Online Appendix for Overconfidence in Political Behavior

By PIETRO ORTOLEVA AND ERIK SNOWBERG

A. Proofs

LEMMA 2: *If $f(y)$ and $g(y)$ are monotone functions, and $f'(y) * g'(y) > 0$, then* $\mathrm{cov}[f(y), g(y)] > 0$.

PROOF:
See Schmidt (2003).

LEMMA 1: *Define* $e_i \equiv \frac{1}{n_i}\sum_{t=1}^{n_i} e_{it}$. *Then* $\kappa = \frac{n_i}{1+(n_i-1)\rho}$, *and* $\kappa_i = \frac{n_i}{1+(n_i-1)\rho_i}$.

PROOF:
The posterior likelihood in the model is proportional to

$$\begin{aligned}
\mathcal{L}(x|\mathbf{e}_i) &\propto \mathcal{L}(\mathbf{e}_i|x)\mathcal{L}_0(x) \\
&\propto \exp\left\{-\frac{1}{2}\begin{pmatrix} x-e_{i1} \\ x-e_{i2} \\ \vdots \\ x-e_{in_i}\end{pmatrix}^T \begin{pmatrix} 1 & \rho_i & \cdots & \rho_i \\ \rho_i & 1 & \cdots & \rho_i \\ \vdots & \vdots & \ddots & \vdots \\ \rho_i & \rho_i & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} x-e_{i1} \\ x-e_{i2} \\ \vdots \\ x-e_{in_i}\end{pmatrix}\right\} \\
&\qquad * \exp\left\{-\frac{1}{2}x^2\tau\right\} \\
&= \exp\left\{-\frac{1}{2}\left(\frac{n_i x^2 - 2x\sum_{t=1}^{n_i} e_{it}}{1+(n_i-1)\rho_i} + C\right)\right\} * \exp\left\{-\frac{1}{2}x^2\tau\right\} \\
&\propto \exp\left\{-\frac{1}{2}\frac{n_i+\tau(1+(n_i-1)\rho_i)}{1+(n_i-1)\rho_i}\left(x - \frac{\sum_{t=1}^{n_i} e_{it}}{n_i+\tau(1+(n_i-1)\rho_i)}\right)^2\right\}
\end{aligned}$$

where C is constant with respect to x. Thus, defining $e_i = \frac{1}{n_i}\sum_{t=1}^{n_i} e_{it}$, the posterior belief of a citizen is distributed according to

$$\mathcal{N}\left[\frac{n_i e_i}{n_i+\tau(1+(n_i-1)\rho_i)}, \frac{n_i+\tau(1+(n_i-1)\rho_i)}{1+(n_i-1)\rho_i}\right]. \tag{1}$$

Substituting $\rho_i = \dfrac{n_i - \kappa_i}{(n_i - 1)\kappa_i}$ the posterior is given by $\mathcal{N}\left[\frac{\kappa_i e_i}{\kappa_i + \tau}, \kappa_i + \tau\right]$, which is the same as the posterior that a citizen would have if they received a single signal $e_i = x + \varepsilon_i$, where the citizen believes $\varepsilon_i \sim \mathcal{N}[0, \kappa_i]$. Finally, note that $E[e_i] = x$, and

$$\mathrm{var}[e_i] = \left(\frac{1}{n_i}\right)^2 \sum_{t=1}^{n_i} \mathrm{var}[\varepsilon_{it}] + 2\left(\frac{1}{n_i}\right)^2 \frac{n_i(n_i - 1)}{2} \mathrm{cov}[\varepsilon_{it}, \varepsilon_{it'}] = \frac{1}{n_i} + \frac{n_i - 1}{n_i}\rho.$$

Thus, $e_i \sim \mathcal{N}\left[x, \frac{n_i}{1 + (n_i - 1)\rho}\right] \equiv \mathcal{N}[x, \kappa]$.

PROPOSITION 1:

(i) Overconfidence is increasing with the number of experiences (signals) n.

(ii) The mean ideology in the population, conditional on n, $E[\mathcal{I}_i|n]$, is increasing in n if and only if $x > 0$, and decreasing in n iff $x < 0$.

(iii) If ρ is large enough, $\mathrm{var}[\mathcal{I}_i|n]$ *is increasing in n.*

PROOF:

A citizen's overconfidence after n_i signals is given by:

$$\frac{n_i + \tau(1 + (n_i - 1)\rho_i)}{1 + (n_i - 1)\rho_i} - \frac{n_i + \tau(1 + (n_i - 1)\rho)}{1 + (n_i - 1)\rho} > 0 \iff \rho_i < \rho$$

The difference in overconfidence between the citizen with $n_i + 1$ and n_i signals is given by

$$\frac{n_i(\rho - \rho_i)(2 + (n_i - 1)(\rho + \rho_i - \rho\rho_i))}{(1 + (n_i - 1)\rho_i)(1 + n_i\rho_i)(1 + (n_i - 1)\rho)(1 + n\rho)} > 0$$

because $0 < \rho_i < \rho < 1$ and $n_i \geq 2$.

The second and third part follow from Lemma 1: using (1) a citizen's mean belief is distributed according to

$$\mathcal{N}\left[\frac{n_i x}{n_i + \tau(1 + (n_i - 1)\rho_i)}, \frac{(n_i + \tau(1 + (n_i - 1)\rho_i))^2}{n_i(1 + (n_i - 1)\rho)}\right]$$

and the difference between the mean of that distribution at $n_i + 1$ and n_i is

$$\frac{x\tau(1 - \rho_i)}{(1 + n_i + \tau(1 + n_i\rho_i))(n_i + \tau(1 + (n_i - 1)\rho_i))}$$

which is positive iff x is positive, and negative iff x is negative.

For the third part, we need to show $\text{var}[\mathcal{I}_i|n+1] - \text{var}[\mathcal{I}_i|n] > 0$. That is:

$$\frac{(n_i+1)(1+n_i\rho)}{(n_i+1+\tau(1+n_i\rho_i))^2} - \frac{n_i(1+(n_i-1)\rho)}{(n_i+\tau(1+(n_i-1)\rho_i))^2} > 0 \tag{2}$$

The LHS of (2) is 0 when:

$$\rho = \frac{n_i(n_i+1)(1+\rho_i\tau)^2 - (1-\rho_i)^2\tau^2}{n_i((n_i+1)(1+2\tau)+\tau^2(2+\rho_i(n_i-1)(2-\rho_i)))}$$

The derivative of the LHS of (2) with respect to ρ is

$$\frac{n_i((n_i+1)(1+2\tau)+\tau^2(2+\rho_i(n_i-1)(2-\rho_i)))}{(n_i+1+\tau(1+n_i\rho_i))^2(n_i+\tau(1+(n_i-1)\rho_i))^2} \geq 0$$

Therefore, as long as $\rho \geq \frac{n_i(n_i+1)(1+\rho_i\tau)^2-(1-\rho_i)^2\tau^2}{n_i((n_i+1)(1+2\tau)+\tau^2(2+\rho_i(n_i-1)(2-\rho_i)))}$, (2) is satisfied. This will hold for all n if $\rho \geq \frac{1+\rho_i\tau}{1+\tau(2-\rho_i)}$, and, thus, for all ρ_i if ρ is large.

PROPOSITION 2: *Overconfidence and ideological extremeness are positively correlated. This is true conditional on n, and independent of n if ρ is large enough.*[1]

PROOF:

Fix $n_i = n$. Using (1) and the distribution of e_i we have that the posterior distribution of the mean of beliefs of citizens with a given κ_i is given by

$$E_i[x]|\kappa_i \sim \mathcal{N}\left[\frac{\kappa_i x}{\tau+\kappa_i}, \kappa\left(\frac{\tau+\kappa_i}{\kappa_i}\right)^2\right],$$

where $\Phi[\cdot]$ denotes the c.d.f. of the standard normal distribution. Using the fact that residuals from an OLS regression are orthogonal to regressors we have:

$$\text{cov}[\mathcal{E} - E[\mathcal{E}|\kappa_i], \kappa_i] = 0 \quad \Rightarrow \quad \text{cov}[\mathcal{E}, \kappa_i] = \text{cov}[E[\mathcal{E}|\kappa_i], \kappa_i] \tag{3}$$

Note that $\mathcal{I}_i|\kappa_i = b_i + E_i[x|\kappa_i]$ is a sum of two independent random normal variables. Thus,

$$\mathcal{I}|\kappa_i \sim \mathcal{N}\left[\mu, \frac{1}{\sigma^2}\right].$$

$$\text{where} \quad \mu = \frac{\kappa_i x}{\tau+\kappa_i} \quad \text{and} \quad \sigma^2 = \frac{1}{\kappa}\left(\frac{\kappa_i}{\tau+\kappa_i}\right)^2 + \frac{1}{\tau_b} \tag{4}$$

[1] Specifically, this holds as long as ρ is large enough that population variance of ideology, conditional on n, $\text{var}[\mathcal{I}_i|n]$, is increasing in n—see Proposition 1.

as $\mathcal{E} = |\mathcal{I}|$, $\mathcal{E}|\kappa_i$ is distributed according to a folded normal distribution, so

$$E[\mathcal{E}|\kappa_i] = 2\sigma\,\phi\left[\frac{\mu}{\sigma}\right] + \mu\left(1 - 2\Phi\left[-\frac{\mu}{\sigma}\right]\right) \tag{5}$$

where $\phi[\cdot]$ is the standard normal p.d.f. Taking the derivative of (5) yields:

$$\frac{dE[\mathcal{E}|\kappa_i]}{d\kappa_i} = \frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\sigma}\cdot\frac{d\sigma}{d\kappa_i} + \frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\sigma}\cdot\frac{d\sigma}{d\kappa_i} \tag{6}$$

$$\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\sigma} = 2\phi\left[\frac{\mu}{\sigma}\right] > 0 \quad \text{and} \quad \frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\mu} = 1 - 2\Phi\left[-\frac{\mu}{\sigma}\right] \tag{7}$$

$$\frac{d\sigma}{d\kappa_i} = \frac{1}{2\sigma\sqrt{\kappa}}\frac{d\sigma^2}{d\kappa_i} = \frac{1}{2\sigma\sqrt{\kappa}}\frac{\tau}{(\tau+\kappa_i)^2} > 0 \quad \text{and} \quad \frac{d\mu}{d\kappa_i} = x\frac{\tau}{(\tau+\kappa_i)^2}$$

so clearly the first term in (6) is positive. Note that $\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\mu}$ is positive iff $x > 0$, so the second term in (6) is ≥ 0. This implies that $E[\mathcal{E}|\kappa_i]$ is increasing in κ_i. Lemma 2 then implies $\text{cov}[E[\mathcal{E}|\kappa_i], \kappa_i] > 0$, and so $\text{corr}[\mathcal{E}, \kappa_i] > 0$.

For the logic in the previous paragraph to hold in the case when $n_i \sim F_n$, $\text{var}[\mathcal{I}|n]$ must be increasing in n, which is exactly the condition in Implication 2.

PROPOSITION 3: *If $x > 0$ overconfidence and ideology are positively correlated, both independent of, and conditional on, n.*

PROOF:

Following the logic of (3) we have $\text{cov}[\mathcal{I}, \kappa_i] = \text{cov}[E[\mathcal{I}|\kappa_i], \kappa_i]$. $E[\mathcal{I}|\kappa_i] = \frac{\kappa_i x}{\tau+\kappa_i}$. As this is increasing in κ_i when $x > 0$, Lemma 2 implies that $\text{corr}[\mathcal{I}, \kappa_i] > 0$.

