Online Supplementary Materials

"Information Quality, Disagreement and Political Polarisation"

R. Emre Aytimur   Richard M. H. Suen

A. Proofs

This section contains the proofs of all the theoretical results mentioned in the main paper, except Proposition 1. The proof of Proposition 1 is shown separately in Section B of this document.

Proof of Lemma 1

We begin by noting that (s, b, m) has a joint normal distribution with mean vector $\boldsymbol{\mu}^\dagger$ and covariance matrix $\boldsymbol{\Sigma}^\dagger$ given by

$$\boldsymbol{\mu}^\dagger = \begin{bmatrix} \mu_s \\ \mu_b \\ \mu_m \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\Sigma}^\dagger = \begin{bmatrix} \sigma_s^2 & \omega & \lambda_s \\ \omega & \sigma_b^2 & \lambda_b \\ \lambda_s & \lambda_b & \sigma_m^2 \end{bmatrix},$$

where $\mu_m = \mu_s + \mu_b$, $\omega \equiv \rho_{s,b}\sigma_s\sigma_b$ and $\sigma_m^2 = var\,(m) = \lambda_s + \lambda_b + \sigma_\varepsilon^2$. By Theorem B.7 in Greene (2012, p.1042), the distribution of (s, b) conditional on m is a bivariate normal distribution with mean vector

$$\boldsymbol{\mu}' = \begin{bmatrix} \mu_s \\ \mu_b \end{bmatrix} + \frac{(m - \mu_m)}{\sigma_m^2} \begin{bmatrix} \lambda_s \\ \lambda_b \end{bmatrix},$$

and covariance matrix

$$\boldsymbol{\Sigma}' = \begin{bmatrix} \sigma_s^2 & \omega \\ \omega & \sigma_b^2 \end{bmatrix} - \frac{1}{\sigma_m^2} \begin{bmatrix} \lambda_s \\ \lambda_b \end{bmatrix} \begin{bmatrix} \lambda_s & \lambda_b \end{bmatrix}.$$

Consequently, the marginal distribution of s in the voters' updated belief is a normal distribution with mean

$$E\,(s \mid m) = \mu_s + \frac{\lambda_s}{\sigma_m^2}\,(m - \mu_m)\,,$$

which is the the first element in $\boldsymbol{\mu}'$; and variance

$$var\,(s \mid m) = \sigma_s^2 - \frac{\lambda_s^2}{\sigma_m^2},$$

which is the $(1, 1)$th element of $\boldsymbol{\Sigma}'$. This completes the proof of Lemma 1.

Proof of Corollary 1

Part (i) By Proposition 1, if $\phi > \gamma h(0)/2$, then the extent of equilibrium polarisation is given by

$$x^*_{eq} = \frac{2\phi - \gamma h(0)}{4h(0)\phi + 2}. \tag{A.1}$$

Substituting $h(0) = 1/\left(\widetilde{\sigma}\sqrt{2\pi}\right)$ into (A.1) and rearranging terms give

$$x^*_{eq} = \frac{2\sqrt{2\pi}\phi\widetilde{\sigma} - \gamma}{2\left(2\phi + \sqrt{2\pi}\widetilde{\sigma}\right)} = \frac{\sqrt{2\pi}\phi\left(\widetilde{\sigma} - \sigma_{\min}\right)}{2\phi + \sqrt{2\pi}\widetilde{\sigma}}, \qquad \text{where } \sigma_{\min} \equiv \frac{\gamma}{2\sqrt{2\pi}\phi}. \tag{A.2}$$

Differentiating this with respect to $\widetilde{\sigma}$ yields

$$\frac{dx^*_{eq}}{d\widetilde{\sigma}} = \frac{\sqrt{2\pi}\phi\left(2\phi + \sqrt{2\pi}\sigma_{\min}\right)}{\left(2\phi + \sqrt{2\pi}\widetilde{\sigma}\right)^2} > 0. \tag{A.3}$$

Part (ii) Differentiating (A.1) with respect to ϕ yields

$$\frac{dx^*_{eq}}{d\phi} = \frac{1 + \gamma\left[h(0)\right]^2}{\left[2h(0)\phi + 1\right]^2} > 0.$$

The upper bound is obtained by considering the limit of x^*_{eq} when $\phi \to \infty$, which is

$$\lim_{\phi\to\infty} x^*_{eq} = \lim_{\phi\to\infty}\left[\frac{2\phi - \gamma h(0)}{4h(0)\phi + 2}\right] = \frac{1}{2h(0)} = \sqrt{\frac{\pi}{2}} \cdot \widetilde{\sigma}.$$

This completes the proof.

Proof of Lemma 2

The coefficient ψ is defined as

$$\psi \equiv \frac{Cov(s,m)}{var(m)} = \frac{\lambda_s}{\lambda_s + \lambda_b + \tau_\varepsilon^{-1}}. \tag{A.4}$$

Using (A.4), we can get

$$\frac{d\psi^2}{d\tau_\varepsilon} = 2\psi \cdot \frac{d\psi}{d\tau_\varepsilon} = \frac{2}{\tau_\varepsilon^2}\frac{\lambda_s^2}{\left(\lambda_s + \lambda_b + \tau_\varepsilon^{-1}\right)^3} > 0.$$

Furthermore, it is clear that $var_p(m) = \widehat{\lambda}_s + \widehat{\lambda}_b + \tau_\varepsilon^{-1}$ is decreasing in τ_ε. This completes the proof.

Proof of Proposition 2

The main ideas of the proof have been explained in the main text. Here we only need to verify certain details. From (A.4), we can get

$$\ln \psi^2 = 2 \ln \lambda_s - 2 \ln \left(\lambda_s + \lambda_b + \tau_\varepsilon^{-1} \right)$$

$$\Rightarrow d \ln \psi^2 = \frac{2}{\lambda_s + \lambda_b + \tau_\varepsilon^{-1}} \frac{d\tau_\varepsilon}{\tau_\varepsilon^2}$$

$$\Rightarrow \frac{d \ln \psi^2}{d \ln \tau \varepsilon} = \frac{2}{\tau_\varepsilon} \cdot \frac{1}{\lambda_s + \lambda_b + \tau_\varepsilon^{-1}} = \frac{2}{\tau_\varepsilon} \cdot \frac{1}{var\,(m)},$$

which is equation (15) in the main text. Similarly,

$$d \ln var_p\,(m) = -\frac{1}{\widehat{\lambda}_s + \widehat{\lambda}_b + \tau_\varepsilon^{-1}} \frac{d\tau_\varepsilon}{\tau_\varepsilon^2}$$

$$\Rightarrow \frac{d \ln var_p\,(m)}{d \ln \tau_\varepsilon} = -\frac{1}{\tau_\varepsilon} \cdot \frac{1}{\widehat{\lambda}_s + \widehat{\lambda}_b + \tau_\varepsilon^{-1}} = -\frac{1}{\tau_\varepsilon} \cdot \frac{1}{var_p\,(m)},$$

which is equation (16) in the paper. Substituting (15) and (16) into (14) gives

$$\frac{d \ln \widetilde{\sigma}^2}{d \ln \tau_\varepsilon} = \frac{1}{\tau_\varepsilon} \cdot \left[\frac{2}{var\,(m)} - \frac{1}{var_p\,(m)} \right]$$

$$\Rightarrow \frac{d\widetilde{\sigma}^2}{d\tau_\varepsilon} = \frac{\widetilde{\sigma}^2}{\tau_\varepsilon^2} \cdot \left[\frac{2var_p\,(m) - var\,(m)}{var\,(m) \cdot var_p\,(m)} \right] \gtrless 0 \quad \text{iff} \quad 2var_p\,(m) \gtrless var\,(m)\,.$$

This completes the proof.

Proof of Corollary 2

According to Proposition 3,

$$\frac{d\widetilde{\sigma}^2}{d\tau_\varepsilon} = 2\widetilde{\sigma} \cdot \frac{d\widetilde{\sigma}}{d\tau_\varepsilon} \gtrless 0 \quad \text{iff} \quad 2var_p\,(m) \gtrless var\,(m)\,.$$

The condition $2var_p\,(m) \gtrless var\,(m)$ can be equivalently expressed as

$$2 \left(\widehat{\lambda}_s + \widehat{\lambda}_b \right) + 2\tau_\varepsilon^{-1} \gtrless \lambda_s + \lambda_b + \tau_\varepsilon^{-1}$$

$$\Leftrightarrow \tau_\varepsilon^{-1} \gtrless \lambda_s + \lambda_b - 2 \left(\widehat{\lambda}_s + \widehat{\lambda}_b \right).$$

If $2\left(\widehat{\lambda}_s+\widehat{\lambda}_b\right) \geq \lambda_s+\lambda_b$, then the above condition implies that for any $\tau_\varepsilon > 0$, $2var_p(m) > var(m)$ must be true. Hence,

$$\frac{d\widetilde{\sigma}^2}{d\tau_\varepsilon} > 0 \qquad \text{for all } \tau_\varepsilon > 0.$$

This establishes the result in part (a).

But if $\lambda_s+\lambda_b > 2\left(\widehat{\lambda}_s+\widehat{\lambda}_b\right)$, then we have $2var_p(m) \gtreqless var(m)$ if and only if

$$\tau_\varepsilon \lesseqgtr \tau'_\varepsilon \equiv \frac{1}{\lambda_s+\lambda_b-2\left(\widehat{\lambda}_s+\widehat{\lambda}_b\right)}.$$

This in turn implies

$$\frac{d\widetilde{\sigma}}{d\tau_\varepsilon} \gtreqless 0 \qquad \text{iff} \qquad \tau_\varepsilon \lesseqgtr \tau'_\varepsilon.$$

This establishes the result in part (b).

Proof of Lemma 3

As is evident from (A.4), $\psi < 0$ if and only if $\lambda_s = \sigma_s^2 + \rho_{s,b}\sigma_s\sigma_b < 0$. Provided that $\sigma_s > 0$, $\sigma_b > 0$ and $\rho_{s,b} \in (-1,1)$, this is true if and only if

$$-1 < \rho_{s,b} < -\frac{\sigma_s}{\sigma_b}.$$

The assumption $\sigma_s < \sigma_b$ is necessary to ensure that this range is nonempty. This completes the proof.

Proof of Proposition 3

Straightforward differentiation based on (A.4) yields

$$\frac{d\psi^2}{d\lambda_b} = 2\psi \cdot \frac{d\psi}{d\lambda_b} = -\frac{2\lambda_s^2}{\left(\lambda_s+\lambda_b+\tau_\varepsilon^{-1}\right)^3} < 0,$$

$$\frac{d\psi^2}{d\lambda_s} = 2\psi \cdot \frac{d\psi}{d\lambda_s} = \frac{2\lambda_s\left(\lambda_b+\tau_\varepsilon^{-1}\right)}{\left(\lambda_s+\lambda_b+\tau_\varepsilon^{-1}\right)^3} \gtreqless 0 \qquad \text{iff} \qquad \lambda_s\left(\lambda_b+\tau_\varepsilon^{-1}\right) \gtreqless 0.$$

Note that the sum $\lambda_s+\lambda_b$ must be positive because $\lambda_s+\lambda_b = var(s+b)$ in the voters' initial belief. Therefore, λ_s and λ_b cannot be both negative. This means $\lambda_b+\tau_\varepsilon^{-1} > 0$ must be true under defiant learning and it is not possible to have $\lambda_s\left(\lambda_b+\tau_\varepsilon^{-1}\right) > 0$ when $\lambda_s < 0$. Therefore, under defiant learning, ψ^2 is strictly decreasing in λ_s. This completes the proof.

We now explain how to relate changes in $\{\lambda_s, \lambda_b\}$ to those in $\{\sigma_s, \sigma_b, \rho_{s,b}\}$. Recall the definition of $\{\lambda_s, \lambda_b\}$, i.e.,

$$\lambda_s = \sigma_s^2 + \rho_{s,b}\sigma_s\sigma_b, \qquad \text{and} \qquad \lambda_b = \sigma_b^2 + \rho_{s,b}\sigma_s\sigma_b.$$

Totally differentiating λ_s and λ_b with respect to $\{\sigma_s, \sigma_b, \rho_{s,b}\}$ gives

$$d\lambda_s = \left(2\sigma_s + \rho_{s,b}\sigma_b\right) d\sigma_s + \rho_{s,b}\sigma_s d\sigma_b + \sigma_s\sigma_b d\rho_{s,b}, \tag{A.5}$$

$$d\lambda_b = \rho_{s,b}\sigma_b d\sigma_s + \left(2\sigma_b + \rho_{s,b}\sigma_s\right) d\sigma_b + \sigma_s\sigma_b d\rho_{s,b}. \tag{A.6}$$

Using these equations, we can derive situations in which λ_s changes but λ_b is held fixed (and vice versa). We will give two specific examples to illustrate this point.

Suppose there is no change in σ_b so that $d\sigma_b = 0$. Then according to (A.6),

$$d\lambda_b = 0 \qquad \text{if and only if} \qquad \sigma_s\sigma_b d\rho_{s,b} = -\rho_{s,b}\sigma_b d\sigma_s.$$

Substituting this into (A.5) gives $d\lambda_s = 2\sigma_s d\sigma_s$. Therefore, when $d\sigma_b = 0$,

$$d\lambda_b = 0 \quad \text{and} \quad d\lambda_s \gtrless 0 \qquad \Leftrightarrow \qquad d\sigma_s \gtrless 0 \quad \text{and} \quad d\rho_{s,b} = -\frac{\rho_{s,b}}{\sigma_s} d\sigma_s.$$

Suppose now there is no change in $\rho_{s,b}$ so that $d\rho_{s,b} = 0$. Then according to (A.6), $d\lambda_b = 0$ if and only if

$$d\sigma_b = -\frac{\rho_{s,b}\sigma_b}{2\sigma_b + \rho_{s,b}\sigma_s} d\sigma_s.$$

Substituting this into (A.5) gives

$$\begin{aligned} d\lambda_s &= \left(2\sigma_s + \rho_{s,b}\sigma_b\right) d\sigma_s - \frac{\rho_{s,b}^2\sigma_s\sigma_b}{2\sigma_b + \rho_{s,b}\sigma_s} d\sigma_s \\ &= \frac{4\sigma_s\sigma_b + 2\rho_{s,b}\left(\sigma_s^2 + \sigma_b^2\right)}{2\sigma_b + \rho_{s,b}\sigma_s} d\sigma_s. \end{aligned}$$

In the special case of $\rho_{s,b} = 0$, $\lambda_s = \sigma_s^2$ and $\lambda_b = \sigma_b^2$. It follows that

$$d\lambda_b = 0 \quad \text{and} \quad d\lambda_s \gtrless 0 \qquad \Leftrightarrow \qquad d\sigma_s \gtrless 0 \quad \text{and} \quad d\sigma_b = 0.$$

This concludes the proof of Proposition 3.

Proof of Lemma 4

Using the definition of $Cov\,(s,m)$ and $var\,(m)\,,$ we can rewrite (A.4) as

$$\psi = \frac{Cov\,(s,m)}{var\,(m)} = \frac{\sigma_s^2 + \rho_{s,b}\sigma_s\sigma_b}{\sigma_s^2 + \sigma_b^2 + \sigma_\varepsilon^2 + 2\rho_{s,b}\sigma_s\sigma_b}. \tag{A.7}$$

Differentiating this with respect to $\rho_{s,b}$ gives

$$\frac{d\psi}{d\rho_{s,b}} = \frac{\sigma_s\sigma_b}{var\,(m)} - \frac{Cov\,(s,m)}{[var\,(m)]^2}\cdot 2\sigma_s\sigma_b = \frac{\sigma_s\sigma_b\,[var\,(m) - 2Cov\,(s,m)]}{[var\,(m)]^2},$$

where

$$var\,(m) - 2Cov\,(s,m) = \sigma_b^2 + \sigma_\varepsilon^2 - \sigma_s^2.$$

Hence,

$$\frac{d\psi^2}{d\rho_{s,b}} = 2\psi\frac{d\psi}{d\rho_{s,b}} \gtrless 0 \qquad \text{iff} \qquad Cov\,(s,m)\left(\sigma_b^2 + \sigma_\varepsilon^2 - \sigma_s^2\right) \gtrless 0.$$

Suppose $\sigma_b \leq \sigma_s$. Then by Lemma 3, defiant learning is not possible, so it must be the case that $\lambda_s = Cov\,(s,m) > 0$. It follows immediately that

$$\frac{d\psi^2}{d\rho_{s,b}} \gtrless 0 \qquad \text{iff} \qquad \sigma_b^2 + \sigma_\varepsilon^2 \gtrless \sigma_s^2.$$

Suppose $\sigma_b > \sigma_s$, which implies $\sigma_b^2 + \sigma_\varepsilon^2 - \sigma_s^2 > 0$. Hence,

$$\frac{d\psi^2}{d\rho_{s,b}} \gtrless 0 \quad \Leftrightarrow \quad Cov\,(s,m) = \sigma_s\left(\sigma_s + \rho_{s,b}\sigma_b\right) \gtrless 0 \quad \Leftrightarrow \quad \rho_{s,b} \geq -\frac{\sigma_s}{\sigma_b}.$$

This completes the proof of Lemma 4.

Proof of Lemma 5

Pick any $\delta_v \in \mathbb{R}$. Conditional on m, voter v's expected utility if R wins is

$$\begin{aligned}
& E\left[U\left(x_{eq}^*;\delta_v,s\right)\mid m\right] \\
= & -E\left[\left(\delta_v+s-x_{eq}^*\right)^2\mid m\right] \\
= & -E\left[\left(\delta_v+\psi m-x_{eq}^*+s-\psi m\right)^2\mid m\right] \\
= & -E\left[\left(\delta_v+\psi m-x_{eq}^*\right)^2-2\left(\delta_v+\psi m-x_{eq}^*\right)(s-\psi m)+(s-\psi m)^2\mid m\right] \\
= & -E\left[\left(\delta_v+\psi m-x_{eq}^*\right)^2\mid m\right]-var\left(s\mid m\right) \\
= & -\left\{\left(\delta_v-x_{eq}^*\right)^2+2\left(\delta_v-x_{eq}^*\right)\psi m+(\psi m)^2+var\left(s\mid m\right)\right\}.
\end{aligned}$$

The second-to-last line uses the fact that $E\left(s\mid m\right)=\psi m$, hence

$$E\left[\left(\delta_v+\psi m-x_{eq}^*\right)(s-\psi m)\mid m\right]=\left(\delta_v+\psi m-x_{eq}^*\right)E\left[(s-\psi m)\mid m\right]=0$$

and $E\left[(s-\psi m)^2\mid m\right]=var\left(s\mid m\right).$

Similarly, the voter's expected utility if L wins is

$$E\left[U\left(-x_{eq}^*;\delta_v,s\right)\mid m\right]=-\left\{\left(\delta_v+x_{eq}^*\right)^2+2\left(\delta_v+x_{eq}^*\right)\psi m+(\psi m)^2+var\left(s\mid m\right)\right\}.$$

Under the voters' prior belief, the signal m follows a normal distribution with mean zero and variance $var\left(m\right)$. Therefore, ψm is a normal random variable with mean zero and variance $var\left(\psi m\right)\equiv\left(\sigma^{\dagger}\right)^2=\psi^2 var\left(m\right)$. Let $G\left(\cdot\right)$ be the corresponding cumulative distribution function.

Before m is realised, the voter's expected utility is

$$\begin{aligned}
W\left(x_{eq}^*;\delta_v\right) &= \int_0^{\infty}E\left[U\left(x_{eq}^*;\delta_v,s\right)\mid m\right]dG\left(\psi m\right)+\int_{-\infty}^{0}E\left[U\left(-x_{eq}^*;\delta_v,s\right)\mid m\right]dG\left(\psi m\right) \\
&= -\left[\frac{1}{2}\left(\delta_v-x_{eq}^*\right)^2+\frac{1}{2}\left(\delta_v+x_{eq}^*\right)^2+var\left(s\mid m\right)\right] \\
&\quad -\left[2\left(\delta_v-x_{eq}^*\right)\int_0^{\infty}\psi m dG\left(\psi m\right)+2\left(\delta_v+x_{eq}^*\right)\int_{-\infty}^{0}\psi m dG\left(\psi m\right)\right] \\
&\quad -\int_{-\infty}^{\infty}(\psi m)^2 dG\left(\psi m\right) \\
&= 4x_{eq}^*\int_0^{\infty}\psi m dG\left(\psi m\right)-\left[\delta_v^2+\left(x_{eq}^*\right)^2+var\left(s\mid m\right)+\psi^2 var\left(m\right)\right]. \qquad \text{(A.8)}
\end{aligned}$$

The last line uses the fact that $G(\cdot)$ is the cdf of a symmetric distribution around zero, hence

$$\int_{-\infty}^{\infty} \psi m dG(\psi m) = 0 \qquad \text{and} \qquad \int_{-\infty}^{0} \psi m dG(\psi m) = -\int_{0}^{\infty} \psi m dG(\psi m).$$

Using the formula,

$$\int_{0}^{\infty} x^{2n+1} \exp\left(-Ax^2\right) dx = \frac{n!}{2A^{n+1}}, \qquad \text{for } A > 0 \text{ and } n = 0, 1, 2, ...,$$

we can get

$$\int_{0}^{\infty} \psi m dG(\psi m) = \frac{1}{\sqrt{2\pi}\sigma^{\dagger}} \int_{0}^{\infty} \psi m \exp\left\{-\left[2\left(\sigma^{\dagger}\right)^2\right]^{-1} (\psi m)^2\right\} d(\psi m) = \frac{\sigma^{\dagger}}{\sqrt{2\pi}}.$$

Substituting this into (A.8) gives

$$W\left(x_{eq}^*; \delta_v\right) = 2\sqrt{\frac{2}{\pi}} x_{eq}^* \sigma^{\dagger} - \left(x_{eq}^*\right)^2 - \left[\delta_v^2 + var(s \mid m) + \psi^2 var(m)\right]. \tag{A.9}$$

Finally, by the law of total variance, we can get

$$var(s \mid m) + \psi^2 var(m) = \tau_s^{-1}. \tag{A.10}$$

To see this, first recall that μ_s in the voters' subjective prior belief is normalised to zero. Hence, the variance of s in their prior belief is given by

$$\tau_s^{-1} = var(s) = E\left(s^2\right) = E\left[E\left(s^2 \mid m\right)\right],$$

where the outer expectation is taken with respect to the distribution of m. It follows that

$$\begin{aligned} \tau_s^{-1} &= E\left\{var(s \mid m) + [E(s \mid m)]^2\right\} \\ &= var(s \mid m) + \psi^2 E\left(m^2\right) \\ &= var(s \mid m) + \psi^2 var(m). \end{aligned}$$

The second line uses the facts that $var(s \mid m)$ is a deterministic constant according to equation (4) in the main text and $E(s \mid m) = \psi m$. The last line uses the fact that $E(m) = 0$. This proves

(A.10). Substituting (A.10) into (A.9) gives

$$W\left(x_{eq}^*;\delta_v\right)=\left[2\sqrt{\frac{2}{\pi}}\sigma^\dagger-x_{eq}^*\right]x_{eq}^*-\left(\delta_v^2+\tau_s^{-1}\right).$$

This completes the proof.

Proof of Proposition 4

As mentioned in the main paper, polarisation is welfare-improving if and only if

$$0\le x_{eq}^*=\frac{2\phi-\gamma h(0)}{4h(0)\phi+2}\le 2\sqrt{\frac{2}{\pi}}\sigma^\dagger. \tag{A.11}$$

The second inequality in (A.11) can be rewritten as

$$2\sqrt{\frac{2}{\pi}}\sigma^\dagger\ge x_{eq}^*=\frac{2\phi-\gamma h(0)}{4h(0)\phi+2}$$

$$\Leftrightarrow 2\sqrt{2}\sigma^\dagger\left[4h(0)\phi+2\right]\ge 2\sqrt{\pi}\phi-\sqrt{\pi}\gamma h(0)$$

$$\Leftrightarrow 4\sqrt{2}\sigma^\dagger+\sqrt{\pi}\gamma h(0)\ge 2\left[\sqrt{\pi}-4\sqrt{2}\sigma^\dagger h(0)\right]\phi.$$

There are two possible cases: If $\sqrt{\pi}-4\sqrt{2}\sigma^\dagger h(0)\le 0$, or equivalently,

$$\frac{\widetilde{\sigma}}{\sigma^\dagger}=\sqrt{\frac{var_p(m)}{var(m)}}\le\frac{4}{\pi},$$

then the second inequality in (A.11) is automatically satisfied for all $\phi\ge 0$. This means $W\left(x_{eq}^*;\delta_v\right)\ge W(0;\delta_v)$ for any $x_{eq}^*\ge 0$.

But if $\sqrt{\pi}-4\sqrt{2}\sigma^\dagger h(0)>0$, or equivalently,

$$\frac{\widetilde{\sigma}}{\sigma^\dagger}=\sqrt{\frac{var_p(m)}{var(m)}}>\frac{4}{\pi},$$

then the second inequality in (A.11) holds if and only if

$$\phi\le\frac{4\sqrt{2}\sigma^\dagger+\sqrt{\pi}\gamma h(0)}{2\left[\sqrt{\pi}-4\sqrt{2}\sigma^\dagger h(0)\right]}=\frac{\sqrt{\pi}\left(8\sigma^\dagger\cdot\widetilde{\sigma}+\gamma\right)}{2\sqrt{2}\left(\pi\widetilde{\sigma}-4\sigma^\dagger\right)}.$$

Obviously, the second part of this proposition is meaningful only if there exists $\left\{\gamma,\phi,\sigma^\dagger,\widetilde{\sigma}\right\}$ that satisfy all the conditions. Specifically, let $\mathcal{S}$ be the set of $\left(\gamma,\sigma^\dagger,\widetilde{\sigma}\right)\in\mathbb{R}_+^3$ such that $\widetilde{\sigma}>4\sigma^\dagger/\pi$

and

$$\frac{\sqrt{\pi}\left(8\sigma^{\dagger}\cdot\widetilde{\sigma}+\gamma\right)}{2\sqrt{2}\left(\pi\widetilde{\sigma}-4\sigma^{\dagger}\right)}\geq\phi_{\min}\equiv\frac{\gamma}{2\sqrt{2\pi}\widetilde{\sigma}}.$$

We now show that this set is non-empty. Pick any $\gamma \geq 0$. For any $\widetilde{\sigma} > 0$ and $\sigma^{\dagger} > 0$ that satisfies $\pi\widetilde{\sigma} > 4\sigma^{\dagger}$, the above inequality can be rewritten as

$$\pi\widetilde{\sigma}\left(8\sigma^{\dagger}\cdot\widetilde{\sigma}+\gamma\right)\geq\gamma\left(\pi\widetilde{\sigma}-4\sigma^{\dagger}\right)$$

$$\Leftrightarrow 8\pi\sigma^{\dagger}\cdot\widetilde{\sigma}^{2}+4\gamma\sigma^{\dagger}\geq 0. \tag{A.12}$$

Note that (A.12) is true for all $\widetilde{\sigma} > 0$ and $\sigma^{\dagger} > 0$, therefore $\mathcal{S}$ is non-empty. This concludes the proof.

