty)$, then for every positive integer N_2 there exists at least one equilibrium $(\mu_2(\cdot), d_1(\cdot), d_2(\cdot), g_1(\cdot))$, where*

i. $\mu_2(m_2 \mid \theta_2)$ *is uniform, supported on* $[a_{2,i-1}, a_{2,i}]$ *if* $\theta_2 \in (a_{2,i-1}, a_{2,i})$,

ii. $g_1(\theta_2 \mid m_2)$ *is uniform supported on* $[a_{2,i-1}, a_{2,i}]$ *if* $m_2 \in (a_{2,i-1}, a_{2,i})$,

iii. $a_{2,i+1} - a_{2,i} = a_{2,i} - a_{2,i-1} + 4b_S a_{2,i}$ *for* $i = 1, ..., N_2 - 1$,

$a_{2,-(i+1)} - a_{2,-i} = a_{2,-i} - a_{2,-(i-1)} + 4b_S a_{2,-i}$ *for* $i = 1, ..., N_j - 1$,

where $b_S \equiv \left((2\lambda-1)\left(\lambda+\delta\right)\left(\lambda^2+\delta\right)\right) / \left(\left(\lambda\left(1-\lambda\right)+\delta\right)\left(\lambda^2+\delta\left(1-\lambda\right)\right)\right)$ *and*

iv. $d_j(m) = d_j^S$, $j = 1, 2$, *where* d_j^S *is given by (S1.1) and (S1.2).*

Moreover, all other finite equilibria have relationships between θ_1 and θ_2 and the managers' choices of d_1 and d_2 that are the same as those in this class for some value of N_2; they are therefore economically equivalent.

Proof: The proof is analogous to the proof of Proposition 1. Details are available from the authors upon request.

PROPOSITION S1.2 (Efficiency). *The limit of strategy profiles and beliefs* $(\mu_2(\cdot)\ , d_1(\cdot), d_2(\cdot), g_1(\cdot))$ *as* $N_2 \to \infty$ *is a Perfect Bayesian Equilibrium of the communication game. In this equilibrium the total expected profits* $\mathrm{E}[\pi_1 + \pi_2]$ *are higher than in any other equilibrium.*

Proof: The proof is analogous to the proof of Proposition 2. Details are available from the authors upon request.

In the remaining analysis we focus on the efficient equilibrium.

LEMMA S1.1. *In the most efficient equilibrium in which* $N_2 \to \infty$ *the residual variance is given by*

$$\mathrm{E}\left[(\theta_2 - \mathrm{E}[\theta_2|m_2])^2\right] = S_S \sigma_2^2,$$

where $S_S = b_S/(3 + 4b_S)$.

Proof: The proof is analogous to the proof of Lemma 1. Details are available from the authors upon request.

PROPOSITION S1.3 (Organizational Performance). *Under Decentralization with sequential decision making the expected profits are given by*

$$\Pi_S = -\left((A_D + X)\left(\sigma_1^2 + \sigma_2^2\right) + (B_D - X)\, S_S \sigma_2^2\right), \tag{S1.3}$$

where A_D *and* B_D *are defined in (17) and*

$$X \equiv \delta^3 (2\lambda - 1)^2 \frac{2\lambda^4 + \lambda^2 (6\lambda + 1)\delta + 2\lambda (2 + \lambda)\delta^2 + 2\delta^3}{(\lambda + 2\delta)^2 \left(\lambda^3 + \delta^2 + 3\lambda^2 \delta\right)^2}.$$

Proof: The proof is analogous to the proof of Proposition 4. Details are available from the authors upon request.

We can now prove the following proposition.

PROPOSITION S1.3 (Sequential Decision Making). *Suppose that* $\sigma_1^2 = \sigma_2^2 = \sigma^2$. *Then,*

- *i.* *For any* $\lambda \in (1/2, 1]$ *Centralization strictly dominates Decentralization with* sequential *decision making when coordination is sufficiently important.*
- *ii.* *For any* $\delta \in (0, \infty)$ *Decentralization with sequential decision making strictly dominates Centralization when the own-division bias* $\lambda > 1/2$ *is sufficiently small.*

