ONLINE SUPPLEMENTAL APPENDIX
Motivated reasoning and democratic accountability
American Political Science Review

Andrew T. Little* Keith E. Schnakenberg† Ian R. Turner‡

September 2021

Contents

*Associate Professor, Department of Political Science, University of California, Berkeley. Contact: andrew.little@berkeley.edu.
†Associate Professor, Department of Political Science, Washington University in St. Louis. Contact: keschnak@wustl.edu.
‡Assistant Professor, Department of Political Science, Yale University. Contact: ian.turner@yale.edu.

A Main examples

In this section we provide the derivations for each of the three examples: polarized partisanship, spatial motivations, and confirmation bias. We also show when each optimal conclusion approaches the mean of the Bayesian belief. In general we have that each voter j forms an optimal conclusion by maximizing the following with respect to $\tilde{\theta}_I$,

$$\begin{aligned}\log f_{\theta_I|s}(\tilde{\theta}_I|s) + \delta v(a_j, \tilde{\theta}_I) &= \log\left(\frac{1}{\overline{\sigma}_\theta\sqrt{2\pi}} e^{-\frac{(\tilde{\theta}_I - \overline{\mu}(s))^2}{2\overline{\sigma}_\theta^2}}\right) + \delta v(a_j, \tilde{\theta}_I) \\ &= -\log(\overline{\sigma}_\theta) - \frac{1}{2}\log(2\pi) - \frac{(\tilde{\theta}_I - \overline{\mu}(s))^2}{2\overline{\sigma}_\theta^2} + \delta v(a_j, \tilde{\theta}_I).\end{aligned}$$

Differentiating yields the general first-order condition:

$$-\frac{\tilde{\theta}_I - \overline{\mu}(s)}{\overline{\sigma}_\theta^2} + \delta \frac{\partial v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I} = 0. \tag{1}$$

For each example we need only plug in the particular functional form for $v(a_j, \tilde{\theta}_I)$.

A.1 Polarized partisanship.

In the first example in which voters are motivated to form 'large' conclusions (in absolute terms), in the direction of their affinities, we set $v(a_j, \tilde{\theta}_I) = \tilde{\theta}_I a_j$. Thus, $\frac{\partial v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I} = a_j$. Plugging this into (1) we recover j's optimal conclusion from the first example:

$$\begin{aligned}-\frac{\tilde{\theta}_I - \overline{\mu}(s)}{\overline{\sigma}_\theta^2} + \delta a_j &= 0, \\ \tilde{\theta}_I &= \overline{\mu}(s) + \delta \overline{\sigma}_\theta^2 a_j.\end{aligned}$$

It is straightforward to see that the optimal conclusion approaches the mean of the Bayesian posterior when δ, $\overline{\sigma}_\theta^2$, and a_j approach 0:

$$\begin{aligned}\lim_{\delta \to 0} \left[\overline{\mu}(s) + \delta \overline{\sigma}_\theta^2 a_j\right] &= \overline{\mu}(s), \\ \lim_{\overline{\sigma}_\theta^2 \to 0} \left[\overline{\mu}(s) + \delta \overline{\sigma}_\theta^2 a_j\right] &= \overline{\mu}(s), \\ \lim_{a_j \to 0} \left[\overline{\mu}(s) + \delta \overline{\sigma}_\theta^2 a_j\right] &= \overline{\mu}(s).\end{aligned}$$

A.2 Confirmation bias.

In the second example we set $v(a_j, \tilde{\theta}_I) = -\tilde{\theta}_I^2$ so voters are motivated to form conclusions near their prior of 0. Thus, $\frac{\partial v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I} = -2\tilde{\theta}_I$. This yields the first-order condition,

$$-\frac{\tilde{\theta}_I - \overline{\mu}(s)}{\overline{\sigma}_\theta^2} - 2\delta\tilde{\theta}_I = 0,$$

$$\tilde{\theta}_I = \frac{\overline{\mu}(s)}{1 + 2\delta\overline{\sigma}_\theta^2},$$

In terms of when the optimal conclusion approaches the fully Bayesian benchmark we have:

$$\lim_{\delta \to 0} \left[\frac{\overline{\mu}(s)}{1 + 2\delta\overline{\sigma}_\theta^2} \right] = \overline{\mu}(s),$$

$$\lim_{\overline{\sigma}_\theta^2 \to 0} \left[\frac{\overline{\mu}(s)}{1 + 2\delta\overline{\sigma}_\theta^2} \right] = \overline{\mu}(s),$$

and that a_j does not impact distortions in this case.