Proof of Proposition 5

Suppose $var\left(m\right) = var_{p}\left(m\right),$ which implies $\sigma^{\dagger} = \widetilde{\sigma}$. Then equation (22) in the main paper can be rewritten as

$$W\left(x_{eq}^{*};\delta_{v}\right)=\left[2\sqrt{\frac{2}{\pi}}\widetilde{\sigma}-x_{eq}^{*}\right]x_{eq}^{*}-\left(\delta_{v}^{2}+\tau_{s}^{-1}\right),$$

for any $x_{eq}^{*} \geq 0$, or equivalently for any $\widetilde{\sigma} \geq \sigma_{\min}$. Differentiating this with respect to τ_{ε} yields

$$\begin{aligned}
\frac{dW\left(x_{eq}^{*};\delta_{v}\right)}{d\tau_{\varepsilon}} &= \left[2\sqrt{\frac{2}{\pi}}\cdot\frac{d\widetilde{\sigma}}{d\tau_{\varepsilon}}-\frac{dx_{eq}^{*}}{d\tau_{\varepsilon}}\right]x_{eq}^{*}+\left[2\sqrt{\frac{2}{\pi}}\widetilde{\sigma}-x_{eq}^{*}\right]\frac{dx_{eq}^{*}}{d\tau_{\varepsilon}}\\
&= 2\sqrt{\frac{2}{\pi}}\cdot x_{eq}^{*}\cdot\frac{d\widetilde{\sigma}}{d\tau_{\varepsilon}}+2\left[\sqrt{\frac{2}{\pi}}\widetilde{\sigma}-x_{eq}^{*}\right]\frac{dx_{eq}^{*}}{d\tau_{\varepsilon}}\\
&= 2\left\{\sqrt{\frac{2}{\pi}}\cdot x_{eq}^{*}+\left[\sqrt{\frac{2}{\pi}}\widetilde{\sigma}-x_{eq}^{*}\right]\frac{dx_{eq}^{*}}{d\widetilde{\sigma}}\right\}\frac{d\widetilde{\sigma}}{d\tau_{\varepsilon}}.
\end{aligned} \tag{A.13}$$

The last line use the decomposition

$$\frac{dx_{eq}^{*}}{d\tau_{\varepsilon}}=\frac{dx_{eq}^{*}}{d\widetilde{\sigma}}\cdot\frac{d\widetilde{\sigma}}{d\tau_{\varepsilon}}.$$

By Proposition 2, $d\widetilde{\sigma}/d\tau_{\varepsilon} > 0$ when $var\left(m\right) = var_{p}\left(m\right).$ Hence, we will focus on the expression inside the curly brackets in (A.13), which involves both x_{eq}^{*} and its derivative with respect to $\widetilde{\sigma}$.

As shown in the proof of Corollary 1, x_{eq}^{*} can also be expressed as

$$x_{eq}^{*}\left(\widetilde{\sigma}\right)=\max\left\{\frac{\sqrt{2\pi}\phi\left(\widetilde{\sigma}-\sigma_{\min}\right)}{2\phi+\sqrt{2\pi}\widetilde{\sigma}},0\right\},\qquad\text{where }\sigma_{\min}\equiv\frac{\gamma}{2\sqrt{2\pi}\phi}.$$

The notation $x_{eq}^{*}(\widetilde{\sigma})$ highlights the dependence of x_{eq}^{*} on $\widetilde{\sigma}$. Also, for any $\widetilde{\sigma} > \sigma_{\min}$,

$$\frac{dx_{eq}^{*}(\widetilde{\sigma})}{d\widetilde{\sigma}} = \frac{\sqrt{2\pi}\phi\left(2\phi + \sqrt{2\pi}\sigma_{\min}\right)}{\left(2\phi + \sqrt{2\pi}\widetilde{\sigma}\right)^{2}} > 0. \tag{A.14}$$

From this equation, it is obvious that

$$\frac{d^{2}x_{eq}^{*}(\widetilde{\sigma})}{d\widetilde{\sigma}^{2}} < 0, \qquad \text{for any } \widetilde{\sigma} > \sigma_{\min}, \tag{A.15}$$

which means $x_{eq}^{*}(\widetilde{\sigma})$ is strictly increasing and strictly concave when it is strictly positive. This also implies that for any $\widetilde{\sigma} > \sigma_{\min} > 0$,

$$\frac{dx_{eq}^{*}(\widetilde{\sigma})}{d\widetilde{\sigma}} \leq \frac{\sqrt{2\pi}\phi}{2\phi + \sqrt{2\pi}\sigma_{\min}} < \frac{\sqrt{2\pi}}{2} = \sqrt{\frac{\pi}{2}}. \tag{A.16}$$

These facts will be useful later on.

Define an auxiliary function $\Lambda : [\sigma_{\min}, \infty) \to \mathbb{R}$ according to

$$\Lambda(\widetilde{\sigma}) \equiv \sqrt{\frac{2}{\pi}} \cdot x_{eq}^{*}(\widetilde{\sigma}) + \left[\sqrt{\frac{2}{\pi}}\widetilde{\sigma} - x_{eq}^{*}(\widetilde{\sigma})\right] \frac{dx_{eq}^{*}(\widetilde{\sigma})}{d\widetilde{\sigma}}, \tag{A.17}$$

which is the expression inside the curly brackets in (A.13). Since $x_{eq}^{*}(\sigma_{\min}) = 0$, we can get

$$\Lambda(\sigma_{\min}) = \sqrt{\frac{2}{\pi}}\sigma_{\min} \cdot \left.\frac{dx_{eq}^{*}(\widetilde{\sigma})}{d\widetilde{\sigma}}\right|_{\widetilde{\sigma}=\sigma_{\min}} > 0.$$

For $\widetilde{\sigma} \geq \sigma_{\min}$, the expression in the squared brackets in (A.17) can be expanded to become

$$\begin{aligned} \sqrt{\frac{2}{\pi}}\widetilde{\sigma} - x_{eq}^{*}(\widetilde{\sigma}) &= \sqrt{\frac{2}{\pi}}\widetilde{\sigma} - \frac{\sqrt{2\pi}\phi(\widetilde{\sigma} - \sigma_{\min})}{2\phi + \sqrt{2\pi}\widetilde{\sigma}} \\ &= \sqrt{\frac{2}{\pi}}\left[\widetilde{\sigma} - \frac{\pi\phi(\widetilde{\sigma} - \sigma_{\min})}{2\phi + \sqrt{2\pi}\widetilde{\sigma}}\right] \\ &= \sqrt{\frac{2}{\pi}}\left[\frac{\sqrt{2\pi}\widetilde{\sigma}^{2} - (\pi - 2)\phi\widetilde{\sigma} + \pi\phi\sigma_{\min}}{2\phi + \sqrt{2\pi}\widetilde{\sigma}}\right]. \end{aligned}$$

We will determine the sign of this expression by considering the quadratic equation:

$$\sqrt{2\pi}\widetilde{\sigma}^{2} - (\pi - 2)\phi\widetilde{\sigma} + \pi\phi\sigma_{\min} = 0. \tag{A.18}$$

There are three possible scenarios, which are shown in Figure A1.

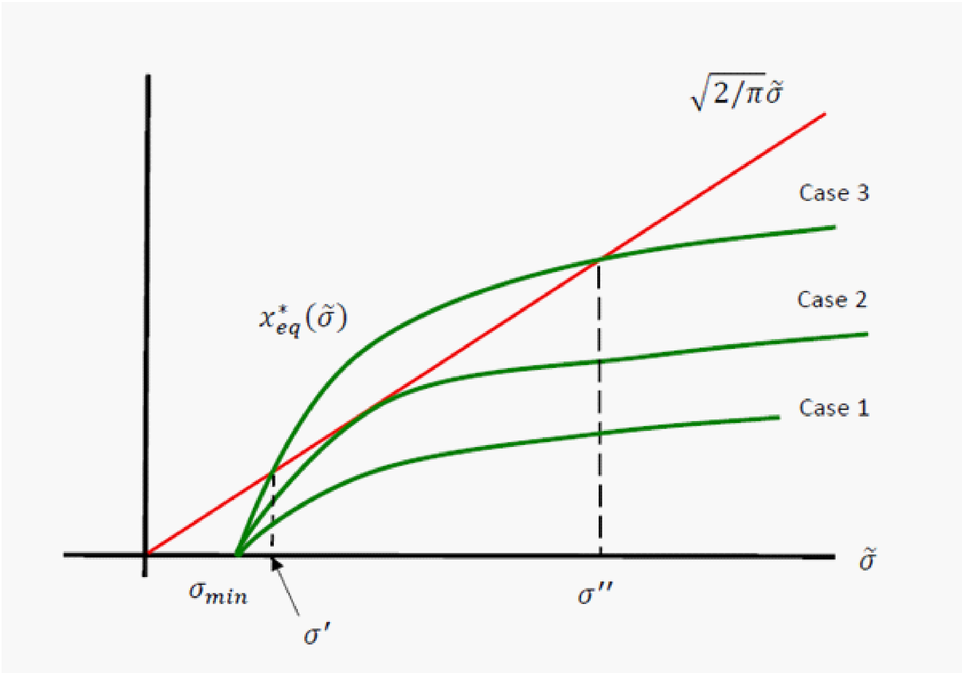


Figure A1.

Case 1 Suppose equation (A.18) has no real roots. This happens when

$$(\pi-2)^2\phi^2 - 4\sqrt{2\pi}\cdot\pi\phi\sigma_{\min} = (\pi-2)^2\phi^2 - 2\pi\gamma < 0.$$

It follows that $\sqrt{2\pi}\widetilde{\sigma}^2 - (\pi-2)\phi\widetilde{\sigma} + \pi\phi\sigma_{\min} > 0$ for any $\widetilde{\sigma} \geq \sigma_{\min}$, which is equivalent to

$$\sqrt{\frac{2}{\pi}}\widetilde{\sigma} - x^*_{eq}(\widetilde{\sigma}) > 0.$$

Combining this with (A.17) yields

$$\Lambda(\widetilde{\sigma}) \equiv \sqrt{\frac{2}{\pi}}\cdot x^*_{eq}(\widetilde{\sigma}) + \left[\sqrt{\frac{2}{\pi}}\widetilde{\sigma} - x^*_{eq}(\widetilde{\sigma})\right]\underbrace{\frac{dx^*_{eq}(\widetilde{\sigma})}{d\widetilde{\sigma}}}_{(+)} > 0 \tag{A.19}$$

$$\Rightarrow \frac{dW\left(x^*_{eq};\delta_v\right)}{d\tau_\varepsilon} = 2\Lambda(\widetilde{\sigma})\cdot\frac{d\widetilde{\sigma}}{d\tau_\varepsilon} > 0, \qquad \text{for any } \widetilde{\sigma} \geq \sigma_{\min}.$$

Case 2 Suppose (A.18) has a repeated real root, which happens when $(\pi-2)^2\phi^2 = 2\pi\gamma$. Let $\sigma_r > \sigma_{\min} > 0$ be the repeated root. Then we have

$$\sqrt{\frac{2}{\pi}}\sigma_r = x^*_{eq}(\sigma_r) \qquad \text{and} \qquad \sqrt{\frac{2}{\pi}}\widetilde{\sigma} - x^*_{eq}(\widetilde{\sigma}) > 0,$$

for any $\widetilde{\sigma} \geq \sigma_{\min}$ and $\widetilde{\sigma} \neq \sigma_r$. When evaluated at $\widetilde{\sigma} = \sigma_r$,

$$\Lambda(\sigma_r) \equiv \sqrt{\frac{2}{\pi}} \cdot x^*_{eq}(\sigma_r) > 0.$$

For any $\widetilde{\sigma} \geq \sigma_{\min}$ and $\widetilde{\sigma} \neq \sigma_r$, (A.19) will continue to hold. This again implies that $W\left(x^*_{eq};\delta_v\right)$ is strictly increasing in τ_ε whenever $x^*_{eq} \geq 0$.

Case 3 Suppose (A.18) has two distinct real roots, denoted by σ' and σ''. This happens when $(\pi-2)^2\phi^2 > 2\pi\gamma$. As shown in Figure A1, both roots must be strictly greater than $\sigma_{\min} > 0$. Without loss of generality, suppose $\sigma'' > \sigma' > \sigma_{\min} > 0$. From Figure A1, it is obvious that for any $\widetilde{\sigma} \in [\sigma_{\min}, \sigma'] \cup [\sigma'', \infty)$, we have

$$\sqrt{\frac{2}{\pi}}\widetilde{\sigma} - x^*_{eq}(\widetilde{\sigma}) \geq 0,$$

with strictly equality holds only at σ' and σ''. Therefore, for these values of $\widetilde{\sigma}$, we have $\Lambda(\widetilde{\sigma}) > 0$.

For any $\widetilde{\sigma} \in (\sigma', \sigma'')$,

$$\sqrt{\frac{2}{\pi}}\widetilde{\sigma} - x^*_{eq}(\widetilde{\sigma}) < 0.$$

Differentiating (A.17) with respect to $\widetilde{\sigma}$ gives

$$\Lambda'(\widetilde{\sigma}) = \left[2\sqrt{\frac{2}{\pi}} - \frac{dx^*_{eq}(\widetilde{\sigma})}{d\widetilde{\sigma}}\right]\underbrace{\frac{dx^*_{eq}(\widetilde{\sigma})}{d\widetilde{\sigma}}}_{(+)} + \underbrace{\left[\sqrt{\frac{2}{\pi}}\widetilde{\sigma} - x^*_{eq}(\widetilde{\sigma})\right]}_{(-)}\underbrace{\frac{d^2x^*_{eq}(\widetilde{\sigma})}{d\widetilde{\sigma}^2}}_{(-)}.$$

Recall that $x^*_{eq}(\widetilde{\sigma})$ is strictly increasing and strictly concave when $\widetilde{\sigma} \geq \sigma_{\min}$, therefore its second-order derivative is strictly negative. By (A.16),

$$\frac{dx^*_{eq}(\widetilde{\sigma})}{d\widetilde{\sigma}} < \sqrt{\frac{\pi}{2}} \simeq 1.253 < 2\sqrt{\frac{2}{\pi}} \simeq 1.596.$$

Hence, the expression insider the first squared bracket is strictly positive. This means $\Lambda(\widetilde{\sigma})$ is strictly increasing within the range (σ', σ''), therefore

$$\Lambda(\widetilde{\sigma}) > \Lambda(\sigma') = \sqrt{\frac{2}{\pi}} \cdot x^*_{eq}(\sigma') > 0, \qquad \text{for any } \widetilde{\sigma} \in (\sigma', \sigma'').$$

Taken together, these prove that when $(\pi - 2)^2 \phi^2 > 2\pi\gamma$, $\Lambda(\widetilde{\sigma}) > 0$ for any $\widetilde{\sigma} \geq \sigma_{\min}$. As a result,

$$\frac{dW\left(x^*_{eq}; \delta_v\right)}{d\tau_\varepsilon} = 2\Lambda(\widetilde{\sigma}) \cdot \frac{d\widetilde{\sigma}}{d\tau_\varepsilon} > 0, \qquad \text{for any } \widetilde{\sigma} \geq \sigma_{\min}.$$

This concludes the proof of Proposition 5.

Proof of Proposition 6

Part (a)

Differentiating equation (22) in the main paper with respect to τ_ε gives

$$\frac{dW\left(x^*_{eq}; \delta_v\right)}{d\tau_\varepsilon} = 2\left[\sqrt{\frac{2}{\pi}}\sigma^\dagger - x^*_{eq}\right]\frac{dx^*_{eq}}{d\tau_\varepsilon} + 2\sqrt{\frac{2}{\pi}} \cdot x^*_{eq}\frac{d\sigma^\dagger}{d\tau_\varepsilon}. \tag{A.20}$$

To derive the derivative of $\sigma^\dagger$ with respect to τ_ε, first recall

$$\left(\sigma^\dagger\right)^2 = \psi^2 var(m) = \left[\frac{Cov(s,m)}{var(m)}\right]^2 var(m) = \frac{\lambda_s^2}{\lambda_s + \lambda_b + \tau_\varepsilon^{-1}}.$$

Differentiating this with respect to τ_ε gives

$$2\sigma^\dagger \cdot \frac{d\sigma^\dagger}{d\tau_\varepsilon} = \frac{1}{\tau_\varepsilon^2}\frac{\lambda_s^2}{\left(\lambda_s + \lambda_b + \tau_\varepsilon^{-1}\right)^2} = \frac{1}{\tau_\varepsilon^2}\frac{\left(\sigma^\dagger\right)^2}{\lambda_s + \lambda_b + \tau_\varepsilon^{-1}}$$

$$\Rightarrow \frac{d\sigma^\dagger}{d\tau_\varepsilon} = \frac{1}{2\tau_\varepsilon^2}\frac{\sigma^\dagger}{\lambda_s + \lambda_b + \tau_\varepsilon^{-1}} = \frac{1}{2\tau_\varepsilon^2}\frac{\sigma^\dagger}{var(m)} > 0. \tag{A.21}$$

Note that this derivative is independent of ϕ. It follows that the second term in (A.20) must be strictly positive when $x^*_{eq} > 0$.

The derivative of x^*_{eq} with respect to τ_ε can be broken down into two parts:

$$\frac{dx^*_{eq}}{d\tau_\varepsilon} = \underbrace{\frac{dx^*_{eq}}{d\widetilde{\sigma}}}_{(+)} \cdot \frac{d\widetilde{\sigma}}{d\tau_\varepsilon}. \tag{A.22}$$

As shown in Corollary 1, x^*_{eq} is strictly increasing in $\widetilde{\sigma}$ when it is strictly positive. The second part can be derived as follows. Recall that

$$\widetilde{\sigma}^2 = \psi^2 var_p(m) = \left[\frac{Cov(s,m)}{var(m)}\right]^2 var_p(m) = \frac{\lambda_s^2\left(\widehat{\lambda}_s + \widehat{\lambda}_b + \tau_\varepsilon^{-1}\right)}{\left(\lambda_s + \lambda_b + \tau_\varepsilon^{-1}\right)^2}.$$

Taking the logarithm of both sides yields

$$2\ln\widetilde{\sigma} = \ln\lambda_s^2 + \ln\left(\widehat{\lambda}_s + \widehat{\lambda}_b + \tau_\varepsilon^{-1}\right) - 2\ln\left(\lambda_s + \lambda_b + \tau_\varepsilon^{-1}\right).$$

Differentiating both sides with respect to τ_ε gives

$$\frac{2}{\widetilde{\sigma}}\frac{d\widetilde{\sigma}}{d\tau_\varepsilon} = -\frac{1}{\tau_\varepsilon^2}\left(\frac{1}{\widehat{\lambda}_s + \widehat{\lambda}_b + \tau_\varepsilon^{-1}} - \frac{2}{\lambda_s + \lambda_b + \tau_\varepsilon^{-1}}\right)$$

$$\Rightarrow \frac{d\widetilde{\sigma}}{d\tau_\varepsilon} = \frac{1}{2\tau_\varepsilon^2}\cdot\frac{\widetilde{\sigma}}{var_p(m)}\left\{\frac{2var_p(m)}{var(m)} - 1\right\}. \tag{A.23}$$

Note that neither (A.21) nor (A.23) depend on ϕ. Therefore, on the right side of (A.20), only x_{eq}^* and its derivative with respect to $\widetilde{\sigma}$ are dependent on ϕ.

In particular, we know from Corollary 1 that

$$\lim_{\phi\to\infty} x_{eq}^* = \sqrt{\frac{\pi}{2}}\widetilde{\sigma}.$$

Using (A.14) we can get

$$\lim_{\phi\to\infty}\left(\frac{dx_{eq}^*}{d\widetilde{\sigma}}\right) = \sqrt{\frac{\pi}{2}}.$$

These imply

$$\begin{aligned}\lim_{\phi\to\infty}\left[\frac{dW\left(x_{eq}^*;\delta_v\right)}{d\tau_\varepsilon}\right] &= 2\left[\sqrt{\frac{2}{\pi}}\sigma^\dagger - \sqrt{\frac{\pi}{2}}\widetilde{\sigma}\right]\sqrt{\frac{\pi}{2}}\cdot\frac{d\widetilde{\sigma}}{d\tau_\varepsilon} + 2\sqrt{\frac{2}{\pi}}\cdot\sqrt{\frac{\pi}{2}}\widetilde{\sigma}\cdot\frac{d\sigma^\dagger}{d\tau_\varepsilon} \\ &= 2\left[\left(\sigma^\dagger - \frac{\pi}{2}\widetilde{\sigma}\right)\cdot\frac{d\widetilde{\sigma}}{d\tau_\varepsilon} + \widetilde{\sigma}\cdot\frac{d\sigma^\dagger}{d\tau_\varepsilon}\right].\end{aligned} \tag{A.24}$$

The two standard deviations $\widetilde{\sigma}$ and $\sigma^\dagger$ can be related using

$$\widetilde{\sigma} = \eta(\tau_\varepsilon,\boldsymbol{\lambda})\sigma^\dagger, \qquad \text{where } \eta(\tau_\varepsilon,\boldsymbol{\lambda}) \equiv \sqrt{\frac{var_p(m)}{var(m)}} = \sqrt{\frac{\widehat{\lambda}_s + \widehat{\lambda}_b + \tau_\varepsilon^{-1}}{\lambda_s + \lambda_b + \tau_\varepsilon^{-1}}} > 0. \tag{A.25}$$

The notation $\eta(\tau_\varepsilon,\boldsymbol{\lambda})$ highlights the dependence of η on τ_ε and $\boldsymbol{\lambda} = \left(\lambda_s, \lambda_b, \widehat{\lambda}_s, \widehat{\lambda}_b\right)$, but for now we will simply write this as η.

Substituting (A.21), (A.23) and (A.25) into (A.24) gives

$$\lim_{\phi\to\infty}\left[\frac{dW\left(x_{eq}^{*};\delta_v\right)}{d\tau_\varepsilon}\right] = \frac{\eta\cdot\left(\sigma^{\dagger}\right)^2}{\tau_\varepsilon^2}\left[\left(1-\frac{\pi}{2}\eta\right)\cdot\frac{1}{var_p\left(m\right)}\left(2\eta^2-1\right)+\frac{1}{var\left(m\right)}\right]$$
$$= \frac{\eta\cdot\left(\sigma^{\dagger}\right)^2}{\tau_\varepsilon^2 var_p\left(m\right)}\left[\left(1-\frac{\pi}{2}\eta\right)\left(2\eta^2-1\right)+\eta^2\right]$$
$$= -\frac{\eta\cdot\left(\sigma^{\dagger}\right)^2}{\tau_\varepsilon^2 var_p\left(m\right)}\left(\pi\eta^3-3\eta^2-\frac{\pi}{2}\eta+1\right).$$

Note that cubic equation

$$\Phi\left(\eta\right)\equiv\pi\eta^3-3\eta^2-\frac{\pi}{2}\eta+1=0$$

has three distinct real roots: -0.6322, 0.4382 and 1.1490; attains a local maximum at $\eta=-0.1994$ and a local minimum at $\eta=0.8360$. In particular, $\Phi\left(\eta\right)$ is strictly positive and strictly increasing for any $\eta>1.1490$. It follows that

$$\lim_{\phi\to\infty}\left[\frac{dW\left(x_{eq}^{*};\delta_v\right)}{d\tau_\varepsilon}\right]<0 \qquad \text{when } \eta\left(\tau_\varepsilon,\boldsymbol{\lambda}\right)>1.149.$$

Using (A.25), we can rewrite the second inequality as

$$\frac{\widehat{\lambda}_s+\widehat{\lambda}_b+\tau_\varepsilon^{-1}}{\lambda_s+\lambda_b+\tau_\varepsilon^{-1}}>\left(1.149\right)^2\simeq 1.320$$

$$\Leftrightarrow \widehat{\lambda}_s+\widehat{\lambda}_b-1.32\left(\lambda_s+\lambda_b\right)>0.32\tau_\varepsilon^{-1}$$

$$\tau_\varepsilon>\frac{0.32}{\widehat{\lambda}_s+\widehat{\lambda}_b-1.32\left(\lambda_s+\lambda_b\right)}\equiv\tau_c>0.$$

This proves that when $\widehat{\lambda}_s+\widehat{\lambda}_b>1.32\left(\lambda_s+\lambda_b\right)$,

$$\lim_{\phi\to\infty}\left[\frac{dW\left(x_{eq}^{*};\delta_v\right)}{d\tau_\varepsilon}\right]<0 \qquad \text{when } \tau_\varepsilon>\tau_c.$$

This establishes the result in part (a).

Part (b)

The desired result can be obtained by establishing the facts below:

Fact 1 For any given $\left(\lambda_s,\lambda_b,\widehat{\lambda}_s,\widehat{\lambda}_s\right)$, the parties' perceived uncertainty $\widetilde{\sigma}^2$ converges to the limit $\widetilde{\sigma}_\infty^2$ when τ_ε increases indefinitely.