Proof: i. Applying l'Hopital's Rule to (15) and (S1.3) and using the assumption that $\sigma_1^2 = \sigma_2^2 = \sigma^2$ we obtain

$$\lim_{\delta\to\infty} \Pi_C - \lim_{\delta\to\infty} \Pi_S = \frac{8\lambda\,(4\lambda-1)\,(2\lambda-1)^2}{(8\lambda-1)\,(5\lambda-1)}\sigma^2$$

which is strictly positive for any $\lambda > 1/2$.

ii. Taking the derivative of (15) and (S1.3) and using the assumption that $\sigma_1^2 = \sigma_2^2 = \sigma^2$ we get that

$$\frac{\mathrm{d}\,(\Pi_S - \Pi_C)}{\mathrm{d}\lambda} = \frac{8\delta}{3\,(1+2\delta)\,(1+4\delta)}\sigma^2 \quad \text{for } \lambda = 1/2$$

which is strictly positive for all finite $\delta > 0$.

Finally, Figures 7 and 8 are drawn using Propositions 4 and S1.3.

Appendix S2: Different Needs for Coordination.

Since allowing for differences in the needs for coordination only requires adding a parameter in the main model, we do not replicate the full analysis here. Instead we merely state the key expressions and use them to prove Proposition S2.1 which summarizes the claims in the main text. The derivation of these expressions and their interpretation are exactly as in the main model. Also, to simplify we assume that $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

Centralization: The decisions are now given by

$$\begin{aligned} d_1^C &\equiv \left(\frac{1}{1+2(\delta_1+\delta_2)} \left((1+\delta_1+\delta_2) \, \mathrm{E}[\theta_1 \mid m] + (\delta_1+\delta_2) \, \mathrm{E}[\theta_2 \mid m] \right) \right) \\ d_2^C &\equiv \left(\frac{1}{1+2(\delta_1+\delta_2)} \left((\delta_1+\delta_2) \, \mathrm{E}[\theta_1 \mid m] + n(1+\delta_1+\delta_2) \, \mathrm{E}[\theta_2 \mid m] \right) \right). \end{aligned}$$

The residual variance of θ_1 is given by $S_{C,1}\sigma_1^2$ and that of θ_2 is given by $S_{C,2}\sigma_2^2$, where $S_{C,j} \equiv b_{C,j}/\left(3+4b_{C,j}\right)$, $j=1,2$, and

$$\begin{aligned} b_{C,1} &= \frac{(2\lambda-1)\left(\delta_2 + (\delta_1+\delta_2)^2\right)}{\delta_2 + (\delta_1+\delta_2)^2 + \lambda(1+3\delta_1+\delta_2)} \\ b_{C,2} &= \frac{(2\lambda-1)\left(\delta_1 + (\delta_1+\delta_2)^2\right)}{\delta_1 + (\delta_1+\delta_2)^2 + \lambda(1+\delta_1+3\delta_2)}. \end{aligned}$$

The expected profits are given by

$$\Pi_C = -\sigma^2 \left(2\frac{\delta_1+\delta_2}{1+2(\delta_1+\delta_2)} + (S_1+S_2) \frac{1+\delta_1+\delta_2}{1+2(\delta_1+\delta_2)} \right). \tag{S2.1}$$

Applying l'Hopital's Rule gives

$$\lim_{\delta_1 \to \infty} \Pi_C = -2\frac{5\lambda-1}{8\lambda-1}\sigma^2. \tag{S2.2}$$

We can also use (S2.1) to evaluate $\mathrm{d}\Pi_C/\mathrm{d}\lambda$:

$$\frac{\mathrm{d}\Pi_C}{\mathrm{d}\lambda} = -\frac{4}{3}\frac{(\delta_1+\delta_2)}{1+2(\delta_1+\delta_2)}\sigma^2 \quad \text{for } \lambda = 1/2. \tag{S2.3}$$