A.3 Spatial motivations.

In the final example where voters are motivated to match their conclusions to their affinity for the incumbent we set $v(a_j, \tilde{\theta}_I) = -(a_j - \tilde{\theta}_I)^2$. Accordingly, $\frac{\partial v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I} = 2(a_j - \tilde{\theta}_I)$. Plugging in to (1) we have,

$$-\frac{\tilde{\theta}_I - \overline{\mu}(s)}{\overline{\sigma}_\theta^2} + \delta 2(a_j - \tilde{\theta}_I) = 0,$$

$$\tilde{\theta}_I = \frac{\overline{\mu}(s) + 2\delta a_j \overline{\sigma}_\theta^2}{1 + 2\delta\overline{\sigma}_\theta^2},$$

which can be rewritten as

$$\tilde{\theta}_I = \frac{1}{1 + 2\delta\overline{\sigma}_\theta^2}\overline{\mu}(s) + \frac{2\delta\overline{\sigma}_\theta^2}{1 + 2\delta\overline{\sigma}_\theta^2} a_j.$$

We can characterize when the optimal conclusion approaches the mean of the Bayesian posterior as follows:

$$\lim_{\delta\to 0}\left[\frac{\overline{\mu}(s)+2\delta a_j\overline{\sigma}_\theta^2}{1+2\delta\overline{\sigma}_\theta^2}\right]=\overline{\mu}(s),$$
$$\lim_{\overline{\sigma}_\theta^2\to 0}\left[\frac{\overline{\mu}(s)+2\delta a_j\overline{\sigma}_\theta^2}{1+2\delta\overline{\sigma}_\theta^2}\right]=\overline{\mu}(s),$$
$$\lim_{a_j\to 0}\left[\frac{\overline{\mu}(s)+2\delta a_j\overline{\sigma}_\theta^2}{1+2\delta\overline{\sigma}_\theta^2}\right]=\frac{\overline{\mu}(s)}{1+2\delta\overline{\sigma}_\theta^2}.$$

Since $1+2\delta\overline{\sigma}_\theta^2>1$, a voter with $a_j=0$ only has a conclusion equal to the Bayesian mean if $\overline{\mu}(s)$ is exactly equal to zero (which happens with probability zero). So, motivated reasoning still manifests in this case even for completely "neutral" voters.

B Proofs of results

B.1 Proposition 1

Proposition 1. *Under Assumption 1:*
(i) there exists a unique optimal conclusion $\tilde{\theta}_I^*(s,a_j,\delta;\hat{e})$ for each voter $j\in N$,
(ii) if $\frac{\partial^2 v(a_j,\tilde{\theta}_I)}{\partial a_j\partial\tilde{\theta}_I}>0$, then the optimal conclusion is strictly increasing in the voter affinity ($\frac{\partial\tilde{\theta}_I^*}{\partial a_j}>0$), and the strength of this relationship is increasing in the directional motive ($\frac{\partial^2\tilde{\theta}_I^*}{\partial a_j\partial\delta}>0$), and
(iii) the optimal conclusion is strictly increasing in the signal of performance ($\frac{\partial\tilde{\theta}_I^*}{\partial s}>0$), and if v is strictly concave in θ, then the strength of this relationship is strictly decreasing in the directional motive ($\frac{\partial^2\tilde{\theta}_I^*}{\partial s\partial\delta}<0$).