Fact 2 Suppose $\lambda_b \geq 0$ and $\widehat{\lambda}_s + \widehat{\lambda}_s > \sigma^2_{\min}$. Then there exists a unique threshold value $\lambda_c \geq 0$ such that $\widetilde{\sigma}^2_\infty > \sigma^2_{\min}$ for any $\lambda_s > \lambda_c$.

These two facts together ensure that, when τ_ε approaches infinity, the extent of policy polarisation $x^*_{eq}(\widetilde{\sigma})$ will tend to the limit $x^*_{eq}(\widetilde{\sigma}_\infty)$ which is strictly positive. According to (A.14) in the proof of Proposition 5, $x^*_{eq}(\widetilde{\sigma}_\infty) > 0$ implies $dx^*_{eq}/d\widetilde{\sigma} > 0$ when evaluated at $\widetilde{\sigma}_\infty$.

Fact 3 For any given $\left(\lambda_s, \lambda_b, \widehat{\lambda}_s, \widehat{\lambda}_s\right)$, $\lim_{\tau_\varepsilon \to \infty} \left(d\sigma^\dagger / d\tau_\varepsilon\right) > 0$.

Fact 4 As shown in Corollary 2 part (b), whenever $\lambda_s + \lambda_b > 2\left(\widehat{\lambda}_s + \widehat{\lambda}_b\right)$, there exists a unique threshold value $\tau'_\varepsilon > 0$ such that

$$\frac{d\widetilde{\sigma}}{d\tau_\varepsilon} < 0 \qquad \text{for any } \tau_\varepsilon > \tau'_\varepsilon.$$

Fact 5 Suppose $\lambda_s > 0$ and $\lambda_s + \lambda_b > (\pi/2)^2\left(\widehat{\lambda}_s + \widehat{\lambda}_b\right)$. Then there exists a unique threshold value $\tau''_\varepsilon > 0$ such that

$$\sqrt{\frac{2}{\pi}}\sigma^\dagger - x^*_{eq}(\widetilde{\sigma}) > 0 \qquad \text{for any } \tau_\varepsilon > \tau''_\varepsilon.$$

These five facts are valid when $\lambda_s > 0$, $\lambda_b \geq 0$, $\widehat{\lambda}_s + \widehat{\lambda}_b > \sigma^2_{\min}$ and $\lambda_s + \lambda_b > (\pi/2)^2\left(\widehat{\lambda}_s + \widehat{\lambda}_b\right)$. Taken together, they imply

$$\begin{aligned} & \lim_{\tau_\varepsilon \to \infty}\left[\frac{dW\left(x^*_{eq};\delta_v\right)}{d\tau_\varepsilon}\right] \\ = \; & \underbrace{2\lim_{\tau_\varepsilon \to \infty}\left(\sqrt{\frac{2}{\pi}}\sigma^\dagger - x^*_{eq}\right)}_{(+)} \cdot \underbrace{\lim_{\tau_\varepsilon \to \infty}\left(\frac{d\widetilde{\sigma}}{d\tau_\varepsilon}\right)}_{(-)} \cdot \underbrace{\lim_{\tau_\varepsilon \to \infty}\left(\frac{dx^*_{eq}}{d\widetilde{\sigma}}\right)}_{(+)} + 2\sqrt{\frac{2}{\pi}} \cdot \underbrace{x^*_{eq}(\widetilde{\sigma}_\infty)}_{(+)} \cdot \underbrace{\lim_{\tau_\varepsilon \to \infty}\left(\frac{d\sigma^\dagger}{d\tau_\varepsilon}\right)}_{=0} < 0 \end{aligned}$$

and therefore establish the desired result.

Fact 1 follows immediately from the definition of $\widetilde{\sigma}^2$, i.e.,

$$\widetilde{\sigma}^2 \equiv \left[\frac{Cov(s,m)}{var(m)}\right]^2 var_p(m) = \frac{\lambda_s^2\left(\widehat{\lambda}_s + \widehat{\lambda}_b + \tau_\varepsilon^{-1}\right)}{\left(\lambda_s + \lambda_b + \tau_\varepsilon^{-1}\right)^2}$$

$$\Rightarrow \lim_{\tau_\varepsilon \to \infty} \widetilde{\sigma}^2 = \frac{\lambda_s^2\left(\widehat{\lambda}_s + \widehat{\lambda}_b\right)}{\left(\lambda_s + \lambda_b\right)^2} \equiv \widetilde{\sigma}^2_\infty.$$

Note that $\widetilde{\sigma}_{\infty}^{2} > \sigma_{\min}^{2}$ if and only if

$$\lambda_s^2 \left(\widehat{\lambda}_s + \widehat{\lambda}_b\right) > \sigma_{\min}^2 \left(\lambda_s + \lambda_b\right)^2$$

$$\Leftrightarrow \left(\widehat{\lambda}_s + \widehat{\lambda}_b - \sigma_{\min}^2\right) \lambda_s^2 - 2\sigma_{\min}^2 \lambda_b \lambda_s - \sigma_{\min}^2 \lambda_b^2 > 0. \tag{A.26}$$

Consider the quadratic equation

$$\left(\widehat{\lambda}_s + \widehat{\lambda}_b - \sigma_{\min}^2\right) \lambda_s^2 - 2\sigma_{\min}^2 \lambda_b \lambda_s - \sigma_{\min}^2 \lambda_b^2 = 0.$$

The discriminant of which is given by $\Delta = 4\left(\widehat{\lambda}_s + \widehat{\lambda}_b\right) \sigma_{\min}^2 \lambda_b^2 > 0$, hence it has two distinct real roots. The assumption $\widehat{\lambda}_s + \widehat{\lambda}_b > \sigma_{\min}^2 > 0$ implies that the product of the two roots is negative. In other words, there is one positive and one negative real root. Since we focus on positive values of λ_s, we will consider the positive root alone which is given by

$$\lambda_c = \frac{\sigma_{\min} \lambda_b \left[\sigma_{\min} + \sqrt{\left(\widehat{\lambda}_s + \widehat{\lambda}_b\right)}\right]}{\widehat{\lambda}_s + \widehat{\lambda}_b - \sigma_{\min}^2} \geq 0, \qquad \text{when } \lambda_b \geq 0. \tag{A.27}$$

It follows that for any $\lambda_s > \lambda_c \geq 0$, (A.26) holds and hence $\widetilde{\sigma}_{\infty}^2 > \sigma_{\min}^2$. This establishes Fact 2.

Fact 3 follows immediately from (A.21), which states that

$$\frac{d\sigma^{\dagger}}{d\tau_{\varepsilon}} = \frac{1}{2\tau_{\varepsilon}^2} \frac{\sigma^{\dagger}}{\lambda_s + \lambda_b + \tau_{\varepsilon}^{-1}} = \frac{1}{2\tau_{\varepsilon}^2} \frac{\lambda_s}{\left(\lambda_s + \lambda_b + \tau_{\varepsilon}^{-1}\right)^{\frac{3}{2}}}$$

$$\Rightarrow \lim_{\tau_{\varepsilon} \to \infty} \left(\frac{d\sigma^{\dagger}}{d\tau_{\varepsilon}}\right) = 0.$$

The last step is to establish Fact 5. According to Corollary 1,

$$x_{eq}^{*}\left(\widetilde{\sigma}\right) \leq \sqrt{\frac{\pi}{2}} \widetilde{\sigma},$$

for any given $\widetilde{\sigma}$. Using this, we can get

$$\begin{aligned}
\sqrt{\frac{2}{\pi}} \sigma^{\dagger} - x_{eq}^{*}\left(\widetilde{\sigma}\right) &\geq \sqrt{\frac{2}{\pi}} \sigma^{\dagger} - \sqrt{\frac{\pi}{2}} \widetilde{\sigma} \\
&= \sqrt{\frac{2}{\pi}} \left(\sigma^{\dagger} - \frac{\pi}{2} \widetilde{\sigma}\right) \\
&= \sqrt{\frac{2}{\pi}} \psi \left(\sqrt{var\left(m\right)} - \frac{\pi}{2} \sqrt{var_p\left(m\right)}\right).
\end{aligned}$$

Note that $\sqrt{var\,(m)} \geq \frac{\pi}{2}\sqrt{var_p\,(m)}$ if and only if

$$var\,(m) = \lambda_s + \lambda_b + \tau_\varepsilon^{-1} \geq \left(\frac{\pi}{2}\right)^2 var_p\,(m) = \left(\frac{\pi}{2}\right)^2 \left(\widehat{\lambda}_s + \widehat{\lambda}_b + \tau_\varepsilon^{-1}\right)$$

$$\Leftrightarrow \lambda_s + \lambda_b - \left(\frac{\pi}{2}\right)^2 \left(\widehat{\lambda}_s + \widehat{\lambda}_b\right) \geq \left[\left(\frac{\pi}{2}\right)^2 - 1\right] \tau_\varepsilon^{-1}$$

$$\Leftrightarrow \tau_\varepsilon \geq \frac{\left(\frac{\pi}{2}\right)^2 - 1}{\lambda_s + \lambda_b - \left(\frac{\pi}{2}\right)^2 \left(\widehat{\lambda}_s + \widehat{\lambda}_b\right)} \equiv \tau_\varepsilon'' > 0.$$

The last line uses the assumption that $\lambda_s + \lambda_b > (\pi/2)^2 \left(\widehat{\lambda}_s + \widehat{\lambda}_b\right)$. Therefore, if $\lambda_s > 0$ so that $\psi > 0$, we have

$$\sqrt{\frac{2}{\pi}}\sigma^\dagger - x_{eq}^*\left(\widetilde{\sigma}\right) > 0 \qquad \text{whenever } \tau_\varepsilon \geq \tau_\varepsilon''.$$

This proves Fact 5.

Note that if $\lambda_b = \widehat{\lambda}_b = 0$, so that $\lambda_s = \sigma_s^2$ and $\widehat{\lambda}_s = \widehat{\sigma}_s^2$, then $\widetilde{\sigma}_\infty^2 = \widehat{\sigma}_s^2$ and the positive root in (A.27) becomes $\lambda_c = 0$. It follows that

$$\lim_{\tau_\varepsilon \to \infty} \left[\frac{dW\left(x_{eq}^*; \delta_v\right)}{d\tau_\varepsilon}\right] < 0 \qquad \text{if} \qquad \widehat{\sigma}_s > \sigma_{\min} \text{ and } \sigma_s > \frac{\pi}{2} \cdot \widehat{\sigma}_s.$$

This establishes the result in part (b) and concludes the proof of Proposition 6.

B. Proof of Proposition 1

B1. Preliminaries

The purpose of this subsection is twofold: (i) To introduce some additional notations that are frequently used in the proof and (ii) to establish an intermediate result. Define an auxiliary function $\widetilde{\mathcal{W}}_R : \mathbb{R}^2 \to \mathbb{R}$ according to

$$\widetilde{\mathcal{W}}_R (x_R; x_L) \equiv \begin{cases} \gamma/2 & \text{if } x_R = x_L, \\ \Phi_R (x_R; x_L) & \text{if } x_R > x_L, \\ \Psi_R (x_R; x_L) & \text{if } x_R < x_L, \end{cases}$$

$$\Phi_R (x_R; x_L) \equiv \left[(x_L - \phi)^2 - (x_R - \phi)^2 + \gamma \right] [1 - H (\overline{x})] ,$$

$$\Psi_R (x_R; x_L) \equiv \left[(x_L - \phi)^2 - (x_R - \phi)^2 + \gamma \right] H (\overline{x}) ,$$

where $\overline{x} = (x_L + x_R)/2$ and $H(\cdot)$ is the cumulative distribution function of $N\left(\widetilde{\mu}, \widetilde{\sigma}^2\right)$ with $\widetilde{\mu}$ normalised to zero. Then party R's expected utility can be expressed as

$$\mathcal{W}_R (x_R; x_L) = \widetilde{\mathcal{W}}_R (x_R; x_L) - (x_L - \phi)^2 ,$$

Obviously the term $-(x_L - \psi_R)^2$ is irrelevant for R's policy choices. Hence, it suffice to focus on $\widetilde{\mathcal{W}}_R (x_R; x_L)$, which we will refer to as R's "effective" expected utility. By the same token, party L's effective expected utility is given by

$$\widetilde{\mathcal{W}}_L (x_L; x_R) = \begin{cases} \gamma/2 & \text{if } x_R = x_L, \\ \Phi_L (x_L; x_R) & \text{if } x_R > x_L, \\ \Psi_L (x_L; x_R) & \text{if } x_R < x_L, \end{cases}$$

where

$$\Phi_L (x_L; x_R) \equiv \left[(x_R + \phi)^2 - (x_L + \phi)^2 + \gamma \right] H (\overline{x}) ,$$

$$\Psi_L (x_L; x_R) \equiv \left[(x_R + \phi)^2 - (x_L + \phi)^2 + \gamma \right] [1 - H (\overline{x})] .$$

It is important to note that both $\Phi_R (x_R; x_L)$ and $\Phi_L (x_L; x_R)$ are continuous at $x_R = x_L$ even when $\widetilde{\mathcal{W}}_R (x_R; x_L)$ and $\widetilde{\mathcal{W}}_L (x_L; x_R)$ are not (due to the discontinuity of the winning probability function at $x_R = x_L$).

Let $\mathcal{B}_R(x_L)$ denote R's best-response correspondence under a given $x_L \in \mathbb{R}$, i.e.,

$$\mathcal{B}_R(x_L) \equiv \arg\max_{x_R \in \mathbb{R}} \left\{ \widetilde{\mathcal{W}}_R(x_R; x_L) \right\}.$$

Party L's best-response correspondence $\mathcal{B}_L(x_R)$ is similarly defined.

Let $(x_L^*, x_R^*) \in \mathbb{R}^2$ be a pure-strategy Nash equilibrium of the voting game, so that $x_R^* \in \mathcal{B}_R(x_L^*)$ and $x_L^* \in \mathcal{B}_L(x_R^*)$. At this stage, we do not confine ourselves to symmetric equilibria. In particular, the results presented in this and the next subsections are valid even if $x_R^* \neq x_L^*$. We begin with an intermediate result which specifies the relevant range of x_R^* and x_L^* in any kind of equilibrium under quadratic utility for the political parties. This result is well-known in the existing literature and is often stated without proof. We include the proof here for the sake of completeness.

Lemma B1 *Any voting equilibrium* (x_L^*, x_R^*), *if exists, must satisfy* $-\phi \leq x_L^* \leq x_R^* \leq \phi$.

Proof of Lemma B1 The proof of Lemma B1 is organised into three main parts.

Part I *For any* $x_L \in \mathbb{R}$, $x_R \in \mathcal{B}_R(x_L)$ *implies*

$$\phi - \sqrt{(x_L - \phi)^2 + \gamma} \leq x_R \leq \phi + \sqrt{(x_L - \phi)^2 + \gamma}. \tag{B.1}$$

Similarly, for any $x_R \in \mathbb{R}$, $x_L \in \mathcal{B}_L(x_R)$ *implies*

$$-\phi - \sqrt{(x_R + \phi)^2 + \gamma} \leq x_L \leq -\phi + \sqrt{(x_R + \phi)^2 + \gamma}. \tag{B.2}$$

Proof of Part I Fix $x_L \in \mathbb{R}$. If R chooses $x_R = x_L$, then its effective expected utility is

$$\widetilde{\mathcal{W}}_R(x_L; x_L) = \frac{\gamma}{2} \geq 0.$$

Hence, it is never optimal for R to choose any x_R that yields a negative effective expected utility. In other words, any $x_R \in \mathcal{B}_R(x_L)$ must satisfy

$$(x_L - \phi)^2 - (x_R - \phi)^2 + \gamma \geq 0 \tag{B.3}$$

which is equivalently to (B.1). The condition in (B.2) can be obtained by applying the same argument on party L. ■

Part II *Any voting equilibrium* (x_L^*, x_R^*), *if exists, must satisfy* $x_R^* \in [-\phi, \phi]$ and $x_L^* \in [-\phi, \phi]$.

Proof of Part II We begin by showing that if $x_R^* > \phi$, then either R or L will have an incentive to deviate. Hence, in equilibrium it must be the case that $x_R^* \leq \phi$. Using this, we can prove that $x_L^* \leq \phi$. The proof that $x_L^* \geq -\phi$ and $x_R^* \geq -\phi$ is largely similar and hence omitted.

Suppose the contrary that $x_R^* > \phi$. Then there are four exhaustive and mutually exclusive scenarios: (i) $x_R^* > \phi > x_L^*$, (ii) $x_R^* > x_L^* \geq \phi$, (iii) $x_R^* = x_L^* > \phi$ and (iv) $x_L^* > x_R^* > \phi$. It should be understood that all (x_L^*, x_R^*) considered below satisfy the inequalities in (B.1) and (B.2), so that (B.3) and

$$(x_R + \phi)^2 - (x_L + \phi)^2 + \gamma \geq 0 \tag{B.4}$$

are satisfied.

Scenario (i) Suppose $x_R^* > \phi > x_L^*$. Then party R's effective expected utility is given by

$$\Phi_R(x_R^*, x_L^*) = \left[(x_L^* - \phi)^2 - (x_R^* - \phi)^2 + \gamma\right] \left\{1 - H\left[\frac{1}{2}(x_R^* + x_L^*)\right]\right\}.$$

If R lowers its policy choice to ϕ, then its effective expected utility becomes

$$\begin{aligned}
\Phi_R(\phi, x_L^*) &= \left[(x_L^* - \phi)^2 + \gamma\right] \left\{1 - H\left[\frac{1}{2}(\phi + x_L^*)\right]\right\} \\
&> \left[(x_L^* - \phi)^2 - (x_R^* - \phi)^2 + \gamma\right] \left\{1 - H\left[\frac{1}{2}(\phi + x_L^*)\right]\right\} \\
&\geq \left[(x_L^* - \phi)^2 - (x_R^* - \phi)^2 + \gamma\right] \left\{1 - H\left[\frac{1}{2}(x_R^* + x_L^*)\right]\right\} \\
&= \Phi_R(x_R^*, x_L^*).
\end{aligned}$$

The third line uses (B.3) and the fact that $\overline{H}(\cdot) = 1 - H(\cdot)$ is strictly decreasing. This shows that R will have an incentive to deviate to $x_R = \phi$. Hence, $x_R^* > \phi > x_L^*$ cannot be an equilibrium.

Scenario (ii) Suppose $x_R^* > x_L^* \geq \phi$, which implies that $\frac{1}{2}(x_R^* + x_L^*) > \phi > \widetilde{\mu} = 0$. Therefore, we have

$$H\left[\frac{1}{2}(x_R^* + x_L^*)\right] > H(0) = \frac{1}{2} \tag{B.5}$$

$$\Rightarrow 1 - H\left[\frac{1}{2}\left(x_R^* + x_L^*\right)\right] < \frac{1}{2}. \tag{B.6}$$

If $x_R^* > x_L^* \geq \phi$, then R's effective expected utility is

$$\Phi_R\left(x_R^*; x_L^*\right) = \left[\left(x_L^* - \phi\right)^2 - \left(x_R^* - \phi\right)^2 + \gamma\right]\left\{1 - H\left[\frac{1}{2}\left(x_R^* + x_L^*\right)\right]\right\}.$$

Note that $x_R^* > x_L^* \geq \phi$ also implies $\left(x_R^* - \phi\right)^2 > \left(x_L^* - \phi\right)^2$. Combining this, (B.6) and $\gamma \geq 0$ gives

$$\Phi_R\left(x_R^*; x_L^*\right) < \gamma\left\{1 - H\left[\frac{1}{2}\left(x_R^* + x_L^*\right)\right]\right\} \leq \frac{\gamma}{2},$$

where $\gamma/2$ is R's effective expected utility when choosing $x_R = x_L^*$. This gives R an incentive to switch to $x_R = x_L^*$.

Scenario (iii) Suppose $x_R^* = x_L^* > \phi$, which again implies (B.5). Suppose now R lowers its policy choice to ϕ. By doing so, R's effective utility becomes

$$\begin{aligned}\Psi_R\left(\phi; x_L^*\right) &= \left[\left(x_L - \phi\right)^2 + \gamma\right] H\left[\frac{1}{2}\left(\phi + x_L^*\right)\right] \\ &> \gamma H\left[\frac{1}{2}\left(\phi + x_L^*\right)\right] \geq \frac{\gamma}{2}.\end{aligned}$$

The last inequality uses (B.5). Thus R will deviate from $x_R = x_L^* > \phi$ to $x_R = \phi$.

Scenario (iv) Suppose $x_L^* > x_R^* > \phi$. We now show that party L will have an incentive to deviate. Party L's effective expected utility under $x_L^* > x_R^* > \phi$ is

$$\Psi_L\left(x_L^*; x_R^*\right) \equiv \left[\left(x_R^* + \phi\right)^2 - \left(x_L^* + \phi\right)^2 + \gamma\right]\left\{1 - H\left[\frac{1}{2}\left(x_R^* + x_L^*\right)\right]\right\}.$$

Note that $x_L^* > x_R^* > \phi > -\phi$ implies $\left(x_L^* + \phi\right)^2 > \left(x_R^* + \phi\right)^2$. Using this and (B.6) gives

$$\Psi_L\left(x_L^*; x_R^*\right) < \gamma\left\{1 - H\left[\frac{1}{2}\left(x_R^* + x_L^*\right)\right]\right\} \leq \frac{\gamma}{2}.$$

This shows that L will be strictly better off by switching to $x_L^* = x_R^* > \phi$.

To summarise, we have shown that either R or L will have an incentive to deviate when $x_R^* > \phi$. Hence, any voting equilibrium must involve $x_R^* \leq \phi$. We now show that x_L^* must also be bounded above by ϕ.

Suppose the contrary that $x_L^* > \phi$. Since we have already ruled out the cases when $x_L^* > x_R^* >$

ϕ and $x_R^* \geq x_L^* > \phi$, there are only two remaining cases to consider: (a) $x_L^* > \phi > x_R^*$ and (b) $x_L^* > \phi = x_R^*$.

Scenario (a) Suppose $x_L^* > \phi > x_R^*$. Then R will prefer to deviate to ϕ. To see this, start with R's effective expected utility under $x_L^* > \phi > x_R^*$, which is

$$\begin{aligned}
\Psi_R\left(x_R^*; x_L^*\right) &= \left[\left(x_L^* - \phi\right)^2 - \left(x_R^* - \phi\right)^2 + \gamma\right] H\left[\frac{1}{2}\left(x_R^* + x_L^*\right)\right] \\
&< \left[\left(x_L^* - \phi\right)^2 + \gamma\right] H\left[\frac{1}{2}\left(x_R^* + x_L^*\right)\right] \\
&< \left[\left(x_L^* - \phi\right)^2 + \gamma\right] H\left[\frac{1}{2}\left(\phi + x_L^*\right)\right] = \Psi_R\left(\phi; x_L^*\right).
\end{aligned}$$

The third line uses the fact that $H(\cdot)$ is strictly increasing. This shows that R will be strictly better off by deviating to ϕ.

Scenario (b) Suppose $x_L^* > \phi = x_R^* > -\phi$, which implies $(\phi + \phi)^2 < \left(x_L^* + \phi\right)^2$ and

$$\overline{H}\left[\frac{1}{2}\left(\phi + x_L^*\right)\right] < \overline{H}(0) = \frac{1}{2}.$$

Using these, we can show that L will be strictly better off by choosing the same policy as R. Formally,

$$\begin{aligned}
\Psi_L\left(x_L^*; \phi\right) &= \left[(\phi + \phi)^2 - \left(x_L^* + \phi\right)^2 + \gamma\right] \left\{1 - H\left[\frac{1}{2}\left(\phi + x_L^*\right)\right]\right\} \\
&< \gamma\left\{1 - H\left[\frac{1}{2}\left(\phi + x_L^*\right)\right]\right\} \leq \frac{\gamma}{2}.
\end{aligned}$$

Hence, $x_L^* > \phi = x_R^*$ cannot be an equilibrium. This proves that both x_R^* and x_L^* must be bounded above by ϕ.

Using a similar line of argument, we can show that both x_R^* and x_L^* must be bounded below by $-\phi$. The details are not shown here. ■

For any $x_L \in \mathbb{R}$, define a subset of $\mathbb{R}$ according to

$$\mathcal{S}_R\left(x_L\right) \equiv \left\{x_R \in \mathbb{R} : \phi - \sqrt{\left(x_L - \phi\right)^2 + \gamma} \leq x_R \leq \phi\right\}.$$

Similarly, for any $x_R \in \mathbb{R}$, define

$$\mathcal{S}_L(x_R) \equiv \left\{ x_L \in \mathbb{R} : -\phi \leq x_L \leq -\phi + \sqrt{(x_R + \phi)^2 + \gamma} \right\}.$$

Taken together, Parts I & II establish that any voting equilibrium must satisfy $x_R^* \in \mathcal{S}_R(x_L^*)$ and $x_L^* \in \mathcal{S}_L(x_R^*)$.

Part III *Any voting equilibrium* (x_L^*, x_R^*), *if exists, must satisfy* $x_L^* \leq x_R^*$.

Proof of Part III Suppose the contrary that $x_R^* < x_L^*$. Then there are two possible scenarios: (A) $\frac{1}{2}(x_R^* + x_L^*) \leq 0$, and (B) $\frac{1}{2}(x_R^* + x_L^*) > 0$.

Scenario (A) In this case, we have

$$H\left[\frac{1}{2}(x_R^* + x_L^*)\right] \leq \frac{1}{2}.$$

Since both x_R^* and x_L^* must be bounded above by ϕ, $x_R^* < x_L^*$ implies $(x_R^* - \phi)^2 > (x_L^* - \phi)^2$. Using these observations, we can show that R is strictly better off by choosing the same policy as L. Starting with R's effective expected utility under $x_R^* < x_L^*$,

$$\begin{aligned} \Psi_R(x_R^*; x_L^*) &= \left[(x_L^* - \phi)^2 - (x_R^* - \phi)^2 + \gamma\right] H\left[\frac{1}{2}(x_R^* + x_L^*)\right] \\ &< \gamma H\left[\frac{1}{2}(x_R^* + x_L^*)\right] \leq \frac{\gamma}{2}. \end{aligned}$$

This proves that R will have an incentive to deviate.