Decentralization: The decisions under Decentralization are now given by

$$
\begin{aligned}
d_1^D &= \frac{\lambda\theta_1}{\lambda+\lambda\delta_1+(1-\lambda)\delta_2}+\frac{((1-\lambda)\delta_1+\lambda\delta_2)(\lambda\delta_1+(1-\lambda)\delta_2)}{(\lambda+\delta_1+\delta_2)(\lambda+\lambda\delta_1+(1-\lambda)\delta_2)}\mathrm{E}[\theta_1 \mid \theta_2, m] \\
&\quad +\frac{\lambda\delta_1+(1-\lambda)\delta_2}{\lambda+\delta_1+\delta_2}\mathrm{E}[\theta_2 \mid \theta_1, m] \\
d_2^D &= \frac{\lambda\theta_2}{\lambda+(1-\lambda)\delta_1+\lambda\delta_2}+\frac{(1-\lambda)\delta_1+\lambda\delta_2}{\lambda+\delta_1+\delta_2}\mathrm{E}[\theta_1 \mid \theta_2, m] \\
&\quad +\frac{((1-\lambda)\delta_1+\lambda\delta_2)(\lambda\delta_1+(1-\lambda)\delta_2)}{(\lambda+\delta_1+\delta_2)(\lambda+(1-\lambda)\delta_1+\lambda\delta_2)}\mathrm{E}[\theta_2 \mid \theta_1, m].
\end{aligned}
$$

The residual variance of θ_1 is given by $S_{D,1}\sigma_1^2$ and that of θ_2 is given by $S_{D,2}\sigma_2^2$, where $S_{D,j} \equiv b_{C,j}/(3+4b_{C,j})$, $j=1,2$, and

$$
\begin{aligned}
b_1 &= \frac{(2\lambda-1)\delta_1(\lambda+\lambda\delta_1+(1-\lambda)\delta_2)}{(\lambda(1-\lambda)+\lambda\delta_1+(1-\lambda)\delta_2)((1-\lambda)\delta_1+\lambda\delta_2)} \\
b_2 &= \frac{(2\lambda-1)\delta_2(\lambda+\lambda\delta_2+(1-\lambda)\delta_1)}{(\lambda(1-\lambda)+\lambda\delta_2+(1-\lambda)\delta_1)((1-\lambda)\delta_2+\lambda\delta_1)}.
\end{aligned}
$$

The expected profits are given by

$$
\Pi_D = -\mathrm{E}\left[\left(d_1^D-\theta_1\right)^2+\left(d_2^D-\theta_2\right)^2+(\delta_1+\delta_2)\left(d_1^D-d_2^D\right)^2\right], \tag{S2.4}
$$

where

$$
\begin{aligned}
&\mathrm{E}\left[\left(d_1^D-\theta_1\right)^2\right] \\
&= \sigma^2\left(2\frac{(\delta_2+\lambda\delta_1-\lambda\delta_2)^2}{(\lambda+\delta_1+\delta_2)^2}-\frac{(\lambda\delta_1+(1-\lambda)\delta_2)^2}{(\lambda+\delta_1+\delta_2)^2}S_2\right. \\
&\quad \left.+\frac{((1-\lambda)\delta_1+\lambda\delta_2)(\lambda\delta_1+(1-\lambda)\delta_2)^2(2\lambda+(1+\lambda)\delta_1+(2-\lambda)\delta_2)}{(\lambda+\delta_1+\delta_2)^2(\lambda+\lambda\delta_1+(1-\lambda)\delta_2)^2}S_1\right) \\
&\mathrm{E}\left[(d_2-\theta_2)^2\right] \\
&= \sigma^2\left(2\frac{(-\delta_1+\lambda\delta_1-\lambda\delta_2)^2}{(\lambda+\delta_1+\delta_2)^2}-\frac{((1-\lambda)\delta_1+\lambda\delta_2)^2}{(\lambda+\delta_1+\delta_2)^2}S_1\right. \\
&\quad \left.+\frac{((1-\lambda)\delta_1+\lambda\delta_2)^2(\lambda\delta_1+(1-\lambda)\delta_2)(2\lambda+(2-\lambda)\delta_1+(1+\lambda)\delta_2)}{(\lambda+\delta_1+\delta_2)^2(\lambda+(1-\lambda)\delta_1+\lambda\delta_2)^2}S_2\right) \\
&\mathrm{E}\left[\left(d_1^D-d_2^D\right)^2\right] \\
&= \sigma^2\left(2\frac{\lambda^2}{(\lambda+\delta_1+\delta_2)^2}+\frac{\lambda^2(2\lambda+(1+\lambda)\delta_1+(2-\lambda)\delta_2)((1-\lambda)\delta_1+\lambda\delta_2)}{(\lambda+\delta_1+\delta_2)^2(\lambda+\lambda\delta_1+(1-\lambda)\delta_2)^2}S_1\right. \\
&\quad \left.+\frac{\lambda^2(2\lambda+(2-\lambda)\delta_1+(1+\lambda)\delta_2)(\lambda\delta_1+(1-\lambda)\delta_2)}{(\lambda+\delta_1+\delta_2)^2(\lambda+(1-\lambda)\delta_1+\lambda\delta_2)^2}S_2\right).
\end{aligned}
$$