Proof of Proposition 1. The first-order condition for an optimal conclusion is

$$-\frac{\tilde{\theta}_I-\overline{\mu}(s)}{\overline{\sigma}_\theta^2}+\delta\frac{\partial v(a_j,\tilde{\theta}_I)}{\partial\tilde{\theta}_I}=0. \tag{2}$$

For part (i). The first term is linear and strictly decreasing in $\tilde{\theta}_I$, and the second term is weakly decreasing in $\tilde{\theta}_I$, and so

$$\lim_{\theta\to\infty}\left[-\frac{\tilde{\theta}_I-\overline{\mu}(s)}{\overline{\sigma}_\theta^2}+\delta\frac{\partial v(a_j,\tilde{\theta}_I)}{\partial\theta}\right]=-\infty$$

and

$$\lim_{\theta\to-\infty}\left[-\frac{\tilde{\theta}_I-\overline{\mu}(s)}{\overline{\sigma}_\theta^2}+\delta\frac{\partial v(a_j,\tilde{\theta}_I)}{\partial\tilde{\theta}_I}\right]=\infty.$$

Thus, we have $-\frac{\theta_I - \overline{\mu}(s)}{\overline{\sigma}^2_\theta} + \delta \frac{\partial v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I} > 0$ for some $\tilde{\theta}_I < 0$ and $-\frac{\tilde{\theta}_I - \overline{\mu}(s)}{\overline{\sigma}^2_\theta} + \delta \frac{\partial v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I} < 0$ for some $\theta > 0$. By continuity, there exists some $\tilde{\theta}_I^*(s, a_j, \delta; \hat{e})$ that solves (2). Strict concavity of the objective function implies that this is the unique maximum.

For part (*ii*), applying the implicit function theorem to (2) gives

$$\frac{\partial \tilde{\theta}_I^*}{\partial a_j} = -\frac{\delta \frac{\partial^2 v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j}}{-\frac{1}{\overline{\sigma}^2_\theta} + \delta \frac{\partial^2 v}{\partial \tilde{\theta}_I^2}} = \frac{\frac{\partial^2 v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j}}{\frac{1}{\delta \overline{\sigma}^2_\theta} - \frac{\partial^2 v}{\partial \tilde{\theta}_I^2}}$$

If $\frac{\partial^2 v(a_j, \tilde{\theta}_I)}{\partial a_j \partial \tilde{\theta}_I} = 0$ then this derivative is zero.

If $\frac{\partial^2 v(a_j, \tilde{\theta}_I)}{\partial a_j \partial \tilde{\theta}_I} > 0$, then the numerator of the right-most expression is strictly positive, and the denominator must be strictly positive since $\frac{1}{\overline{\sigma}^2_\theta} > 0$ and $\delta \frac{\partial^2 v}{\partial \tilde{\theta}_I^2} \leq 0$, which implies $\frac{\partial \tilde{\theta}^*}{\partial a_j} > 0$. Further, the denominator is decreasing in δ, and hence $\frac{\partial \tilde{\theta}_I^*}{\partial a_j}$ is strictly increasing in δ.

For part (*iii*), implicitly differentiating the first-order condition with respect to s gives

$$\frac{\partial \tilde{\theta}_I^*}{\partial s} = -\frac{\frac{\overline{\mu}'(s)}{\overline{\sigma}^2_\theta}}{-\frac{1}{\overline{\sigma}^2_\theta} + \delta \frac{\partial^2 v}{\partial \tilde{\theta}_I^2}} = \frac{\frac{\overline{\mu}'(s)}{\overline{\sigma}^2_\theta}}{\frac{1}{\overline{\sigma}^2_\theta} - \delta \frac{\partial^2 v}{\partial \tilde{\theta}_I^2}}.$$

The numerator and denominator in the right-most expression are both strictly positive, hence $\frac{\partial \tilde{\theta}^*}{\partial s} > 0$. If $\frac{\partial^2 v}{\partial \tilde{\theta}_I^2} = 0$ then this derivative is not a function of δ, and if $\frac{\partial^2 v}{\partial \tilde{\theta}_I^2} < 0$ it is decreasing in δ. ■

B.2 Lemma 1

Lemma 1. *Under Assumptions 1 and 2,*
(i) the optimal conclusion is linear in a_j and s. In particular, it can be written:

$$\tilde{\theta}_I^*(s, a_j, \delta; \hat{e}) = \alpha_0 + \alpha_1 a_j + \beta(s - \hat{e}), \tag{3}$$

where $\alpha_1 \geq 0$ and $\beta \geq 0$.
(ii) α_1 is strictly increasing in δ if and only if $\frac{\partial^2 v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j} > 0$.
(iii) β is strictly decreasing in δ if and only if v is strictly concave in $\tilde{\theta}_I$.