Scenario (B) In this case, we can show that L will be strictly better off by choosing the same policy as R. The proof is similar to Scenario (A). In this case, we have

$$\overline{H}\left[\frac{1}{2}(x_R^* + x_L^*)\right] < \frac{1}{2}.$$

Since both x_R^* and x_L^* must be bounded below by $-\phi$, $x_R^* < x_L^*$ implies $(x_L^* + \phi)^2 > (x_R^* + \phi)^2$. Party L's effective expected utility in this scenario is

$$\begin{aligned} \Psi_L(x_L^*; x_R^*) &= \left[(x_R^* + \phi)^2 - (x_L^* + \phi)^2 + \gamma\right] \left\{1 - H\left[\frac{1}{2}(x_R^* + x_L^*)\right]\right\} \\ &< \gamma \left\{1 - H\left[\frac{1}{2}(x_R^* + x_L^*)\right]\right\} \leq \frac{\gamma}{2}. \end{aligned}$$

Therefore L will have an incentive to deviate from $x_R^* < x_L^*$ to $x_L^* = x_R^*$.

This rules out $x_R^* < x_L^*$ as a voting equilibrium, which proves the statement in Part III. Taken together, Parts I to III establish that $-\phi \leq x_L^* \leq x_R^* \leq \phi$. This completes the proof of Lemma B1. ■

B2. Some Preliminary Results

In light of Lemma B1, it is clear that only $\Phi_R(x_R; x_L)$, $\Phi_L(x_L; x_R)$ [i.e., the part of $\widetilde{\mathcal{W}}_R(x_R; x_L)$ and $\widetilde{\mathcal{W}}_L(x_L; x_R)$ when $x_R > x_L$] and $\gamma/2$ are relevant for equilibrium analysis. In this subsection, we present three preliminary results that are related to the maximiser of $\Phi_R(x_R; x_L)$ and $\Phi_L(x_L; x_R)$. These results are the main ingredients in the proof of Proposition 1.

Lemma B2 *For any $x_L \in \mathbb{R}$, there exists at most one value of x_R in $[x_L, \phi]$ that solves the first-order condition*

$$\frac{d\Phi_R(x_R; x_L)}{dx_R} = 0. \tag{B.7}$$

Proof of Lemma B2 Fix $x_L \in \mathbb{R}$. The first-order and second-order derivatives of $\Phi_R(x_R; x_L)$ with respect to x_R are, respectively, given by

$$\frac{d\Phi_R(x_R; x_L)}{dx_R} = \frac{1}{2}\left[(x_L - \phi)^2 - (x_R - \phi)^2 + \gamma\right] \overline{H}'(\overline{x}) - 2(x_R - \phi)\overline{H}(\overline{x}), \tag{B.8}$$

$$\begin{aligned} \frac{d^2\Phi_R(x_R; x_L)}{dx_R^2} &= \frac{1}{4}\left[(x_L - \phi)^2 - (x_R - \phi)^2 + \gamma\right] \overline{H}''(\overline{x}) \\ &\quad -2(x_R - \phi)\overline{H}'(\overline{x}) - 2\overline{H}(\overline{x}). \end{aligned} \tag{B.9}$$

Evaluating the first-order derivative at $x_R = \phi$ gives

$$\left.\frac{d\Phi_R(x_R; x_L)}{dx_R}\right|_{x_R = \phi} = \frac{1}{2}\left[(x_L - \phi)^2 + \gamma\right] \overline{H}'\left[\frac{1}{2}(\phi + x_L)\right] < 0.$$

This shows that it is never optimal for R to choose $x_R = \phi$.

Let x'_R be a solution of (B.7), or a stationary point. Then we have

$$\left[(x_L - \phi)^2 - \left(x'_R - \phi\right)^2 + \gamma\right] \overline{H}'(\overline{x}) = 4\left(x'_R - \phi\right) \overline{H}(\overline{x})$$

$$\Rightarrow (x_L - \phi)^2 - \left(x'_R - \phi\right)^2 + \gamma = 4\left(x'_R - \phi\right) \frac{\overline{H}(\overline{x})}{\overline{H}'(\overline{x})}, \tag{B.10}$$

where $\overline{x} = \left(x_L + x'_R\right)/2$. Substituting (B.10) into (B.9) and rearranging terms gives

$$\left.\frac{d^2\Phi_R(x_R; x_L)}{dx_R^2}\right|_{x_R = x'_R} = \left(x'_R - \phi\right)\left[\frac{\overline{H}(\overline{x})}{\overline{H}'(\overline{x})}\overline{H}''(\overline{x}) - 2\overline{H}'(\overline{x})\right] - 2\overline{H}(\overline{x}).$$

It is known that both $H(\cdot)$ and $\overline{H}(\cdot)$ of a normal distribution are log-concave functions [see Bagnoli and Bergstrom (2005) for details]. This means

$$\frac{d^2 \ln \overline{H}(\overline{x})}{d\overline{x}^2} = \frac{\overline{H}''(\overline{x})}{\overline{H}(\overline{x})} - \left[\frac{\overline{H}'(\overline{x})}{\overline{H}(\overline{x})}\right]^2 \leq 0, \qquad \text{for any } \overline{x} \in \mathbb{R},$$

$$\Rightarrow \frac{\overline{H}'(\overline{x})}{\left[\overline{H}(\overline{x})\right]^2}\left\{\frac{\overline{H}(\overline{x})}{\overline{H}'(\overline{x})}\overline{H}''(\overline{x}) - \overline{H}'(\overline{x})\right\} \leq 0.$$

Since $\overline{H}'(\cdot) < 0$, logconcavity of $\overline{H}(\cdot)$ implies

$$\frac{\overline{H}(\overline{x})}{\overline{H}'(\overline{x})}\overline{H}''(\overline{x}) \geq \overline{H}'(\overline{x}) \geq 2\overline{H}'(\overline{x}), \qquad \text{for any } \overline{x} \in \mathbb{R}. \tag{B.11}$$

This, together with $x'_R < \phi$, guarantees that

$$\left.\frac{d^2\Phi_R(x_R; x_L)}{dx_R^2}\right|_{x_R = x'_R} < 0.$$

The shows that any solution of the first-order condition in (B.7) will also satisfy the second-order condition for maximisation.

Finally, we show that there exists at most one solution to (B.7). Suppose the contrary that there are two distinct stationary points, say x'_R and x''_R. Without loss of generality, suppose $x_L \leq x'_R < x''_R < \phi$. Then by the above result, both x'_R and x''_R will satisfy the second-order condition for maximisation. This means there exists $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that (i) $x_L < x'_R + \varepsilon_1 < x''_R - \varepsilon_2$, (ii) $\Phi_R(x_R; x_L)$ is strictly decreasing over the range $\left(x'_R, x'_R + \varepsilon_1\right)$, and (iii) $\Phi_R(x_R; x_L)$ is strictly increasing over the range $\left(x''_R - \varepsilon_2, x''_R\right)$. Since $\Phi_R(x_R; x_L)$ is continuously differentiable in x_R, there must exist at least one other stationary point x_R''' between x_R' and x_R'' that is a (local) minimum. This contradicts the fact that any stationary point must satisfy the second-order condition for maximisation. Hence, x_R' and x_R'' must be the same point. This completes the proof of Lemma B2. ■

Lemma B3

(i) *For any $x_L \in \mathbb{R}$, if the following condition holds*

$$\frac{\gamma}{2}\overline{H}'(x_L) - 2(x_L - \phi)\overline{H}(x_L) \leq 0, \tag{B.12}$$

then $\Phi_R(x_L; x_L) > \Phi_R(x_R; x_L)$ for all $x_R \in (x_L, \phi]$.

(ii) *For any $x_L \in \mathbb{R}$, if the following condition holds,*

$$\frac{\gamma}{2}\overline{H}'(x_L) - 2(x_L - \phi)\overline{H}(x_L) > 0, \tag{B.13}$$

then there exists a unique value $h_R(x_L) \in (x_L, \phi)$ such that $\Phi_R[h_R(x_L); x_L] > \Phi_R(x_R; x_L)$ for all $x_R \in [x_L, \phi]$.

Proof of Lemma B3 Fix $x_L \in \mathbb{R}$. Consider the first-order derivative of $\Phi_R(x_R; x_L)$ in (B.8), which is strictly negative when evaluated at $x_R = \phi$. Hence, $\Phi_R(x_R; x_L)$ is strictly decreasing as x_R approaches ϕ. When evaluated at $x_R = x_L$, the same derivative becomes

$$\left.\frac{d\Phi_R(x_R; x_L)}{dx_R}\right|_{x_R = x_L} = \frac{\gamma}{2}\overline{H}'(x_L) - 2(x_L - \phi)\overline{H}(x_L).$$

The sign of this expression depends on the shape of $\overline{H}(\cdot)$ and model parameters. There are two possible scenarios, which are stated in (B.12) and (B.13).

First consider the case when (B.12) holds with equality. This means $x_R = x_L$ is a stationary point that satisfies the first-order condition in (B.7). By Lemma B2, $x_R = x_L$ is the unique maximiser within the range $[x_L, \phi]$ so that $\Phi_R(x_L; x_L) > \Phi_R(x_R; x_L)$ for all $x_R \in (x_L, \phi]$. Next, suppose (B.12) holds as a strict inequality, which means $\Phi_R(x_R; x_L)$ is strictly decreasing at $x_R = x_L$. We now show that $\Phi_R(x_R; x_L)$ must be strictly decreasing over the entire range of $[x_L, \phi]$ so that it has a single peak at $x_R = x_L$. Suppose the contrary that the first-order derivative of $\Phi_R(x_R; x_L)$ in (B.8) is strictly positive at some $\widehat{x}_R \in (x_L, \phi)$. Since $\Phi_R(x_R; x_L)$ is continuously differentiable in x_R, there must exist two other values, x'_R and x''_R, in (x_L, ϕ) such that (i) $x'_R < \widehat{x}_R < x''_R$ and (ii) both x'_R and x''_R are stationary points, i.e., (B.7) holds. This, however, contradicts the result in Lemma B2. Hence, $\Phi_R(x_R; x_L)$ must be strictly decreasing in x_R over the entire range of $[x_L, \phi]$ when (B.12) is a strict inequality. This scenario is depicted in Figure B1 Panel (a). This proves the statement in part (i).

Next, consider the case when (B.13) is valid, which means $\Phi_R(x_R; x_L)$ is strictly increasing at $x_R = x_L$. Since $\Phi_R(x_R; x_L)$ is continuously differentiable in x_R, there exists a value $h_R(x_L) \in (x_L, \phi)$ that solves the first-order condition in (B.7). By Lemma B2, $h_R(x_L)$ is unique and satisfies the second-order condition for maximisation. Hence, $\Phi_R[h_R(x_L); x_L] > \Phi_R(x_R; x_L)$ for all $x_R \in [x_L, \phi]$. This scenario is shown in Figure B1 Panel (b). This completes the proof of Lemma B3. ■

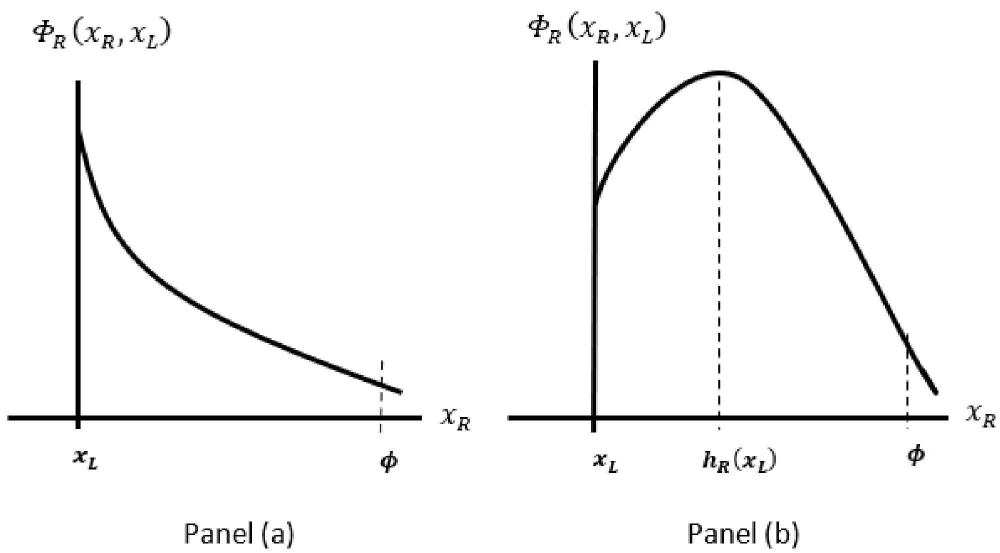


Figure B1: The Shape of $\Phi_R(x_R; x_L)$.

The third preliminary result establishes similar properties of $\Phi_L(x_L; x_R)$. The proof uses the same line of arguments as in Lemmas B2 and B3, hence it is omitted.

Lemma B4

(i) *For any* $x_R \in \mathbb{R}$, *if the following condition holds*

$$\frac{\gamma}{2} H'(x_R) - 2(x_R + \phi) H(x_R) \geq 0, \tag{B.14}$$

then $\Phi_L(x_R; x_R) > \Phi_L(x_L; x_R)$ *for all* $x_L \in [-\phi, x_R)$.

(ii) *For any* $x_R \in \mathbb{R}$, *if the following condition holds*

$$\frac{\gamma}{2} H'(x_R) - 2(x_R + \phi) H(x_R) < 0, \tag{B.15}$$

then there exists a unique value $h_L(x_R) \in (-\phi, x_R)$ *such that* $\Phi_L[h_L(x_R); x_R] > \Phi_L(x_L; x_R)$ *for all* $x_L \in [-\phi, x_R]$.

B3. Main Proof

We now proceed to the proof of Proposition 1.

Part (a) If we set $x_L = 0$ in (B.12), then we can get

$$\frac{\gamma}{2} \overline{H}'(0) - 2(-\phi) \overline{H}(0) = -\frac{\gamma}{2} h(0) + \phi \leq 0 \Rightarrow \phi \leq \frac{\gamma h(0)}{2}.$$

It follows from part (i) of Lemma B3 that

$$\Phi_R(0; 0) = \frac{\gamma}{2} > \Phi_R(x_R; 0), \qquad \text{for all } x_R \in (0, \phi].$$

This means choosing $x_R = 0$ is R's unique best response to $x_L = 0$. Likewise, if we set $x_R = 0$ in (B.14), then we can get

$$\frac{\gamma}{2} H'(0) \geq 2\phi H(0) \Rightarrow \frac{\gamma h(0)}{2} \geq \phi.$$

By the first part of Lemma B4,

$$\Phi_L(0; 0) = \frac{\gamma}{2} > \Phi_L(x_L; 0), \qquad \text{for all } x_L \in [-\phi, 0),$$

which means choosing $x_L = 0$ is L's unique best response to $x_R = 0$. Hence, $x_R^* = x_L^* = 0$ is a (symmetric) voting equilibrium when $\phi \leq \gamma h(0)/2$.

To prove that it is a unique symmetric voting equilibrium, suppose the contrary that there exists another one with $x_R^* = -x_L^* = \widetilde{x}_{eq} \in (0, \phi)$. In other words, $x_R^* = \widetilde{x}_{eq}$ is an interior solution that maximises $\Phi_R(x_R; -\widetilde{x}_{eq})$ over the range $[-\widetilde{x}_{eq}, \phi]$. The first-order condition for this maximisation problem is given by

$$\frac{d\Phi_R\left(x_R;-\widetilde{x}_{eq}\right)}{dx_R}=\frac{1}{2}\left[\left(-\widetilde{x}_{eq}-\phi\right)^2-\left(x_R-\phi\right)^2+\gamma\right]\overline{H}'\left(\overline{x}\right)-2\left(x_R-\phi\right)\overline{H}\left(\overline{x}\right)=0.$$

Since this is a symmetric equilibrium, $\overline{x}=0$. Substituting this and $x_R=\widetilde{x}_{eq}$ into the above condition gives

$$\left[\left(-\widetilde{x}_{eq}-\phi\right)^2-\left(\widetilde{x}_{eq}-\phi\right)^2+\gamma\right]\overline{H}'\left(0\right)=4\left(\widetilde{x}_{eq}-\phi\right)\overline{H}\left(0\right)$$

$$\Rightarrow-\left(4\phi\widetilde{x}_{eq}+\gamma\right)h\left(0\right)=2\left(\widetilde{x}_{eq}-\phi\right)$$

$$\Rightarrow\widetilde{x}_{eq}=\frac{2\phi-\gamma h\left(0\right)}{4\phi h\left(0\right)+2}\leq 0.$$

The negative sign follows from $\phi\leq\gamma h\left(0\right)/2$. The last line contradicts the presumption that $\widetilde{x}_{eq}>0$, hence, $x_R^*=x_L^*=0$ is the unique symmetric voting equilibrium when $\phi\leq\gamma h\left(0\right)/2$. This establishes part (a) of Proposition 1.

Part (b) Suppose $\phi>\gamma h\left(0\right)/2$. Define

$$x_{eq}^*=\frac{2\phi-\gamma h\left(0\right)}{4\phi h\left(0\right)+2}>0.$$

Following the argument in part (a), $x_R=x_{eq}^*$ is the unique value that satisfies the first-order condition

$$\frac{d\Phi_R\left(x_R;-x_{eq}^*\right)}{dx_R}=0.$$

As shown in the proof of Lemma B2, this means $x_R=x_{eq}^*$ is the unique value that maximises $\Phi_R\left(x_R;-x_{eq}^*\right)$ over the range $\left[-x_{eq}^*,\phi\right]$. Consequently,

$$\Phi_R\left(x_{eq}^*;-x_{eq}^*\right)>\Phi_R\left(-x_{eq}^*;-x_{eq}^*\right)=\gamma\left[1-H\left(-x_{eq}^*\right)\right]\geq\frac{\gamma}{2}=\widetilde{\mathcal{W}}_R\left(-x_{eq}^*;-x_{eq}^*\right).\qquad\text{(B.16)}$$

The last inequality follows from the facts that $\gamma\geq 0$ and $H\left(-x_{eq}^*\right)<1/2$ as $x_{eq}^*>0$. Note that $\widetilde{\mathcal{W}}_R\left(x_R;x_L\right)$ is discontinuous at $x_R=x_L$ (except when $x_L=0$), i.e.,

$$\widetilde{\mathcal{W}}_R\left(-x_{eq}^*;-x_{eq}^*\right)\neq\Phi_R\left(-x_{eq}^*;-x_{eq}^*\right),$$

due to the discontinuity in R's winning probability. The last part of (B.16) thus ensures that R has no incentive to deviate from $x_R=x_{eq}^*$ to $x_R=-x_{eq}^*$. This proves that choosing $x_R=x_{eq}^*$

is R's unique best response to $x_L = -x^*_{eq}$. Using the same line of argument, we can show that choosing $x_L = -x^*_{eq}$ is L's unique best response to $x_R = x^*_{eq}$. It follows that $x^*_R = -x^*_L = x^*_{eq} > 0$ is the unique symmetric polarised equilibrium when $\phi > \gamma h(0)/2$.

To see that the convergent equilibrium $x^*_R = x^*_L = 0$ cannot emerge in this case, we will use the result in part (ii) of Lemma B3. Note that $\phi > \gamma h(0)/2$ can be obtained by setting $x_L = 0$ in (B.13). Then this result states that there exists a unique value $h_R(0) \in (0, \phi)$ such that $\Phi_R[h_R(0); 0] > \Phi_R(x_R; 0)$ for all $x_R \in [0, \phi]$. In other words, choosing $x_R = 0$ is not R's best response to $x_L = 0$ when $\phi > \gamma h(0)/2$. Hence, $x^*_R = x^*_L = 0$ cannot be an equilibrium. This establishes part (b) of Proposition 1 and concludes the whole proof.

C. Extended Model with Multiple Signals

Belief and Information

In this section, we present an extended model in which voters receive imperfect information about the hidden state $s \in \mathbb{R}$ from $n \geq 1$ different sources prior the election. Each information channel $i \in \{1, 2, ..., n\}$ produces a noisy public signal m_i which is potentially biased. Specifically, let $m_i = b_i + s + \varepsilon_i$, where $b_i \in \mathbb{R}$ is an unknown parameter that captures the inherent bias of the ith information channel and $\varepsilon_i \in \mathbb{R}$ is the error term in m_i, for all $i \in \{1, 2, ..., n\}$. Voters share the same subjective prior belief about the state variable s and the biases $\mathbf{b} = (b_1, ..., b_n)^T$. This is assumed to take the form of a joint multivariate normal distribution $\mathbf{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where

$$\boldsymbol{\mu}_0 = \begin{bmatrix} \mu_s \\ \boldsymbol{\mu}_b \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\Sigma}_0 = \begin{bmatrix} \sigma_s^2 & \boldsymbol{\Omega}^T \\ \boldsymbol{\Omega} & \boldsymbol{\Sigma}_b \end{bmatrix}.$$

In the above expressions, μ_s and σ_s^2 are scalars representing the mean and variance of the marginal distribution of s; whereas $\boldsymbol{\mu}_b$ and $\boldsymbol{\Sigma}_b$ are the mean vector and covariance matrix of the marginal distribution of $\mathbf{b}$.[1] The covariances between s and $\mathbf{b}$ are captured by the 1-by-n row vector $\boldsymbol{\Omega}^T = (\omega_{1,...,}\omega_n)$, where $\omega_i \equiv Cov(s, b_i)$. A positive ω_i means that b_i is expected to exaggerate or complement the effect of the hidden state. A negative value, on the other hand, means that voters expect b_i to contradict or subdue the effect of s.

The error terms, $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_n)^T$, are drawn from a normal distribution $\mathbf{N}(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$. Each ε_i is independent of the distribution of political attitudes δ_v and the voters' prior belief about $(s, \mathbf{b})$. The statistical properties of $\boldsymbol{\varepsilon}$ are known to both voters and political parties.

Given the voters' prior belief, the signals $\mathbf{m} = (m_1, ..., m_n)^T$ have a joint normal distribution with mean vector

$$\boldsymbol{\mu}_m = \mu_s \cdot \mathbf{1}_n + \boldsymbol{\mu}_b,$$

where $\mathbf{1}_n$ is an n-by-1 column of ones, and covariance matrix

$$\begin{aligned} \boldsymbol{\Sigma}_m &= E\left[(\mathbf{m} - \boldsymbol{\mu}_m)(\mathbf{m} - \boldsymbol{\mu}_m)^T\right] \\ &= \boldsymbol{\Sigma}_b + \sigma_s^2 \cdot \mathbf{1}_n \mathbf{1}_n^T + \boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Omega} \mathbf{1}_n^T + \boldsymbol{\Omega}^T \mathbf{1}_n. \end{aligned} \tag{C.1}$$

Equation (C.1) suggests that the quality of the signals (as measured by the inverse of $\boldsymbol{\Sigma}_m$) is

[1] All the covariance matrices appeared in this document are assumed to be (at least) positive semidefinite.

determined by three groups of factors:[2] (i) the precision of the voters' subjective prior belief, as captured by $\tau_s \equiv \sigma_s^{-2}$ and the inverse of $\mathbf{\Sigma}_b$, (ii) the precision of the signal errors, as captured by the inverse of $\mathbf{\Sigma}_\varepsilon$, and (iii) the covariances between s and $\mathbf{b}$ contained in $\mathbf{\Omega}$.

After observing the public signals, voters update their belief about $(s, \mathbf{b})$ using Bayes' rule. The marginal distribution of s in the posterior belief is characterised in Lemma C1. In order to state this result, we need to introduce two additional notations: Define $\mathbf{\Lambda} \equiv E\left[(s - \mu_s)(\mathbf{m} - \boldsymbol{\mu}_m)^T\right]$, which is a 1-by-$n$ row vector containing the covariance between s and $\mathbf{m}$. The ith element of $\mathbf{\Lambda}$ is $\lambda_i \equiv Cov(s, m_i) = \sigma_s^2 + \omega_i$. Let $\kappa_{i,j}$ be the element on the ith row and jth column of the precision matrix $\mathbf{\Sigma}_m^{-1}$. The proof of Lemma C1 is shown later in this section.

Lemma C1 *The marginal distribution of s in the voters' posterior belief is a normal distribution with mean*

$$E(s \mid \mathbf{m}) = \mu_s + \sum_{j=1}^{n} \alpha_j \left(m_j - \mu_s - \mu_{b_j}\right), \tag{C.2}$$

and variance

$$var(s \mid \mathbf{m}) = \sigma_s^2 - \sum_{j=1}^{n} \lambda_j \alpha_j, \tag{C.3}$$

where $\alpha_j \equiv \sum_{i=1}^{n} \lambda_i \kappa_{i,j}$ for all $j \in \{1, 2, ..., n\}$.

Similarly to the model in the main text, voter v's ideal policy is given by

$$\delta_v^* = \delta_v + E(s \mid \mathbf{m}).$$

If $x_R \neq x_L$, then this voter will support R if either (i) $x_R > x_L$ and $\delta_v^* > \overline{x}$, or (ii) $x_R < x_L$ and $\delta_v^* < \overline{x}$.