Applying l'Hopital's Rule gives

$$\lim_{\delta_1 \to \infty} \Pi_D = -2\frac{8\lambda^3 - 9\lambda^2 + 6\lambda - 1}{5\lambda - 1}\sigma^2. \tag{S2.5}$$

We can also use (S2.4) to evaluate $\mathrm{d}\Pi_C/\mathrm{d}\lambda$ for $\lambda = 1/2$:

$$\frac{\mathrm{d}\Pi_D}{\mathrm{d}\lambda} = -\frac{8}{3}\frac{(\delta_1 + \delta_2)^2}{(1 + 2(\delta_1 + \delta_2))^2}\sigma^2. \tag{S2.6}$$

We can now prove the following proposition.

PROPOSITION S2.1 (Different Needs for Coordination).

i. *For any* $\lambda \in (1/2, 1]$ and $\delta_j \in [0, \infty)$, $j = 1, 2$, *Centralization strictly dominates Decentralization when coordination is sufficiently important for Division* $k \neq j$.

ii. *For any* $\delta_1, \delta_2 \in (0, \infty)$ *Decentralization strictly dominates Centralization dominates Centralization when the own-division bias* $\lambda > 1/2$ *is sufficiently small.*

Proof: i. Using (S2.2) and (S2.5) we obtain

$$\lim_{\delta \to \infty} \Pi_C - \lim_{\delta \to \infty} \Pi_S = 8\lambda(4\lambda - 1)\frac{(2\lambda - 1)^2}{(8\lambda - 1)(5\lambda - 1)}\sigma^2$$

which is strictly positive for any $\lambda > 1/2$.

ii. Using (S2.3) and (S2.6) we find that the difference in the derivatives at $\lambda = 1/2$ is given by

$$\frac{\mathrm{d}(\Pi_C - \Pi_S)}{\mathrm{d}\lambda} = \frac{4}{3}\frac{(\delta_1 + \delta_2)}{(1 + 2(\delta_1 + \delta_2))^2}\sigma^2 \quad \text{for } \lambda = 1/2$$

which is strictly positive for all finite $\delta_1, \delta_2 > 0$.

Appendix S3: Different Division Sizes.

Since allowing for different division sizes only requires adding a parameter in the main model, we do not replicate the full analysis here. Instead we merely state the key expressions and use them to prove Proposition S3.1 which summarizes the claims in the main text. The derivation of these expressions and their interpretation are exactly as in the main model. Also, to simplify we assume that $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

Centralization: Let $\beta \equiv (1-\alpha)$. Then decisions are given by

$$\begin{aligned} d_1^C &\equiv \left(\frac{1}{\alpha\beta+\delta}\left(\alpha\left(\beta+\delta\right)\mathrm{E}\left(\theta_1 \mid m\right)+\beta\delta\mathrm{E}\left(\theta_2 \mid m\right)\right)\right) \\ d_2^C &\equiv \left(\frac{1}{\alpha\beta+\delta}\left(\alpha\delta\mathrm{E}\left(\theta_1 \mid m\right)+\beta\left(\alpha+\delta\right)\mathrm{E}\left(\theta_2 \mid m\right)\right)\right). \end{aligned}$$

The residual variance of θ_1 is given by $S_{C,1}\sigma_1^2$ and that of θ_2 is given by $S_{C,2}\sigma_2^2$, where $S_{C,j} \equiv b_{C,j}/\left(3+4b_{C,j}\right)$, $j=1,2$, and

$$\begin{aligned} b_{C,1} &= \frac{\beta\delta\left(2\lambda-1\right)\left(\beta^2+\delta\right)}{\alpha\lambda\left(\beta^2+\delta^2\right)+\beta\left(\delta+\beta^2\right)\left(1-\lambda\right)\delta+\alpha\beta\lambda\left(2+\beta\right)\delta} \\ b_{C,2} &= \frac{\alpha\delta\left(2\lambda-1\right)\left(\alpha^2+\delta\right)}{\alpha\left(\delta+\alpha^2\right)\left(1-\lambda\right)\delta+\beta\lambda\left(\alpha^2+\delta^2\right)+\alpha\beta\lambda\left(2+\alpha\right)\delta}. \end{aligned}$$