Proof of Lemma 1. Recall that the first-order condition for an optimal conclusion is

$$-\frac{\tilde{\theta}_I - \overline{\mu}(s)}{\overline{\sigma}^2_\theta} + \delta \frac{\partial v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I} = 0.$$

With assumption 2, we can write $\frac{\partial v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I} = \gamma_0 + \gamma_\theta \tilde{\theta}_I + \gamma_a a_j$. Solving for $\tilde{\theta}_I$ yields

$$\begin{aligned}
\tilde{\theta}_I &= \frac{\overline{\mu}(s) + \gamma_0 \delta \overline{\sigma}_\theta^2 + a_j \delta \gamma_a \overline{\sigma}_\theta^2}{1 - \delta \gamma_\theta \overline{\sigma}_\theta^2} \\
&= \frac{\frac{\sigma_\varepsilon^2 (s-\hat{e})}{\sigma_\varepsilon^2 + \sigma_\theta^2} + \gamma_0 \delta \overline{\sigma}_\theta^2 + a_j \delta \gamma_a \overline{\sigma}_\theta^2}{1 - \delta \gamma_\theta \overline{\sigma}_\theta^2} \\
&= \underbrace{\frac{\gamma_0 \delta \overline{\sigma}_\theta^2}{1 - \delta \gamma_\theta \overline{\sigma}_\theta^2}}_{=\alpha_0} + \underbrace{\frac{\delta \gamma_a \overline{\sigma}_\theta^2}{1 - \delta \gamma_\theta \overline{\sigma}_\theta^2}}_{=\alpha_1} a_j + \underbrace{\frac{\sigma_\varepsilon^2}{(\sigma_\varepsilon^2 + \sigma_\theta^2)(1 - \delta \gamma_\theta \overline{\sigma}_\theta^2)}}_{=\beta} (s - \hat{e})
\end{aligned}$$

Since $\delta \geq 0$, $\gamma_\theta \leq 0$, and $\overline{\sigma}_\theta^2 > 0$, $1 - \delta \gamma_\theta \overline{\sigma}_\theta^2 \geq 0$, i.e., the denominators of all three fractions in this expression are positive. This implies that β is strictly positive, and α_1 is weakly positive (and strictly positive if the numerator is strictly positive).

For part (ii), given the linear specification $\frac{\partial^2 v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j} = 0$ if and only if $\gamma_a = 0$, which implies $\alpha_1 = 0$. $\frac{\partial^2 v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j} > 0$ when $\gamma_a > 0$, in which case $\frac{\partial \alpha_1}{\partial \delta} > 0$.

For part (iii), given the linear specification $\frac{\partial^2 v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I^2} = \gamma_\theta$. The weak concavity assumption in this specification is that $\gamma_\theta \leq 0$, with $\gamma_\theta < 0$ capturing strict concavity. If $\gamma_\theta = 0$, then $\beta = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\theta^2}$, which is not a function of δ. If $\gamma_\theta < 0$, then $\frac{\partial \beta}{\partial \delta} < 0$. ■

B.3 Proposition 2

Proposition 2. *Under assumption 2:*
(i) If $a_m = 0$ or the election is a dead heat ($\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m = 0$), then polarization has no impact on incumbent effort.
(ii) If $a_m \neq 0$, then increasing polarization (i.e. increases in α_1) increases effort when the incumbent is behind ($\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m < 0$) and $a_m > 0$ or the incumbent is ahead ($\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m > 0$) and $a_m < 0$, and decreases effort otherwise.