PROPOSITION 4: *If $x > 0$ then $\text{cov}[\mathcal{E}, \kappa_i|\mathcal{I}_i \geq 0] > \text{cov}[\mathcal{E}, \kappa_i|\mathcal{I}_i \leq 0]$ both independent of, and conditional on, n.*

PROOF:

Fix $n_i = n$, and assume that $b_i = 0$ for all i. (We shall relax this assumption later.) By the definition of covariance, and as $\mathcal{E} = |\mathcal{I}_i|$

$$\begin{aligned}
\text{cov}[\mathcal{E}_i, \kappa_i|\mathcal{I}_i \geq 0] &= E[\mathcal{I}_i\kappa_i|\mathcal{I}_i \geq 0] - E[\mathcal{I}_i|\mathcal{I}_i \geq 0]E[\kappa_i|\mathcal{I}_i \geq 0] \\
\text{cov}[\mathcal{E}_i, \kappa_i|\mathcal{I}_i \leq 0] &= E[-\mathcal{I}_i\kappa_i|\mathcal{I}_i \leq 0] - E[-\mathcal{I}_i|\mathcal{I}_i \leq 0]E[\kappa_i|\mathcal{I}_i \leq 0] \\
&= -E[\mathcal{I}_i\kappa_i|\mathcal{I}_i \leq 0] + E[\mathcal{I}_i|\mathcal{I}_i \leq 0]E[\kappa_i|\mathcal{I}_i \leq 0].
\end{aligned}$$

So

$$\text{cov}[\mathcal{E}, \kappa_i|\mathcal{I}_i \geq 0] > \text{cov}[\mathcal{E}, \kappa_i|\mathcal{I}_i \leq 0] \tag{8}$$

holds if and only if

$$E[\mathcal{I}_i\kappa_i|\mathcal{I}_i \geq 0] - E[\mathcal{I}_i|\mathcal{I}_i \geq 0]E[\kappa_i] \; + \; E[\mathcal{I}_i\kappa_i|\mathcal{I}_i \leq 0] - E[\mathcal{I}_i|\mathcal{I}_i \leq 0]E[\kappa_i] > 0$$
$$E[\mathcal{I}_i(\kappa_i - E[\kappa_i])|\mathcal{I}_i \geq 0] \; + \; E[\mathcal{I}_i(\kappa_i - E[\kappa_i])|\mathcal{I}_i \leq 0] > 0, \tag{9}$$

where the first line follows from the definition of covariance, and the fact that $\text{sign}(\mathcal{I}_i) = \text{sign}(e_i)$ and $e_i \perp \kappa_i$, which implies that $E[\kappa_i|\mathcal{I}_i \geq 0] = E[\kappa_i|\mathcal{I}_i \leq 0] = E[\kappa_i]$. It then follows that (8) holds iff

$$\int_{\underline{\kappa}}^{\infty}\int_0^{\infty} \frac{e_i}{\Pr[\mathcal{I}_i \geq 0]} \frac{\kappa_i}{\kappa_i + \tau}(\kappa_i - E[\kappa_i]) dF_{e_i} dF_{\kappa}$$
$$+ \int_{\underline{\kappa}}^{\infty}\int_{-\infty}^{0} \frac{e_i}{\Pr[\mathcal{I}_i \leq 0]} \frac{\kappa_i}{\kappa_i + \tau}(\kappa_i - E[\kappa_i]) dF_{e_i} dF_{\kappa} > 0$$

$$\iff \int_{\underline{\kappa}}^{\infty} \frac{\kappa_i}{\kappa_i + \tau}(\kappa_i - E[\kappa_i]) \left[\int_0^{\infty} \frac{e_i}{\Pr[\mathcal{I}_i \geq 0]} dF_{e_i} + \int_{-\infty}^{0} \frac{e_i}{\Pr[\mathcal{I}_i \leq 0]} dF_{e_i}\right] dF_{\kappa} > 0$$
$$\iff \int_{\underline{\kappa}}^{\infty} \frac{\kappa_i}{\kappa_i + \tau}(\kappa_i - E[\kappa_i]) \Big[E[e_i|e_i \geq 0] + E[e_i|e_i \leq 0]\Big] dF_{\kappa} > 0$$
$$\iff \Big[E[e_i|e_i \geq 0] + E[e_i|e_i \leq 0]\Big] \int_{\underline{\kappa}}^{\infty} \frac{\kappa_i}{\kappa_i + \tau}(\kappa_i - E[\kappa_i]) dF_{\kappa} > 0$$
$$\iff \Big[E[e_i|e_i \geq 0] + E[e_i|e_i \leq 0]\Big] \text{cov}\left[\frac{\kappa_i}{\kappa_i + \tau}, \kappa_i\right] > 0$$

where the last line follows by inverting the steps used to get to (9). As $\frac{\kappa_i}{\kappa_i+\tau}$ is increasing in κ_i, Lemma 2 gives that $\text{cov}\left[\frac{\kappa_i}{\kappa_i+\tau}, \kappa_i\right] > 0$. Thus, (8) holds iff

$$E[e_i|e_i \geq 0] + E[e_i|e_i \leq 0] > 0. \tag{10}$$

$e_i \sim \mathcal{N}[x, \kappa]$ implies

$$x \;=\; \Phi[-x\sqrt{\kappa}]E[e_i|e_i \leq 0] + \big(1 - \Phi\big[-x\sqrt{\kappa}\big]\big)E[e_i|e_i \geq 0]$$
$$\Rightarrow \quad E[e_i|e_i \geq 0] \;=\; \frac{x - \Phi\big[-x\sqrt{\kappa}\big]E[e_i|e_i \leq 0]}{1 - \Phi\big[-x\sqrt{\kappa}\big]},$$

where $\Phi[\cdot]$ is the standard normal c.d.f. Thus, (10) can be re-written as

$$\frac{\big(1 - 2\Phi\big[-x\sqrt{\kappa}\big]\big)E[e_i|e_i \leq 0] + x}{1 - \Phi\big[-x\sqrt{\kappa}\big]} > 0, \tag{11}$$

which holds as long as the numerator is positive. Note that

$$E[|e_i|] = x - 2\Phi\big[-x\sqrt{\kappa}\big]E[e_i|e_i \leq 0],$$

and, from using the expectation of the folded normal we have that

$$E[e_i|e_i \leq 0] = \frac{x\Phi[-x\sqrt{\kappa}] - \frac{1}{\sqrt{\kappa}}\phi\big[x\sqrt{\kappa}\big]}{\Phi\big[-x\sqrt{\kappa}\big]},$$

where $\phi[\cdot]$ is the standard normal p.d.f. Thus, the numerator of (11) can be re-written as

$$\frac{2x\Phi[-x\sqrt{\kappa}] - \frac{\phi\left[x\sqrt{\kappa}\right]}{\sqrt{\kappa}} - 2x\Phi[-x\sqrt{\kappa}]^2 + \frac{2\phi\left[x\sqrt{\kappa}\right]\Phi[-x\sqrt{\kappa}]}{\sqrt{\kappa}}}{\Phi[-x\sqrt{\kappa}]} \equiv \frac{Z(x\sqrt{\kappa})}{\Phi[-x\sqrt{\kappa}]} > 0,$$

which holds iff $Z(x\sqrt{\kappa}) > 0$ for $x \in (0, \infty)$ and $\kappa \in [1, \infty)$. Integration by parts gives:

$$\Phi[-x\sqrt{\kappa}] = \frac{1}{2} - \phi\big[x\sqrt{\kappa}\big]\left(x + \frac{x^3}{3} + \frac{x^5}{5\cdot 3} + \frac{x^7}{7\cdot 5\cdot 3} + \ldots\right) = \frac{1}{2} - \phi\big[x\sqrt{\kappa}\big]q(x)$$

which can be applied to $Z(x\sqrt{\kappa})$ to yield

$$\begin{aligned} Z(x\sqrt{\kappa}) &= \frac{x}{2} - \frac{2}{\sqrt{\kappa}}\phi\big[x\sqrt{\kappa}\big]^2 q(x) - 2x\phi\big[x\sqrt{\kappa}\big]^2 q(x)^2 \\ &= \frac{x}{2} - \frac{q(x)e^{-\kappa x^2}}{\pi}\left(\frac{1}{\sqrt{\kappa}} - xq(x)\right). \end{aligned}$$

Note that as $\kappa \in [1, \infty)$, then for a fixed x, $Z(x\sqrt{\kappa})$ is minimized at $\kappa = 1$. Thus, it is sufficient to show that $Z(x\sqrt{1}) = Z(x) > 0, \forall x \in (0, \infty)$. We will now show there is a single inflection point of $Z(x)$, that is $Z''(x^*) = 0$ for a unique value of $x^* \in [0, \infty)$, and then use this to show that $Z(x) > 0, \forall x \in (0, \infty)$.

Using the fact that $\phi\big[x\big]q(x) = \Phi\big[x\big] - \frac{1}{2}$, we can re-write

$$\begin{aligned} Z(x) &= \frac{x}{2} - 2\phi\big[x\big]\left(\Phi\big[x\big] - \frac{1}{2}\right) - 2x\left(\Phi\big[x\big] - \frac{1}{2}\right)^2, \\ Z'(x) &= \frac{1}{2} - 2\phi\big[x\big]^2 - 2x\phi\big[x\big]\left(\Phi\big[x\big] - \frac{1}{2}\right) - 2\left(\Phi\big[x\big] - \frac{1}{2}\right)^2, \\ Z''(x) &= 2(x^2 - 3)\phi\big[x\big]\left(\Phi\big[x\big] - \frac{1}{2}\right) + 2x\phi\big[x\big]^2, \text{ and} \\ Z'''(x) &= 2\phi\big[x\big]\left(x(5 - x^2)\left(\Phi\big[x\big] - \frac{1}{2}\right) - (2 + x^2)\phi\big[x\big]\right) = 2\phi\big[x\big]g(x). \end{aligned}$$

We need bounds on $\Phi\big[x\big] - \frac{1}{2}$. As $x > 0$ we have

$$\Phi[x] - \frac{1}{2} = \int_0^x \phi\big[y\big]dy > \int_0^x \phi\big[x\big]dy = x\phi\big[x\big] \tag{12}$$

which implies

$$Z''(x) < 2x(x^2-2)\phi\big[x\big]^2 < 0 \text{ if } x \in (0, \sqrt{2}]$$

and it is clear that $Z''(x) > 0, \forall x \in [\sqrt{3}, \infty)$. Together this implies that $Z''(x) = 0$ somewhere in $[\sqrt{2}, \sqrt{3}]$. Note from the statement of $Z'''(x)$ above that $\text{sign}(Z'''(x)) = \text{sign}(g(x))$. Further, applying (12) we have that

$$g''(x) = 10\phi\big[x\big] - 8x^2\phi\big[x\big] - 6x\left(\Phi\big[x\big] - \frac{1}{2}\right) < -14\phi\big[x\big]\left(x^2 - \frac{5}{7}\right) < 0 \text{ on } [\sqrt{2}, \sqrt{3}].$$

As $g(x)$ is concave on $[\sqrt{2}, \sqrt{3}]$, this implies that

$$\inf_{x \in [\sqrt{2}, \sqrt{3}]} g(x) = \min\left\{g(\sqrt{2}), g(\sqrt{3})\right\}.$$

Using (12) we have

$$\begin{aligned} g(\sqrt{2}) &= 3\sqrt{2}\left(\Phi\big[\sqrt{2}\big] - \frac{1}{2}\right) - 4\phi\big[\sqrt{2}\big] > 2\phi\big[\sqrt{2}\big] > 0, \text{ and} \\ g(\sqrt{3}) &= 2\sqrt{3}\left(\Phi\big[\sqrt{3}\big] - \frac{1}{2}\right) - 5\phi\big[\sqrt{3}\big] > \phi\big[\sqrt{3}\big] > 0. \end{aligned}$$

As both are positive, this implies that $Z'''(x) > 0, \forall x \in [\sqrt{2}, \sqrt{3}]$, and this combined with $Z''(x) < 0, \forall x \in [0, \sqrt{2}]$ and $Z''(x) > 0, \forall x \in [\sqrt{3}, \infty)$ implies there is a unquie $x^* \in (0, \infty)$ such that $Z''(x^*) = 0$.