The expected profits are given by

$$\Pi_C = -\mathrm{E}\left[\left(d_1^C-\theta_1\right)^2+\left(d_2^C-\theta_2\right)^2+2\delta\left(d_1^C-d_2^C\right)^2\right],$$

where

$$\begin{aligned} \mathrm{E}\left[\left(d_1-\theta_1\right)^2\right] &= \sigma^2\left(\frac{2\delta^2\beta^2}{\left(\alpha\beta+\delta\right)^2}+\alpha\left(\beta+\delta\right)\frac{\alpha\beta+\left(2-\alpha\right)\delta}{\left(\alpha\beta+\delta\right)^2}S_1-\frac{\delta^2\beta^2}{\left(\alpha\beta+\delta\right)^2}S_2\right) \\ \mathrm{E}\left[\left(d_2-\theta_2\right)^2\right] &= \sigma^2\left(\frac{2\alpha^2\delta^2}{\left(\alpha\beta+\delta\right)^2}-\frac{\alpha^2\delta^2}{\left(\alpha\beta+\delta\right)^2}S_1+\beta\left(\alpha+\delta\right)\frac{\alpha\beta+\left(1+\alpha\right)\delta}{\left(\alpha\beta+\delta\right)^2}S_2\right) \\ \mathrm{E}\left[\left(d_1-d_2\right)^2\right] &= \sigma^2\left(\frac{2\alpha^2\beta^2}{\left(\alpha\beta+\delta\right)^2}-\frac{\alpha^2\beta^2}{\left(\alpha\beta+\delta\right)^2}\left(S_1+S_2\right)\right). \end{aligned}$$

Applying l'Hopital's Rule we find that

$$\lim_{\delta\to\infty}\Pi_C = \frac{-2\alpha\left(1-\alpha\right)\left(8\lambda-1\right)\left(5\lambda-1\right)}{\left(5\lambda-1-\alpha\left(2\lambda-1\right)\right)\left(3\lambda+\left(2\lambda-1\right)\alpha\right)}\sigma^2. \tag{S3.1}$$

Also, differentiating we find that

$$\frac{\mathrm{d}\Pi_C}{\mathrm{d}\lambda} = -\frac{8}{3}\frac{\alpha(1-\alpha)\delta}{\alpha(1-\alpha)+\delta}\sigma^2 \quad \text{for } \lambda = 1/2. \tag{S3.2}$$

Decentralization: The decisions are now given by

$$\begin{aligned}
d_1^D &= \frac{\left(\alpha\lambda\theta_1 + \delta\left(\alpha\lambda + \beta\left(1-\lambda\right)\right)\mathrm{E}\left(d_2^D \mid \theta_1, m\right)\right)}{\alpha\lambda\left(1+\delta\right) + \beta\delta\left(1-\lambda\right)} \\
d_2^D &= \frac{\left(\beta\lambda\theta_2 + \delta\left(\alpha\left(1-\lambda\right) + \beta\lambda\right)\mathrm{E}\left(d_1^D \mid \theta_2, m\right)\right)}{\beta\lambda\left(1+\delta\right) + \alpha\delta\left(1-\lambda\right)}
\end{aligned}$$

where $\beta \equiv (1-\alpha)$ and

$$\begin{aligned}
\mathrm{E}\left[d_1^D \mid \theta_2, m\right] &= \frac{\left(\alpha\left(\alpha\delta\left(1-\lambda\right) + \beta\lambda\left(1+\delta\right)\right)\mathrm{E}\left[\theta_1 \mid \theta_2, m\right] + \beta\delta\left(\alpha\lambda + \beta\left(1-\lambda\right)\right)\mathrm{E}\left[\theta_2 \mid \theta_1, m\right]\right)}{\left(\alpha^2+\beta^2\right)\delta\left(1-\lambda\right) + \alpha\beta\lambda\left(1+2\delta\right)} \\
\mathrm{E}\left[d_2^D \mid \theta_1, m\right] &= \frac{\left(\alpha\delta\left(\alpha\left(1-\lambda\right) + \beta\lambda\right)\mathrm{E}\left[\theta_1 \mid \theta_2, m\right] + \beta\left(\alpha\lambda\left(1+\delta\right) + \beta\delta\left(1-\lambda\right)\right)\mathrm{E}\left[\theta_2 \mid \theta 1 m\right]\right)}{\left(\alpha^2+\beta^2\right)\delta\left(1-\lambda\right) + \alpha\beta\lambda\left(1+2\delta\right)}.
\end{aligned}$$