In the presence of multiple signals, the two political parties' common belief about $(s, \mathbf{b})$ is given by a normal distribution $\mathbf{N}\left(\widehat{\boldsymbol{\mu}}_0, \widehat{\mathbf{\Sigma}}_0\right)$ with

$$\widehat{\boldsymbol{\mu}}_0 = \begin{bmatrix} \widehat{\mu}_s \\ \widehat{\boldsymbol{\mu}}_b \end{bmatrix} \qquad \text{and} \qquad \widehat{\mathbf{\Sigma}}_0 = \begin{bmatrix} \widehat{\sigma}_s^2 & \widehat{\mathbf{\Omega}}^T \\ \widehat{\mathbf{\Omega}} & \widehat{\mathbf{\Sigma}}_b \end{bmatrix}.$$

The elements of $\widehat{\boldsymbol{\mu}}_0$ and $\widehat{\mathbf{\Sigma}}_0$ can be interpreted similarly as those of $\boldsymbol{\mu}_0$ and $\mathbf{\Sigma}_0$. Under this belief,

[2] Except for some special cases (such as those considered in Sections D and E), there is no general formula for $\mathbf{\Sigma}_m^{-1}$. Hence, the discussion here should be considered as heuristic in nature.

each signal m_i has an expected value $E_p(m_i) = \widehat{\mu}_s + \widehat{\mu}_{b_i}$. The covariance structure among the n signals is determined by

$$Cov_p(m_i, m_j) = Cov_p(b_i, b_j) + \widehat{\sigma}_s^2 + Cov_p(s, b_i) + Cov_p(s, b_j),$$

where $Cov_p(b_i, b_j)$ is the (i,j)th element of $\widehat{\mathbf{\Sigma}}_b$ and $Cov_p(s, b_i)$ is the ith element of $\widehat{\mathbf{\Omega}}$, for all i, $j \in \{1, 2, ..., n\}$. We use the subscript "p" to indicate that these moments are derived from the parties' belief. From the parties' perspective, $E(s \mid \mathbf{m})$ is a normal random variable with mean

$$E_p[E(s \mid \mathbf{m})] \equiv \widetilde{\mu} = \mu_s + \sum_{j=1}^{n} \alpha_j \left[(\widehat{\mu}_s - \mu_s) + \left(\widehat{\mu}_{b_j} - \mu_{b_j} \right) \right] \tag{C.4}$$

and variance

$$var_p\{E(s \mid \mathbf{m})\} \equiv \widetilde{\sigma}^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j Cov_p(m_i, m_j). \tag{C.5}$$

To derive (C.5), we first rewrite (C.2) as

$$E(s \mid \mathbf{m}) = \left(1 - \sum_{j=1}^{n} \alpha_j \right) \mu_s - \sum_{j=1}^{n} \alpha_j \mu_{b_j} + \sum_{j=1}^{n} \alpha_j m_j.$$

The first two terms in the above equation are deterministic constants. Hence, we can write

$$var_p\{E(s \mid \mathbf{m})\} = var_p\left(\sum_{j=1}^{n} \alpha_j m_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j Cov_p(m_i, m_j).$$

The rest of the model is the same as described in Section 2 of the main text. In particular, the characterisations of x_{eq}^* in Proposition 1 and Corollary 1 are unchanged.

Proof of Lemma C1

The proof is based on a well-known result concerning conditional multivariate normal distributions. This result is as follows [see, for instance, Greene (2012, p.1042, Theorem B.7)]. Suppose $[\mathbf{X}_1, \mathbf{X}_2]$ has a joint multivariable normal distribution $\mathbf{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$, where

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}.$$

The marginal distribution of $\mathbf{X}_i$ is given by $\mathbf{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii})$ for $i \in \{1,2\}$. Then the conditional distribution of $\mathbf{X}_1$ given $\mathbf{X}_2$ is normal with mean vector

$$\boldsymbol{\mu}_{1,2} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2),$$

and covariance matrix

$$\boldsymbol{\Sigma}_{11,2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.$$

In order to apply this result, first note that $(s, \mathbf{b}, \mathbf{m})$ has a joint multivariate normal distribution with mean vector $\boldsymbol{\mu}^\dagger$ and covariance matrix $\boldsymbol{\Sigma}^\dagger$ given by

$$\boldsymbol{\mu}^\dagger = \begin{bmatrix} \mu_s \\ \boldsymbol{\mu}_b \\ \boldsymbol{\mu}_m \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\Sigma}^\dagger = \begin{bmatrix} \sigma_s^2 & \boldsymbol{\Omega} & \boldsymbol{\Lambda} \\ \boldsymbol{\Omega}^T & \boldsymbol{\Sigma}_b & \boldsymbol{\Theta} \\ \boldsymbol{\Lambda}^T & \boldsymbol{\Theta}^T & \boldsymbol{\Sigma}_m \end{bmatrix}.$$

The meaning of $\boldsymbol{\Lambda}$ in the covariance matrix has been explained in the main text. The covariances between $\mathbf{b}$ and $\mathbf{m}$ are captured by the n-by-n matrix $\boldsymbol{\Theta} \equiv E\left[(\mathbf{b} - \boldsymbol{\mu}_\mathbf{b})(\mathbf{m} - \boldsymbol{\mu}_m)^T\right]$. The (i,j)th element of $\boldsymbol{\Theta}$ is denoted by $\theta_{i,j} \equiv Cov(b_i, m_j) = \omega_i + Cov(b_i, b_j)$. Note that this notation is different from the one in the main text where we use λ_b to represent the covariance between b and the signal.

Using the theorem mentioned above, the posterior distribution of $(s, \mathbf{b})$ after observing $\mathbf{m}$ is a normal distribution with mean vector

$$\boldsymbol{\mu}' = \begin{bmatrix} \mu_s \\ \boldsymbol{\mu}_b \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Theta} \end{bmatrix} \boldsymbol{\Sigma}_m^{-1}(\mathbf{m} - \boldsymbol{\mu}_m), \tag{C.6}$$

and covariance matrix

$$\boldsymbol{\Sigma}' = \begin{bmatrix} \sigma_s^2 & \boldsymbol{\Omega} \\ \boldsymbol{\Omega}^T & \boldsymbol{\Sigma}_b \end{bmatrix} - \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Theta} \end{bmatrix} \boldsymbol{\Sigma}_m^{-1} \begin{bmatrix} \boldsymbol{\Lambda}^T & \boldsymbol{\Theta}^T \end{bmatrix}. \tag{C.7}$$

It follows that the marginal distribution of s in the voters' posterior belief is also normal. To derive the posterior mean and posterior variance of s, we first define $\kappa_{i,j}$ as the element on the

ith row and jth column of $\mathbf{\Sigma}_m^{-1}$. Then

$$\begin{aligned}
\begin{bmatrix} \mathbf{\Lambda} \\ \mathbf{\Theta} \end{bmatrix} \mathbf{\Sigma}_m^{-1} (\mathbf{m}-\boldsymbol{\mu}_m) &= \begin{bmatrix} \lambda_1 & \cdots & \lambda_n \\ \theta_{1,1} & \cdots & \theta_{1,n} \\ \vdots & & \vdots \\ \theta_{n,1} & \cdots & \theta_{n,n} \end{bmatrix} \begin{bmatrix} \kappa_{1,1} & \cdots & \kappa_{1,n} \\ \vdots & \ddots & \\ \kappa_{n,1} & \cdots & \kappa_{n,n} \end{bmatrix} \begin{bmatrix} m_1 - \mu_{m_1} \\ \vdots \\ m_n - \mu_{m_n} \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} \lambda_1 & \cdots & \lambda_n \\ \theta_{1,1} & \cdots & \theta_{1,n} \\ \vdots & & \vdots \\ \theta_{n,1} & \cdots & \theta_{n,n} \end{bmatrix}}_{(n+1)\text{-by-}n} \underbrace{\begin{bmatrix} \sum_{j=1}^n \kappa_{1,j} \left(m_j - \mu_{m_j}\right) \\ \vdots \\ \sum_{j=1}^n \kappa_{n,j} \left(m_j - \mu_{m_j}\right) \end{bmatrix}}_{n\text{-by-}1}.
\end{aligned}$$

The first entry in the resulting $(n+1)$-by-1 vector is

$$\mathbf{\Lambda}\mathbf{\Sigma}_m^{-1} (\mathbf{m}-\boldsymbol{\mu}_m) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \kappa_{i,j} \left(m_j - \mu_{m_j}\right).$$

Substituting this into (C.6) gives the posterior mean of s,

$$\begin{aligned}
E(s \mid \mathbf{m}) &= \mu_s + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \kappa_{i,j} \left(m_j - \mu_{m_j}\right) \\
&= \mu_s + \sum_{j=1}^n \underbrace{\left(\sum_{i=1}^n \lambda_i \kappa_{i,j}\right)}_{\alpha_j} \left(m_j - \mu_{m_j}\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\begin{bmatrix} \mathbf{\Lambda} \\ \mathbf{\Theta} \end{bmatrix} \mathbf{\Sigma}_m^{-1} \begin{bmatrix} \mathbf{\Lambda}^T & \mathbf{\Theta}^T \end{bmatrix} &= \begin{bmatrix} \lambda_1 & \cdots & \lambda_n \\ \theta_{1,1} & \cdots & \theta_{1,n} \\ \vdots & & \vdots \\ \theta_{n,1} & \cdots & \theta_{n,n} \end{bmatrix} \begin{bmatrix} \kappa_{1,1} & \cdots & \kappa_{1,n} \\ \vdots & \ddots & \\ \kappa_{n,1} & \cdots & \kappa_{n,n} \end{bmatrix} \begin{bmatrix} \lambda_1 & \theta_{1,1} & \cdots & \theta_{n,1} \\ \vdots & & & \vdots \\ \lambda_n & \theta_{1,n} & \cdots & \theta_{n,n} \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 & \cdots & \lambda_n \\ \theta_{1,1} & \cdots & \theta_{1,n} \\ \vdots & & \vdots \\ \theta_{n,1} & \cdots & \theta_{n,n} \end{bmatrix} \begin{bmatrix} \sum_{j=1}^n \kappa_{1,j}\lambda_j & \sum_{j=1}^n \kappa_{1,j}\theta_{1,j} & \cdots & \sum_{j=1}^n \kappa_{1,j}\theta_{n,j} \\ \vdots & & & \vdots \\ \sum_{j=1}^n \kappa_{n,j}\lambda_j & \sum_{j=1}^n \kappa_{n,j}\theta_{1,j} & \cdots & \sum_{j=1}^n \kappa_{n,j}\theta_{n,j} \end{bmatrix}.
\end{aligned}$$

The $(1,1)$th element of the resulting $(n+1)$-by-$(n+1)$ matrix is

$$\mathbf{\Lambda}\mathbf{\Sigma}_m^{-1}\mathbf{\Lambda}^T = \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\kappa_{i,j}\lambda_j.$$

Substituting this into (C.7) gives the posterior variance of s,

$$var\left(s \mid \mathbf{m}\right) = \sigma_s^2 - \sum_{j=1}^{n}\left(\sum_{i=1}^{n}\lambda_i\kappa_{i,j}\right)\lambda_j.$$

This completes the proof of Lemma C1. ■

D. Unbiased Independent Signals

In this and the next section, we consider two special cases of the multiple signal model. In both cases, voters' posterior expectation of s and parties' perceived uncertainty can be expressed as

$$E(s \mid \mathbf{m}) = \psi\widehat{m} \qquad \text{and} \qquad \widetilde{\sigma}^2 = \psi^2 var_p(\widehat{m}),$$

where $\widehat{m}$ is a sufficient statistic of the observed signals $\{m_1, ..., m_n\}$ and ψ captures the responsiveness of voters' posterior expectation to $\widehat{m}$. Similar to the model in the main text, the coefficient ψ captures the learning effect while $var_p(\widehat{m})$ captures the uncertainty effect.

In the current special case, it is assumed that (i) both voters and politicians believe with certainty that all n signals are unbiased so that each b_i is a deterministic constant and normalised to zero, and (ii) the error terms $\{\varepsilon_1, ..., \varepsilon_n\}$ are independently drawn from different probability distributions. Specifically, each ε_i is assumed to be drawn from a normal distribution $N\left(0, \tau_{\varepsilon_i}^{-1}\right)$, where τ_{ε_i} is the precision of m_i. The expressions of $E(s \mid \mathbf{m})$, $var(s \mid \mathbf{m})$ and $\widetilde{\sigma}^2$ are shown in Lemma D1. The proof of Lemma D1 and Proposition C1 will be shown later in this section.

Lemma D1 *Suppose all the signals are unbiased and each ε_i is independently drawn from the distribution $N\left(0, \tau_{\varepsilon_i}^{-1}\right)$ for all i. Define ψ and $\widehat{m}$ according to*

$$\psi \equiv \frac{\sum_{i=1}^{n} \tau_{\varepsilon_i}}{\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}} > 0 \quad \text{and} \quad \widehat{m} \equiv \sum_{i=1}^{n} \zeta_i m_i, \tag{D.1}$$

where $\tau_s \equiv \sigma_s^{-2}$ and $\zeta_i \equiv \tau_{\varepsilon_i} / \sum_{i=1}^{n} \tau_{\varepsilon_i}$ for all i. Then the mean and variance of s in the voters' posterior beliefs are given by

$$E(s \mid \mathbf{m}) = \psi\widehat{m} \qquad \textit{and} \qquad var(s \mid \mathbf{m}) = \frac{1}{\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}}. \tag{D.2}$$

The political parties' perceived uncertainty is given by $\widetilde{\sigma}^2 = \psi^2 var_p(\widehat{m})$, where

$$var_p(\widehat{m}) \equiv \frac{\widehat{\tau}_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}}{\widehat{\tau}_s \left(\sum_{i=1}^{n} \tau_{\varepsilon_i}\right)}, \tag{D.3}$$

and $\widehat{\tau}_s \equiv \widehat{\sigma}_s^{-2}$.

In this special case, the summary measure $\widehat{m}$ is a weighted average of all the signals whereby more precise signals are weighted more heavily. If the error terms $\{\varepsilon_1, ..., \varepsilon_n\}$ are i.i.d. normal random variables, so that $\tau_{\varepsilon_i} = \tau_\varepsilon$ for all i, then the summation $\sum_{i=1}^{n} \tau_{\varepsilon_i}$ in (D.1)-(D.3) will be replaced by $n\tau_\varepsilon$. The parties' perceived uncertainty then becomes

$$\widetilde{\sigma}^2 = \frac{n\tau_\varepsilon \left(\widehat{\tau}_s + n\tau_\varepsilon\right)}{\widehat{\tau}_s \left(\tau_s + n\tau_\varepsilon\right)^2}.$$

On the other hand, if voters and parties share the same subjective prior beliefs about s so that $\widehat{\tau}_s = \tau_s$, then the parties' perceived uncertainty becomes

$$\widetilde{\sigma}^2 = \frac{\sum_{i=1}^{n} \tau_{\varepsilon_i}}{\tau_s \left(\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}\right)}.$$

Recall that policy polarisation will emerge in a symmetric equilibrium if and only if $\widetilde{\sigma}^2$ exceeds a certain threshold value $\sigma^2_{\min}$. Thus, understanding the relations between $\{\tau_s, \widehat{\tau}_s, \tau_{\varepsilon_1}, ..., \tau_{\varepsilon_n}\}$ and $\widetilde{\sigma}^2$ is essential in understanding how quality of information and disagreement will affect policy polarisation.[3] To this end, we first examine the effects of changing $\{\tau_s, \widehat{\tau}_s, \tau_{\varepsilon_1}, ..., \tau_{\varepsilon_n}\}$ on $\widetilde{\sigma}^2$ in Proposition D1.

Proposition D1 *Suppose all the signals are unbiased and each ε_i is independently drawn from the distribution $N\left(0, \tau_{\varepsilon_i}^{-1}\right)$ for all i.*

(a) *Holding other factors constant, an increase in either τ_s or $\widehat{\tau}_s$ will lower the value of $\widetilde{\sigma}^2$.*

(b) *Holding other factors constant, an increase in τ_{ε_i}, for any $i \in \{1, 2, ..., n\}$, will raise the value of ψ but lower the value of $var_p(\widehat{m})$.*

(c) *Holding other factors constant,*

$$\frac{d\widetilde{\sigma}^2}{d\tau_{\varepsilon_i}} \gtrless 0 \qquad \textit{if and only if} \qquad 2var_p(\widehat{m}) \gtrless var(\widehat{m}), \tag{D.4}$$

for any $i \in \{1, 2, ..., n\}$.

The first part of Proposition D1 states that policy polarisation is less likely to emerge and less severe when either voters or parties are more certain about the hidden state in their prior beliefs. This result can be easily explained through the learning effect and the uncertainty effect. As voters become more certain about s, they will be less reliant on the signals in the learning process. Consequently, their posterior expectation will be less responsive to $\widehat{m}$ (i.e., ψ decreases).

[3] The threshold value $\sigma^2_{\min}$ itself is independent of the precision parameters $\{\tau_s, \widehat{\tau}_s, \tau_{\varepsilon_1}, ..., \tau_{\varepsilon_n}\}$.

From the parties' perspective, this means less *ex ante* uncertainty in the median voter's ideal policy $E(s \mid \mathbf{m})$, hence a lower value of $\widetilde{\sigma}^2$.[4] As explained in Section 2 in the main text, this will strengthen the parties' office motivation and incentivise them to move closer to their opponent's position in order to boost their winning probability. Hence, an increase in τ_s will lower polarisation by weakening the learning effect. An increase in $\widehat{\tau}_s$, on the other hand, has no impact on the voters' learning process. But as the parties' become more certain about the hidden state, they also perceive the signals as less uncertain. This suppresses the uncertainty effect and reduces the extent of polarisation.

The other parts of Proposition D1 analyse the effects of changing a single τ_{ε_i} on $\widetilde{\sigma}^2$. Part (b) shows that such a change will have opposite effects on ψ and $var_p(\widehat{m})$. Firstly, having more precise signals will encourage voters to become more reliant on them when updating their beliefs. This will enhance polarisation by strengthening the learning effect. An increase in τ_{ε_i} also means that the signal m_i becomes more precise which will curb the uncertainty effect and lower polarisation. As in Section 3 of the main text, the overall effect on $\widetilde{\sigma}^2$ can be determined by considering

$$\frac{d\ln\widetilde{\sigma}^2}{d\ln\tau_{\varepsilon_i}} = 2\frac{d\ln\psi}{d\ln\tau_{\varepsilon_i}} + \frac{d\ln var_p(\widehat{m})}{d\ln\tau_{\varepsilon_i}}.$$

The first term on the right captures the contribution of the learning effect, while the second term represents the uncertainty effect. As shown in the proof of Proposition D1,

$$\frac{d\ln\psi}{d\ln\tau_{\varepsilon_i}} = \frac{\tau_{\varepsilon_i}}{\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)^2}\frac{1}{var(\widehat{m})} > 0.$$

$$\frac{d\ln var_p(\widehat{m})}{d\ln\tau_{\varepsilon_i}} = -\frac{\tau_{\varepsilon_i}}{\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)^2}\frac{1}{var_p(\widehat{m})} < 0.$$

The interpretations are essentially the same as in the main text. The result in (D.4) can be obtained by combining these equations.

We can also express this condition in terms of the precision parameters. In the current special case, $2var_p(\widehat{m}) \gtreqless var(\widehat{m})$ if and only if

$$\tau_s \gtreqless \frac{\widehat{\tau}_s\sum_{i=1}^n \tau_{\varepsilon_i}}{\widehat{\tau}_s + 2\sum_{i=1}^n \tau_{\varepsilon_i}}. \tag{D.5}$$

[4] In the extreme case when τ_s is arbitrarily large, $var(s \mid \mathbf{m})$ will converge to zero and $E(s \mid \mathbf{m})$ will converge to the expected value of s in the prior distribution, which is $\mu_s = 0$. The median voter's ideal policy then converges to the median value of δ_v, which is a known constant. This eliminates the uncertainty faced by the parties and paves the way for policy convergence.

Thus, improving the precision of the noisy signals will increase [resp., reduce] perceived uncertainty and polarisation if and only if τ_s is greater [resp., less] than a threshold that is determined by $\widehat{\tau}_s$ and $\sum_{i=1}^{n} \tau_{\varepsilon_i}$. Notice that if there is no disagreement between voters and parties so that $\tau_s = \widehat{\tau}_s$ and $var\left(\widehat{m}\right) = var_p\left(\widehat{m}\right)$, then more precise signals will *always* lead to an increase in $\widetilde{\sigma}^2$ and polarisation. This is no longer the case when voters and political parties disagree. In particular, if the political parties are sufficiently more certain or more knowledgeable on the policy issue (the hidden state) so that $2var_p\left(\widehat{m}\right) < var\left(\widehat{m}\right)$, then more precise signal(s) will reduce polarisation.[5]

Proof of Lemma D1

Suppose each b_i, $i \in \{1, 2, ..., n\}$, is a deterministic constant normalised to zero, and suppose $\mu_s = 0$. Then $(s, \mathbf{m})$ has a joint multivariate normal distribution with zero mean vector and covariance matrix $\boldsymbol{V}$ given by

$$\boldsymbol{V} = \begin{bmatrix} \sigma_s^2 & \boldsymbol{\Lambda}^T \\ \boldsymbol{\Lambda} & \boldsymbol{\Sigma}_m \end{bmatrix},$$

where $\boldsymbol{\Lambda} = \sigma_s^2 \times \mathbf{1}_n$. Given that each ε_i is drawn from the distribution $N\left(0, \sigma_{\varepsilon_i}^2\right)$, where $\sigma_{\varepsilon_i}^2 = \tau_{\varepsilon_i}^{-1}$, then the covariance structure of $\{m_1, ..., m_n\}$ is given by

$$Cov\left(m_i, m_j\right) = \begin{cases} \sigma_s^2 + \sigma_{\varepsilon_i}^2 & \text{for } i = j, \\ \sigma_s^2 & \text{for } i \neq j. \end{cases}$$

Hence, $\boldsymbol{\Sigma}_m$ can be expressed as the sum of two n-by-n matrices,

$$\boldsymbol{\Sigma}_m = \mathbf{A} + \sigma_s^2 \mathbf{1}_n \mathbf{1}_n^T,$$

where $\mathbf{A}$ is a diagonal matrix with diagonal elements $\left(\sigma_{\varepsilon_1}^2, ..., \sigma_{\varepsilon_n}^2\right)$. The inverse of $\boldsymbol{\Sigma}_m$ can be derived using equation (3) in Henderson and Searle (1981, p.53). Specifically, this equation states that for any matrix $\mathbf{M} = \mathbf{A} + r\mathbf{u}\mathbf{v}^T$, where $\mathbf{A}$ can be any invertible matrix, r is a scalar, $\mathbf{u}$ is a

[5] Note that the expression on the right side of (D.5) is strictly lower than $\widehat{\tau}_s$. Hence, part (c) of Proposition D1 implies

$$\frac{d\widetilde{\sigma}^2}{d\tau_{\varepsilon_i}} < 0 \quad \text{iff} \quad \tau_s < \frac{\widehat{\tau}_s \sum_{i=1}^{n} \tau_{\varepsilon_i}}{\widehat{\tau}_s + 2\sum_{i=1}^{n} \tau_{\varepsilon_i}} < \widehat{\tau}_s.$$

column vector and $\mathbf{v}^T$ is a row vector, the inverse can be expressed as

$$\mathbf{M}^{-1} = \mathbf{A}^{-1} - \xi \mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}, \tag{D.6}$$

where

$$\xi = \frac{r}{1 + r\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}.$$

See also the "updating formula" in Greene (2012, p.992). Hence, by setting $r = \sigma_s^2$, $\mathbf{u} = \mathbf{1}_n$ and $\mathbf{v}^T = \mathbf{1}_n^T$, we can get

$$\mathbf{\Sigma}_m^{-1} = \mathbf{A}^{-1} - \xi \mathbf{A}^{-1}\mathbf{1}_n\mathbf{1}_n^T\mathbf{A}^{-1}, \tag{D.7}$$

where

$$\xi = \frac{\sigma_s^2}{1 + \sigma_s^2\mathbf{1}_n^T\mathbf{A}^{-1}\mathbf{1}_n}.$$

Since $\mathbf{A}$ is a diagonal matrix, its inverse is simply

$$\mathbf{A}^{-1} = \begin{bmatrix} \tau_{\varepsilon_1} & 0 & \cdots & 0 \\ 0 & \tau_{\varepsilon_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \tau_{\varepsilon_n} \end{bmatrix}. \tag{D.8}$$

It follows that $\mathbf{1}_n^T\mathbf{A}^{-1}\mathbf{1}_n = \sum_{i=1}^n \tau_{\varepsilon_i}$, and

$$\xi = \frac{\sigma_s^2}{1 + \sigma_s^2\sum_{i=1}^n \tau_{\varepsilon_i}} = \frac{1}{\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}}, \tag{D.9}$$

where $\tau_s \equiv \sigma_s^{-2}$. In addition,

$$\mathbf{A}^{-1}\mathbf{1}_n\mathbf{1}_n^T\mathbf{A}^{-1} = \begin{bmatrix} \tau_{\varepsilon_1} \\ \tau_{\varepsilon_2} \\ \vdots \\ \tau_{\varepsilon_n} \end{bmatrix} \begin{bmatrix} \tau_{\varepsilon_1} & \tau_{\varepsilon_2} & \cdots & \tau_{\varepsilon_n} \end{bmatrix} = \begin{bmatrix} \tau_{\varepsilon_1}^2 & \tau_{\varepsilon_1}\tau_{\varepsilon_2} & \cdots & \tau_{\varepsilon_1}\tau_{\varepsilon_n} \\ \tau_{\varepsilon_1}\tau_{\varepsilon_2} & \tau_{\varepsilon_2}^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \tau_{\varepsilon_1}\tau_{\varepsilon_n} & \cdots & \cdots & \tau_{\varepsilon_n}^2 \end{bmatrix}. \tag{D.10}$$

Using (D.7)-(D.10), we can express the elements on any jth column of $\mathbf{\Sigma}_m^{-1}$ as

$$\kappa_{i,j} = \begin{cases} \tau_{\varepsilon_j} - \xi{\tau_{\varepsilon_j}}^2 & \text{for } i = j, \\ -\xi\tau_{\varepsilon_i}\tau_{\varepsilon_j} & \text{for } i \neq j, \end{cases}$$

$$\Rightarrow \alpha_j = \sum_{i=1}^{n} \lambda_i \kappa_{i,j} = \sigma_s^2 \tau_{\varepsilon_j} \left(1 - \xi \sum_{i=1}^{n} \tau_{\varepsilon_i}\right) = \frac{\tau_{\varepsilon_j}}{\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}}. \tag{D.11}$$