There are now two cases: $Z'(x^*) \geq 0$ or $Z'(x^*) < 0$. First, consider $Z'(x^*) \geq 0$. Note that $Z(0) = 0$, $Z'(0) = \frac{\pi-2}{2\pi} > 0$, so if $Z'(x^*) \geq 0$, this implies that $Z'(x) > 0, \forall x \in [0, \infty)$, and thus $Z(x) > 0, \forall x \in (0, \infty)$, as desired.

If, on the other hand, $Z'(x^*) < 0$, then the fact that $\lim_{x\to\infty} Z'(x) \to 0$ implies $Z'(x) < 0, \forall x \in [x^*, \infty)$ as $Z''(x) > 0, \forall x \in (x^*, \infty)$. This, coupled with the fact that $\lim_{x\to\infty} Z(x) \to 0$, implies that $Z(x^*) > 0$ and $Z(x) > 0, \forall x \in [x^*, \infty)$. In addition, $Z''(x) < 0, \forall x \in [0, x^*)$ implies that $\inf_{x \in [0, x^*]} Z(x) = \min\{Z(0), Z(x^*)\} = Z(0) = 0$. Thus, $Z(x) > 0, \forall x \in (0, \infty)$, as desired.

Thus (10) holds, and so too does (8), when $b_i = 0$ for any n. The fact that the covariance of a random variable that is a sum of random variables is linear in the covariances of the individual random variables means that this will also hold when $b_i \neq 0$ as b_i is independent of κ_i, and independently of n.

PROPOSITION 5: *If $x > 0$ and ρ is large enough then* $\text{cov}[\mathcal{E}, n_i | \mathcal{I}_i \geq 0] > \text{cov}[\mathcal{E}, n_i | \mathcal{I}_i \leq 0]$.[2]

[2]Simulations indicate this holds for all $\rho > \rho_i$. ρ close to one is needed for tractability.

PROOF:

Set $\rho = 1$, which implies that $e_i \perp n_i$. We will use this fact to allow us to map this proposition onto the proof of Proposition 4.

Fix n in that proof to $n' > \max n$. Then each n here maps to a unique ρ_i in the proof of Proposition 4. As nothing in that proof depends on the distribution of ρ_i, the result follows immediately. As this proposition holds for $\rho = 1$, it will continue to hold as long as ρ is close to one.

PROPOSITION 6: *If ρ is large enough, then $\operatorname{cov}[\mathcal{E}, \kappa_i - \kappa | n_i]$ is increasing in n_i.*

PROOF:

Using (3), we have $\operatorname{cov}[\mathcal{E}, \kappa_i - \kappa | n_i] = \operatorname{cov}\big[\kappa_i - \kappa, E[\mathcal{E}|\kappa_i] | n_i\big]$. Note that $\kappa_i - \kappa$ and $E[\mathcal{E}|\kappa_i]$ are both functions of n_i and ρ_i. Then, as ρ_i is distributed according to F_{ρ_i} on $[\underline{\rho}, \overline{\rho}]$ (with $\underline{\rho} \geq 0$, and $\overline{\rho} < \rho$), we have:

$$
\begin{aligned}
&\frac{d\operatorname{cov}\big[\kappa_i - \kappa, E[\mathcal{E}|\kappa_i] | n_i\big]}{dn_i} \\
&= \frac{d}{dn_i}\int_{\underline{\rho}}^{\overline{\rho}} \big((\kappa_i - \kappa) E[\mathcal{E}|\kappa_i] | n_i\big) dF_{\rho_i} - E\big[E[\mathcal{E}|\kappa_i] | n_i\big] \frac{d}{dn_i}\int_{\underline{\rho}}^{\overline{\rho}} \big(\kappa_i - \kappa | n_i\big) dF_{\rho_i} \\
&\qquad - E[\kappa_i - \kappa | n_i] \frac{d}{dn_i}\int_{\underline{\rho}}^{\overline{\rho}} \big(E[\mathcal{E}|\kappa_i] | n_i\big) dF_{\rho_i} \\
&= \int_{\underline{\rho}}^{\overline{\rho}} \left(\frac{d(\kappa_i - \kappa)}{dn_i} E[\mathcal{E}|\kappa_i] + (\kappa_i - \kappa) \frac{dE[\mathcal{E}|\kappa_i]}{dn_i} \bigg| n_i \right) dF_{\rho_i} \\
&\qquad - E\big[E[\mathcal{E}|\kappa_i] | n_i\big] \int_{\underline{\rho}}^{\overline{\rho}} \left(\frac{d(\kappa_i - \kappa)}{dn_i} \bigg| n_i \right) dF_{\rho_i} \\
&\qquad - E[\kappa_i - \kappa | n_i] \int_{\underline{\rho}}^{\overline{\rho}} \left(\frac{dE[\mathcal{E}|\kappa_i]}{dn_i} \bigg| n_i \right) dF_{\rho_i} \\
&= \int_{\underline{\rho}}^{\overline{\rho}} \left(\frac{d(\kappa_i - \kappa)}{dn_i} E[\mathcal{E}|\kappa_i] \bigg| n_i \right) dF_{\rho_i} - E\big[E[\mathcal{E}|\kappa_i] | n_i\big] \int_{\underline{\rho}}^{\overline{\rho}} \left(\frac{d(\kappa_i - \kappa)}{dn_i} \bigg| n_i \right) dF_{\rho_i} \\
&\qquad + \int_{\underline{\rho}}^{\overline{\rho}} \left((\kappa_i - \kappa) \frac{dE[\mathcal{E}|\kappa_i]}{dn_i} \bigg| n_i \right) dF_{\rho_i} - E[\kappa_i - \kappa | n_i] \int_{\underline{\rho}}^{\overline{\rho}} \left(\frac{dE[\mathcal{E}|\kappa_i]}{dn_i} \bigg| n_i \right) dF_{\rho_i} \\
&= \operatorname{cov}\left[\frac{d(\kappa_i - \kappa)}{dn_i}, E[\mathcal{E}|\kappa_i] \bigg| n_i \right] + \operatorname{cov}\left[\kappa_i - \kappa, \frac{dE[\mathcal{E}|\kappa_i]}{dn_i} \bigg| n_i \right].
\end{aligned}
$$

Note that the proof of Proposition 2 gives that $E[\mathcal{E}|\kappa_i]|n_i$ is increasing in κ_i, which is decreasing in ρ_i, so $E[\mathcal{E}|\kappa_i]|n_i$ is decreasing in ρ_i. Using the definition of $\kappa_i - \kappa$ from

Lemma 1 we also have that

$$\frac{d^2(\kappa_i - \kappa)}{dn_i d\rho_i} = \frac{1 - \rho_i + n_i(\rho_i - 2)}{(1 + (n_i - 1)\rho_i)^3} < 0.$$

So $E[\mathcal{E}|\kappa_i]$ and $\frac{d(\kappa_i - \kappa)}{dn_i}$ are decreasing in ρ_i. By Lemma 2,

$$\operatorname{cov}\left[\frac{d(\kappa_i - \kappa)}{dn_i}, E[\mathcal{E}|\kappa_i]\middle| n_i\right] > 0.$$

For the second covariance above: from the definition of κ_i and κ in Lemma 1, we have that $\kappa_i - \kappa$ is decreasing in ρ_i. Substituting these definitions into (4), we have

$$\mu = \frac{nx}{n(1 + \rho_i\tau) + \tau(1 - \rho_i)} \quad \text{and} \quad \sigma^2 = \frac{\mu^2}{\kappa x^2} + \frac{1}{\tau_b}$$

we have that $\mathcal{I}|\kappa_i, n_i \sim \mathcal{N}\left[\mu, \frac{1}{\sigma^2}\right]$, and $\mathcal{E}|\kappa_i, n_i$ is a folded normal with mean given by (5), which is a function of μ and σ. We thus can write:

$$\begin{aligned}
\frac{dE[\mathcal{E}|\kappa_i]}{dn_i} &= \frac{\partial E[\mathcal{E}|\kappa_i]}{\partial \sigma} \cdot \frac{d\sigma}{dn_i} + \frac{\partial E[\mathcal{E}|\kappa_i]}{\partial \mu} \cdot \frac{d\mu}{dn_i} \\
\frac{d^2E[\mathcal{E}|\kappa_i]}{dn_i d\rho_i} &= \frac{d}{d\rho_i}\left(\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial \sigma}\right) \cdot \frac{d\sigma}{dn_i} + \frac{\partial E[\mathcal{E}|\kappa_i]}{\partial \sigma} \cdot \frac{d^2\sigma}{dn_i d\rho_i} \qquad (13) \\
&\quad + \frac{d}{d\rho_i}\left(\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial \mu}\right) \cdot \frac{d\mu}{dn_i} + \frac{\partial E[\mathcal{E}|\kappa_i]}{\partial \mu} \cdot \frac{d^2\mu}{dn_i d\rho_i}.
\end{aligned}$$

We will now show that (13) is negative. Starting with the second term note that when $\rho = 1$:

$$\begin{aligned}
\frac{d^2\sigma^2}{dn_i d\rho_i} &= 2\sigma\frac{d^2\sigma}{dn_i d\rho_i} + 2\frac{d\sigma}{d\rho_i} \cdot \frac{d\sigma}{dn_i}, \text{ thus,} \\
\frac{d^2\sigma}{dn_i d\rho_i} &= \frac{1}{2\sigma}\left(\frac{d^2\sigma^2}{dn_i d\rho_i} - 2\frac{d\sigma}{d\rho_i} \cdot \frac{d\sigma}{dn_i}\right) \\
&= \frac{1}{2\sigma}\left(-\frac{2n\tau(2\tau(n-1)(1-\rho_i) + n(1+\tau))}{(n(1+\rho_i\tau) + \tau(1-\rho_i))^4}\right. \\
&\qquad \left. + 2\frac{n^2\tau(n-1) \cdot n\tau(1-\rho_i)}{\sigma^2(n(1+\rho_i\tau) + \tau(1-\rho_i))^6}\right) \\
&< \frac{n\tau}{\sigma}\left(\frac{-n^2(2\tau(n-1)(1-\rho_i) + n(1+\tau)) + n^2\tau(n-1)(1-\rho_i)}{\sigma^2(n(1+\rho_i\tau) + \tau(1-\rho_i))^6}\right) \\
&= -\frac{n\tau}{\sigma}\left(\frac{n^2(\tau(n-1)(1-\rho_i) + n(1+\tau))}{\sigma^2(n(1+\rho_i\tau) + \tau(1-\rho_i))^6}\right) < 0,
\end{aligned}$$

where the inequality comes from setting $1/\tau_b = 0$ in the definition of σ^2. Further, from (7) we have $\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\sigma} > 0$, so the second term of (13) is negative. Next for the fourth term:

$$\frac{d^2\mu}{dn_i d\rho_i} = -\frac{\tau(n-1)(1-\rho_i) + n(1+\tau)}{(n(1+\rho_i\tau) + \tau(1-\rho_i))^3}x$$

and from (7) we have that $\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\mu} > 0$ if $x > 0$, and negative if $x < 0$. Therefore the fourth term of (13) is negative (or zero when $x = 0$).