The residual variance of θ_1 is given by $S_{D,1}\sigma_1^2$ and that of θ_2 is given by $S_{D,2}\sigma_2^2$, where $S_{D,j} \equiv b_{D,j}/\left(3+4b_{C,j}\right)$, $j=1,2$, and

$$\begin{aligned}
b_{D,1} &= \frac{\alpha\beta\left(2\lambda-1\right)\left(\alpha\lambda\left(1+\delta\right) + \beta\left(1-\lambda\right)\delta\right)}{\left(\alpha\left(1-\lambda\right)+\beta\lambda\right)\left(\beta^2\left(1-\lambda\right)^2\delta + \alpha^2\lambda^2\delta + \alpha\beta\lambda\left(1+2\delta\right)\left(1-\lambda\right)\right)} \\
b_{D,2} &= \frac{\alpha\beta\left(2\lambda-1\right)\left(\alpha\left(1-\lambda\right)\delta + \beta\lambda\left(1+\delta\right)\right)}{\left(\beta+\alpha\lambda-\beta\lambda\right)\left(\alpha^2\left(1-\lambda\right)^2\delta + \beta^2\lambda^2\delta + \alpha\beta\lambda\left(2\delta+1\right)\left(1-\lambda\right)\right)}
\end{aligned}$$

The expected profits are given by

$$\Pi_D = -\mathrm{E}\left[\left(d_1^D-\theta_1\right)^2 + \left(d_2^D-\theta_2\right)^2 + 2\delta\left(d_1^D-d_2^D\right)^2\right],$$

where

$$\begin{aligned}
&\mathrm{E}\left[\left(d_1^D-\theta_1\right)^2\right] \\
= &\frac{\sigma^2}{\left(\left(\alpha^2+\beta^2\right)\left(1-\lambda\right)\delta + \alpha\beta\lambda\left(1+2\delta\right)\right)^2}\left(\frac{2\delta^2\beta^2\left(\alpha\lambda\left(1+\delta\right)+\beta\left(1-\lambda\right)\delta\right)^2\left(\alpha\lambda+\beta\left(1-\lambda\right)\right)^2}{\left(\alpha\lambda\left(1+\delta\right)+\beta\left(1-\lambda\right)\delta\right)^2}\right. \\
&+\frac{\alpha\delta^3\left(\alpha\left(1-\lambda\right)+\beta\lambda\right)\left(\alpha\lambda+\beta\left(1-\lambda\right)\right)^2\left(\left(\alpha^2+2\beta^2\right)\left(1-\lambda\right)\delta + \alpha\beta\lambda\left(2+3\delta\right)\right)S_1}{\left(\alpha\lambda\left(1+\delta\right)+\beta\left(1-\lambda\right)\delta\right)^2} \\
&\left.-\beta^2\delta^2\left(\alpha\lambda+\beta\left(1-\lambda\right)\right)^2 S_2\right)
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\left[\left(d_2^D-\theta_2\right)^2\right] \\
= \quad &\frac{\sigma^2}{\left(\left(\alpha^2+\beta^2\right)(1-\lambda)\delta+\alpha\beta\lambda(1+2\delta)\right)^2}\left(2\delta^2\alpha^2\left(\alpha(1-\lambda)+\beta\lambda\right)^2-\alpha^2\delta^2\left(\alpha(1-\lambda)+\beta\lambda\right)^2 S_1\right. \\
&\left.+\frac{\beta\delta^3\left(\alpha(1-\lambda)+\beta\lambda\right)^2\left(\alpha\lambda+\beta(1-\lambda)\right)\left(\delta\left(2\alpha^2+\beta^2\right)(1-\lambda)+\alpha\beta\lambda(2+3\delta)\right)S_2}{\left(\alpha(1-\lambda)\delta+\beta\lambda(1+\delta)\right)^2}\right)
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\left[\left(d_1^D-d_2^D\right)^2\right] \\
= \quad &\frac{\sigma^2\lambda^2}{\left(\left(\alpha^2+\beta^2\right)(1-\lambda)\delta+\alpha\beta\lambda(1+2\delta)\right)^2}\left(2\alpha^2\beta^2\right. \\
&+\frac{\alpha^3\delta\left(\left(\alpha^2+2\beta^2\right)(1-\lambda)\delta+\alpha\beta\lambda(3\delta+2)\right)\left(\alpha(1-\lambda)+\beta\lambda\right)S_1}{\left(\alpha\lambda(1+\delta)+\beta(1-\lambda)\delta\right)^2} \\
&\left.+\frac{\beta^3\delta\left(\alpha\lambda(1+\delta)+\beta\delta(1-\lambda)\right)\left(\delta\left(2\alpha^2+\beta^2\right)(1-\lambda)+\alpha\beta\lambda(2+3\delta)\right)\left(\alpha\lambda+\beta(1-\lambda)\right)S_2}{\left(\alpha\lambda(1+\delta)+\beta\delta(1-\lambda)\right)\left(\alpha(1-\lambda)\delta-\beta\lambda(1+\delta)\right)^2}\right).
\end{aligned}$$