Substituting these and $\lambda_i = \sigma_s^2$ into Equations (C.2) and (C.3) in Section C gives

$$E(s \mid \mathbf{m}) = \sum_{i=1}^{n} \alpha_i m_i = \frac{\sum_{i=1}^{n} \tau_{\varepsilon_i} m_i}{\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}} = \underbrace{\frac{\sum_{i=1}^{n} \tau_{\varepsilon_i}}{\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}}}_{\psi} \cdot \underbrace{\sum_{i=1}^{n} \zeta_i m_i}_{\widehat{m}},$$

where $\zeta_i \equiv \tau_{\varepsilon_i} / \sum_{i=1}^{n} \tau_{\varepsilon_i}$ for all i, and

$$var(s \mid \mathbf{m}) = \frac{1}{\tau_s} - \sum_{i=1}^{n} \lambda_i \alpha_i = \frac{1}{\tau_s} - \frac{1}{\tau_s} \frac{\sum_{i=1}^{n} \tau_{\varepsilon_i}}{\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}} = \frac{1}{\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}}.$$

What remains is to derive the formula for $\widetilde{\sigma}^2$. Under the parties' belief, the covariance structure of $\{m_1, ..., m_n\}$ is given by

$$Cov_p(m_i, m_j) = \begin{cases} \widehat{\sigma}_s^2 + \sigma_{\varepsilon_j}^2 & \text{for } i = j, \\ \widehat{\sigma}_s^2 & \text{for } i \neq j, \end{cases}$$

for any i and j. Substituting these into (C.5) gives

$$\begin{aligned} \widetilde{\sigma}^2 &= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j Cov_p(m_i, m_j) = \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} \alpha_j Cov_p(m_i, m_j) \\ &= \sum_{i=1}^{n} \alpha_i \left[\alpha_i \left(\widehat{\sigma}_s^2 + \sigma_{\varepsilon_i}^2\right) + \sum_{j \neq i} \alpha_j \widehat{\sigma}_s^2 \right] \\ &= \sum_{i=1}^{n} \alpha_i^2 \sigma_{\varepsilon_i}^2 + \sum_{i=1}^{n} \alpha_i \cdot \sum_{j=1}^{n} \alpha_j \cdot \widehat{\sigma}_s^2. \end{aligned} \tag{D.12}$$

Using (D.11), the first term in (D.12) can be simplified as follows:

$$\sum_{i=1}^{n} \alpha_i^2 \sigma_{\varepsilon_i}^2 = \sum_{i=1}^{n} \frac{\tau_{\varepsilon_i}^2 \sigma_{\varepsilon_i}^2}{\left(\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}\right)^2} = \frac{\sum_{i=1}^{n} \tau_{\varepsilon_i}}{\left(\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}\right)^2},$$

since $\tau_{\varepsilon_i} \sigma_{\varepsilon_i}^2 = 1$ for all i. The second term in (D.12) can be simplified as follows:

$$\begin{aligned} \sum_{i=1}^{n} \alpha_i \cdot \sum_{j=1}^{n} \alpha_j \cdot \widehat{\sigma}_s^2 &= \left(\sum_{i=1}^{n} \alpha_i\right)^2 \widehat{\sigma}_s^2 \\ &= \left(\frac{\sum_{i=1}^{n} \tau_{\varepsilon_i}}{\tau_s + \sum_{i=1}^{n} \tau_{\varepsilon_i}}\right)^2 \widehat{\sigma}_s^2. \end{aligned}$$

Hence, we can now rewrite (D.12) as

$$\begin{aligned}
\widetilde{\sigma}^2 &= \frac{\sum_{i=1}^n \tau_{\varepsilon_i}}{\left(\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}\right)^2} + \left(\frac{\sum_{i=1}^n \tau_{\varepsilon_i}}{\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}}\right)^2 \widehat{\sigma}_s^2 \\
&= \underbrace{\left(\frac{\sum_{i=1}^n \tau_{\varepsilon_i}}{\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}}\right)^2}_{\psi^2} \left[\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)^{-1} + \widehat{\sigma}_s^2\right].
\end{aligned}$$

The last step is to show that

$$var_p\left(\widehat{m}\right) = \left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)^{-1} + \widehat{\sigma}_s^2.$$

First, note that

$$\begin{aligned}
var_p\left(\widehat{m}\right) &= var_p\left(\sum_{i=1}^n \zeta_i m_i\right) \\
&= \sum_{i=1}^n \zeta_i \sum_{j=1}^n \zeta_j Cov_p\left(m_i, m_j\right) = \sum_{i=1}^n \zeta_i \left(\widehat{\sigma}_s^2 \sum_{j=1}^n \zeta_j + \zeta_i \sigma_{\varepsilon_i}^2\right).
\end{aligned}$$

Since $\sum_{j=1}^n \zeta_j = 1$ and $\zeta_i \sigma_{\varepsilon_i}^2 = \left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)^{-1}$ for all i, we can simplify the above expression to become

$$var_p\left(\widehat{m}\right) = \widehat{\sigma}_s^2 + \left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)^{-1}.$$

This proves that $\widetilde{\sigma}^2 = \psi^2 var_p\left(\widehat{m}\right).$

Using $\tau_s \equiv \widehat{\sigma}_s^{-2}$, we can show that

$$var_p\left(\widehat{m}\right) = \frac{\left(\widehat{\tau}_s + \sum_{i=1}^n \tau_{\varepsilon_i}\right)}{\widehat{\tau}_s\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)}. \tag{D.13}$$

Using the same line of argument and replacing $\widehat{\sigma}_s^2$ with σ_s^2, we can show that

$$var\left(\widehat{m}\right) = \frac{\left(\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}\right)}{\tau_s\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)}, \tag{D.14}$$

which is the unconditional variance of $\widehat{m}$ under the voters' subjective prior belief. This completes the proof of Lemma D1. ■

Proof of Proposition D1

As shown in the previous proof, the parties' perceived uncertainty $\widetilde{\sigma}^2$ can be expressed as

$$\widetilde{\sigma}^2 = \underbrace{\left(\frac{\sum_{i=1}^n \tau_{\varepsilon_i}}{\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}}\right)^2}_{\psi^2} \cdot \underbrace{\frac{\widehat{\tau}_s + \sum_{i=1}^n \tau_{\varepsilon_i}}{\widehat{\tau}_s \sum_{i=1}^n \tau_{\varepsilon_i}}}_{var_p(\widehat{m})}.$$

It is clear that any changes in τ_s will only affect ψ but not $var_p\left(\widehat{m}\right)$. Likewise, any changes in $\widehat{\tau}_s$ will only affect $var_p\left(\widehat{m}\right)$ but not ψ. Consider the logarithm of ψ,

$$\ln \psi = \ln\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right) - \ln\left(\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}\right).$$

Totally differentiating this with respect to $\{\psi, \tau_s, \tau_{\varepsilon_i}\}$ gives

$$\frac{d\psi}{\psi} = -\frac{\tau_s}{\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}} \frac{d\tau_s}{\tau_s} + \frac{\tau_s \tau_{\varepsilon_i}}{\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)\left(\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}\right)} \frac{d\tau_{\varepsilon_i}}{\tau_{\varepsilon_i}}.$$

Suppose $d\tau_{\varepsilon_i} = 0$, then we have

$$\frac{d\psi}{d\tau_s} = -\frac{\psi}{\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}} < 0 \qquad \Rightarrow \qquad \frac{d\widetilde{\sigma}^2}{d\tau_s} < 0.$$

On the other hand, if $d\tau_s = 0$, then

$$\frac{d\psi}{d\tau_{\varepsilon_i}} = \frac{\tau_s \psi}{\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)\left(\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}\right)} > 0, \tag{D.15}$$

$$\Rightarrow \frac{\tau_{\varepsilon_i}}{\psi}\frac{d\psi}{d\tau_{\varepsilon_i}} = \frac{\tau_{\varepsilon_i}}{\sum_{i=1}^n \tau_{\varepsilon_i}} \frac{\tau_s}{\tau_s + \sum_{i=1}^n \tau_{\varepsilon_i}} = \frac{\tau_{\varepsilon_i}}{\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)^2} \frac{1}{var\left(\widehat{m}\right)}.$$

The second equality follows from (D.14).

Similarly, totally differentiating $\ln\left[var_p\left(\widehat{m}\right)\right]$ with respect to $\{var_p\left(\widehat{m}\right), \widehat{\tau}_s, \tau_{\varepsilon_i}\}$ gives

$$\ln\left[var_p\left(\widehat{m}\right)\right] = \ln\left[\widehat{\tau}_s + \sum_{i=1}^n \tau_{\varepsilon_i}\right] - \ln \widehat{\tau}_s - \ln\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)$$

$$\frac{dvar_p\left(\widehat{m}\right)}{var_p\left(\widehat{m}\right)} = -\frac{\sum_{i=1}^n \tau_{\varepsilon_i}}{\widehat{\tau}_s + \sum_{i=1}^n \tau_{\varepsilon_i}} \frac{d\widehat{\tau}_s}{\widehat{\tau}_s} - \frac{\widehat{\tau}_s \tau_{\varepsilon_i}}{\left(\sum_{i=1}^n \tau_{\varepsilon_i}\right)\left(\widehat{\tau}_s + \sum_{i=1}^n \tau_{\varepsilon_i}\right)} \frac{d\tau_{\varepsilon_i}}{\tau_{\varepsilon_i}}.$$

When all other factors except $\widehat{\tau}_s$ are kept constant,

$$\frac{dvar_p\left(\widehat{m}\right)}{d\widehat{\tau}_s} = -\frac{\sum_{i=1}^{n}\tau_{\varepsilon_i}}{\widehat{\tau}_s + \sum_{i=1}^{n}\tau_{\varepsilon_i}}\frac{var_p\left(\widehat{m}\right)}{\widehat{\tau}_s} < 0 \qquad \Rightarrow \qquad \frac{d\widetilde{\sigma}^2}{d\widehat{\tau}_s} < 0. \tag{D.16}$$

If $d\tau_s = 0$, then

$$\frac{dvar_p\left(\widehat{m}\right)}{d\tau_{\varepsilon_i}} = -\frac{\widehat{\tau}_s var_p\left(\widehat{m}\right)}{\left(\sum_{i=1}^{n}\tau_{\varepsilon_i}\right)\left(\widehat{\tau}_s + \sum_{i=1}^{n}\tau_{\varepsilon_i}\right)} < 0, \tag{D.17}$$

$$\Rightarrow \frac{\tau_{\varepsilon_i}}{var_p\left(\widehat{m}\right)}\frac{dvar_p\left(\widehat{m}\right)}{d\tau_{\varepsilon_i}} = -\frac{\tau_{\varepsilon_i}}{\sum_{i=1}^{n}\tau_{\varepsilon_i}}\frac{\widehat{\tau}_s}{\left(\widehat{\tau}_s + \sum_{i=1}^{n}\tau_{\varepsilon_i}\right)} = -\frac{\tau_{\varepsilon_i}}{\left(\sum_{i=1}^{n}\tau_{\varepsilon_i}\right)^2}\frac{1}{var_p\left(\widehat{m}\right)}.$$

The second equality follows from (D.13). Equations (D.15) and (D.17) together prove the statement in part (b).

Holding τ_s and $\widehat{\tau}_s$ constant, the overall effect of changing τ_{ε_i} on $\widetilde{\sigma}^2$ can be determined by

$$\begin{aligned}\frac{\tau_{\varepsilon_i}}{\widetilde{\sigma}^2}\frac{d\widetilde{\sigma}^2}{d\tau_{\varepsilon_i}} &= 2\frac{\tau_{\varepsilon_i}}{\psi}\frac{d\psi}{d\tau_{\varepsilon_i}} + \frac{\tau_{\varepsilon_i}}{var_p\left(\widehat{m}\right)}\frac{dvar_p\left(\widehat{m}\right)}{d\tau_{\varepsilon_i}} \\ &= \left[\frac{2}{var\left(\widehat{m}\right)} - \frac{1}{var_p\left(\widehat{m}\right)}\right]\frac{\tau_{\varepsilon_i}}{\left(\sum_{i=1}^{n}\tau_{\varepsilon_i}\right)^2}.\end{aligned}$$

Hence,

$$\frac{d\widetilde{\sigma}^2}{d\tau_{\varepsilon_i}} \gtrless 0 \Leftrightarrow 2var_p\left(\widehat{m}\right) \gtrless var\left(\widehat{m}\right).$$

Using (D.13) and (D.14), we can show that

$$2var_p\left(\widehat{m}\right) \gtrless var\left(\widehat{m}\right) \quad \text{ if and only if } \quad \frac{2\tau_s}{\tau_s + \sum_{i=1}^{n}\tau_{\varepsilon_i}} \gtrless \frac{\widehat{\tau}_s}{\widehat{\tau}_s + \sum_{i=1}^{n}\tau_{\varepsilon_i}},$$

which is equivalent to

$$\tau_s \gtrless \frac{\widehat{\tau}_s\sum_{i=1}^{n}\tau_{\varepsilon_i}}{\widehat{\tau}_s + 2\sum_{i=1}^{n}\tau_{\varepsilon_i}}.$$

This completes the proof of Proposition D1. ■

E. Unbiased, Correlated and Exchangeable Signals

In this section we maintain the assumption that all signals are (believed to be) unbiased so that $b_i = 0$ for all i, but the error terms $\{\varepsilon_1, ..., \varepsilon_n\}$ are now assumed to be exchangeable normal random variables. Specifically, this means each ε_i has the same marginal distribution with mean zero and precision τ_ε, and each pair $(\varepsilon_i, \varepsilon_j)$, $i \neq j$, has the same covariance. The covariance matrix $\mathbf{\Sigma}_\varepsilon$ is now given by

$$\mathbf{\Sigma}_\varepsilon = \frac{1}{\tau_\varepsilon} \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & & \ddots & \vdots \\ \rho & \cdots & \rho & 1 \end{bmatrix}, \tag{E.1}$$

where $\rho \geq -1/(n-1)$ is the correlation coefficient between any pair $(\varepsilon_i, \varepsilon_j)$, $i \neq j$. The lower bound of ρ is necessary for $\mathbf{\Sigma}_\varepsilon$ to be positive semi-definite. The resulting expressions of $E(s \mid \mathbf{m})$, $var(s \mid \mathbf{m})$ and $\widetilde{\sigma}^2$ are shown in Lemma E1. The proof of Lemma E1 and Proposition E1 are shown later in this section.

Lemma E1 *Suppose all the signals are unbiased and the error terms $\{\varepsilon_1, ..., \varepsilon_n\}$ are exchangeable normal random variables with zero mean vector and covariance matrix $\mathbf{\Sigma}_\varepsilon$ as shown in (E.1). Define* ψ and $\widehat{m}$ according to

$$\psi \equiv \frac{n\tau_\varepsilon}{n\tau_\varepsilon + \tau_s [1 + (n-1)\rho]} > 0 \qquad \textit{and} \qquad \widehat{m} \equiv \frac{1}{n}\sum_{i=1}^{n} m_i.$$

Then the mean and variance of s in the voters' posterior beliefs are given by

$$E(s \mid \mathbf{m}) = \psi \widehat{m} \qquad \textit{and} \qquad var(s \mid \mathbf{m}) = \frac{1 + (n-1)\rho}{n\tau_\varepsilon + \tau_s [1 + (n-1)\rho]}.$$

The political parties' perceived uncertainty is given by $\widetilde{\sigma}^2 = \psi^2 var_p(\widehat{m})$, where

$$var_p(\widehat{m}) = \frac{n\tau_\varepsilon + \widehat{\tau}_s [1 + (n-1)\rho]}{n\tau_\varepsilon \widehat{\tau}_s}.$$

The results in Proposition D1 can be readily extended to the current case with only minor changes. These are formally stated in the first three parts of Proposition E1. The interpretations are essentially the same as before, hence they are not repeated here.

Proposition E1 *Suppose all the signals are unbiased and the error terms $\{\varepsilon_1, ..., \varepsilon_n\}$ are exchangeable normal random variables with zero mean vector and covariance matrix $\mathbf{\Sigma}_\varepsilon$ as shown in (E.1).*

(a) *Holding other factors constant, an increase in either τ_s or $\widehat{\tau}_s$ will lower the value of $\widetilde{\sigma}^2$.*

(b) *Holding other factors constant, an increase in τ_ε will raise the value of ψ but lower the value of $var_p(\widehat{m})$.*

(c) *Holding other factors constant,*

$$\frac{d\widetilde{\sigma}^2}{d\tau_\varepsilon} \gtrless 0 \qquad \textit{if and only if} \qquad 2var_p(\widehat{m}) \gtrless var(\widehat{m}). \tag{E.2}$$

(d) *Holding other factors constant, an increase in ρ will lower the value of ψ but raise the value of $var_p(\widehat{m})$.*

(e) *Holding other factors constant,*

$$\frac{d\widetilde{\sigma}^2}{d\rho} \gtrless 0 \qquad \textit{if and only if} \qquad 2var_p(\widehat{m}) \lessgtr var(\widehat{m}). \tag{E.3}$$

The last two parts of Proposition E1 concern the effects of ρ on $\widetilde{\sigma}^2$. A higher value of ρ means that the signals $\{m_1, ..., m_n\}$ are more correlated. In the extreme case when $\rho = 1$, all the signals are essentially echoing each other. From the voters' perspective, observing $n > 1$ perfectly correlated signals is no better than observing a single one in terms of learning the hidden state. Thus, a more positive value of ρ will erode the voters' confidence on the signals and weaken the learning effect.[6] The same increase in ρ, however, also raises the parties' perceived variance of $\widehat{m}$, strengthening the uncertainty effect. The overall effect on $\widetilde{\sigma}^2$ again depends on the relative magnitude between $2var_p(\widehat{m})$ and $var(\widehat{m})$. Interestingly, the condition in (E.3) is the exact opposite of the one in (E.2). This suggests that, for any given set of $\{\tau_s, \widehat{\tau}_s, \tau_\varepsilon, n, \rho\}$, τ_ε and ρ tend to have opposite effects on $\widetilde{\sigma}^2$.

In the current special case, $2var_p(\widehat{m}) \gtrless var(\widehat{m})$ if and only if

$$\tau_s \gtrless \frac{n\tau_\varepsilon\widehat{\tau}_s}{2n\tau_\varepsilon + \widehat{\tau}_s[1 + (n-1)\rho]}.$$

[6]The same idea has been put forward by Ortoleva and Snowberg (2015, p.518), but they have not explored the relation between perceived sigal correlation and policy polarisation.

Similar to the findings in Proposition D1, if there is no disagreement between voters' and politicians' beliefs so that $var_p(\widehat{m}) = var(\widehat{m})$, then an increase in the precision of the signals or a decrease in the correlation between signals will raise the parties' perceived uncertainty. However, when voters and politicians disagree, it is possible that an increase in τ_ε or a decrease in ρ will lead to a lower degree of perceived uncertainty. This happens when $\widehat{\tau}_s$ is sufficiently higher than τ_s or when ρ is sufficiently low.

Proof of Lemma E1

Suppose $\rho \geq -1/(n-1)$. The inverse of $\mathbf{\Sigma}_\varepsilon$ can be shown to take the following form

$$\mathbf{\Sigma}_\varepsilon^{-1} = \frac{\tau_\varepsilon}{1+(n-2)\rho-(n-1)\rho^2}\begin{bmatrix} 1+(n-2)\rho & -\rho & \cdots & -\rho \\ -\rho & 1+(n-2)\rho & \cdots & -\rho \\ \vdots & & \ddots & \vdots \\ -\rho & \cdots & -\rho & 1+(n-2)\rho \end{bmatrix}. \tag{E.4}$$

To see this, note that all diagonal entries of $\mathbf{\Sigma}_\varepsilon\mathbf{\Sigma}_\varepsilon^{-1}$ are given by

$$\frac{1}{1+(n-2)\rho-(n-1)\rho^2}\left[1+(n-2)\rho-(n-1)\rho^2\right] = 1,$$

and all off-diagonal elements of $\mathbf{\Sigma}_\varepsilon\mathbf{\Sigma}_\varepsilon^{-1}$ are given by

$$\frac{1}{1+(n-2)\rho-(n-1)\rho^2}\left\{-\rho+[1+(n-2)\rho]\rho+(n-2)\rho^2\right\} = 0.$$

Define the notation ν according to

$$\nu \equiv \frac{\tau_\varepsilon}{1+(n-2)\rho-(n-1)\rho^2} = \frac{\tau_\varepsilon}{(1-\rho)[1+(n-1)\rho]}.$$

The covariances among the signals $\{m_1, ..., m_n\}$ are given by $Cov(m_i, m_j) = \sigma_s^2 + Cov(\varepsilon_i, \varepsilon_j)$, which implies

$$\mathbf{\Sigma}_m = \mathbf{\Sigma}_\varepsilon + +\sigma_s^2\mathbf{1}_n\mathbf{1}_n^T.$$

Using the same formula in (D.6), we can get

$$\mathbf{\Sigma}_m^{-1} = \mathbf{\Sigma}_\varepsilon^{-1} - \xi\mathbf{\Sigma}_\varepsilon^{-1}\mathbf{1}_n\mathbf{1}_n^T\mathbf{\Sigma}_\varepsilon^{-1}, \tag{E.5}$$

where

$$\xi = \frac{\sigma_s^2}{1 + \sigma_s^2 \mathbf{1}_n^T \mathbf{\Sigma}_\varepsilon^{-1} \mathbf{1}_n} = \frac{1}{\tau_s + \mathbf{1}_n^T \mathbf{\Sigma}_\varepsilon^{-1} \mathbf{1}_n}.$$

It is straightforward to show that

$$\mathbf{1}_n^T \mathbf{\Sigma}_\varepsilon^{-1} \mathbf{1}_n = n\nu\left(1-\rho\right) = \frac{n\tau_\varepsilon}{1+\left(n-1\right)\rho}.$$

$$\Rightarrow \xi = \frac{1+\left(n-1\right)\rho}{n\tau_\varepsilon + \tau_s\left[1+\left(n-1\right)\rho\right]} \tag{E.6}$$

On the other hand,

$$\mathbf{\Sigma}_\varepsilon^{-1} \mathbf{1}_n \mathbf{1}_n^T \mathbf{\Sigma}_\varepsilon^{-1} = \nu^2 \left(1-\rho\right)^2 \mathbf{1}_n \mathbf{1}_n^T. \tag{E.7}$$

Using (E.5)-(E.7), we can write the elements on any jth column of $\mathbf{\Sigma}_m^{-1}$ as

$$\kappa_{i,j} = \begin{cases} \nu\left[1+\left(n-2\right)\rho\right] - \xi\nu^2\left(1-\rho\right)^2 & \text{for } i = j, \\ -\nu\rho - \xi\nu^2\left(1-\rho\right)^2 & \text{for } i \neq j. \end{cases}$$

Using these and $\lambda_i = \sigma_s^2 = \tau_s^{-1}$, we can get

$$\begin{aligned} \alpha_j &= \sum_{i=1}^n \lambda_i \kappa_{i,j} = \frac{1}{\tau_s}\left[\nu\left(1-\rho\right) - n\xi\nu^2\left(1-\rho\right)^2\right] \\ &= \frac{\tau_\varepsilon}{\tau_s\left[1+\left(n-1\right)\rho\right]}\left[1 - \frac{n\tau_\varepsilon\xi}{1+\left(n-1\right)\rho}\right] \\ &= \frac{\tau_\varepsilon}{n\tau_\varepsilon + \tau_s\left[1+\left(n-1\right)\rho\right]}. \end{aligned}$$

Hence, the posterior mean and posterior variance of s are given by

$$E\left(s \mid \mathbf{m}\right) = \underbrace{\frac{n\tau_\varepsilon}{n\tau_\varepsilon + \tau_s\left[1+\left(n-1\right)\rho\right]}}_{\psi} \cdot \underbrace{\frac{1}{n}\sum_{i=1}^n m_i}_{\widehat{m}},$$

$$var\left(s \mid \mathbf{m}\right) = \frac{1}{\tau_s}\left(1 - \sum_{i=1}^n \alpha_j\right) = \frac{1+\left(n-1\right)\rho}{n\tau_\varepsilon + \tau_s\left[1+\left(n-1\right)\rho\right]}.$$

From the parties' perspective, the covariance structure of $\{m_1, ..., m_n\}$ is now given by

$$Cov_p\left(m_i, m_j\right) = \begin{cases} \widehat{\tau}_s^{-1} + \tau_\varepsilon^{-1} & \text{for } i = j, \\ \widehat{\tau}_s^{-1} + \tau_\varepsilon^{-1}\rho & \text{for } i \neq j, \end{cases} \tag{E.8}$$

and the perceived uncertainty is given by

$$\widetilde{\sigma}^2 = \left\{ \frac{n\tau_\varepsilon}{n\tau_\varepsilon + \tau_s \left[1 + (n-1)\rho\right]} \right\}^2 var_p\left(\widehat{m}\right),$$

where

$$\begin{aligned} var_p\left(\widehat{m}\right) &= \frac{1}{n^2} \sum_{j=1}^{n} \sum_{i=1}^{n} Cov_p\left(m_i, m_j\right) \\ &= \frac{1}{n^2} \sum_{j=1}^{n} \left\{ n\widehat{\tau}_s^{-1} + \tau_\varepsilon^{-1} \left[1 + (n-1)\rho\right] \right\} \\ &= \frac{n\tau_\varepsilon + \widehat{\tau}_s \left[1 + (n-1)\rho\right]}{n\tau_\varepsilon \widehat{\tau}_s}. \end{aligned} \tag{E.9}$$

Using the same steps, with $\widehat{\tau}_s^{-1}$ replaced by τ_s^{-1} in (E.8), we can show that

$$var\left(\widehat{m}\right) = \frac{n\tau_\varepsilon + \tau_s \left[1 + (n-1)\rho\right]}{n\tau_\varepsilon \tau_s}. \tag{E.10}$$

This completes the proof of Lemma E1. ■

Proof of Proposition E1

Part (a) As shown above,

$$\widetilde{\sigma}^2 = \underbrace{\left\{ \frac{n\tau_\varepsilon}{n\tau_\varepsilon + \tau_s \left[1 + (n-1)\rho\right]} \right\}^2}_{\psi^2} \cdot \underbrace{\frac{n\tau_\varepsilon + \widehat{\tau}_s \left[1 + (n-1)\rho\right]}{n\tau_\varepsilon \widehat{\tau}_s}}_{var_p(\widehat{m})}.$$

It is clear that any changes in τ_s will only affect ψ but not $var_p\left(\widehat{m}\right)$. In particular, ψ (and hence $\widetilde{\sigma}^2$) is strictly decreasing in τ_s when $\rho > -1/(n-1)$. If $\rho = -1/(n-1)$, then ψ, $var_p\left(\widehat{m}\right)$ and $\widetilde{\sigma}^2$ are all independent of τ_s. On the other hand, an increase in $\widehat{\tau}_s$ will lower $\widetilde{\sigma}^2$ because

$$var_p\left(\widehat{m}\right) = \frac{1}{\widehat{\tau}_s} + \frac{\left[1 + (n-1)\rho\right]}{n\tau_\varepsilon},$$

which is strictly decreasing in $\widehat{\tau}_s$, and ψ is independent of $\widehat{\tau}_s$.