We now show that the first and third terms of (13) are, together, negative. First, we examine the first term, defining $c = \mu/x$, and using the expression for $\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\sigma}$ from (7):

$$\begin{aligned}
\frac{d\sigma}{d\rho_i} &= \frac{d\sigma}{d\mu}\cdot\frac{d\mu}{d\rho_i} = \frac{\mu}{\kappa\sigma x^2}\cdot\frac{d\mu}{d\rho_i} \\
\frac{d}{d\rho_i}\left(\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\sigma}\right) &= \frac{\partial^2 E[\mathcal{E}|\kappa_i]}{(\partial\sigma)^2}\cdot\frac{d\sigma}{d\rho_i} + \frac{\partial^2 E[\mathcal{E}|\kappa_i]}{\partial\sigma\partial\mu}\cdot\frac{d\mu}{d\rho_i} \\
&= 2\phi\left[\frac{\mu}{\sigma}\right]\cdot\frac{\mu}{\sigma^2}\cdot\left(\frac{\mu}{\sigma}\frac{d\sigma}{d\rho_i} - \frac{d\mu}{d\rho_i}\right) \\
&= -2\phi\left[\frac{\mu}{\sigma}\right]\cdot\frac{\mu}{\sigma^2}\cdot\frac{d\mu}{d\rho_i}\cdot\left(1 - \frac{\mu^2}{\sigma^2 x^2}\right) \\
&= -2\phi\left[\frac{\mu}{\sigma}\right]\cdot\frac{\mu}{\sigma^2}\cdot\frac{d\mu}{d\rho_i}\cdot\left(\frac{\kappa}{c^2\tau_b + \kappa}\right).
\end{aligned}$$

And now the third term using the expression for $\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\mu}$ from (7):

$$\begin{aligned}
\frac{d}{d\rho_i}\left(\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\mu}\right) &= \frac{\partial^2 E[\mathcal{E}|\kappa_i]}{(\partial\mu)^2}\cdot\frac{d\mu}{d\rho_i} + \frac{\partial^2 E[\mathcal{E}|\kappa_i]}{\partial\sigma\partial\mu}\cdot\frac{d\sigma}{d\rho_i} \\
&= 2\phi\left[\frac{\mu}{\sigma}\right]\cdot\frac{1}{\sigma}\cdot\left(\frac{d\mu}{d\rho_i} - \frac{\mu}{\sigma}\frac{d\sigma}{d\rho_i}\right) \\
&= 2\phi\left[\frac{\mu}{\sigma}\right]\cdot\frac{1}{\sigma}\cdot\frac{d\mu}{d\rho_i}\cdot\left(\frac{\kappa}{c^2\tau_b + \kappa}\right).
\end{aligned}$$

Deriving a few more quantities:

$$\begin{aligned}
\frac{d\sigma}{dn_i} &= \frac{d\sigma}{d\mu}\cdot\frac{d\mu}{dn_i} = \frac{\mu}{\sigma x^2}\cdot\frac{d\mu}{dn_i}, \\
\frac{d\mu}{d\rho_i} &= -\frac{n\tau(n-1)x}{(n(1+\rho_i\tau) + \tau(1-\rho_i))^2} = -\frac{c^2\tau(n-1)x}{n}, \text{ and} \\
\frac{d\mu}{dn_i} &= \frac{\tau(1-\rho_i)x}{(n(1+\rho_i\tau) + \tau(1-\rho_i))^2} = \frac{c^2\tau(1-\rho_i)x}{n^2},
\end{aligned}$$

which we plug in to show that the first and third term of (13) together are:

$$\begin{aligned}
&\frac{d}{d\rho_i}\bigg(\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\sigma}\bigg)\cdot\frac{d\sigma}{dn_i}+\frac{d}{d\rho_i}\bigg(\frac{\partial E[\mathcal{E}|\kappa_i]}{\partial\mu}\bigg)\cdot\frac{d\mu}{dn_i}\\
&\qquad = 2\phi\Big[\frac{\mu}{\sigma}\Big]\cdot\frac{1}{\sigma}\cdot\frac{d\mu}{d\rho_i}\cdot\bigg(\frac{\kappa}{c^2\tau_b+\kappa}\bigg)\cdot\bigg(\frac{d\mu}{dn_i}-\frac{\mu}{\sigma}\frac{d\sigma}{dn_i}\bigg)\\
&\qquad = 2\phi\Big[\frac{\mu}{\sigma}\Big]\cdot\frac{1}{\sigma}\cdot\bigg(\frac{\kappa}{c^2\tau_b+\kappa}\bigg)^2\cdot\frac{d\mu}{d\rho_i}\cdot\frac{d\mu}{dn_i}\\
&\qquad = -2\phi\Big[\frac{\mu}{\sigma}\Big]\cdot\bigg(\frac{\kappa}{c^2\tau_b+\kappa}\bigg)^2\cdot\frac{c^4\tau^2x^2(n-1)(1-\rho_i)}{\sigma n^3}<0.
\end{aligned}$$

This implies that $\frac{d^2E[\mathcal{E}|\kappa_i]}{dn_i d\rho_i}<0$. Thus, both $\kappa_i-\kappa$ and $\frac{dE[\mathcal{E}|\kappa_i]}{dn_i}$ are decreasing in ρ_i, and so Lemma 2 gives that $\text{cov}\left[\kappa_i-\kappa,\frac{dE[\mathcal{E}|\kappa_i]}{dn_i}\Big|n_i\right]>0$. This implies that $\frac{d}{dn_i}\text{cov}[\kappa_i-\kappa,E[\mathcal{E}|\kappa_i]|n_i]>0$, as desired. Note that $\rho=1$ is needed only to ensure the negativity of the second term, so if $\rho<1$, the overall expression will still be negative. (Simulations indicate that any parameter values compatible with Implication 2 are sufficient to guarantee the negativity of the second term.)

PROPOSITION 7: *Conditional on n:*

(i) More ideologically extreme citizens are more likely to turn out to vote.

(ii) Conditional on overconfidence, more ideologically extreme citizens are more likely to turn out.

(iii) More overconfident citizens are more likely to turn out, both conditional on, and independent of, ideological extremeness.

If ρ is large then these predictions also hold independent of n.

COROLLARY 1: *Conditional on n, strength of partisan identification is increasing in overconfidence, both conditional on, and independent of, ideological extremeness. Moreover, strength of partisan identification is increasing in ideological extremeness, both conditional on, and independent of, overconfidence. If ρ is large enough, these results hold independent of n.*

PROOF:

Fix $n_i=n$ and consider an individual i with ideology $\mathcal{I}$, overconfidence κ_i, and preference bias b_i. Suppose, without loss of generality that $\mathcal{I}_i>0$. Note that $E_i[x]=\mathcal{I}_i-b_i$. This means that we have $U_R(b_i|x)>U_L(b_i|x)$ if and only if $x>-b_i$. Thus,

$\text{Prob}_i[U_R(b_i|x) > U_L(b_i|x)] = \text{Prob}_i[x > -b_i] = 1 - \text{Prob}_i[x < -b_i]$. By construction this is equal to

$$1 - \Phi\left[(-b_i - (\mathcal{I}_i - b_i))\sqrt{\tau + \kappa_i}\right] = \Phi\left[\mathcal{I}_i\sqrt{\tau + \kappa_i}\right]. \tag{14}$$

As $\mathcal{I}_i > 0$, $\mathcal{I}_i\sqrt{\tau + \kappa_i}$ must be strictly increasing in κ_i conditional on $\mathcal{I}_i$, and in $\mathcal{I}_i$ conditional on κ_i. The same must therefore hold for $\Phi[\mathcal{I}_i\sqrt{\tau + \kappa_i}]$, and hence for $\text{Prob}_i[U_R(b_i|x) > U_L(b_i|x)]$. Note that specular results hold conditional on $\mathcal{I}_i < 0$. Thus, we can replace $\mathcal{I}_i$ with $\mathcal{E}_i = |\mathcal{I}_i|$ in (14).

Finally, $F_c(\cdot)$ and $F_c'(\cdot)$ are c.d.f.s and thus increasing in their arguments. This, together with the previous argument gives the second and third parts of the proposition and corollary, conditional on n. This, combined with Proposition 2 gives the first part of Proposition 7 and Corollary 1, and the third part of 7 independent of n.

Consider $n \sim F_n$. Suppose $\rho = 1$, so $\kappa = 1$ for all citizens. Then, more overconfident citizens, those with greater $\kappa_i - \kappa$, will be more confident (greater κ_i). This gives that, conditional on ideology, more overconfident citizens are more likely to turn out. Moreover, those with the same level of overconfidence will have the same confidence, so more ideological citizens will be more likely to turn out, conditional on overconfidence. This, combined with Proposition 2 gives the first part of Proposition 7 and Corollary 1 independent of n. These would continue to hold for ρ large enough, as in Implication 2.

PROPOSITION 8:

(i) $V_i(\kappa_i|n_i)$ *is increasing in* κ_i.

(ii) n_i^* *is increasing in the degree of correlational neglect.*

PROOF:

The utility of taking action a after receiving n signals is:

$$-\int (a - b_i - x)^2 dF_i[x] = -(a - b_i - E_i[x])^2 - \text{var}_i[x|n]$$

so the value of an additional signal is:

$$\begin{aligned} V_i(\kappa_i|n) &= \text{var}_i[x|n] - \text{var}_i[x|n{+}1] \\ &= \frac{1 + (n-1)\rho_i}{n + \tau(1 + (n-1)\rho_i)} - \frac{1 + n\rho_i}{n + 1 + \tau(1 + n\rho_i)} \end{aligned}$$

and plugging in for $\rho_i = \frac{n-\kappa_i}{(n-1)\kappa_i}$ from Lemma 1 we have:

$$\begin{aligned} V_i(\kappa_i|n) &= \frac{\kappa_i(\kappa_i-1)}{(\kappa_i+\tau)(n^2(\kappa_i+\tau)-\kappa_i(1+\tau))} \\ \frac{dV_i(\kappa_i|n)}{d\kappa_i} &= \frac{n^2\tau^2(\kappa_i-1)+\kappa_i^2(n^2-1)(1+2\tau)+\kappa_i\tau^2(n^2-\kappa_i)}{(\kappa_i+\tau)^2(n^2(\kappa_i+\tau)-\kappa_i(1+\tau))^2} > 0 \end{aligned}$$

where the sign of the last line follows from $1 < \kappa_i \leq n$, as $n \geq 2$ and $\rho_i \in [0,1)$.

We have that n_i^* is determined by

$$\underset{n_i}{\arg\max} \quad -(a-b_i-E_i[x])^2 - \mathrm{var}_i[x|n_i] - cn_i \quad \Rightarrow \quad \frac{d\mathrm{var}_i[x|n_i^*]}{dn} = -c$$

and thus n_i^* is defined implicitly by

$$\frac{1-\rho_i}{(n_i^*+\tau(1+(n_i^*-1)\rho_i))^2} - c = 0$$

and using the implicit function theorem we have

$$\frac{dn_i^*}{d\rho_i} = -\frac{\tau(n_i^*-1)(1-\rho_i)+n_i^*(1+\tau)}{2(1-\rho_i)(1+\rho_i\tau)} < 0$$

so n_i^* is increasing in correlational neglect $\rho - \rho_i$.

PROPOSITION 9: *When citizen i is told the ideology of citizen j, and she believes $\kappa_j = \kappa$:*

(i) The ideology of citizen i after communication is $\alpha_i\mathcal{I}_i + \beta_i\mathcal{I}_j$ for some $\alpha_i, \beta_i \in \mathbb{R}_{++}$, where α_i is increasing in κ_i and β_i is decreasing in κ_i.

(ii) If $\mathcal{I}_j \neq (\mathcal{I}_i - b_i)\frac{\kappa}{\kappa+\tau}$, then i's mean belief about the extremeness of j's preferences is increasing in i's level of overconfidence $\frac{d|E_i[b_j]|}{d\kappa_i} > 0$.