Applying l'Hopital's Rule we find that

$$\lim_{\delta\to\infty}\Pi_D=\frac{-2\alpha(1-\alpha)\left(2\alpha(1-\alpha)+2\lambda^2(2\alpha-1)^2(3\lambda-7)-26\alpha\lambda(1-\alpha)+9\lambda-1\right)}{\left(1-2\alpha(1-\alpha)-\lambda(2\alpha-1)^2\right)\left(3\lambda(1-\lambda)+\alpha(1-\alpha)(6\lambda+1)(2\lambda-1)\right)}\sigma^2. \tag{S3.3}$$

Also, differentiating we find that at $\lambda=1/2$

$$\frac{\mathrm{d}\Pi_D}{\mathrm{d}\lambda}=-\frac{16}{3}\alpha(1-\alpha)\delta^2\frac{1-2\alpha(1-\alpha)}{\left(\alpha(1-\alpha)+\delta\right)^2}\sigma^2. \tag{S3.4}$$

We can now prove the following proposition.

PROPOSITION S3.1 (Different Division Sizes).

i. *For any $\lambda>1/2$ and $\alpha>1/2$ Centralization strictly dominates Decentralization when coordination is sufficiently important.*

ii. *For any $\delta\in(0,\infty)$ Decentralization strictly dominates Centralization when the own-division bias $\lambda>1/2$ and the difference in the division sizes $\alpha>1/2$ are sufficiently small.*

Proof: (i.) Using (S3.1) and (S3.3) gives

$$\begin{aligned}
&\lim_{\delta\to\infty}\Pi_C-\lim_{\delta\to\infty}\Pi_D \\
= \quad &\frac{6\alpha\lambda(1-\alpha)(2\alpha-1)^2(2\lambda-1)}{\left(1-2\alpha(1-\alpha)-\lambda(2\alpha-1)^2\right)\left(5\lambda-1-\alpha(2\lambda-1)\right)} \\
&\times\frac{\left(\alpha(1-\alpha)(2\lambda-1)\left(42\lambda^2-11\lambda+1\right)+\lambda(1-\lambda)(5\lambda-1)\right)}{\left((2\lambda-1)\alpha+3\lambda\right)\left(\alpha(1-\alpha)(6\lambda+1)(2\lambda-1)+3\lambda(1-\lambda)\right)}\sigma^2
\end{aligned}$$

which is strictly positive if $\lambda > 1/2$ and $\alpha > 1/2$.

(ii.) Using (S3.2) and (S3.4) we find that the difference in the derivatives at $\lambda = 1/2$ is given by

$$\frac{\mathrm{d}\left(\Pi_D - \Pi_C\right)}{\mathrm{d}\lambda} = \frac{8}{3}\alpha\left(1-\alpha\right)\delta\frac{\alpha\left(1-\alpha\right)\left(1+4\delta\right)-\delta}{\left(\alpha\left(1-\alpha\right)+\delta\right)^2}\sigma^2 \quad \text{for } \lambda = 1/2$$

which is strictly positive if

$$\alpha < \frac{1}{2}\left(1+\sqrt{\frac{1}{1+4\delta}}\right) \equiv \overline{\alpha}.$$