Part (b) Consider the logarithm of ψ and $var_p\left(\widehat{m}\right)$,

$$\ln \psi = \ln n + \ln \tau_\varepsilon - \ln \left\{ n\tau_\varepsilon + \tau_s \left[1 + (n-1)\rho\right] \right\},$$

$$\ln\left[var_p\left(\widehat{m}\right)\right] = \ln\left\{n\tau_\varepsilon + \widehat{\tau}_s\left[1+(n-1)\rho\right]\right\} - \ln n - \ln\tau_\varepsilon - \ln\widehat{\tau}_s.$$

Holding $\{\tau_s, \widehat{\tau}_s, \rho, n\}$ constant, consider the total derivatives of ψ and $var_p\left(\widehat{m}\right)$ with respect to τ_ε, i.e.,

$$\frac{d\psi}{\psi} = \frac{\tau_s\left[1+(n-1)\rho\right]}{n\tau_\varepsilon + \tau_s\left[1+(n-1)\rho\right]}\frac{d\tau_\varepsilon}{\tau_\varepsilon} = \underbrace{\frac{1+(n-1)\rho}{n\tau_\varepsilon var\left(\widehat{m}\right)}}_{(+)}\frac{d\tau_\varepsilon}{\tau_\varepsilon}, \tag{E.11}$$

$$\frac{dvar_p\left(\widehat{m}\right)}{var_p\left(\widehat{m}\right)} = -\frac{\widehat{\tau}_s\left[1+(n-1)\rho\right]}{n\tau_\varepsilon + \widehat{\tau}_s\left[1+(n-1)\rho\right]}\frac{d\tau_\varepsilon}{\tau_\varepsilon} = -\frac{1+(n-1)\rho}{n\tau_\varepsilon var_p\left(\widehat{m}\right)}\frac{d\tau_\varepsilon}{\tau_\varepsilon}. \tag{E.12}$$

These show that an increase in τ_ε will raise the value of ψ but lower $var_p\left(\widehat{m}\right)$.

Part (c) The overall effect on $\widetilde{\sigma}^2$ is determined by

$$\frac{\tau_\varepsilon}{\widetilde{\sigma}^2}\frac{d\widetilde{\sigma}^2}{d\tau_\varepsilon} = 2\frac{\tau_\varepsilon}{\psi}\frac{d\psi}{d\tau_\varepsilon} + \frac{\tau_\varepsilon}{var_p\left(\widehat{m}\right)}\frac{dvar_p\left(\widehat{m}\right)}{d\tau_\varepsilon}.$$

Using (E.11) and (E.12), it can be shown that

$$\frac{d\widetilde{\sigma}^2}{d\tau_\varepsilon} \gtrless 0 \quad \Leftrightarrow \quad 2var_p\left(\widehat{m}\right) \gtrless var\left(\widehat{m}\right).$$

The condition on the right side is equivalent to

$$\frac{2\tau_s}{n\tau_\varepsilon + \tau_s\left[1+(n-1)\rho\right]} \gtrless \frac{\widehat{\tau}_s}{n\tau_\varepsilon + \widehat{\tau}_s\left[1+(n-1)\rho\right]},$$

which can be simplified to become

$$\tau_s \gtrless \frac{n\tau_\varepsilon\widehat{\tau}_s}{2n\tau_\varepsilon + \widehat{\tau}_s\left[1+(n-1)\rho\right]}.$$

This establishes the condition in part (c).

Part (d) Holding $\{\tau_s, \widehat{\tau}_s, \tau_\varepsilon, n\}$ constant, consider the total derivatives of ψ and $var_p\left(\widehat{m}\right)$ with respect to ρ, i.e.,

$$\frac{d\psi}{\psi} = -\frac{\tau_s(n-1)\rho}{n\tau_\varepsilon + \tau_s\left[1+(n-1)\rho\right]}\frac{d\rho}{\rho},$$

$$\frac{dvar_p\left(\widehat{m}\right)}{var_p\left(\widehat{m}\right)} = \frac{\widehat{\tau}_s(n-1)\rho}{n\tau_\varepsilon + \widehat{\tau}_s\left[1+(n-1)\rho\right]}\frac{d\rho}{\rho}.$$

Note that these equations are essentially the same as (E.11) and (E.12) but with opposite sides. The desired result can be obtained by using the same steps as in part (c). This completes the proof of Proposition E1. ■

F. Further Results on Learning Effect

In this section, we analyse the effects of changing $\{\tau_s, \tau_b\}$ on ψ^2 within the single-signal model in the main text. Lemma F1 can be viewed as an extension of Lemma 4 in the main paper. When voters and politicians disagree, i.e., when $(\lambda_s, \lambda_b) \neq \left(\widehat{\lambda}_s, \widehat{\lambda}_b\right)$, any changes in $\{\tau_s, \tau_b\}$ will not affect $var_p(m)$. As a result, their effects on ψ^2 will translate directly to $\widetilde{\sigma}^2$ through the equation

$$\frac{d\widetilde{\sigma}^2}{d\rho_{s,b}} = var_p(m) \cdot \frac{d\psi^2}{d\rho_{s,b}},$$

Lemma F1

(a) *Holding other factors constant,*

$$\frac{d\psi^2}{d\tau_s} \gtreqless 0 \qquad \textit{if and only if} \qquad \left(\rho_{s,b} + \frac{\sigma_s}{\sigma_b}\right)\left[\rho_{s,b} + \frac{2\left(\sigma_b^2 + \sigma_\varepsilon^2\right)}{\left(\sigma_s^2 + \sigma_b^2 + \sigma_\varepsilon^2\right)}\frac{\sigma_s}{\sigma_b}\right] \lesseqgtr 0. \tag{F.1}$$

(b) *Holding other factors constant,*

$$\frac{d\psi^2}{d\tau_b} \gtreqless 0 \qquad \textit{if and only if} \qquad \left(\rho_{s,b} + \frac{\sigma_s}{\sigma_b}\right)\left[\rho_{s,b}\left(\sigma_s^2 + \sigma_b^2 - \sigma_\varepsilon^2\right) + 2\sigma_s\sigma_b\right] \gtreqless 0. \tag{F.2}$$

Proof of Lemma F1

Part (a) Given that $Cov(s, m) = \sigma_s^2 + \rho_{s,b}\sigma_s\sigma_b$ and $var(m) = \sigma_s^2 + \sigma_b^2 + \sigma_\varepsilon^2 + 2\rho_{s,b}\sigma_s\sigma_b$, we can write

$$\psi = \frac{Cov(s, m)}{var(m)} = \frac{\sigma_s^2 + \rho_{s,b}\sigma_s\sigma_b}{\sigma_s^2 + \sigma_b^2 + \sigma_\varepsilon^2 + 2\rho_{s,b}\sigma_s\sigma_b}. \tag{F.3}$$

Differentiating this with respect to τ_s gives

$$\frac{d\psi}{d\tau_s} = \frac{\left\{\left(2\sigma_s + \rho_{s,b}\sigma_b\right) var(m) - 2\sigma_s\left(\sigma_s + \rho_{s,b}\sigma_b\right)^2\right\}}{\left[var(m)\right]^2} \underbrace{\frac{d\sigma_s}{d\tau_s}}_{(-)}.$$

The expression inside the curly brackets can be simplified as follows

$$\begin{aligned}
& \left(2\sigma_s + \rho_{s,b}\sigma_b\right) var\left(m\right) - 2\sigma_s\left(\sigma_s + \rho_{s,b}\sigma_b\right)^2 \\
= \quad & \left(2\sigma_s + \rho_{s,b}\sigma_b\right)\left(\sigma_s^2 + \sigma_b^2 + \sigma_\varepsilon^2 + 2\rho_{s,b}\sigma_s\sigma_b\right) - 2\sigma_s\left(\sigma_s + \rho_{s,b}\sigma_b\right)^2 \\
= \quad & 2\sigma_s\left(\sigma_b^2 + \sigma_\varepsilon^2\right) + \rho_{s,b}\sigma_b\left(\sigma_s^2 + \sigma_b^2 + \sigma_\varepsilon^2\right).
\end{aligned}$$

Hence,

$$\frac{d\psi}{d\tau_s} = \frac{\sigma_b\left(\sigma_s^2 + \sigma_b^2 + \sigma_\varepsilon^2\right)}{\left[var\left(m\right)\right]^2}\left[\rho_{s,b} + \frac{2\sigma_s\left(\sigma_b^2 + \sigma_\varepsilon^2\right)}{\sigma_b\left(\sigma_s^2 + \sigma_b^2 + \sigma_\varepsilon^2\right)}\right]\underbrace{\frac{d\sigma_s}{d\tau_s}}_{(-)}.$$

This, together with

$$\psi = \frac{Cov\left(s,m\right)}{var\left(m\right)} = \frac{\sigma_s\sigma_b}{var\left(m\right)}\left(\rho_{s,b} + \frac{\sigma_s}{\sigma_b}\right),$$

implies that

$$\frac{d\psi^2}{d\tau_s} = 2\psi \cdot \frac{d\psi}{d\tau_s} \gtrless 0 \qquad \text{iff} \qquad \left(\rho_{s,b} + \frac{\sigma_s}{\sigma_b}\right)\left[\rho_{s,b} + \frac{2\left(\sigma_b^2 + \sigma_\varepsilon^2\right)}{\left(\sigma_s^2 + \sigma_b^2 + \sigma_\varepsilon^2\right)}\frac{\sigma_s}{\sigma_b}\right] \lessgtr 0.$$

Part (b) Differentiating the expression in (F.3) with respect to τ_b gives

$$\frac{d\psi}{d\tau_b} = \frac{\sigma_s\left\{\rho_{s,b} var\left(m\right) - 2\left(\sigma_s + \rho_{s,b}\sigma_b\right)\left(\sigma_b + \rho_{s,b}\sigma_s\right)\right\}}{\left[var\left(m\right)\right]^2}\underbrace{\frac{d\sigma_b}{d\tau_b}}_{(-)}.$$

The term inside the curly brackets can be simplified as follows:

$$\begin{aligned}
& \rho_{s,b} var\left(m\right) - 2\left(\sigma_s + \rho_{s,b}\sigma_b\right)\left(\sigma_b + \rho_{s,b}\sigma_s\right) \\
= \quad & -\left[\rho_{s,b}\left(\sigma_s^2 + \sigma_b^2 - \sigma_\varepsilon^2\right) + 2\sigma_s\sigma_b\right].
\end{aligned}$$

Hence,

$$\frac{d\psi}{d\tau_b} = -\frac{-\sigma_s}{\left[var\left(m\right)\right]^2}\left[\rho_{s,b}\left(\sigma_s^2 + \sigma_b^2 - \sigma_\varepsilon^2\right) + 2\sigma_s\sigma_b\right]\underbrace{\frac{d\sigma_b}{d\tau_b}}_{(-)}$$

It follows that

$$\frac{d\psi^2}{d\tau_b} = 2\psi \cdot \frac{d\psi}{d\tau_b} \gtrless 0 \qquad \text{iff} \qquad \left(\rho_{s,b} + \frac{\sigma_s}{\sigma_b}\right)\left[\rho_{s,b}\left(\sigma_s^2 + \sigma_b^2 - \sigma_\varepsilon^2\right) + 2\sigma_s\sigma_b\right] \gtrless 0.$$

This completes the proof.

G. Additional Numerical Examples

In this section we provide additional numerical examples that can help illustrate the theoretical results in Proposition 6 in the main text. Examples 1-4 concern the result in part (a) of the proposition, i.e., when $\widehat{\lambda}_s + \widehat{\lambda}_b$ is sufficiently larger than $\lambda_s + \lambda_b$. Example 5 concerns the result in part (b) of the proposition, i.e., when $\lambda_s + \lambda_b$ is sufficiently larger than $\widehat{\lambda}_s + \widehat{\lambda}_b$.

Example 1 This is a continuation of the first example in the main text. To show that a large polarisation-enhancing disagreement is needed for the results shown in Figures 5(a) and 5(b) in the main text, we consider two other cases: In the first one, we lower the values of $\widehat{\lambda}_s$ and $\widehat{\lambda}_b$ to 0.25, while keeping $\lambda_s = \lambda_b = 0.2$. In the second case, we set $\lambda_s = \lambda_b = 0.25$ and $\widehat{\lambda}_s = \widehat{\lambda}_b = 0.20$ so that disagreement is polarisation-reducing. Other parameters are kept unchanged, i.e., $\gamma = 1$ and $\phi \in \{3, 5, 10, 25, 35\}$. The welfare gains in these two cases are shown in Figures G1(a) and (b), respectively. In both cases, the welfare gain from polarisation is a concave graph in τ_ε but strictly increasing. This is true over a much wide range of τ_ε.

Example 2 In this example, we examine the effects of changing γ, which captures the benefits of holding office for the political parties. As in the first example in the main text, we set $(\lambda_s, \lambda_b) = (0.2, 0.2)$ and $\left(\widehat{\lambda}_s, \widehat{\lambda}_b\right) = (0.5, 0.5)$. We consider two different values of ϕ, which are 5 and 25, and five different values of γ, which are $\{0.5, 1, 2, 5, 10\}$. For each pair (ϕ, γ), we compute the degree of policy polarisation x^*_{eq} and the welfare gain from polarisation, $\left[W\left(x^*_{eq}, \delta_v\right) - W\left(0, \delta_v\right)\right]$, over a range of values of τ_ε. The results obtained under $\phi = 5$ are shown in Figure G2, and those obtained under $\phi = 25$ are presented in Figure G3.

In general, increasing the value of γ will intensify the political parties' office motivation and encourage them to converge. In terms of our notation, this will raise the threshold value of ϕ for polarisation to emerge under a given value of $\widetilde{\sigma}$, i.e.,

$$\phi_{\min} \equiv \frac{\gamma}{\sqrt{2\pi}\widetilde{\sigma}}.$$

For instance, when $\phi = 5$ and $\gamma = 10$, the office motivation is sufficiently strong so that convergent equilibrium will emerge (i.e., $x^*_{eq} = 0$) under the specified range of signal precision τ_ε. On the other hand, a lower positive value of γ has the effect of lowering the critical value of τ_ε beyond which

$$\frac{dW\left(x_{eq}^{*},\delta_{v}\right)}{d\tau_{\varepsilon}}<0.$$

In other words, when γ is strictly positive but small, it is more likely that an improvement in signal precision is welfare-reducing. The same pattern is also observed in Figure G3. This happens because political parties have a strong incentive to diverge when the office motivation is weak. This is thus consistent with the intuition related to Figure 3 in the main text .

Note that in the proof of part (a) of Proposition 6, the critical value of τ_{ε} is independent of γ. The results in Figures G2 and G3 suggest that this is true only when $\phi\rightarrow\infty$.

Example 3 In this example, we examine the effects of changing (λ_s,λ_b) while holding their sum constant. As in the first example in the main text, we set $\gamma=1$, $\left(\widehat{\lambda}_s,\widehat{\lambda}_b\right)=(0.5,0.5)$ and $\lambda_s+\lambda_b=0.4$. But now we consider five different combinations of (λ_s,λ_b), which are $\{(0.4,0),(0.3,0.1),(0.2,0.2),(0.1,0.3),(0,0.4)\}$. The results obtained under $\phi=25$ are shown in Figure G4. Recall that the coefficient ψ is defined as

$$\psi\equiv\frac{\lambda_s}{\lambda_s+\lambda_b+\tau_{\varepsilon}^{-1}}.$$

Therefore, holding τ_{ε} and $(\lambda_s+\lambda_b)$ constant, a decrease in the magnitude of λ_s will weaken the learning effect and suppress polarisation. In the extreme case when $\lambda_s=0$ (but $\lambda_s+\lambda_b>0$), $\psi=\widetilde{\sigma}=0$ and a convergent equilibrium will emerge. These effects are observed in the upper panel of Figure G4. The lower panel shows that a higher positive value of λ_s tends to amplify the magnitude of welfare gain (or loss) from policy polarisation, i.e., increase the value of $\left|W\left(x_{eq}^{*},\delta_v\right)-W\left(0,\delta_v\right)\right|$, as well as the curvature of the graph. The diagram clearly shows that voters suffer a welfare loss when τ_{ε} is high and that any further increase will lead to even greater loss.

Note that changing the value of $\left(\widehat{\lambda}_s,\widehat{\lambda}_b\right)$ while holding their sum fixed will have no effect on the degree of polarisation and welfare gain. This is because $\left(\widehat{\lambda}_s,\widehat{\lambda}_b\right)$ affects x_{eq}^{*} and $\left[W\left(x_{eq}^{*},\delta_v\right)-W\left(0,\delta_v\right)\right]$ only through the variance term $var_p(m)$ in $\widetilde{\sigma}^2$, and $var_p(m)\equiv\widehat{\lambda}_s+\widehat{\lambda}_b+\tau_{\varepsilon}^{-1}$.

Example 4 In this example, we consider values of $\left\{\lambda_s,\lambda_b,\widehat{\lambda}_s,\widehat{\lambda}_b\right\}$ that are significantly different from those used in the first example in the main text. Specifically, we now set $\gamma=1$, $\phi=25$, $(\lambda_s,\lambda_b)=(2,2)$ and $\widehat{\lambda}_b=0$. We then consider five different values of $\widehat{\lambda}_s$, which are $\{2,4,6,8,10\}$.

The results are shown in Figure G5. The lower panel shows that when $\widehat{\lambda}_s + \widehat{\lambda}_b$ is sufficiently greater than $\lambda_s + \lambda_b = 4$, improvement in signal precision can be welfare-reducing as predicted by part (a) of Proposition 6.

Example 5 In this example, we set $\gamma = 1$, $\phi = 1$ and $\lambda_b = 0.1$ as in Figures 6(a) and (b) in the main text. But we now consider a much larger value of $\widehat{\lambda}_s$ and $\widehat{\lambda}_b$, which is $\widehat{\lambda}_s = \widehat{\lambda}_b = 0.4$. Figure G6 shows the results obtained under five different values of λ_s, which are $\{1, 2.5, 5, 7.5, 10\}$. The lower panel shows that when λ_s is large, so that $\lambda_s + \lambda_b$ is sufficiently greater than $\widehat{\lambda}_s + \widehat{\lambda}_b$, then improvement in signal precision can be welfare-reducing.

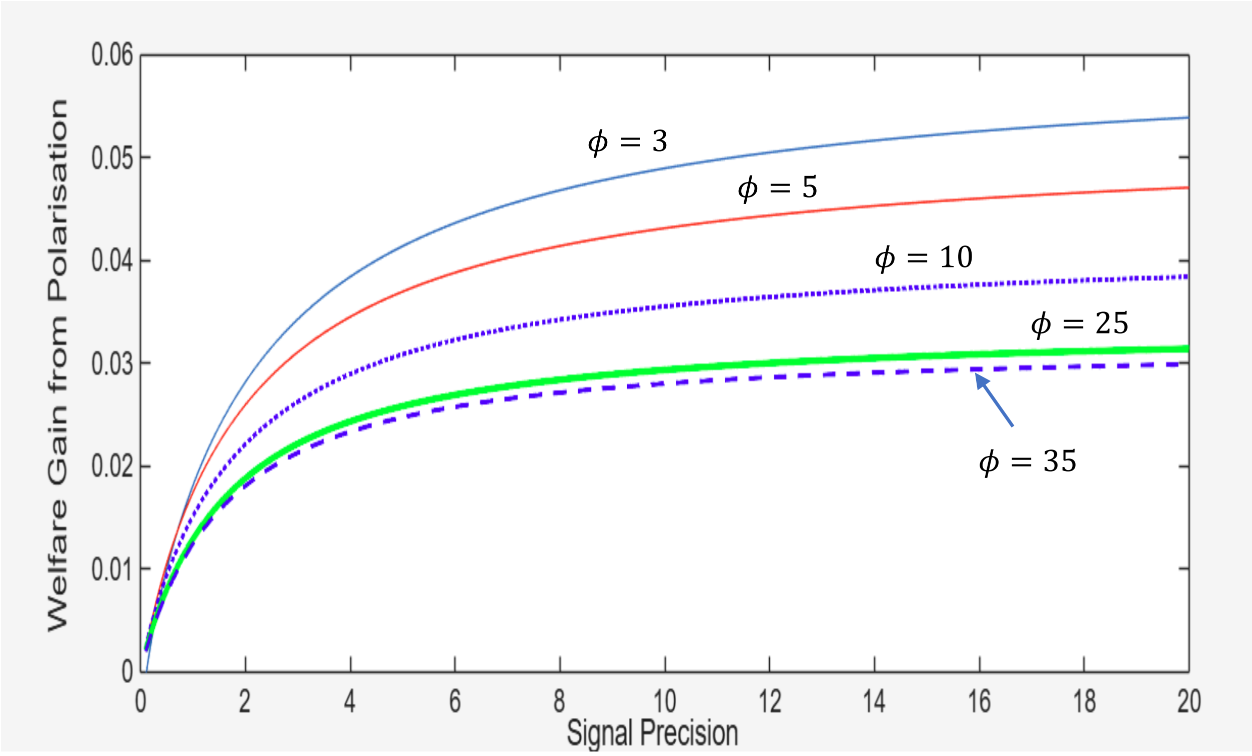


Figure G1(a) Numerical Results of Example 1

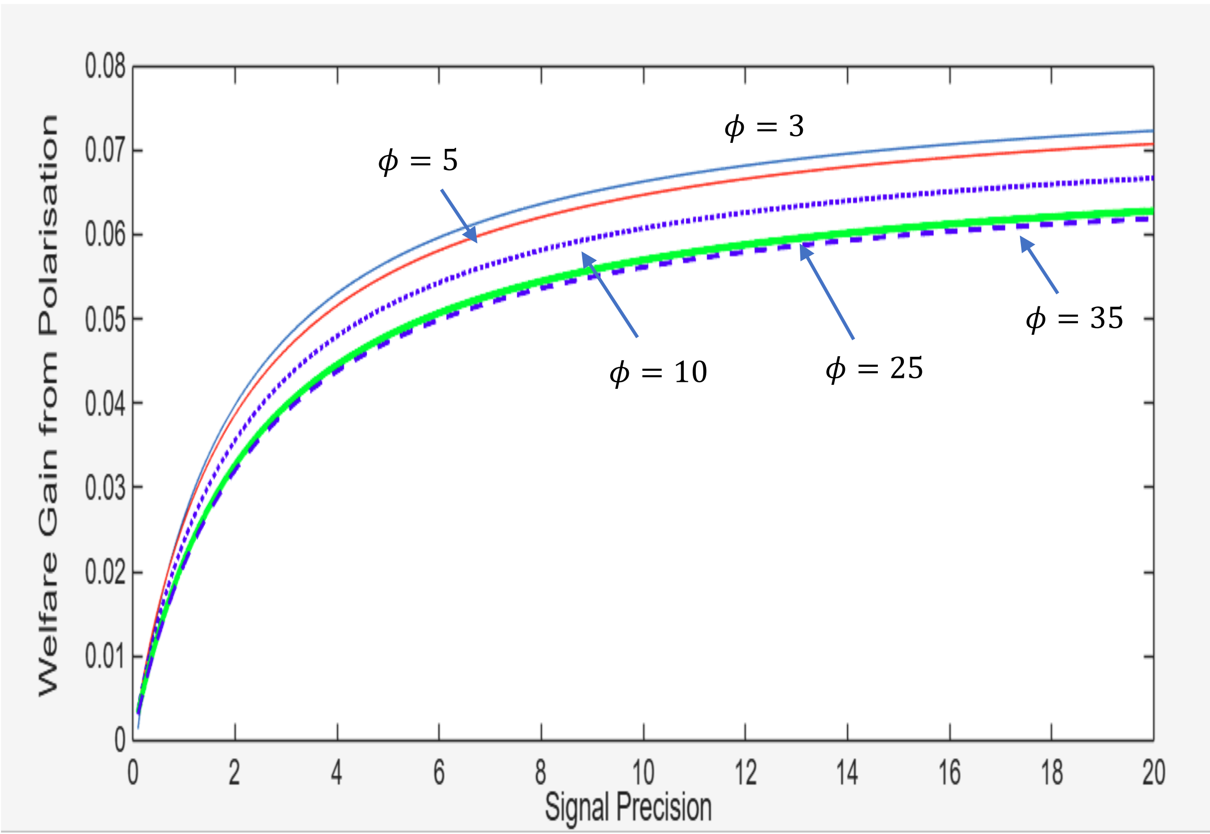
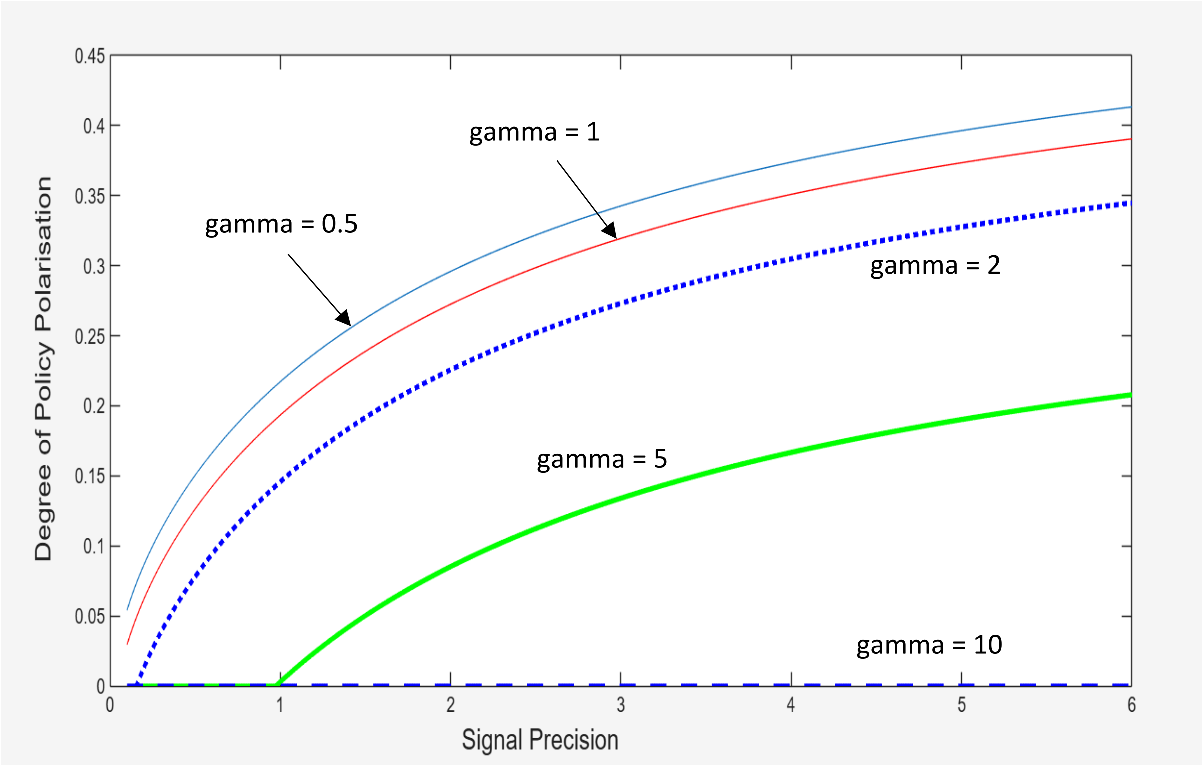