PROOF:

The posterior of citizen i about the bias of citizen j after observing $\mathcal{I}_j$ is:

$$\begin{aligned} \mathcal{L}(b_j|\mathcal{I}_j) &\propto \mathcal{L}(\mathcal{I}_j|b_j)\mathcal{L}(b_j) \\ &\propto \left[\int_{-\infty}^{+\infty} \exp\left\{-\frac{\kappa}{2}\left(x - \frac{\kappa+\tau}{\kappa}(\mathcal{I}_j - b)\right)^2\right\} * \exp\left\{-\frac{\kappa_i+\tau}{2}(x-\mathcal{I}_i)^2\right\} \mathrm{d}x\right] \\ &\quad * \exp\left\{-\frac{\tau_b b_j^2}{2}\right\}. \end{aligned}$$

This is a normal distribution with mean

$$\left(\mathcal{I}_j \frac{\kappa+\tau}{\kappa} - (\mathcal{I}_i - b_i)\right) * \frac{\kappa(\kappa_i+\tau)(\kappa+\tau)}{(\kappa+\tau)\left(\tau^2+\tau\kappa+\tau_b\kappa\right)+\kappa_i\left(\tau^2+2\tau\kappa+\kappa\tau_b+\kappa^2\right)}$$

where the second term is positive and increasing in κ_i. Thus, if $\mathcal{I}_j > (\mathcal{I}_i - b_i)\left(\frac{\kappa}{\kappa+\tau}\right)$, then $E_i[b_j] > 0$ and $E_i[b_j]$ is increasing in κ_i. If $\mathcal{I}_j < (\mathcal{I}_i - b_i)\left(\frac{\kappa}{\kappa+\tau}\right)$, then $E_i[b_j] < 0$ and $E_i[b_j]$ is decreasing in κ_i. Thus, $|E_i[b_j]|$ is increasing in κ_i.

The existence of $\alpha_i, \beta_i \in \mathbb{R}_{++}$ s.t. the ideology of citizen i after communication is $\alpha_i\mathcal{I}_i + \beta_i\mathcal{I}_j$ is a standard result of Bayesian updating. $\alpha_i + \beta_i \neq 1$ because ideology is a signal of both bias and beliefs. The fact that α_i increases and β_i decreases in κ_i is a direct consequence of the standard result that citizens with a high prior precision update less, and also because here they will tend to assign a higher probability to the fact that the differences in ideologies are due to differences in preference biases. In particular, solving for $\alpha_i, \beta_i \in \mathbb{R}_+$:

$$\alpha_i = \frac{(\kappa_i+\tau)\left(\tau^2+2\tau\kappa+\kappa\tau_b+\kappa^2\right)}{(\kappa+\tau)\left(\tau^2+\tau\kappa+\tau_b\kappa\right)+\kappa_i\left(\tau^2+2\tau\kappa+\kappa\tau_b+\kappa^2\right)}, \quad \beta_i = \frac{\kappa(1-\alpha_i)}{\kappa+\tau}$$

and thus

$$\frac{d\alpha_i}{d\kappa_i} = \frac{\tau_b\kappa^2(\tau_b\kappa+(\kappa+\tau)^2)}{(\kappa_i+\tau)(\kappa+\tau)^2+\tau_b\kappa(\kappa_i+\kappa+\tau)^2} > 0.$$

Thus, α_i is increasing in κ_i, so β_i is decreasing in κ_i.

PROPOSITION 10: *Suppose $b_i = 0$, $\forall i$. When citizen i is told the ideology of citizen j:*

(i) The ideology of citizen i after communication is $\gamma_i\mathcal{I}_i + \delta_i\mathcal{I}_j$ for some $\gamma_i, \delta_i \in \mathbb{R}_{++}$, where γ_i is increasing in κ_i and δ_i is decreasing in κ_i.

(ii) $E_i[\kappa_j]$ is increasing in κ_i if i and j are on opposite sides of the aisle, $(\mathcal{I}_i * \mathcal{I}_j < 0)$ or if j is more ideological extreme than i $(\mathcal{E}_j > \mathcal{E}_i)$.

(iii) $E_i[\kappa_j]$ is decreasing in κ_i if i and j are on the same side of the aisle $(\mathcal{I}_i * \mathcal{I}_j > 0)$, and $\mathcal{E}_i > \frac{\tau+\underline{\kappa}}{\underline{\kappa}}\mathcal{E}_j$.

PROOF:

We begin with the second and third parts of the proposition. By Bayes' rule: $\mathcal{L}(\kappa_j|\mathcal{I}_j) \propto \mathcal{L}(\mathcal{I}_j|\kappa_j)\mathcal{L}(\kappa_j)$. Note that $\mathcal{L}(\mathcal{I}_j|\kappa_j) = \phi_{\mathcal{I}_i,\kappa_i+\tau}\left(\mathcal{I}_j(\frac{\tau+\kappa_j}{\kappa_j})\right)$, where $\phi_{\mu,\tau}(\cdot)$ denotes the p.d.f. of a normal distribution with mean μ and precision τ. To prove that $E_i[\kappa_j]$ is increasing in κ_i, it is sufficient to prove that, for any $\kappa_j, \kappa_j' \in \text{supp}(F)$, $\kappa_j < \kappa_j'$, the ratio

$$\begin{aligned}
\frac{\mathcal{L}(\mathcal{I}_j|\kappa_j')}{\mathcal{L}(\mathcal{I}_j|\kappa_j)} &= \frac{\sqrt{\frac{\kappa_i+\tau}{2\pi}}\exp\left\{-\frac{(\kappa_i+\tau)}{2}\left(\mathcal{I}_j\left(\frac{\tau+\kappa_j'}{\kappa_j'}\right)-\mathcal{I}_i\right)^2\right\}}{\sqrt{\frac{\kappa_i+\tau}{2\pi}}\exp\left\{-\frac{(\kappa_i+\tau)}{2}\left(\mathcal{I}_j\left(\frac{\tau+\kappa_j}{\kappa_j}\right)-\mathcal{I}_i\right)^2\right\}} \\
&= \exp\left\{-\frac{\kappa_i+\tau}{2}\left(\left(\mathcal{I}_j\left(\frac{\tau+\kappa_j'}{\kappa_j'}\right)-\mathcal{I}_i\right)^2-\left(\mathcal{I}_j\left(\frac{\tau+\kappa_j}{\kappa_j}\right)-\mathcal{I}_i\right)^2\right)\right\}
\end{aligned}$$

is increasing in κ_i. This holds if and only if

$$\left(\mathcal{I}_j\left(\frac{\tau+\kappa_j'}{\kappa_j'}\right)-\mathcal{I}_i\right)^2<\left(\mathcal{I}_j\left(\frac{\tau+\kappa_j}{\kappa_j}\right)-\mathcal{I}_i\right)^2 \tag{15}$$

for all $\kappa_j,\kappa_j' \in \text{supp}(F)$, $\kappa_j < \kappa_j'$. If the converse of (15) holds for all $\kappa_j,\kappa_j' \in \text{supp}(F)$, $\kappa_j < \kappa_j'$, this is sufficient for $E_i[\kappa_j]$ to be decreasing in κ_i.

As $\frac{\tau+\kappa_j}{\kappa_j}$ is decreasing in κ_j, $\mathcal{E}_j(\frac{\tau+\kappa_j'}{\kappa_j'}) < \mathcal{E}_j(\frac{\tau+\kappa_j}{\kappa_j})$ since $\kappa_j < \kappa_j'$. This implies (15) holds if $\mathcal{I}_i * \mathcal{I}_j < 0$ or $\mathcal{E}_j > \mathcal{E}_i$. By contrast, the converse holds if $\kappa_j,\kappa_j' \in \text{supp}(F)$, $\kappa_j < \kappa_j'$ if $\mathcal{I}_i * \mathcal{I}_j > 0$, and $\mathcal{E}_i > \frac{\tau+\overline{\kappa}}{\underline{\kappa}}\mathcal{E}_j$.

Finally, as in the Proof of Proposition 9, the first part follows from standard properties of Bayesian Updating.

B. Survey Details

The typical way psychologists measure overconfidence is not well suited to surveys. They often use a very large number of questions—up to 150 (see, for example, Alpert and Raiffa, 1969/1982; Soll and Klayman, 2004)—and elicit confidence using confidence intervals, which may be difficult for the average survey respondent to understand (see, for example, Juslin, Wennerholm and Olsson, 1999; Rothschild, 2012).

Our methodology for measuring overconfidence on surveys uses three innovations. The first two are due to Ansolabehere, Meredith and Snowberg (2014). First, the questions we use are about either quantities that everyone knows the scale of, such as dates, or the scale is provided, as in the case of unemployment or inflation. That is, when asking about unemployment rates, the question gives respondents the historical minimum, maximum, and median of that rate. This has been shown to reduce the number of incorrect answers simply due to a respondent not knowing the appropriate scale (Ansolabehere, Meredith and Snowberg, 2013). Second, confidence is elicited on a qualitative scale, which is easily understandably by survey respondents and allows for more conservative controls for actual knowledge.

The third innovation is a modification of the second, and was only utilized on the 2011 CCES. For our general knowledge questions—the year the telephone was invented, the population of Spain, the year Shakespeare was born, and the percent of the U.S. population that lives in California—we elicited confidence using an inverted confidence interval.[3] That is, rather than asking for a confidence interval directly, which we felt may have been too challenging for survey respondents, we asked them to give their estimates of the probability that the true answer was in some interval around their answer. So, for example, after giving their best guess as to the date of Shakespeare's birth, respondents were asked:

> What do you think the percent chance is that your best guess, entered above, is within 50 years of the actual answer?

Given a two-parameter distribution, such as a normal, this is enough to pin down the variance of a respondent's belief.

The sum total of these innovations is that overconfidence can be elicited using a small number of questions that are understandable to most survey respondents, rather than just to university undergraduates.

B.1. Survey Questions

We next present the text of the questions used to construct our overconfidence measure on the 2010 and 2011 CCES, as described in Section IIB. Instructions in brackets indicate limitations on possible answers implemented by the survey company—these were not displayed to respondents. If a survey respondent tried to enter, say, text where only a positive number was allowed, they would be told to edit their entry to conform with the limitations placed on the response field. If a respondent tried to skip a question, the survey would request that the respondent give an answer. If the respondent tried to skip the same question a second time, they were allowed to do so.

1. The unemployment rate is the percent of people actively searching work but not presently employed. Since World War II it has ranged from a low of 2 percent to a high of 11 percent.

   What is your best guess about the unemployment rate in the United States today? Even if you are uncertain, please provide us with your best estimate of the percent of people seeking work but currently without a job in the United States.

   ____% [only allow a positive number]

2. How confident are you of your answer to this question?

   - No confidence at all
   - Not very confident

[3]Note that these general knowledge questions were all from previous research on overconfidence.

- Somewhat unconfident
- Somewhat confident
- Very confident
- Certain

3. The inflation rate is the annual percentage change in prices for basic goods like food, clothing, housing, and energy. Since World War II it has ranged from a high of 14 percent (a 14% increase in prices over the previous year) to a low of -2 percent (a 2% decline in prices over the previous year).

   What is your best guess about the inflation rate in the United States today? Even if you are uncertain, please provide us with your best estimate of about what percent do you think prices went up or down in the last 12 months.

   Do you think prices went up or down?

   - Up
   - Down

4. By what percent do you think prices went up or down?

   ____% [only allow a positive number]

5. How confident are you of your answer to this question?

   - No confidence at all
   - Not very confident
   - Somewhat unconfident
   - Somewhat confident
   - Very confident
   - Certain

6. The unemployment rate is the percent of people actively searching work but not presently employed. Since World War II it has ranged from a low of 2 percent to a high of 11 percent.

   What do you expect the unemployment rate to be a year from now? Even if you are uncertain, please provide us with your best estimate of the percent of people who will be seeking but without a job in the United States in November, 2011.