Proof of Proposition 2. By Corollary 1 the incumbent is reelected if

$$\tilde{\theta}_I^*(s, a_m, \delta; \hat{e}) + a_m + \eta_I \geq \mu_C + \eta_C.$$

Substituting the linear form of θ^* from Lemma 1 and the definition of the signal s we can express this conditions as

$$\alpha_0 + \alpha_1 a_m + \beta(\theta_I + \varepsilon + e - \hat{e}) + a_m + \eta_I \geq \mu_C + \eta_C.$$

Re-arranging to place all random variables on the same side gives:

$$\beta\theta_I + \beta\varepsilon + \eta_I - \eta_c \geq \mu_c - a_m - \alpha_0 - \alpha_1 a_m - \beta(e - \hat{e}).$$

Since θ_I, ε, and $\eta_I - \eta_C$ are all normal (and independent), the sum of the left-hand side is normal with mean μ_η and variance $\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2$. The probability of reelection given an effort level e from the Incumbent's perspective is then

$$\Pr[R = 1|e] = 1 - \Phi\left(\frac{\mu_c - a_m - \alpha_0 - \alpha_1 a_m - \beta(e - \hat{e}) - \mu_\eta}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\right).$$

If the incumbent could exert no effort and the voter knew this ($e = \hat{e} = 0$), the re-election probability is less than $1/2$ if and only if $\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m < 0$, which is why we refer to this condition as indicating when the incumbent is "behind". (This property also implies the equilibrium probability of re-election when voters correctly infer $e = \hat{e}$ will be less than $1/2$.) Conversely, if $\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m > 0$ then the re-election probability is above $1/2$, and we say the incumbent is "ahead".

The marginal effect of effort on reelection is

$$\frac{\partial \Pr[R = 1|e]}{\partial e} = \frac{\beta}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\phi\left(\frac{\mu_c - a_m - \alpha_0 - \alpha_1 a_m - \beta(e - \hat{e}) - \mu_\eta}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\right)$$

The equilibrium effort level depends on this marginal return evaluated at the point where the voter expectation is correct, i.e., $e = \hat{e}$:

$$\left.\frac{\partial \Pr[R = 1|e]}{\partial e}\right|_{e=\hat{e}} = \frac{\beta}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\phi\left(\frac{\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\right)$$

If the marginal return to effort at $e = \hat{e}$ is increasing in a_m, then Incumbent's expected utility satisfies increasing differences in (e, a_m) which implies that effort is monotone increasing in a_m. (Milgrom and Shannon 1994). Conversely, if the marginal return to effort is decreasing in a_m, then equilibrium effort must be decreasing in a_m. Thus, the marginal effect of divergence on effort has the same sign as $\frac{\partial^2 \Pr[R=1|e]}{\partial \alpha_1 \partial e}$ at $e = \hat{e}$. This derivative is:

$$\left.\frac{\partial^2 \Pr[R = 1|e]}{\partial \alpha_1 \partial e}\right|_{e=\hat{e}} = -a_m \frac{\beta}{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}\phi'\left(\frac{\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\right)$$

We now consider six cases:

1. $a_m = 0$. In this case we clearly have $\left.\frac{\partial^2 \Pr[R=1|e]}{\partial \alpha_1 \partial e}\right|_{e=\hat{e}} = 0$.

2. The incumbent is behind and $a_m > 0$. The incumbent is behind if $\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m > 0$. This implies that $\phi'\left(\frac{\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\right) < 0$ since the standard normal distribution is strictly decreasing at strictly positive values. Thus, for $a_m > 0$ we have $\left.\frac{\partial^2 \Pr[R=1|e]}{\partial \alpha_1 \partial e}\right|_{e=\hat{e}} > 0$.

3. The incumbent is behind and $a_m < 0$. For $a_m < 0$ the sign is reversed and $\left.\frac{\partial^2 \Pr[R=1|e]}{\partial \alpha_1 \partial e}\right|_{e=\hat{e}} < 0$.

4. The incumbent is ahead and $a_m > 0$. The incumbent is ahead if $\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m < 0$. This implies that $\phi'\left(\frac{\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\right) > 0$. Thus, for $a_m > 0$ we have $\left.\frac{\partial^2 \Pr[R=1|e]}{\partial \alpha_1 \partial e}\right|_{e=\hat{e}} < 0$.

5. The incumbent is ahead and $a_m < 0$. For $a_m < 0$ the sign is reversed and $\left.\frac{\partial^2 \Pr[R=1|e]}{\partial \alpha_1 \partial e}\right|_{e=\hat{e}} > 0$.