Figure G1(b): Numerical Results of Example 1



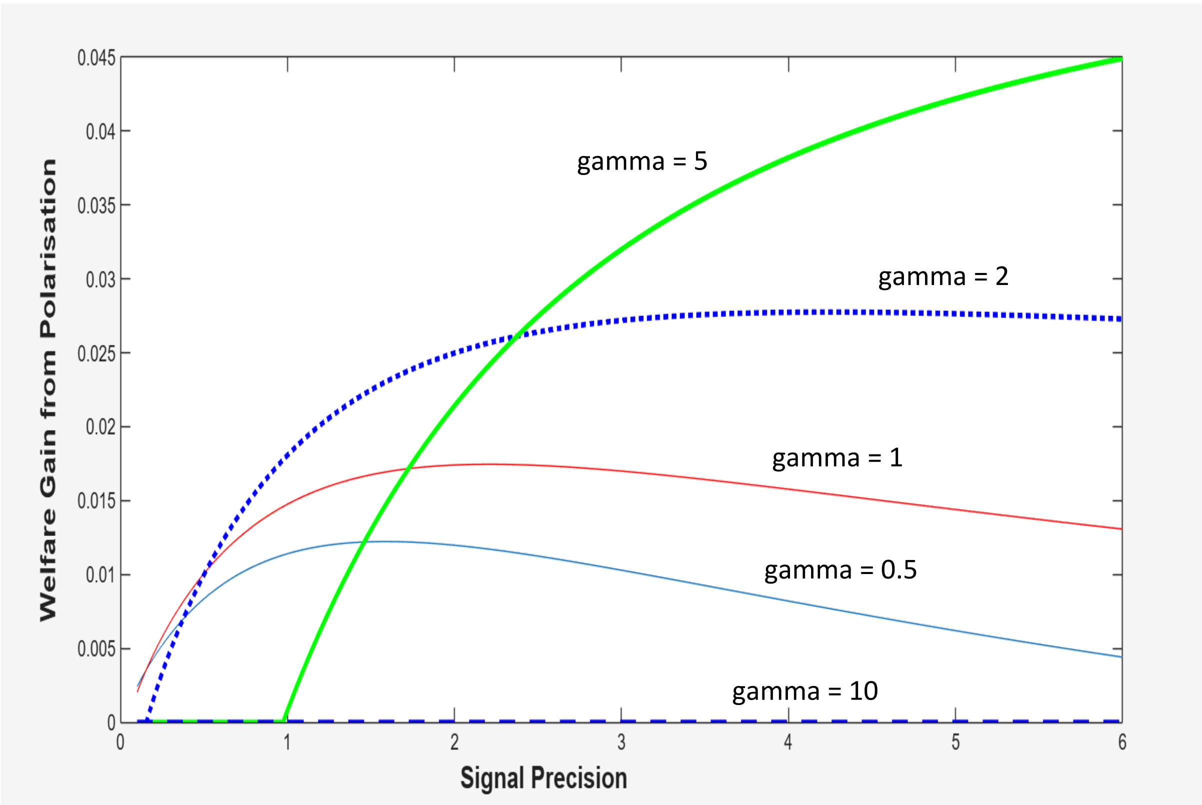
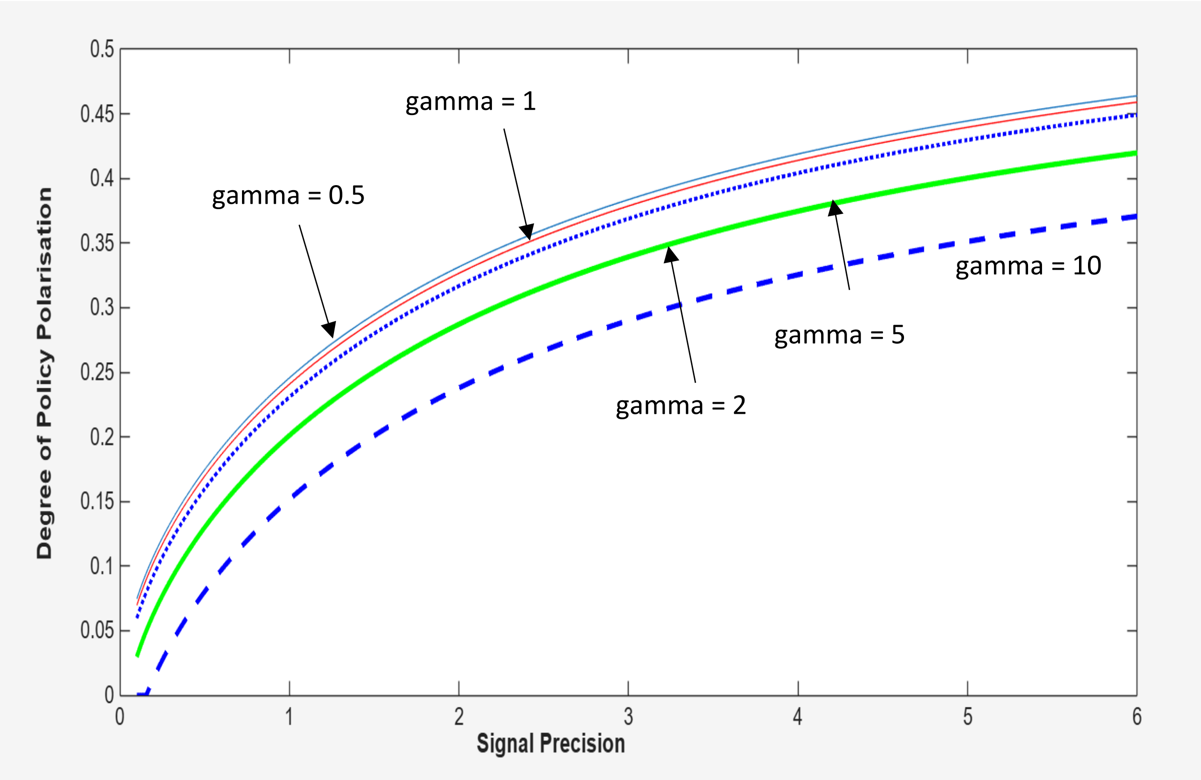


Figure G2: Numerical Results of Example 2 when ϕ = 5.



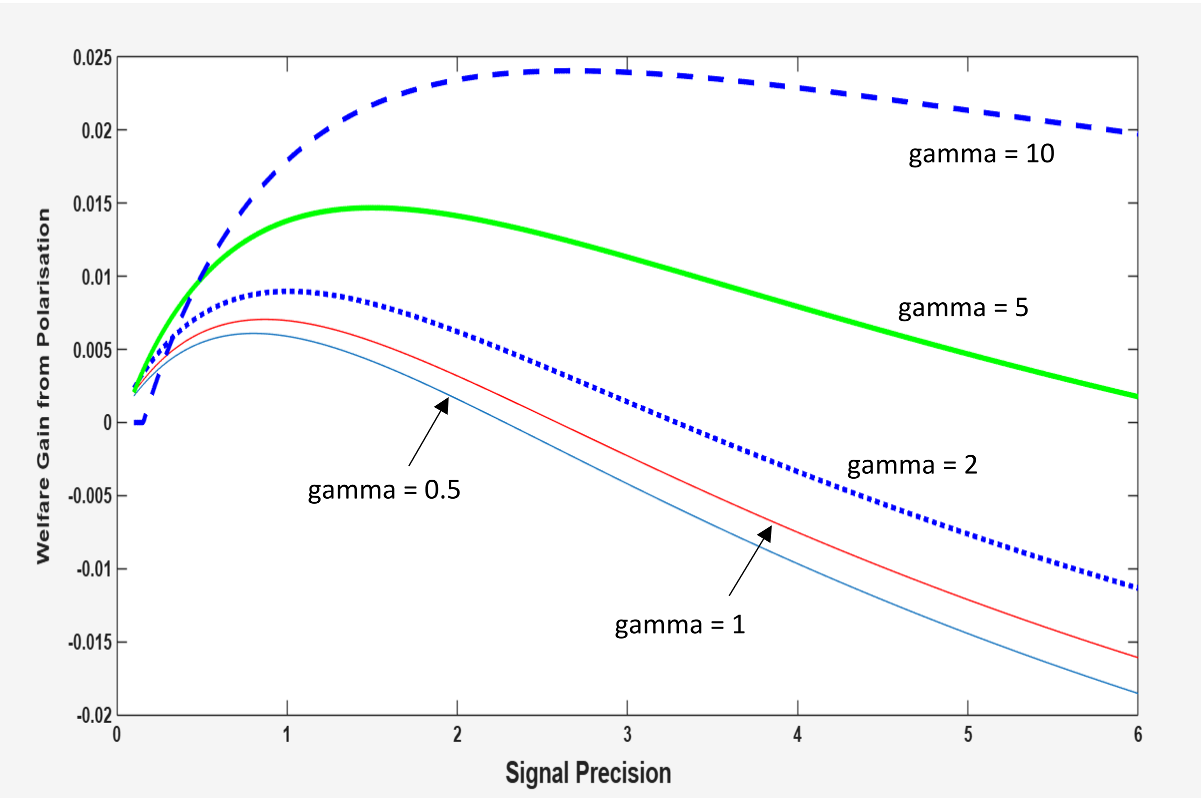
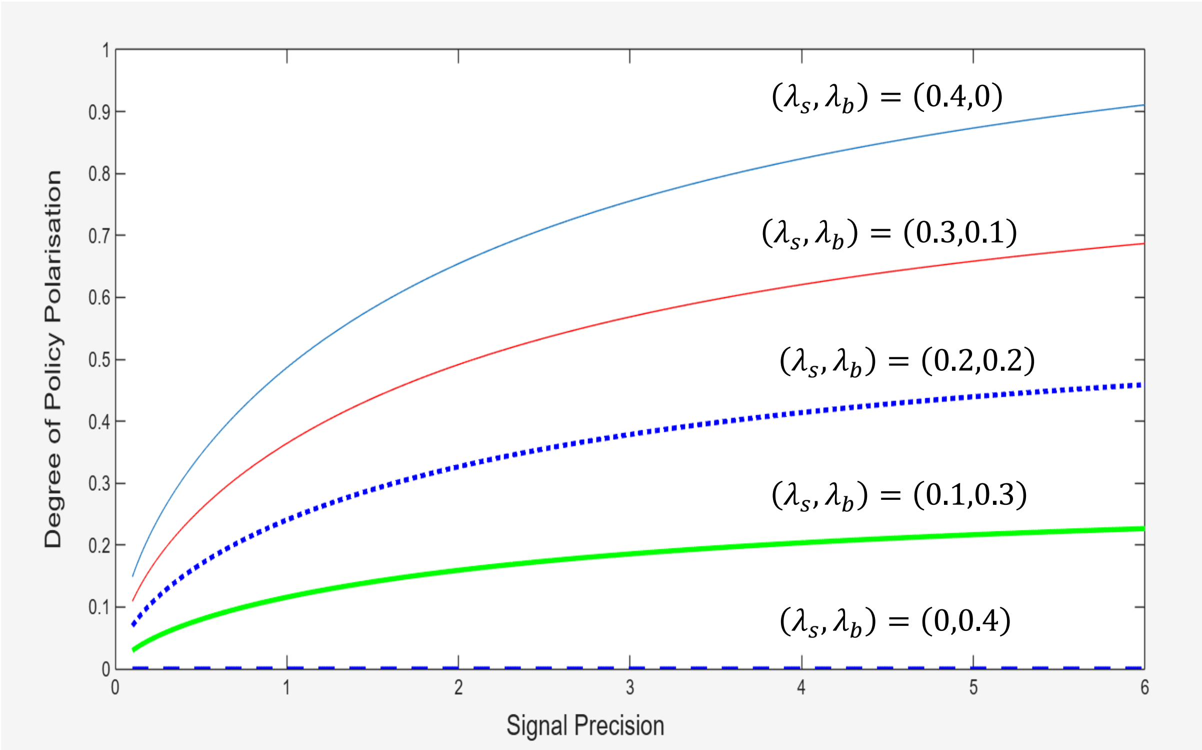


Figure G3: Numerical Results of Example 2 when $\phi = 25$.



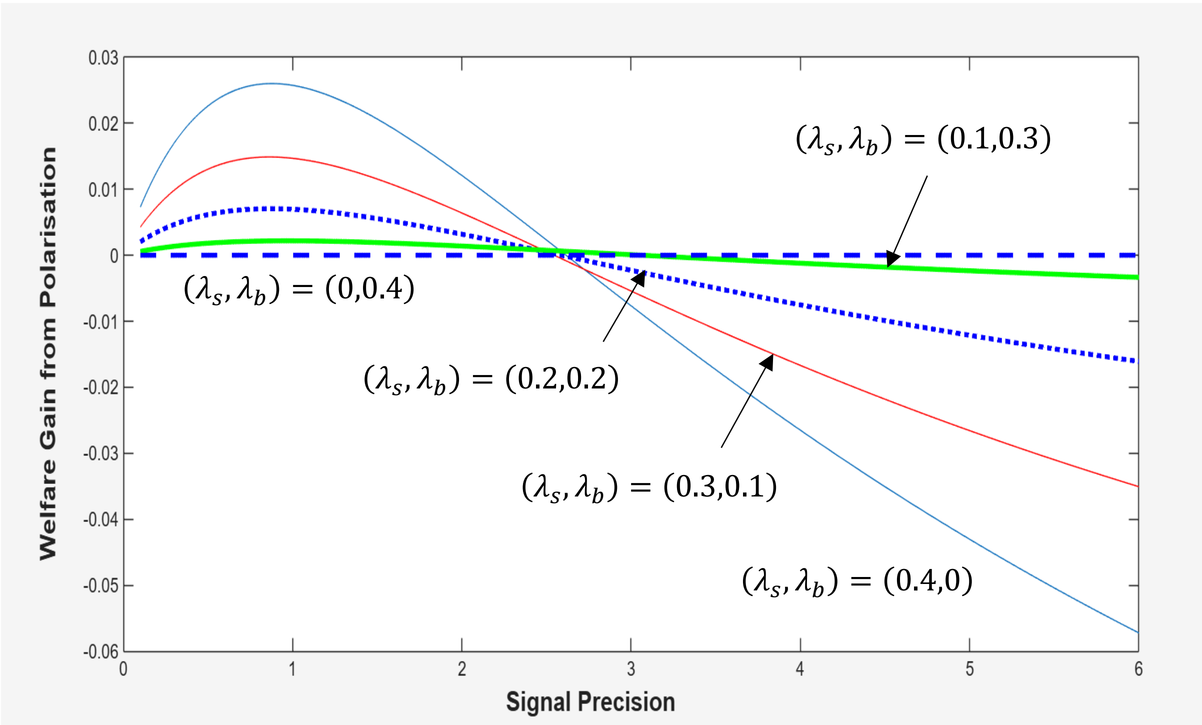


Figure G4: Numerical Results of Example 3.

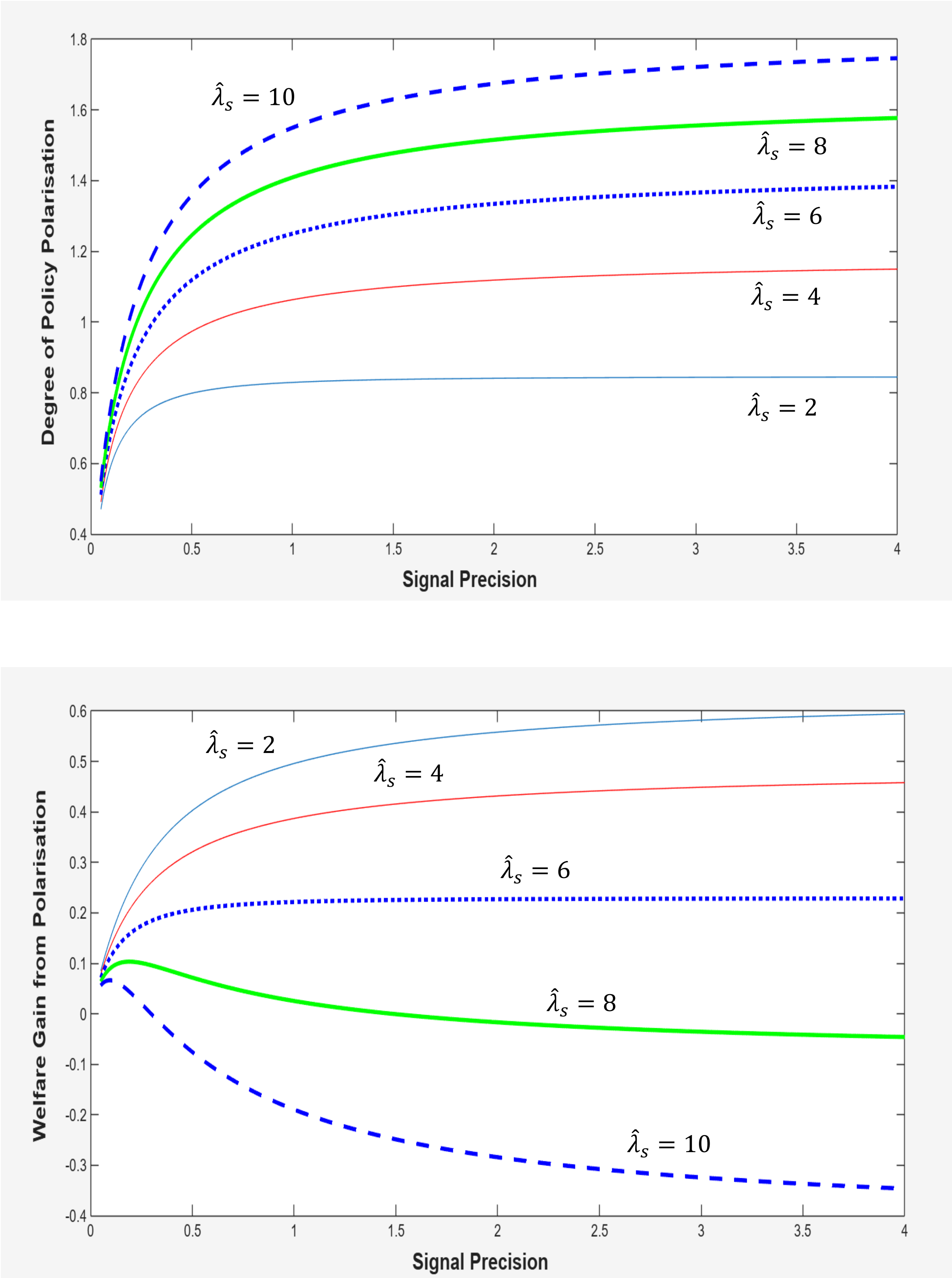


Figure G5: Numerical Results of Example 4

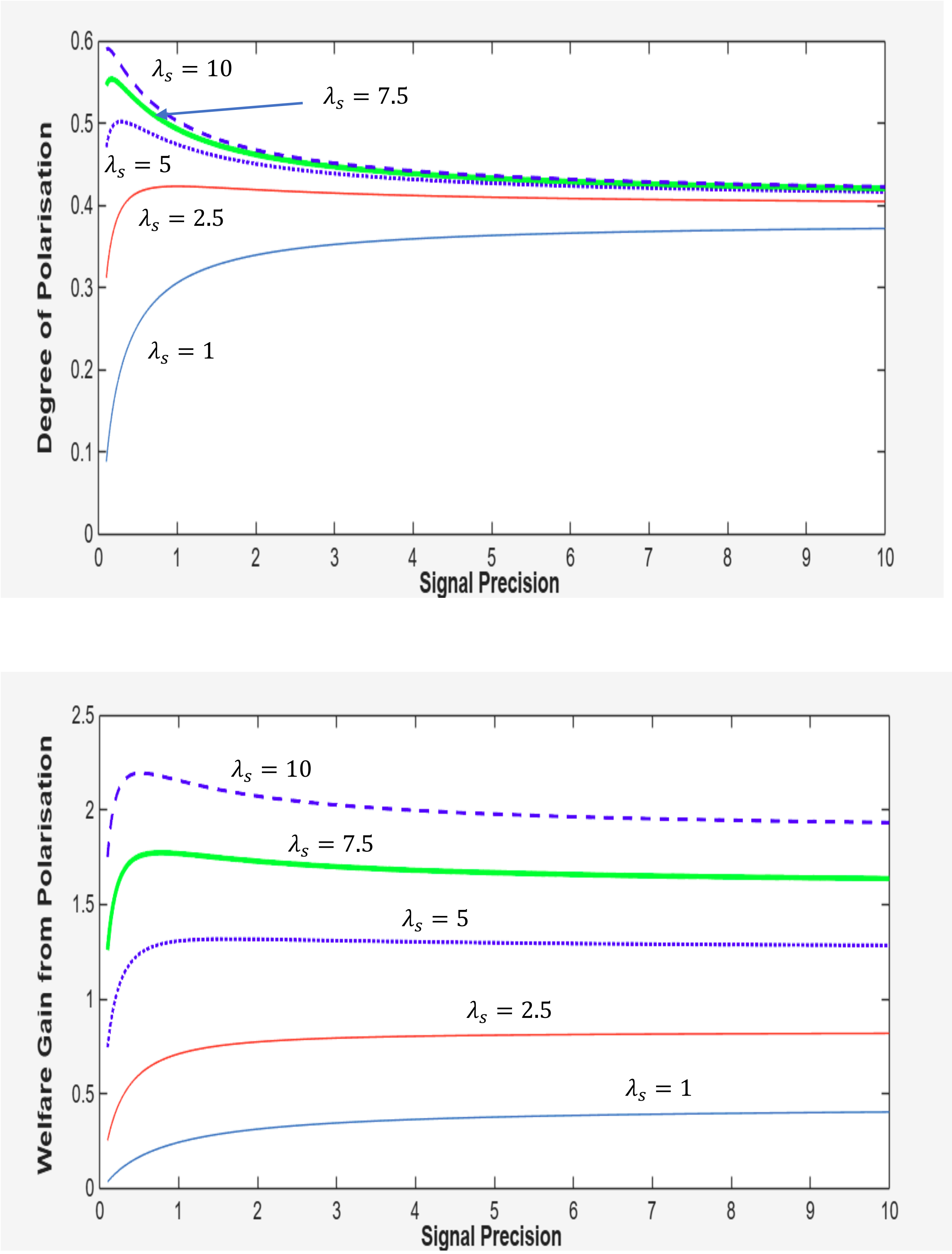


Figure G6: Numerical Results of Example 5

References

[1] Bagnoli, M., and T. Bergstrom (2005) "Log-Concave Probability and its Applications," *Economic Theory*, vol. 26(2), pp. 445-469.

[2] Greene. W. H. (2012) *Econometric Analysis*, 7th edition, New York: Prentice Hall.

[3] Henderson, H. V., and S. R. Searle (1981) "On Deriving the Inverse of a Sum of Matrices," *SIAM Review*, vol. 23(1), pp. 53-60.

[4] Ortoleva, P., and E. Snowberg (2015) "Overconfidence in Political Behavior," *American Economic Review*, vol. 105 (2), pp. 504-535.