   ____% [only allow a positive number]

7. How confident are you of your answer to this question?

   - No confidence at all
   - Not very confident

- Somewhat unconfident
- Somewhat confident
- Very confident
- Certain

8. The inflation rate is the annual percentage change in prices for basic goods like food, clothing, housing, and energy. Since World War II it has ranged from a high of 14 percent (a 14% increase in prices over the previous year) to a low of -2 percent (a 2% decline in prices over the previous year).

   What do you expect the inflation rate to be a year from now? Even if you are uncertain, please provide us with your best estimate of about what percent do you expect prices to go up or down in the next 12 months.

   Do you expect prices to go up or down?

   - Up
   - Down

9. By what percent do you expect prices to go up or down?

   ____% [only allow a positive number]

10. How confident are you of your answer to this question?

    - No confidence at all
    - Not very confident
    - Somewhat unconfident
    - Somewhat confident
    - Very confident
    - Certain

Next, we list the questions from the 2011 CCES used to construct the overconfidence measures discussed in Section VIA. Note that the unemployment questions were changed from 2010, in accordance with the evolving research agenda of Ansolabehere, Meredith and Snowberg.

1. In what year was the telephone invented? Even if you are not sure, please give us your best guess.

   ____

2. How confident are you of your answer to this question?

   - No confidence at all

- Not very confident
- Somewhat unconfident
- Somewhat confident
- Very confident
- Certain

3. As a different way of answering the previous question, what do you think the percent chance is that your best guess, entered above, is within 25 years of the actual answer?

   ____%

4. What is the population of Spain, in millions? Even if you are not sure, please give us your best guess.

   ____

5. How confident are you of your answer to this question?

   - No confidence at all
   - Not very confident
   - Somewhat unconfident
   - Somewhat confident
   - Very confident
   - Certain

6. As a different way of answering the previous question, what do you think the percent chance is that your best guess, entered above, is within 15 million of the actual answer?

   ____%

7. In what year was the playwright William Shakespeare born? Even if you are not sure, please give us your best guess.

   ____%

8. How confident are you of your answer to this question?

   - No confidence at all
   - Not very confident
   - Somewhat unconfident
   - Somewhat confident
   - Very confident
   - Certain

9. As a different way of answering the previous question, what do you think the percent chance is that your best guess, entered above, is within 50 years of the actual answer?

   ____%

10. What percent of the US population lives in California? Even if you are not sure, please give us your best guess.

   ____

11. How confident are you of your answer to this question?

    - No confidence at all
    - Not very confident
    - Somewhat unconfident
    - Somewhat confident
    - Very confident
    - Certain

12. As a different way of answering the previous question, what do you think the percent chance is that your best guess, entered above, is within 5 percentage points of the actual answer?

   ____%

13. According to the Bureau of Labor Statistics, since World War II the most non-agricultural jobs the US economy has lost in a year is 5.4 million. The most jobs gained in a year has been 4.2 million. Over the same period, the US economy has gained an average of 1.4 million jobs a year.

    What is your best guess about the number of jobs gained or lost in the last year?

    Over the past year, I think the US economy has overall

    - Lost jobs
    - Gained jobs

14. How many jobs do you think have been lost or gained over the past year?

    ____ million jobs [only allow a positive number]

15. How confident are you of your answer to this question?

    - No confidence at all
    - Not very confident
    - Somewhat unconfident
    - Somewhat confident

- Very confident
- Certain

16. The inflation rate is the annual percentage change in prices for basic goods like food, clothing, housing, and energy. Since World War II it has ranged from a high of 14.4. percent (a 14.4% increase in prices over the previous year) to a low of -1.2 percent (a 1.2% decline in prices over the previous year).

    What is your best guess about the inflation rate in the United States today?

    Do you think prices went up or down?

    - Up
    - Down

17. By what percent do you think prices went up or down?

    ____% [only allow a positive number]

18. How confident are you of your answer to this question?

    - No confidence at all
    - Not very confident
    - Somewhat unconfident
    - Somewhat confident
    - Very confident
    - Certain

19. According to the Bureau of Labor Statistics, since World War II the most non-agricultural jobs the US economy has lost in a year is 5.4 million. The most jobs gained in a year has been 4.2 million. Over the same period, the US economy has gained an average of 1.4 million jobs a year.

    What is your best guess about the number of jobs that will be gained or lost over the next year?

    Over the next year, I think the US economy will overall

    - Lose jobs
    - Gain jobs

20. How many jobs do you think the US economy will lose or gain over the next year?

    ____ million jobs [only allow a positive number]

21. How confident are you of your answer to this question?

    - No confidence at all

- Not very confident
- Somewhat unconfident
- Somewhat confident
- Very confident
- Certain

22. The inflation rate is the annual percentage change in prices for basic goods like food, clothing, housing, and energy. Since World War II it has ranged from a high of 14.4. percent (a 14.4% increase in prices over the previous year) to a low of -1.2 percent (a 1.2% decline in prices over the previous year).

    What do you expect the inflation rate to be a year from now?

    Do you expect prices to go up or down?

    - Up
    - Down

23. By what percent do you expect prices to go up or down?

    ____% [only allow a positive number]

24. How confident are you of your answer to this question?

    - No confidence at all
    - Not very confident
    - Somewhat unconfident
    - Somewhat confident
    - Very confident
    - Certain

C. Additional Theoretical Specifications

This section addresses, in a casual way, a number of theoretical questions that have been posed to us. While the result of our inquiry into these questions did not produce results that merit a discussion in the main text, we thought it would be useful to record the results.

Distributional Assumptions.—Throughout the paper we make heavy use of normal distributions. This has advantages for both tractability and interpretation. In particular, tractability is helped by the fact that a normal is a self-conjugate prior, and that properties of the normal are well studied in statistics. The advantage in interpretation comes from the fact that the normal is a two-parameter distribution (the mean and precision), so it is straightforward to implement and interpret overconfidence as a function of precision without worrying about the effects of higher (or lower) order moments.

However, this leads to questions about how much our results are driven by the use of normal distributions. Or, conversely, many seminar attendees have conjectured that it would be straight-forward to extend our results to well-behaved distributions. Here we give some guidance on this subject.

We start by discussing how our results might generalize to other distributions. Without the normal distribution, the correlational neglect model becomes intractable. The value of this model is that it allows us to make predictions about the role of the number of signals that could not be obtained under any fully Bayesian model, as discussed in Section IIIA.

However if one is willing to put aside these predictions, it is possible to discuss the role of the normal when citizens receive uncorrelated signals they over-interpret (as in the "model" of Lemma 1). The proof of Proposition 2 relies on the fact that both overconfidence and extremeness are increasing in correlational neglect. Intuitively this seems as though it would hold for a wide range of distributions (at least when $x = 0$), but it is quite difficult to show this analytically. Using the normal distribution, then, gives two advantages: tractability, as just discussed, and clarity, as unique among commonly used multi-variate conjugate distributions the variance-covariance matrix is a parameter. This makes the definition of correlational neglect very clear as it does not require tweaking other parameters of the distribution. Indeed, while we have verified that our primary result holds when priors are distributed according to a Beta (or uniform) distribution, and signals are Bernoulli, the interpretation of even this simple model is much more difficult.

If one uses a support with only two possible states, then our results may not always hold. However, it is known that such a setup (without overconfidence) produces perverse results: see McMurray (2013). In particular, with only two states, the precision of beliefs may decrease, rather than increase with more signals. However, this would be inconsistent empirical results in Section IIIA.

Multi-Dimensional Issue Spaces.—Our theory has implications for how ideology on different dimensions would be related to overconfidence. For example, if the information on a given dimension were all public, with agreed upon correlational structure, then there should be no relationship between ideology and overconfidence on that dimension. While this implication is straight-forward to work out, we did not feel that it was testable with current data.

In particular, in order to test this, one would need to know quite a bit about where citizens get their data from, and how citizens infer about how this data affects them. For example, even if most economic information is public, how that information relates to a citizen's permanent income is more opaque. Learning about that relationship would entail seeing how nationwide economic performance seemed to affect a citizen's own employment situation. As these very personal signals would have an unknown correlational structure, there is plenty of room for correlational neglect.

Likewise, positions on a social issue like gay marriage may appear to have no informational content at all, and hence, there should be no relationship between overconfidence and ideology on this dimension. However, it is perfectly reasonable that one's position on gay marriage may depend on beliefs about the likelihood that a loved one, say a child, is gay. This likelihood may be drawn, in part, from the number of openly gay people in a citizen's social environment. If a citizen neglects the fact that they live in a religious community where others are not open about their sexuality, then they will tend to underestimate the probability that a loved may turn out to be gay. This will lead to both overconfidence and more extreme positions, as before.

We believe that applying our theory to multi-dimensional spaces would be interesting, and possibly fruitful. We refrain from doing so in this paper because it does not add to the predictions we can test in our data.

Other Dimensions of Personality.—We treat correlational neglect as akin to a personality trait, which has raised questions of how this might be related to other personality traits. In particular, previous research has found that overconfidence is related to the extraversion of the "Big Five" personality inventory (Schaefer et al., 2004), although Moore and Healy (2007) has found that it is orthogonal to all traits in the Big Five. Regardless, extraversion does not have any significant explanatory power for the political behaviors we consider here (Gerber et al., 2010, 2012). Other studies have noted a link between overconfidence and narcissism. Little is known about the relationship between narcissism and political behavior, nor are there formal theories (that we are aware of) that relate narcissism to decision making more generally.

C.1. Rational Models of Voting

Our model of voter turnout, and partisan identification, is based on a specific form of expressive voting (Fiorina, 1976; Brennan and Hamlin, 1998). In particular a citizen i votes if and only if

$$\left|\text{Prob}_i[U_R(b_i|x) > U_L(b_i|x)] - \frac{1}{2}\right| - c_i > 0, \tag{C1}$$

where c_i is an i.i.d. draw from some distribution F_c, which is strictly increasing on $\left(0, \frac{1}{2}\right)$. In addition $c_i \perp (\rho_i, b_i, e_{it})$.

While any political economy model where turnout is exogenous implicitly uses an expressive voting model (and others use it more explicitly, see Knight, 2013), there are a number of other approaches in the political economy literature. As each approach has its partisans, we thought it worthwhile to discuss those models, and show, where possible, how our model relates to them.

Before discussing alternative models, we should note that we focused on the expressive approach because we believe it is correct, and because it is compatible (as shown below) with a promising approach in the literature, that voters are choice- or regret-avoidant (Matsusaka, 1995; Degan and Merlo, 2011; Degan, 2013).

In addition, this modeling approach allows for both non-trivial turnout and strong partisan identification even if the policies proposed by political parties are similar to each other, as seems to be the case in reality (Snowberg, Wolfers and Zitzewitz, 2007*a*,*b*). This is generally not possible in more traditional models. To make this specific, suppose that both parties propose very similar platforms, and consider a citizen who is very confident that the best policy for her is proposed by party R. According to our model, this citizen would strongly identify with, and turn out to vote, for party R. However, if these behaviors were rooted in expected utility, and the parties espoused similar platforms, this would not hold. For any reasonably smooth utility function there is a small difference in utility between the two parties—and hence no reason to strongly identify with one party or the other, or turnout.

Pivotal Voting.—In these models, the turnout decision is driven largely by whether or not a voter is likely to be pivotal—that is, change the outcome of the election (Riker and Ordeshook, 1968). In this model a citizen turns out to vote if and only if

$$pB_i - C_i + D_i > 0 \tag{C2}$$

where p is the probability an individual citizen's vote is pivotal—that is, changes the winner of the election—and B_i is the benefit to the citizen of the citizen's favored candidate winning over the other candidate. The remaining terms, C_i and D_i, are the instrumental costs and benefits of voting, which are unrelated to the outcome of the election.