6. The remaining cases are the knife-edged case where $a_m \neq 0$ but the election is *ex ante* "tied", i.e., $\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m = 0$. The marginal effect of increasing α_1 is equal to zero since $\phi'\left(\frac{\mu_c - \mu_\eta - a_m - \alpha_0 - \alpha_1 a_m}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\right) = 0$, but once α_1 increases this will push the incumbent to be ahead if $a_m > 0$ and behind if $a_m < 0$, and so this folds into the cases 3 and 4 hold.

Putting this together, we have shown that effort is increasing in α_1 when the incumbent is behind and $a_m > 0$ or ahead and $a_m < 0$, has no effect on effort when $a_m = 0$, and decreases effort otherwise. This completes the proof. ■

B.4 Proposition 3

Proposition 3. *Under assumption 2, equilibrium incumbent effort is reduced by desensitization effects of motivated reasoning (e^* is increasing in β).*

Proof of Proposition 3. From the proof of Proposition 2 we have the following marginal effect of effort on reelection:

$$\left.\frac{\partial \Pr[R=1|e]}{\partial e}\right|_{e=\hat{e}} = \frac{\beta}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\phi\left(\frac{\mu_c - a_m - \alpha_0 - \alpha_1 a_m}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\right). \quad (4)$$

To save on notation, let

$$
\begin{aligned}
g(\beta) &:= \frac{1}{\sqrt{\beta^2\sigma_\theta^2+\beta^2\sigma_\varepsilon^2+\sigma_\eta^2}} \\
h(\beta) &:= \beta g(\beta) \text{ and} \\
\chi &:= \mu_c-\mu_\eta-a_m-\alpha_0-\alpha_1 a_m.
\end{aligned}
$$

We can now rewrite the marginal effect of effort as

$$\left.\frac{\partial \Pr[R=1|e]}{\partial e}\right|_{e=\hat{e}} = h(\beta)\phi(g(\beta)\chi).$$

The effect of increasing β has the same sign as $\left.\frac{\partial^2 \Pr[R=1|e]}{\partial\beta\partial e}\right|_{e=\hat{e}}$. Evaluating this derivative gives us:

$$\left.\frac{\partial^2 \Pr[R=1|e]}{\partial\beta\partial e}\right|_{e=\hat{e}} = h'(\beta)\phi(g(\beta)\chi)+h(\beta)\phi'(g(\beta)\chi)\chi g'(\beta)$$

We will show that this expression is always strictly positive by separately showing that (I) $h'(\beta)\phi(g(\beta)\chi)>0$ and (II) $h(\beta)\phi'(g(\beta)\chi)\chi g'(\beta)\geq 0$ which implies that the sum is positive.

For (I), dividing the numerator and denominator by β gives that:

$$h(\beta) = \frac{1}{\sqrt{\sigma_\theta^2+\sigma_\varepsilon^2+\sigma_\eta^2/\beta^2}}.$$

Since the denominator is strictly decreasing in β, it is immediate that $h'(\beta)>0$, and since $\phi(\cdot)>0$ (as this is the pdf of a standard normal random variable), we have $h'(\beta)\phi(g(\beta)\chi)>0$.

For (II), first note that $h(\beta)>0$ and $g'(\beta)<0$, so the claim is equivalent to $\phi'(g(\beta)\chi)\chi\leq 0$. Since χ can take on any real value there are three cases:

1. $\chi=0$. In this case $\phi'(g(\beta)\chi)\chi=0$.
2. $\chi>0$. Then $\phi'(g(\beta)\chi)<0$ since $g(\beta)>0$ and the normal distribution is increasing up to its mode at 0 (i.e., $\phi(x)>0$ for $x<0$). So $\phi'(g(\beta)\chi)\chi<0$.
3. $\chi<0$. Then $\phi'(g(\beta)\chi)>0$ since $\phi(x)<0$ for $x>0$. So, $\phi'(g(\beta)\chi)\chi<0$.