It seems reasonable to assume that more-overconfident citizens would over-estimate their probability of being pivotal. This would lead to the prediction that more overconfident citizens would be more likely to turnout.

However, whether or not more ideologically extreme people are more likely to turn out will depend on their utility function. It is well known in the literature on pivotal voting than in order for more ideologically extreme people to be more likely to turn out, utilities need to be very concave: that is, they care much more about small differences in policy when those policies are very far away from their ideal, than when those policies are close to them. Adding overconfidence adds some additional issues: in particular, in order to have more extreme citizens be more likely to turn out the utility function has to be more concave than a quadratic loss function. We have examined a quartic loss function, and even this degree of concavity will not guarantee the result: it holds only for specific parameters and values of the fourth moment of the distribution of beliefs.

Finally, we do not know if it is possible to replicate our conditional predictions about the role of overconfidence and extremeness using a pivotal voter model. As such, it seems that turning in our model for a pivotal voter model would be a poor choice.

Group Utilitarian.—In the group-utilitarian framework a citizen votes not just because voting may improve her utility, but because it will improve the utility of others like her as well (Coate and Conlin, 2004; Feddersen and Sandroni, 2006). In these models there is heterogeneity in the costs of voting, and this selects who, from a group, actually turns out. In order to use our model of overconfidence, there needs to be a mapping from beliefs to the cost of voting. An expression for the cost of voting like the left-hand-side of (C1) works, and once this is nested in the group-utilitarian framework will produce the same comparative statics as in Proposition 7. This occurs because in the group utilitarian framework those with the lowest costs of voting vote (up to some threshold), and the overconfident, and ideologically extreme, have the lowest costs according to (C1). While it would have been possible to use the full group-utilitarian framework in Section IVB, we felt that, for concision, it was best to avoid that machinery and show directly the important assumption that gives the predictions in that section.

C.2. Behavioral Models of Voting

The remaining two models we discuss—like the expressive voting model—focus on the idiosyncratic costs and benefits of turning out to vote. In particular, they focus on large elections where the number of voters grows large, and hence, $p_i \to 0$.

Regret-Avoidance.—Matsusaka (1995) argues that voter turnout is driven in part by whether citizens anticipate they will regret their vote. We view this theory as descriptively accurate: indeed, we ran a survey on a convenience sample using Mechanical Turk, and found that over 60 percent of respondents reported that they took into account whether they might regret their vote when deciding whether or not to vote. Almost 40 percent could name someone they regretted voting for.[4]

It is straightforward to show that our model is consistent with a model of regret-avoidant voting. In particular, as $p_i \to 0$, a citizen's turnout decision depends only on the idiosyncratic, instrumental costs and benefits of voting in (C1), C_i and D_i. We decompose the instrumental cost into two parts: direct costs C_i', such as the opportunity cost of going to vote, and a regret penalty $\mathcal{R}_i$ that accrues if the citizen votes for a candidate whose platform turns out to be worse for the citizen, given the state. That is

$$D_i - C_i \equiv D_i - \mathcal{R}_i \mathbb{I}_{\text{vote=wrong}} - C_i'$$

with D_i, $\mathcal{R}_i$ and C_i' i.i.d. draws from some (possibly different) distributions.[5] We then have:

PROPOSITION C1: *In large elections when $D_i - C_i \equiv D_i - \mathcal{R}_i \mathbb{I}_{\text{vote=wrong}} - C_i'$, comparative statics on voter turnout and partisan identification are the same as comparative statics on*

$$\left| Prob_i[U_R(b_i|x) > U_L(b_i|x)] - \frac{1}{2} \right| - c_i > 0.$$

PROOF:

[4]For more on regret-avoidance, see Connolly and Zeelenberg (2002), Zeelenberg (1999), Zeelenberg, Inman and Pieters (2001). Models of regret have then been frequently used to explain behavioral patterns which are not compatible with standard, expected-utility, models (Bell, 1982; Loomes and Sugden, 1982; Loomes and Sugden, 1987; Sugden, 1993; and Sarver, 2008). Indeed, Matsusaka's approach is a direct instantiation of Sugden (1993), applied to politics.

[5]We emphasize that, although we pick a particular formalization, (expected) regret can be seen as either a reduction in the benefit of voting, or an increase in the cost of voting.

When elections are large $p \to 0$ in (C2). Supposing citizen i favors candidate R if he or she were to vote, citizen i will vote if and only if

$$\begin{aligned} D_i - \mathcal{R}_i E[\mathbb{I}_{\text{vote=wrong}}] - C_i' &> 0 \\ \text{Prob}[\text{vote} = \text{wrong}] &< \frac{D_i - C_i'}{\mathcal{R}_i} \\ 1 - \text{Prob}[U_R(b_i|x) > U_L(b_i|x)] &< \frac{D_i - C_i'}{\mathcal{R}_i} \\ \text{Prob}[U_R(b_i|x) > U_L(b_i|x)] - \frac{1}{2} &> \frac{1}{2} - \frac{D_i - C_i'}{\mathcal{R}_i} \equiv c_i. \end{aligned}$$

The absolute value follows from considering the case where i favors candidate L.

We chose to display this chain of logic here to simplify and shorten exposition in the text.

Choice-Avoidance.—Degan and Merlo (2011) use the same idea as Matsusaka (1995). However, they note that as it is unlikely that a citizen will discover the actual state, they will not anticipate regretting their decision; instead, they discuss their model in terms of choice-avoidance. It should be clear from the form of (C1) that citizens who make their voting decision in this way are choice-avoidant. In particular, a citizen avoids choice unless the choice is clear.[6]

C.3. *Strength of Partisan Identification*

Our initial model of strength of partisan identification assumed that citizens would invest in a partisan identity only if they believed there was a sufficiently high probability that they would stay on the same side of the ideological spectrum as they received more signals.

This yields the same predictions as Corollary 1. More overconfident citizens would believe that, with high-probability, future signals would just confirm what they already knew. As such, there is little chance that they would end up on the opposite side of the ideological spectrum. Thus, more-overconfident citizens would be more likely to strongly identify with a party.

More ideologically extreme citizens would know that they would need a more extreme signal that the state is on the other side of the ideological spectrum in order to cross-over to that side. As such, there is little chance they would end up on the opposite side, and they would thus be more likely to strongly identify with a party.

We removed this additional model from the text of the paper in order to simplify and shorten the exposition.

[6]For examples of choice avoidance in other contexts see Iyengar, Huberman and Jiang (2004), Iyengar and Lepper (2000), Boatwright and Nunes (2001), Shah and Wolford (2007), Schwartz (2004), Choi, Laibson and Madrian (2009), DellaVigna (2009), Reutskaja and Hogarth (2009), and Bertrand et al. (2010).

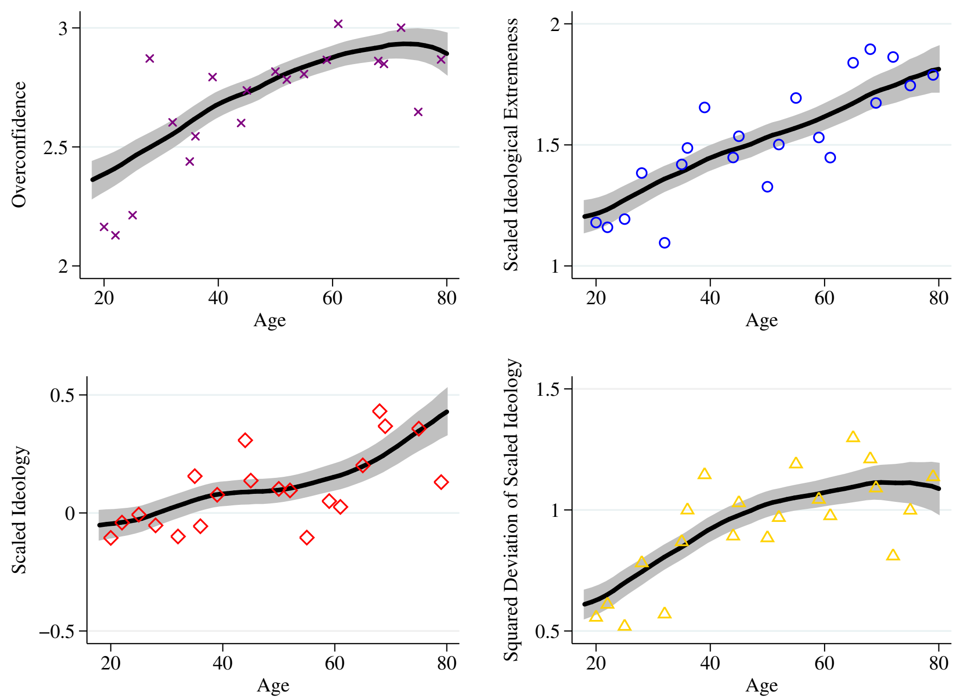

Figure D1. Age, Overconfidence, and Ideology

Note: Each point is the average for three years of age. Trendiness, in black, and 95 percent confidence intervals, in gray, use an Epanechnikov kernel with a bandwidth of 8.

D. Additional Empirical Specifications

In the text we present our preferred specifications. Here we provide additional specifications that we excluded from the text for concision. Please note that our working paper, Ortoleva and Snowberg (2013), contains even more alternative specifications, including alternative ideology measures, extremism measures that are constructed directly from ideology without first controlling for economic variables, demographic controls, and unweighted specifications. The results in all cases are similar.

A feature of the theory that we do not emphasize in the text, is that our model predicts that many of the relationships in the data should be monotonic. For example, the increase in overconfidence with media exposure and age should be monotonic. Moreover, the relationship between overconfidence and (average) extremeness, and overconfidence and (average) ideology (when $x > 0$) should both be monotonically increasing relationships. We do not discuss those results in the text because the econometrics of testing for monotonicity is still in its infancy (see Patton and Timmermann, 2010, 2012). As such, we prefer to show graphical representations of some of these monotonic relationships.

The relationships in Figure 1 appear monotonic. The same graphs for age are found in Figure D1. Figure D2 shows the relationship between overconfidence and ideology, and overconfidence and extremeness. Once again, these relationships appear monotonic. Figure D3 shows the results in Table 5 graphically. These results also appear monotonic, and you can see the lower slope in the left-hand graphs (for the covariance to the left of center) than those in the right-hand graphs.

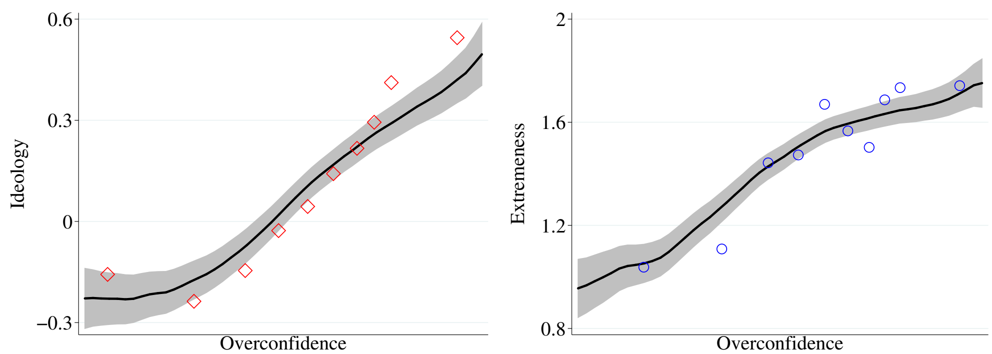

FIGURE D2. THE RELATIONSHIP BETWEEN OVERCONFIDENCE AND EXTREMENESS, AND OVERCONFIDENCE AND IDEOLOGY APPEARS TO BE MONOTONIC.

Note: Each point is the average for a decile of overconfidence, or three years of age. Trendiness, in black, and 95 percent confidence intervals, in gray, use an Epanechnikov kernel with a bandwidth of 0.5.

Finally, we show the regression results for all three—including the two alternate—measures of ideology that Table 4 is based on. As can be seen, the patterns in Table 3 are also found in the alternative measures. Indeed, in most cases, the results using the alternative measures are stronger.

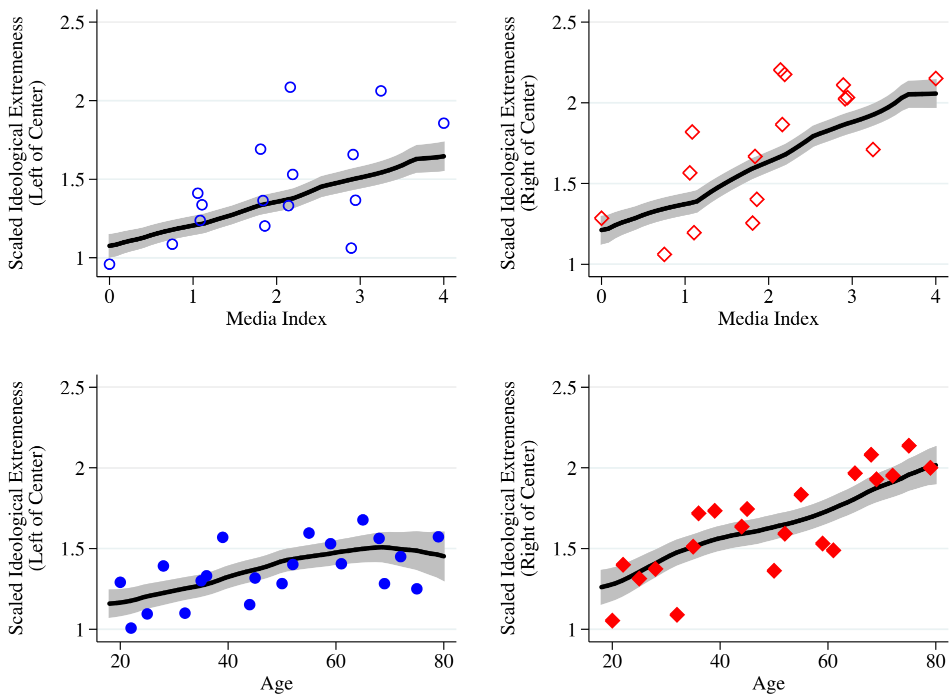

Figure D3. There is a greater covariance between extremeness and overconfidence for right-of-center citizens than left of center citizens.

Note: Each point is the average for a value of the media index, or three years of age. Trendiness, in black, and 95 percent confidence intervals, in gray, use an Epanechnikov kernel with a bandwidth of 0.8 for media figures, and 8 for age figures.

TABLE D1—IDEOLOGY, AND IDEOLOGICAL EXTREMENESS IS ROBUSTLY RELATED TO OVERCONFIDENCE.

Ideology Measure:	Scaled			Self-Reported Treatment of "Don't Know"					
				Centrist (0)			Missing (.)		
Panel A: Ideology (Right-Left)									
Overconfidence	0.22***	0.22***	0.20***	0.22***	0.22***	0.17***	0.25***	0.25***	0.21***
	(.028)	(.023)	(.024)	(.022)	(.019)	(.021)	(.025)	(.020)	(.022)
Economic Controls		Y	Y		Y	Y		Y	Y
Number of Signals			Y			Y			Y
R^2	0.047	0.16	0.23	0.048	0.15	0.23	0.057	0.18	0.27
N		2,868			2,910			2,754	
Panel B: Extremism (Generated from Right-Left Ideology Purged of Economic Controls)									
Overconfidence	0.23***	0.17***	0.12***	0.26***	0.24***	0.20***	0.28***	0.26***	0.23***
	(.028)	(.027)	(.026)	(.024)	(.024)	(.025)	(.026)	(.026)	(.026)
Economic Controls		Y	Y		Y	Y		Y	Y
Number of Signals			Y			Y			Y
R^2	0.05	0.19	0.29	0.067	0.084	0.16	0.069	0.083	0.17
N		2,868			2,910			2,754	

Note: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors, clustered by age (73 clusters), in parentheses. All specifications estimated using WLS with CCES sampling weights.

REFERENCES

Alpert, M., and H. Raiffa. 1969/1982. "A Progress Report on the Training of Probability Assessors." In *Judgement Under Uncertainty: Heuristics and Biases*. , ed. Daniel Kahneman, Paul Slovic and Amos Tversky, 294–305. Cambridge and New York:Cambridge University Press.

Ansolabehere, Stephen, Marc Meredith, and Erik Snowberg. 2013. "Asking About Numbers: Why and How." *Political Analysis*, 21(1): 48–69.

Ansolabehere, Stephen, Marc Meredith, and Erik Snowberg. 2014. "Mecro-Economic Voting: Local Information and Micro-Perceptions of the Macro-Economy." *Economics & Politics*, 26(3): 380–410.

Bell, David E. 1982. "Regret in Decision Making Under Uncertainty." *Operations Research*, 30(5): 961–981.

Bertrand, Marianne, Dean S. Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman. 2010. "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment." *The Quarterly Journal of Economics*, 125(1): 263–306.

Boatwright, Peter, and Joseph C. Nunes. 2001. "Reducing Assortment: An Attribute-Based Approach." *The Journal of Marketing*, 65(1): 50–63.

Brennan, Geoffrey, and Alan Hamlin. 1998. "Expressive Voting and Electoral Equilibrium." *Public Choice*, 95(1-2): 149–175.

Choi, James, David Laibson, and Brigitte Madrian. 2009. "Reducing the Complexity Costs of 401 (k) Participation Through Quick Enrollment." In *Developments in the Economics of Aging*. , ed. David A. Wise, 57–82. Chicago, IL:University of Chicago Press.

Coate, Stephen, and Michael Conlin. 2004. "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence." *The American Economic Review*, 94(5): 1476–1504.

Connolly, Terry, and Marcel Zeelenberg. 2002. "Regret in Decision Making." *Current Directions in Psychological Science*, 11(6): 212–216.

Degan, Arianna. 2013. "Civic duty and political advertising." *Economic Theory*, 52(2): 531–564.

Degan, Arianna, and Antonio Merlo. 2011. "A Structural Model of Turnout and Voting in Multiple Elections." *Journal of the European Economic Association*, 9(2): 209–245.

DellaVigna, Stefano. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315–372.

Feddersen, Timothy, and Alvaro Sandroni. 2006. "A Theory of Participation in Elections." *The American Economic Review*, 96(4): 1271–1282.

Fiorina, Morris P. 1976. "The Voting Decision: Instrumental and Expressive Aspects." *Journal of Politics*, 38(2): 390–413.

Gerber, Alan S., Gregory A. Huber, David Doherty, and Conor M. Dowling. 2012. "Personality and the Strength and Direction of Partisan Identification." *Political Behavior*, 34(4): 653–688.

Gerber, Alan S., Gregory A. Huber, David Doherty, Conor M. Dowling, and Shang E. Ha. 2010. "Personality and Political Attitudes: Relationships Across Issue Domains and Political Contexts." *American Political Science Review*, 104(1): 111–133.

Iyengar, Sheena S., and Mark L. Lepper. 2000. "When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?" *Journal of Personality and Social Psychology*, 79(6): 995–1006.

Iyengar, Sheena S., Grub Huberman, and Wei Jiang. 2004. "How Much Choice is Too Much? Contributions to 401 (k) Retirement Plans." In *Pension Design and Structure: New Lessons from Behavioral Finance*. , ed. Olivia S. Mitchell and Stephen P. Utkus, Chapter 5, 83–95. Oxford University Press.

Juslin, Peter, Pia Wennerholm, and Henrik Olsson. 1999. "Format Dependence in Subjective Probability Calibration." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4): 1038–1052.

Knight, Brian. 2013. "Evaluating Competing Explanations for the Midterm Gap: A Unified Econometric Approach with Microfoundations." Brown University, *mimeo.*

Loomes, Graham, and Robert Sugden. 1982. "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty." *The Economic Journal*, 92(368): 805–824.

Loomes, Graham, and Robert Sugden. 1987. "Some Implications of a More General Form of Regret Theory." *Journal of Economic Theory*, 41(2): 270–287.

Matsusaka, John G. 1995. "Explaining Voter Turnout Patterns: An Information Theory." *Public Choice*, 84(1): 91–117.

McMurray, Joseph C. 2013. "Ideology as Opinion: A Spatial Model of Common-value Elections." Bringam Young University, *mimeo.*

Moore, Don A., and Paul J. Healy. 2007. "The Trouble with Overconfidence." Carnegie Mellon University, *mimeo.*

Ortoleva, Pietro, and Erik Snowberg. 2013. "Overconfidence in Political Behavior." NBER Working Paper Series # 19250.

Patton, Andrew J, and Allan Timmermann. 2010. "Why do Forecasters Disagree? Lessons from the Term Structure of Cross-sectional Dispersion." *Journal of Monetary Economics*, 57(7): 803–820.

Patton, Andrew J, and Allan Timmermann. 2012. "Forecast Rationality Tests Based on Multi-Horizon Bounds." *Journal of Business & Economic Statistics*, 30(1): 1–17.

Reutskaja, Elena, and Robin M. Hogarth. 2009. "Satisfaction in Choice as a Function of the Number of Alternatives: When "Goods Satiate"." *Psychology and Marketing*, 26(3): 197–203.

Riker, William H., and Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *The American Political Science Review*, 62(1): 25–42.

Rothschild, David. 2012. "Expectations: Point Estimates, Probability Distributions, Confidence, and Forecasts." Microsoft Research, *mimeo*.

Sarver, Todd. 2008. "Anticipating Regret: Why Fewer Options May Be Better." *Econometrica*, 76(6): 263–305.

Schaefer, Peter S., Cristina C. Williams, Adam S. Goodie, and W. Keith Campbell. 2004. "Overconfidence and the Big Five." *Journal of Research in Personality*, 38(5): 473–480.

Schmidt, Klaus D. 2003. "On the Covariance of Monotone Functions of a Random Variable." Lehrstuhl für Versicherungsmathematik Technische Universität Dresden, *mimeo.*

Schwartz, Barry. 2004. *The Paradox of Choice: Why More is Less.* New York, NY:Harper Collins.

Shah, Avni M., and George Wolford. 2007. "Buying Behavior as a Function of Parametric Variation of Number of Choices." *Psychological Science*, 18(5): 369–370.

Snowberg, Erik, Justin Wolfers, and Eric Zitzewitz. 2007*a*. "Partisan Impacts on the Economy: Evidence from Prediction Markets and Close Elections." *The Quarterly Journal of Economics*, 122(2): 807–829.

Snowberg, Erik, Justin Wolfers, and Eric Zitzewitz. 2007*b*. "Party Influence in Congress and the Economy." *Quarterly Journal of Political Science*, 2(3): 277–286.

Soll, Jack B., and Joshua Klayman. 2004. "Overconfidence in Interval Estimates." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2): 299–314.

Sugden, Robert. 1993. "An Axiomatic Foundation for Regret Theory." *Journal of Economic Theory*, 60(1): 159–180.

Zeelenberg, Marcel. 1999. "Anticipated Regret, Expected Feedback and Behavioral Decision Making." *Journal of Behavioral Decision Making*, 12(2): 93–106.

Zeelenberg, Marcel, J. Jeffrey Inman, and Rik G.M. Pieters. 2001. "What we do when Decisions go Awry: Behavioral Consequences of Experienced Regret." In *Conflict and Tradeoffs in Decision Making*. , ed. Elke U. Weber, Jonathan Baron and Graham Loomes, 136–155. Cambridge University Press.