Thus, we have $\beta\chi g'(\beta)\phi'(g(\beta)\chi)\geq 0$ in every case. This shows that $\left.\frac{\partial^2 \Pr[R=1|e]}{\partial\beta\partial e}\right|_{e=\hat{e}}>0$, which implies that desensitization reduces effort. ■

B.5 Corollary 2

Corollary 2. *Under Assumptions 1 and 2, if there is desensitization* $\left(\frac{\partial^2 v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I^2} < 0\right)$ *or divergence affects the median voter* $\left(a_m \neq 0 \text{ and } \frac{\partial^2 v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j} > 0\right)$*, then as* $\delta \to \infty$*,* $e^* \to 0$*.*

Proof of Corollary 2. Recall the equilibrium marginal return to effort is:

$$\left.\frac{\partial \Pr[R=1|e]}{\partial e}\right|_{e=\hat{e}} = \frac{\beta}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\phi\left(\frac{\mu_c - a_m - \alpha_0 - \alpha_1 a_m}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\right). \tag{5}$$

To prove the result we need to show that either of the two stated conditions implies that this term approaches 0 as $\delta \to \infty$.

As shown in the proof of Lemma 1, in the linear case the β term is given by:

$$\beta = \frac{\sigma_\varepsilon^2}{(\sigma_\varepsilon^2 + \sigma_\theta^2)(1 - \delta\gamma_\theta\overline{\sigma}_\theta^2)}$$

The proof of Lemma 1 also shows that $\frac{\partial^2 v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I^2} < 0$ if and only if $\gamma_\theta < 0$, and if this holds then as $\delta \to \infty$, $\beta \to 0$. Equation 5 is bounded above by:

$$\left.\frac{\partial \Pr[R=1|e]}{\partial e}\right|_{e=\hat{e}} \leq \frac{\beta}{\sqrt{\beta^2\sigma_\theta^2 + \beta^2\sigma_\varepsilon^2 + \sigma_\eta^2}}\phi(0).$$

(This follows from the fact that ϕ is maximized at 0.). From this it follows that as $\beta \to 0$, $\left.\frac{\partial \Pr[R=1|e]}{\partial e}\right|_{e=\hat{e}} \to 0$ (regardless of how δ affects α_1), and hence $e^* \to 0$. This completes the proof for any case with desensitization.

For the remaining case, recall that:

$$\alpha_1 = \frac{\delta\gamma_a\overline{\sigma}_\theta^2}{1 - \delta\gamma_\theta\overline{\sigma}_\theta^2}$$

Since we have already proven the result when there is desensitization, it is sufficient to show the result for the case where divergence affects the median voter ($a_m \neq 0$ and $\frac{\partial^2 v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j} > 0$), under the assumption of no desensitization, or $\gamma_\theta = 0$. Plugging in $\gamma_\theta = 0$ we have $\alpha_1 = \delta\gamma_a\overline{\sigma}_\theta^2$. $\frac{\partial^2 v(a_j,\tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j} > 0$ implies $\gamma_a > 0$, so as $\delta \to \infty$, $\alpha_1 \to \infty$.

If $a_m \neq 0$, then as $\delta \to \infty$ the right-hand side of equation (5) approaches zero, since $\alpha_1 \to \infty$ as $\delta \to \infty$ and $\lim_{x\to\pm\infty}\phi(x) = 0$. (Since there is no desensitization, δ does not affect the outer term.) Hence $\left.\frac{\partial \Pr[R=1|e]}{\partial e}\right|_{e=\hat{e}} \to 0$, and $e^* \to 0$. ■

B.6 Remark 1

Remark 1. *Suppose Assumption 2 holds and voter affinities are normally distributed with mean* μ_a *and variance* σ_a^2*. Further, let* $a_m = 0$*,* $\mu_a = 0$*, and* $\frac{\partial \beta}{\partial \delta} = 0$ *so that there is divergence but no desensitization. Then motivated reasoning can affect incumbent vote share even when it does not affect equilibrium effort.*

Proof of Remark 1. Follows from argument/derivations in text given Proposition 2 showing that when $a_m = 0$ belief divergence does not affect effort and the fact that there is no belief desensitization effects when $\frac{\partial \beta}{\partial \delta} = 0$. ■

References

Milgrom, Paul and Chris Shannon. 1994. "Monotone Comparative Statics." *Econometrica* 62(1):157–180.